



The Influence of Vegetation Height Heterogeneity on Forest and Woodland Bird Species Richness across the United States

Qiongyu Huang^{1*}, Anu Swatantran¹, Ralph Dubayah¹, Scott J. Goetz²

¹ Department of Geographical Sciences, University of Maryland, College Park, Maryland, United States of America, ² Woods Hole Research Center, Falmouth, Massachusetts, United States of America

Abstract

Avian diversity is under increasing pressures. It is thus critical to understand the ecological variables that contribute to large scale spatial distribution of avian species diversity. Traditionally, studies have relied primarily on two-dimensional habitat structure to model broad scale species richness. Vegetation vertical structure is increasingly used at local scales. However, the spatial arrangement of vegetation height has never been taken into consideration. Our goal was to examine the efficacies of three-dimensional forest structure, particularly the spatial heterogeneity of vegetation height in improving avian richness models across forested ecoregions in the U.S. We developed novel habitat metrics to characterize the spatial arrangement of vegetation height using the National Biomass and Carbon Dataset for the year 2000 (NBCD). The height-structured metrics were compared with other habitat metrics for statistical association with richness of three forest breeding bird guilds across Breeding Bird Survey (BBS) routes: a broadly grouped woodland guild, and two forest breeding guilds with preferences for forest edge and for interior forest. Parametric and non-parametric models were built to examine the improvement of predictability. Height-structured metrics had the strongest associations with species richness, yielding improved predictive ability for the woodland guild richness models ($r^2 = \sim 0.53$ for the parametric models, 0.63 the non-parametric models) and the forest edge guild models ($r^2 = \sim 0.34$ for the parametric models, 0.47 the non-parametric models). All but one of the linear models incorporating height-structured metrics showed significantly higher adjusted- r^2 values than their counterparts without additional metrics. The interior forest guild richness showed a consistent low association with height-structured metrics. Our results suggest that height heterogeneity, beyond canopy height alone, supplements habitat characterization and richness models of forest bird species. The metrics and models derived in this study demonstrate practical examples of utilizing three-dimensional vegetation data for improved characterization of spatial patterns in species richness.

Citation: Huang Q, Swatantran A, Dubayah R, Goetz SJ (2014) The Influence of Vegetation Height Heterogeneity on Forest and Woodland Bird Species Richness across the United States. PLoS ONE 9(8): e103236. doi:10.1371/journal.pone.0103236

Editor: Ricardo Bomfim Machado, University of Brasilia, Brazil

Received: March 20, 2014; **Accepted:** June 28, 2014; **Published:** August 7, 2014

Copyright: © 2014 Huang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All BBS files are available from the FTP database (<ftp://ftpext.usgs.gov/pub/er/md/laurel/BBS/DataFiles/>) All NBCD2000 files are available from the online database (<http://www.whrc.org/mapping/nbcd/>).

Funding: The research received support from the NASA Applied Sciences programs on Ecological Forecasting and Biological Diversity (<http://www.nasa.gov/applied-sciences/eco-forecasting.html>) (grant # NNX09AK20G to SJG and RD). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: qhuang@umd.edu

Introduction

Avian diversity has been under increasing pressure from anthropogenic disturbances such as habitat loss and fragmentation [1]. Successful conservation planning relies upon understanding how the distribution of avian richness responds to existing and potential changes in environmental conditions which influence their distributions. Discovering the drivers of large-scale spatial variation of species richness has been a central debate in ecology [2–5], and many hypotheses have been proposed to address this issue [6–11]. One major hypothesis suggests that habitat heterogeneity is a key factor because it leads to greater spatial variability of habitat physical conditions, and therefore permits greater niche specialization resulting in more species richness [12–15]. Particularly in North America, habitat heterogeneity theory

predicted the richness of some faunas significantly better than the species-energy theory [7,14,15]. This latter theory also has widespread support, and hypothesizes that productive energy through food webs or species physiological constraints to ambient energy determines species richness [4,6,8,16].

Traditionally large scale habitat heterogeneity has been quantified mostly as topographical variability [7,14,17] or two dimensional habitat characteristics derived from remote sensing products [18,19]. Vertical habitat structure may also lead to niche generalization, and as such be an important element of habitat heterogeneity affecting biodiversity [20]. However, it has rarely been used to explain species richness at broad scales. The incorporation of vertical heterogeneity is especially important for avian richness models where vertical habitat structure at local

scales has long been recognized as a critical factor influencing bird life history [21–23] and abundance [24,25].

Until recently, there have been relatively few studies utilizing three-dimensional habitat information due to difficulties of acquiring measurements of vertical vegetation structure beyond the plot scale over extended geographical areas [20]. This has changed significantly since the emergence of active remote sensing systems such as Light Detection and Ranging (lidar) and Radio Detection and Ranging (radar) which provide capability to map the vertical dimension of vegetation at local to regional scales [20,26]. There is an increasing number of studies using lidar and radar derived three-dimensional vegetation structure to model biodiversity, many of which have revealed significant association between vegetation vertical structure, habitat quality, species richness and abundance [27–32]. However, none of the existing habitat metrics sufficiently characterize the spatial arrangement of vegetation height (i.e. the heterogeneity of height), nor its potential for predicting avian richness distributions over large geographical extents.

Related advances have been made in the development of statistical fusion models that provide a means to effectively combine remotely sensed data from radar, lidar, optical remote sensing systems and forest inventory data, yielding wall-to-wall high resolution vegetation structure maps at the continental scale [33–37]. The production of these maps not only enables the creation of habitat metrics that capture rich vegetation height heterogeneity, but also the comparison of the predictive abilities in various forms of these metrics. Our study is designed to embrace these opportunities by examining the relationship between forest bird richness, height-structured habitat metrics and avian richness models involving various degrees of forest height heterogeneity.

The overall goal of our study is to examine the potential of three-dimensional habitat structure in improving avian richness models at broad geographical scales. In doing so we hope to expand our understanding of the relationship between habitat structure and the spatial distribution of avian species richness, and to lay the foundation for constructing habitat metrics that better utilize increasingly available three-dimensional habitat data. Specifically we address the following questions:

1. How do the height-structured metrics compare with traditional habitat metrics in their ability to associate and predict forest bird richness in the forested ecoregion of the U.S.? Does incorporating the height-structured metrics improve the explanatory ability of avian richness models that use traditional habitat metrics?
2. How do the predictive abilities of richness models vary among forest bird guilds with contrasting preferences to habitat edges?

First, we introduce the conceptual similarities and differences between traditional habitat metrics and two types of height-structured metrics. Next, we describe the data and the methods we used to create the habitat metrics in this study. We then use correlation analysis and multivariate regression models to examine the relationships between different combinations of metrics and the species richness of three forest breeding bird guilds. Lastly, we examine the models' explanatory abilities and the importance of individual metrics in predicting the richness of the three guilds.

Background

Traditional habitat metrics are based primarily on two-dimensional habitat structure, such as land cover types, patch size and shape statistics. Developing such metrics generally depends on two steps: a) classifying scene space into binary habitat and non-

habitat land cover types; b) delineating habitat patches based on the rule of contiguity (Figure 1) [38]. There have been numerous studies using habitat patch metrics and derivative habitat edge and contrast metrics to associate with ecological attributes such as species richness, reproductive success and individual fitness of birds [39–41]. However vegetation height information generally plays little role in the process of delineating habitat patches and characterizing their properties.

Some studies have applied three-dimensional habitat information in habitat quality and species diversity models [28,30,31,42–44]. Usually, these applications rely on simple summary statistics such as mean, maximum, minimum and standard deviation to characterize three-dimensional vegetation structure. Summary statistics are straightforward and easy to obtain, but they cannot fully capture the heterogeneity of vegetation vertical structure. To give an example, one can have two forested landscapes with the same mean, maximum, minimum and standard deviation of tree height but with greatly different spatial arrangements of trees (e.g. tall trees can cluster in a few locations or can randomly distribute over the landscape which would have very different ecological implications for bird communities).

To account for more height heterogeneity, we created two groups of height-structured habitat metrics, the first of which integrates vegetation height information into the habitat patch framework while the second one characterizes canopy height distribution directly using second-order texture algorithms.

At the canopy level, vertical differences in vegetation create boundaries that segment contiguous habitats into smaller patches, each with similar height values (Figure 1). We first classified height pixels into a few height classes to characterize vertical edges and patches. Next, we grouped adjacent pixels from the same height class into patches. We treated the boundaries dividing those vertical patches as vertical edges (Figure 1). We also weighted the vertical edges by their depth (the height difference between two sides of a vertical edge) to capture the contrast of the height values of neighboring patches. By doing so, we could adapt a wide range of conventional habitat patch and edge metrics to account for complex spatial variability of canopy height.

Besides utilizing habitat patch and edge metrics to capture vegetation height heterogeneity, the second approach we introduce here involves calculation of the second-order (co-occurrence) texture statistics [45] directly from the gridded vegetation height maps. Second-order texture measures indicate the probabilities of each combination of pixel values co-occurring in a specific direction and distance [45]. These metrics can quantify spatial heterogeneity in terms of the spatial distribution and dependencies of height values [46] through grey level co-occurrence matrix. Texture measures are conventionally extracted from individual band of remotely sensed imagery and aerial photographs to assist object or land cover type discriminations [46,47]. Normally a small moving window is used to calculate the grey level co-occurrence matrix in specified neighborhoods. Texture measures extracted from optical remote sensing imageries have been used to infer broadly defined habitat heterogeneity that includes various environmental factors (e.g. land cover type, vegetation type, soil condition as well as vertical structure). This type of habitat structural information has been linked to avian species richness in many studies [48–50]. Here we derived the second-order texture metrics from gridded canopy height maps to directly characterize habitat height structure and to associate them with variation in avian richness.

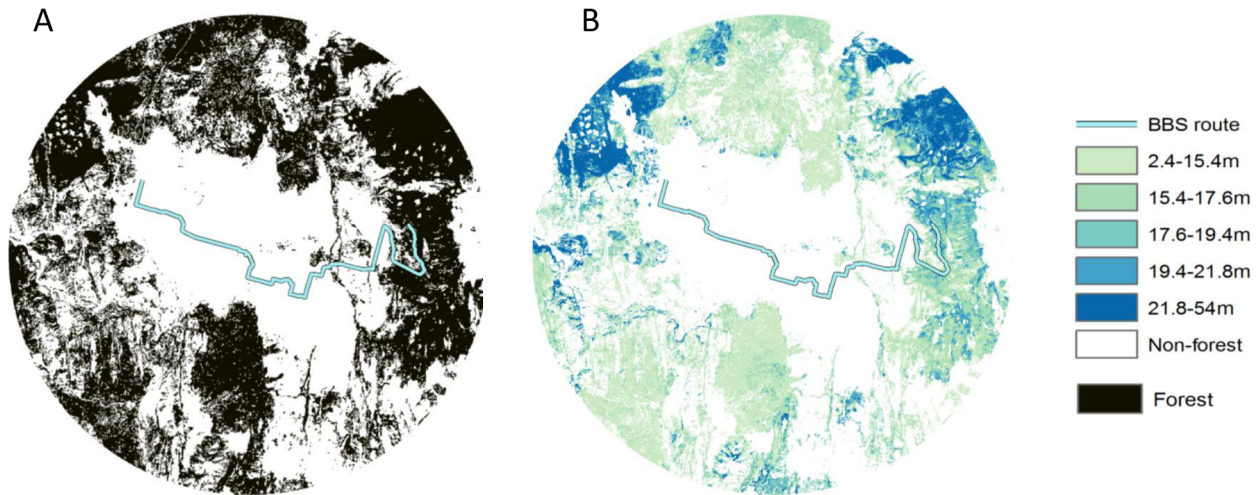


Figure 1. An example of the delineation of habitat patches at one BBS location. A two-dimensional vegetation map (A) and a vegetation map segmented by height structure (B) are shown. The pixel-based segmentation method (supporting information S1) is used to segment two dimensional habitat maps by using height thresholds.
doi:10.1371/journal.pone.0103236.g001

Data Sets and Methods

Avian Data

The study area includes 21 predominately forested ecological regions (provinces) [51] across the conterminous U.S. (Table S1) (Figure 2). We used avian records from the Breeding Bird Survey (BBS) to model species richness over the entire study range. BBS is an annual road side survey organized by U.S. Geological Survey (USGS) [40,52]. Initiated in 1966, BBS has over 4000 survey routes located on secondary roads across the continental U.S. and Canada. Each survey route is 39.4 km long. Every year, during the avian breeding season, surveys are conducted by competent volunteers using the protocol of three-minute point count at 50 stops at 0.8 km intervals. All birds seen or heard within 0.37 km radius are recorded [53]. We removed the records whose survey procedures or associated data are not acceptable by BBS standard. We also removed the records surveyed by first year observers to minimize observer bias [54]. We selected 134 broadly grouped woodland breeding birds species (here after “woodland guild”) based on the USGS species groupings [55]. We also selected 26 and 49 bird species as the forest breeding guilds with preference for interior forest habitat and forest edge habitat respectively (here after “interior forest guild” and “forest edge guild”) based on the classification of Boulinier et al.1998 [56]. A complete list of birds involved in this study and their guild assignment are given in Table S2. Because most of the interior forest and forest edge bird species are distributed in the Eastern U.S., we limited our analysis on these two guilds to the 10 forested ecoregions in the east (Figure 2).

Adjustments were made to take into account the detection probability bias [57]. We used the “fossil” package [58] in the R statistical program [59] to calculate the adjusted species richness using a first-order jackknife estimator [60,61]. This estimator is based on multiple recapture studies in closed populations, which allows detection probability to vary among species. It is also the basic estimator underlying the species richness adjustments used by a USGS-developed BBS pre-processing program called COMDYN [62]. We averaged the available first-order jackknife richness within the five year period between 1998 and 2002 to temporally approximate the acquisition time of the radar data

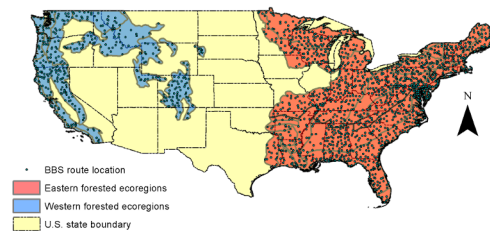


Figure 2. Distribution of BBS routes through the primarily forested ecoregions in the U.S. The richness models for the woodland guild were built using data from both eastern and western forested ecoregions. The forest edge and interior forest bird richness was modeled in the eastern forested ecoregions only.
doi:10.1371/journal.pone.0103236.g002

which played a key role in developing the vegetation height maps as discussed in the following section. The resulting mean avian richness is the richness we refer to in the rest of the study.

Forest height data and habitat metrics

The National Biomass and Carbon Database of the year 2000 (NBCD) [33] provides an estimate of vegetation height distribution and variation at fine resolution for the conterminous U.S. The dataset is based on combined information from U.S. Department of Agriculture’s Forest Service Forest Inventory and Analysis data, high-resolution Interferometric Synthetic Aperture Radar data acquired from 2000 Shuttle Radar Topography Mission and optical remote sensing data from the Landsat ETM+ sensor. Products from the USGS’ National Land Cover Dataset 2001 and the Landscape Fire and Resource Management Planning Tools Project were also used during the process as input to build the empirical model for tree height estimation. The basal area weighted tree height (hereafter, “tree height”) maps produced by the model gives spatially explicit vegetation vertical structure maps over the conterminous U.S. of 30 m-resolution.

We adapted a method to use 19 km (~half the length of a BBS route) radius buffers placed on the centroid of each BBS route, encompassing ~1100 km² areas to characterize the surrounding

habitat around BBS locations [31,63,64]. We created habitat metrics on 1751 such circular landscapes where there are available BBS species richness data. (A), (B), (C), and (D), four metric sets incorporating a total of 26 metrics were calculated for each landscape (Table 1). The methods to produce each set of metrics are described in more details in the supporting information S1. The first two metric sets, embedded with little to no vegetation height heterogeneity, included (A) summary height statistics (hereafter “summary statistics”) and (B) traditional patch-based metrics. The other two metric sets incorporated height heterogeneity: (C) patch metrics characterizing vertical patches and edges (hereafter, “height-structured patch-based metrics”), and (D) second-order texture metrics capturing vertical heterogeneity of height distributions (Table 1). The metric sets (A) and (B) were created as baselines to compare with the height-structured metric sets (C) and (D).

All the metrics created are listed in Table 1, and the detailed formula and descriptions for each metric are presented in Table S3. In order to differentiate the metrics with the same name from metrics set (B) and (C), capital letter “B” or “C” were given as prefixes to acronyms to indicate metric set membership (Table 1).

Species Richness Models

We first explored the statistical correlation between richness of the three avian guilds and the habitat metrics to evaluate the

association between individual habitat metrics and the richness of different guilds. The woodland species richness models were based on data of all 21 forested ecoregions, and the interior forest and forest edge guild models were limited to data of the 10 forested ecoregions from Eastern U.S. as noted earlier (Figure 2).

We selected 2 metrics from each of metric set (C) and (D) that on average had the best association with the richness of the three guilds as the best performing height-structured metrics (BPHMs). These four BPHMs were later combined with the traditional habitat metrics in multivariable models for comparisons of improvement. We limited our choice to only the four best metrics to avoid subsequent overfitting of our multivariate models while still maintaining enough representativeness.

We next constructed 6 multivariate linear models to explain each guild’s richness. The first 4 models were created using the complete list of metrics from set (A), (B), (C), and (D) respectively. They served to compare the explanatory abilities of models that characterize habitat condition with very different approaches. The two other models combined metric set (A) and (B) individually with the 4 BPHMs. We created the combined models to examine the impacts of adding spatial arrangement of height in richness models characterizing habitat in traditional ways.

We used a bootstrapping technique to provide the mean value and confidence intervals for the richness models’ adjusted- r^2 values and AIC values to assess models’ explanatory ability and goodness

Table 1. List of all metrics developed in the study.

Data	Metrics type (Set)	Metric name (Abbreviation)
NBCD vegetation height map	Summary height statistics (A)	Mean height (MEAN)
		Standard deviation of height (SD)
		Minimum height (Min)
		Maximum height (Max)
Two-dimensional vegetation cover map	Traditional patch-based metrics (B)	Number of patches(B.NP)
		Mean patch area (B.Area.MN)
		Standard deviation of patch area (B.Area.SD)
		Edge density (B. ED)
		Total edge (B.TE)
		Mean fractal dimension index (B.FRAC.MN)
Vegetation cover map segmented by height structure	Height-structured patch-based metrics (C)	Number of patches(C.NP)
		Mean patch area (C.Area.MN)
		Standard deviation of patch area (C.Area.SD)
		Total edge (C.TE)
		Mean fractal dimension index (C.FRAC.MN)
		Standard deviation of fractal dimension index (C.FRAC.SD)
		Contrast weighted edge density (C.CWED)
		Mean of edge contrast index(C.ECON.MN)
		Standard deviation of edge contrast index (C.ECON.SD)
		Shannon’s diversity index (C.SHDI)
NBCD vegetation height map	Second-order texture metrics (D)	Entropy
		Contrast
		Angular second moment (ASM)
		Homogeneity
		Dissimilarity

doi:10.1371/journal.pone.0103236.t001

of fit, as well as the variability of these measures. The bootstrap resampling was repeated 3000 times for each model. To examine the significance level of model improvements the 95% confidence interval of adjusted- r^2 values and AIC values were obtained with the bias-corrected and accelerated (BCA) bootstrap algorithm [65] to make the interval's median unbiased and adjusted for skewness.

Lastly we explored the effect of combining the 26 metrics from all four metric sets using a non-parametric Random Forest (RF) model. The RF model [66] is known for being able to handle large number of input variables without overfitting [67]. It is also well-suited for our study because the model allows for covariance between predictor variables, which commonly exists between different habitat metrics. The RF model also provides a mechanism for assessing predictor variable importance using a measure of cross-validated mean square error (out of bag mean square error (OOB MSE)). The higher the increase of OOB MSE (IncMSE) is, the more important a specific metric is. More detailed introduction of RF model is described in the supporting information S1. We also ran 6 RF models on the same combinations of metrics used by the linear models to compare the differences between linear and RF models. We set the number of trees to be 2000 for all models to allow for the mean residual error to converge. In our study the RF models were built with Random Forests package [68] in the R statistical program [59].

Results

Predictor metric correlation

The predictor metrics that correlate best with bird species richness varied among guilds. For woodland species richness, (D) the second order texture metrics generally had the greatest predictive ability, followed by (B) the traditional patch-based metrics, and (C) the height-structured patch metrics. For forest edge bird richness, the metrics with the strongest correlation were (C) the height-structured patch-based metrics followed by (D) the second order texture metrics and (A) the summary height statistics (Figure 3). Interior forest bird species richness in general had consistently low correlation with any metrics. Among all the metrics developed (Table 1, Table S3), the metric with the greatest predictive capability for interior guild richness was mean vegetation height. ASM had the strongest average predictive capability over the richness models for three guilds, followed by entropy, C.TE, homogeneity, and C.CWED (Figure 3), all of which are height-structured metrics. We selected ASM, entropy, C.TE and C.CWED as the four BPHMs to be combined with models relied on the traditional metrics.

The direction of the correlation between metrics and the bird richness was generally consistent across three guild types except for metrics with weak correlation (Table S4). Among the variables with highest average correlation, ASM and homogeneity both had negative correlation with the richness of all three guilds. Conversely, entropy, C.TE, C.CWED all showed strong positive correlation for each guild's richness (Figure 4, Table S4).

After incorporating vegetation height heterogeneity in patch-based metrics, the metrics characterizing patch number and area (AREA.MN, AREA.SD, and NP) showed a decreased correlation with the woodland guild richness. Conversely, the strength of the correlation between edge metrics (ED, TE) and the woodland guild richness increased. For the forest edge species both the patch and edge related metrics showed a prominent increase of correlation after incorporating vegetation height heterogeneity. The direction of the correlation for some patch-based metrics also changed. The NP metric showed an exceptionally large change for

the woodland guild richness: from -0.45 to 0.25 after incorporating vertical patches (Table S4, Figure 3).

Predictive models

The non-parametric RF models combining all 26 metrics (all-inclusive models) from the four metric sets were the ones with greatest ability to predict species richness for each guild (Figure 4, Table S5). Among those models the lowest species richness variability was explained for the interior forest guild ($r^2 = 0.11$), but the forest edge guild richness was predicted moderately well ($r^2 = 0.47$) and the predictive model was strong for the woodland guild ($r^2 = 0.63$) (Figure 5). The most important variable for predicting the woodland guild richness were two traditional patch-based metrics (B.AREA.MN and B.AREA.SD) followed by two second order texture metrics (entropy and ASM). The forest edge species richness model was most dependent on two height-structured patch metrics (C.CWED and C.NP) followed by two summary height statistics (MAX and MEAN). The most important predictive metrics for the interior forest guild model were MEAN followed by B.AREA.MN and B.AREA.SD (Figure 5).

For the RF models, our results consistently showed that adding height-structured metrics improved the model predictive ability. Specifically, the explained variance of the all-inclusive RF models for woodland and forest edge guild were up to 0.27 and 0.21 higher respectively than the RF models with only traditional habitat metrics. In addition for these two guilds, when the RF models were combined with the four BPHMs, the improvement for explained variance were up to 0.21 (woodland guild) and 0.13 (forest edge guild). The interior forest guild however showed only minor improvements when combined with any height-structured metrics. In general for woodland and forest edge guild, RF models' predictabilities were higher than the comparable linear models by a prominent margin. (Figure 4 and Table S5).

The linear models had a lower explanatory ability than their RF counterparts. For a specific combination of habitat metrics, the linear models explained the most amount of variation in the woodland guild richness and the least in the interior forest guild richness. The one exception was the model using summary statistics of height (set A), which showed the highest predictability for forest edge guild richness, followed by woodland guild richness, and then the interior forest guild richness (Figure 4, Table S5). In every guild, the models incorporating the four BPHMs showed consistently higher predictability than the models without (Figure 4).

Combining the BPHMs with the summary height statistics resulted in significantly higher adjusted- r^2 values in the woodland and forest edge models (Figure 4). The AIC value for the woodland richness model also improved significantly. In comparison, when combined with traditional patch-based metrics, the BPHMs significantly increased the adjusted- r^2 for the forest edge guild model, while significantly improving the AIC values for both the forest edge and woodland guild models (Table S5).

Discussion

A large number of hypotheses have been proposed to explain the spatial patterns of species richness over broad geographical scales [2,4,6,69–71]. While it is unlikely that one single mechanism can explain species richness patterns completely, a large portion of the literature testing habitat heterogeneity hypothesis has focused on the association between species richness and two dimensional habitat structure, often combined with land cover type composition and distribution [63,64,72,73]. On the other hand other studies testing species-energy hypothesis have relied on covariates

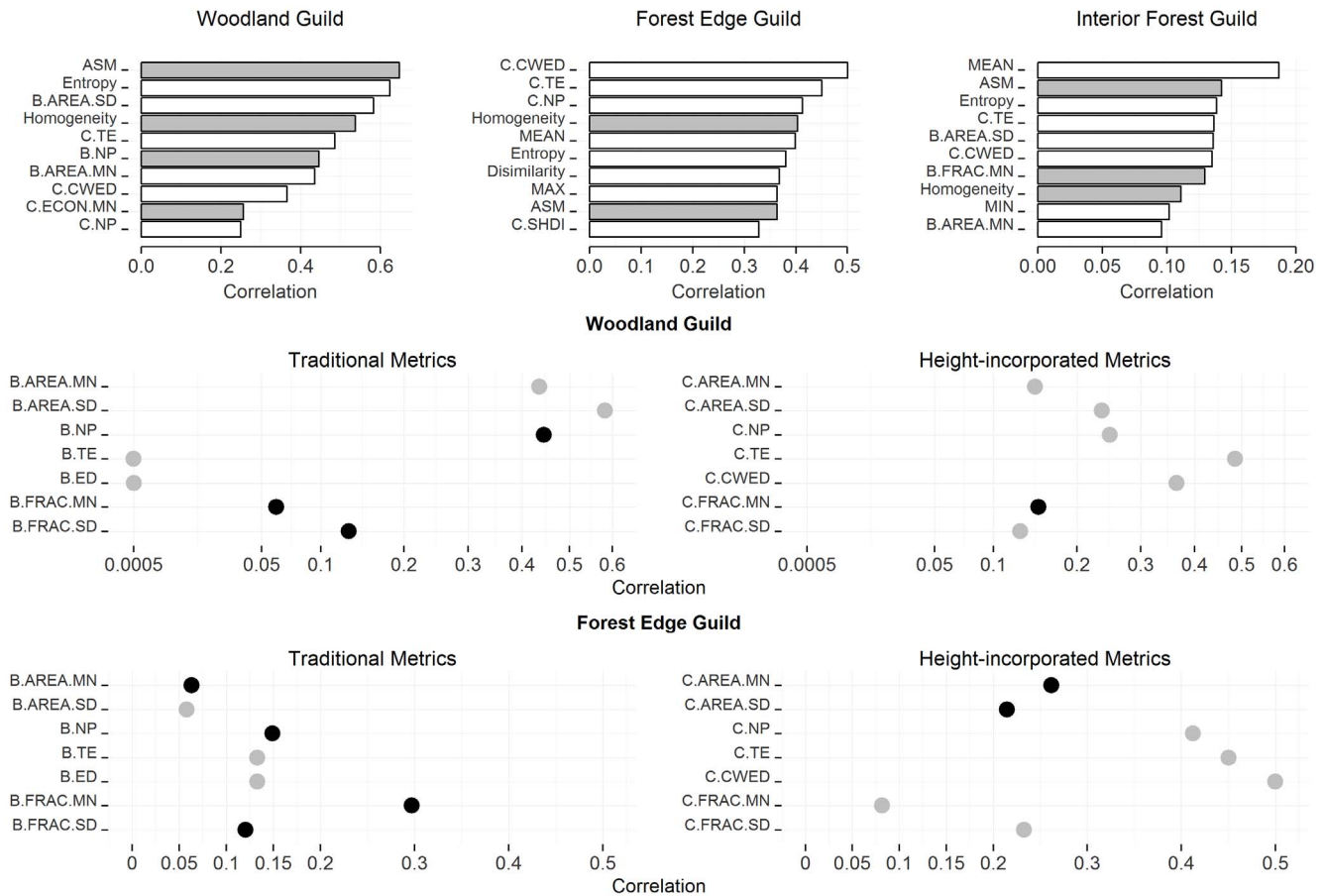


Figure 3. Guild richness associations with various metrics. (Top row): correlation bar plots of the most predictive metrics of species richness by guild. White bars represent a positive correlation and grey indicate a negative correlation. (Bottom rows): correlation comparisons between comparable patch-based metrics with and without considering the vertical patches and edges for the woodland and forest edge guild. The left panels show traditional metrics without accounting for height-heterogeneity; the right panels are height-structured counterparts. The black dots indicate a negative correlation and the grey ones indicate a positive correlation. doi:10.1371/journal.pone.0103236.g003

related to ecosystem productivity and energy such as evapotranspiration and photosynthetic capacity indices like the normalized difference vegetation index (NDVI) [74–76] to explain large scale species richness patterns. Studies to associate habitat vertical structure with species richness are, however, often focused at local scale [31], which limited the efficacies of habitat heterogeneity models to explain species richness at broad scale.

Only recently was vegetation height information assessed as a predictor of avian species richness across the conterminous U.S. in two studies [31,77]. One of these [31] used the same NBCD data we employed here, but they explored only summary statistics of vegetation height and biomass combined with land cover type composition and distribution. The other used sparsely sampled height metrics from a satellite lidar system that is no longer operating, and included climatic data as predictive variables [77]. Although our models employed only the distribution vegetation structures, with no input from other land cover type or climatic data, their explanatory ability for the woodland guild was comparable to these recent results ($r^2 = 0.70$ for the forest guild model [31], and $r^2 = 0.60$ for the open woodland model [77]). We found that models combining only vegetation vertical and horizontal structure can explain a significant amount of species richness for the broadly grouped woodland guild and the forest breeding guild with preferences for the forest edge habitat. More

importantly, our results showed that incorporating vegetation vertical heterogeneity, and not just mean and standard deviation of height, greatly improves the ability to explain variability in avian richness for the two guilds. The spatial arrangement of vegetation height plays an important role in associating the quality of habitat condition and diversity of ecological niches for bird species within the two groups.

Traditionally habitat edges are thought to affect species movement, interaction, mortality and community dynamics [78]. The summary height statistics are considered indicators of habitat diversity and forest successional stage [20,79,80]. The traditional way of characterizing habitat through two-dimensional habitat patch distribution and summary height statistics still play important roles in our multivariate richness models. The large pool of traditional patch-based metrics provides a well-known framework to readily incorporate vertical height distribution once habitat patches are segmented by height. Both traditional and our height-structured metrics contribute to explanation of the variance of avian richness, although the importance of individual metrics in the models varies from guild to guild. Thus, our study shows that for the woodland avian guild and forest edge guild, the species richness is highly sensitive to the vegetation height heterogeneity, and the addition of the spatial arrangement of vegetation height provides significantly improved estimates of species richness for the

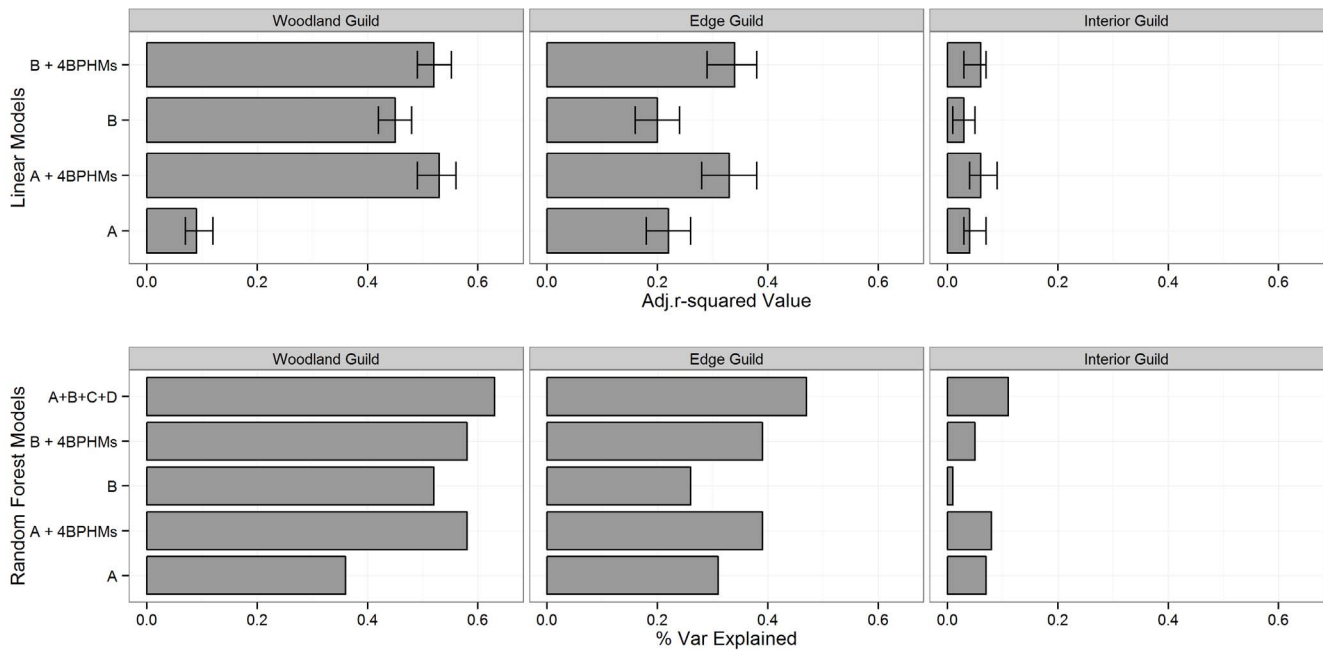


Figure 4. Predictive ability of multivariable models. A, B, C, and D are the four habitat metric sets, and 4BPHMs are the four best predictive height-structured metrics. Each of the top panels shows four linear models with whiskers giving 95% confidence interval of adjusted- r^2 values. The length of the bar represents the mean adjusted- r^2 for these models. The lower panels show the explained variance of the comparable random forest (RF) models. Uniquely the top bars at lower panels are the results from the models employed all metrics from the four metric sets. doi:10.1371/journal.pone.0103236.g004

two guilds. The patch-based height-structured metrics and the second order texture metrics thereby supplement and extend common methods of characterizing habitat condition and predicting avian species richness.

We note that the BBS data is collected along roadways where volunteers can easily and regularly traverse, thus the areas along the survey routes could be subject to disturbances such as motor vehicle traffic or habitat conversion [81,82]; i.e. they may not be representative samplings of forest spatial and vertical variability. This characteristic of the data set could pose a challenge for systematic sampling of interior forest bird species in the surrounding areas and is likely one of the contributing factors for the consistently low species richness and weak correlations with our metrics and models in the case of the interior forest guild. Alternatively, forest edge habitats are relatively more exposed to stressors such as wind damages and human disturbances. They can exhibit higher vertical structure diversity than the interior forest areas [83]. It may be that interior birds are less adapted to habitat structure heterogeneity, and thus exhibit limited sensibility to habitat structure metrics. Lastly the results could also be attributed to the different ways members of avian guilds utilize habitat. Forest edge and majority of woodland bird species tend to use a wide range of habitat, and their degree of co-existence can vary in a broad spectrum over space. In comparison the interior forest guild, composed mostly of forest specialists that avoid other habitat types [84] with overlapping ecological niches, are more likely to face greater interspecific competition which limits species richness despite diverse height structure across landscapes [85]. However while the models see low association between height heterogeneity metrics and interior forest guild richness, there are still likely more specific vertical structure preferences associated with individual bird species [30,86].

While four BPHMs highlighted in our study showed a good ability to associate with species richness and to improve broad

scale avian richness modeling, it is reasonable to assume that height-structured metrics have potential to be improved further given the large number of options that remained unexplored. First, the pixel-based segmentation method used in our study (supporting information S1) is one of the simplest algorithms to delineate vertical patches and edges. The method is based on a set of global threshold values while not considering neighboring heterogeneity [87]. The process of setting up the threshold values and weight matrix (for contrast metrics) inevitably involves somewhat arbitrary decisions. More complex segmentation methods such as edge and region-based methods can be performed readily with commercial and open-source software packages that potentially may produce more efficacious vertical patches and be less arbitrary [88]. Secondly, there are many untested texture measures [45]. The relationship between texture metrics and the avian richness varies as the size of moving window changes [48]. More work is needed to understand the impact of those methodological options for further improving species richness models.

Conclusions

As active remote sensing technologies like radar and lidar mature and become more widely available, data sets characterizing vegetation vertical structure should become increasingly useful for biodiversity applications and management. Our study showed that vegetation height heterogeneity is associated with habitat diversity and species richness for some forest avian guilds. Thus, while recognizing the advances conveyed by incorporating height information, there is an imperative to explore in more depth the role of such heterogeneity. Furthermore we suggest not just height, but vertical canopy heterogeneity, e.g. foliar profiles and layering, will provide an even richer source of information from which to develop new metrics and models [30,83].

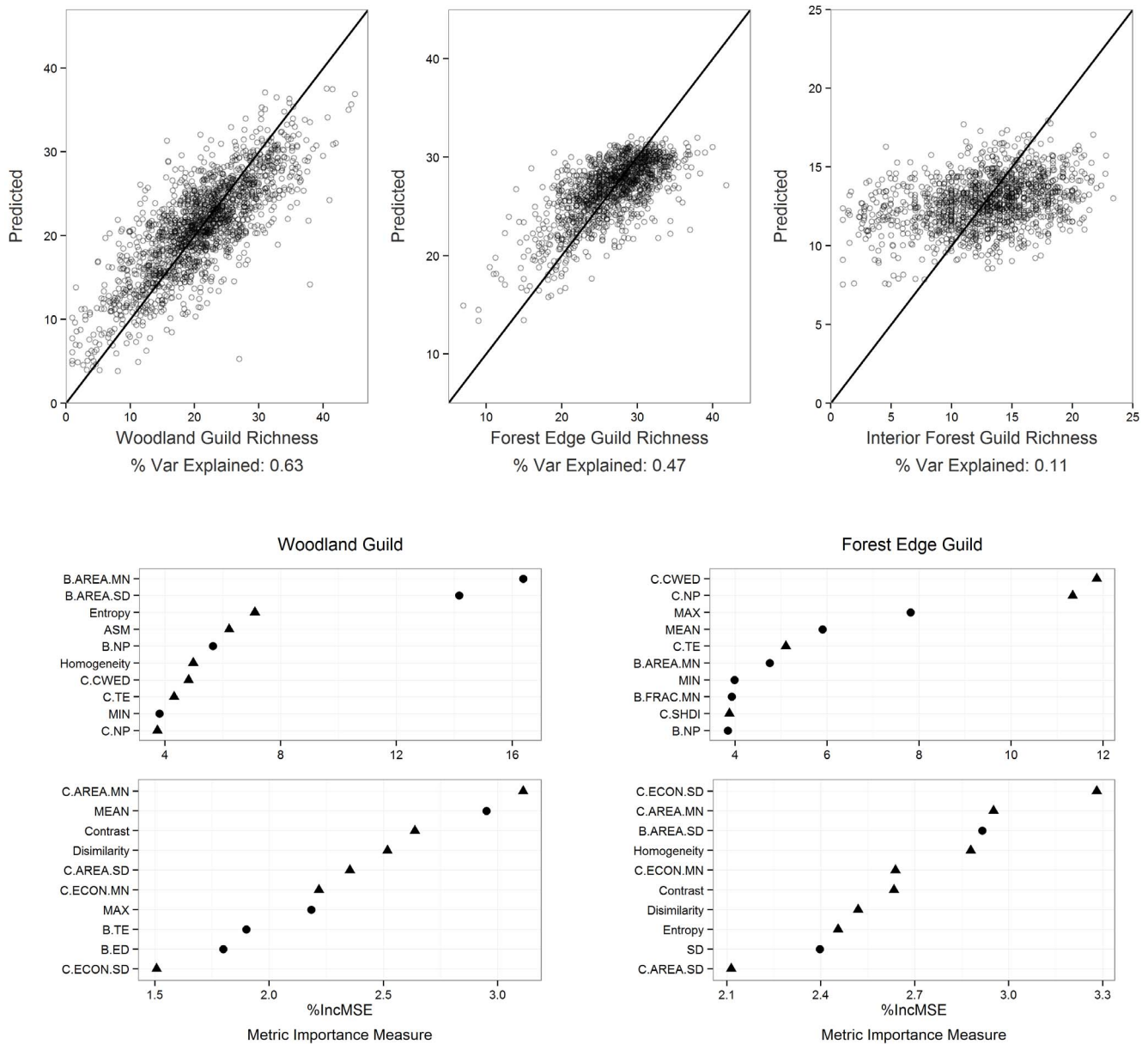


Figure 5. Random Forest model results. (Top row): Modeled vs. actual species richness for three guilds using all-inclusive random forest models. (Below the scatter plots): variable importance plots show the percent increase in mean square error (%IncMSE) of the top 20 most influential metrics in the woodland guild richness model and the forest edge guild richness model (note different scales on X-axes). The metrics characterizing vegetation height heterogeneity are plotted with triangles and the rest of the metrics are circles.
doi:10.1371/journal.pone.0103236.g005

Incorporating such information will require data on not only canopy height but canopy vertical structure, the latter of which is unavailable at continental scales. Nonetheless, the metrics and models used in our analyses provide a means to incorporate and utilize three-dimensional habitat information, with the goal of better understanding the controls on avian species richness and habitat use.

Supporting Information

Table S1 Ecological provinces involved and the sub-region assignment. (XLSX)

Table S2 Species list and guild classification. (XLSX)

Table S3 Metric descriptions. (DOCX)

Table S4 Metrics correlation with species richness. (XLSX)

Table S5 Model performances for different multivariable models. (DOCX)

Supporting Information S1 Contrast matrix for weighted-edge metrics. (PDF)

Acknowledgments

We thank Dr. John Sauer and Dr. James Kellner for valuable discussions on the project, as well as the academic editor and the peer reviewers for their input in helping to improve the manuscript. Lastly we also thank the BBS staff and volunteers for their effort in conducting the surveys and making their data widely available.

References

- Gaston KJ, Blackburn TM, Goldewijk KK (2003) Habitat conversion and global avian biodiversity loss. *Proc R Soc Lond B Biol Sci* 270: 1293–1300. doi:10.1098/rspb.2002.2303.
- Palmer MW (1994) Variation in species richness: Towards a unification of hypotheses. *Folia Geobot Phytotaxon* 29: 511–530. doi:10.1007/BF02883148.
- Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge; New York: Cambridge University Press. 460 p.
- Gaston KJ (2000) Global patterns in biodiversity. *Nature* 405: 220–227. doi:10.1038/35012228.
- Gaston KJ, Spicer JI (2004) *Biodiversity: An Introduction*. 2 edition. Malden, MA: Wiley-Blackwell. 208 p.
- Waide RB, Willig MR, Steiner CF, Mittelbach G, Gough L, et al. (1999) The Relationship Between Productivity and Species Richness. *Annu Rev Ecol Syst* 30: 257–300. doi:10.1146/annurev.ecolsys.30.1.257.
- Rahbek C, Graves GR (2001) Multiscale assessment of patterns of avian species richness. *Proc Natl Acad Sci* 98: 4534–4539. doi:10.1073/pnas.071034898.
- Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, et al. (2003) Energy, Water, and Broad-Scale Geographic Patterns of Species Richness. *Ecology* 84: 3105–3117. doi:10.1890/03-8006.
- Willig MR, Kaufman DM, Stevens RD (2003) Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annu Rev Ecol Syst* 34: 273–309.
- Currie DJ, Mittelbach GG, Cornell HV, Field R, Guégan JF, et al. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol Lett* 7: 1121–1134. doi:10.1111/j.1461-0248.2004.00671.x.
- Colwell RK, Rahbek C, Gotelli NJ (2004) The mid-domain effect and species richness patterns: what have we learned so far? *Am Nat* 163: E1–23. doi:10.1086/382056.
- Davies RG, Orme CDL, Storch D, Olson VA, Thomas GH, et al. (2007) Topography, energy and the global distribution of bird species richness. *Proc R Soc B Biol Sci* 274: 1189–1197. doi:10.1098/rspb.2006.0061.
- Koh CN, Lee PF, Lin RS (2006) Bird species richness patterns of northern Taiwan: primary productivity, human population density, and habitat heterogeneity. *Divers Distrib* 12: 546–554. doi:10.1111/j.1366-9516.2006.00238.x.
- Kerr JT, Packer L (1997) Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature* 385: 252–254.
- Kerr JT, Southwood TRE, Cihlar J (2001) Remotely sensed habitat diversity predicts butterfly species richness and community similarity in Canada. *Proc Natl Acad Sci* 98: 11365–11370. doi:10.1073/pnas.201398398.
- Evans KL, Warren PH, Gaston KJ (2005) Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biol Rev Camb Philos Soc* 80: 1–25.
- Richerson PJ, Lum K (1980) Patterns of plant species diversity in California: relation to weather and topography. *Am Nat*: 504–536.
- Turner JRG, Lennon JJ, Lawrenson JA (1988) British bird species distributions and the energy theory. *Nature* 335: 539–541. doi:10.1038/335539a0.
- Duro DC, Coops NC, Wulder MA, Han T (2007) Development of a large area biodiversity monitoring system driven by remote sensing. *Prog Phys Geogr* 31: 235–260. doi:10.1177/0309133307079054.
- Bergen KM, Goetz SJ, Dubayah RO, Henebry GM, Hunsaker CT, et al. (2009) Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *J Geophys Res* 114:GE00E06. doi:10.1029/2008JG000883.
- Halaj J, Ross D, Moldenke A (2000) Importance of habitat structure to the arthropod food-web in Douglas-fir canopies. *Oikos* 90: 139–152. doi:10.1034/j.1600-0706.2000.900114.x.
- Kelly J (1993) The Effect of Nest Predation on Habitat Selection by Dusky Flycatchers. *Condor* 95: 83–93. doi:10.2307/1369389.
- Robinson SK, Holmes RT (1984) Effects of plant species and foliage structure on the foraging behavior of forest birds. *The Auk* 101: 672–684.
- MacArthur RH, MacArthur JW (1961) On Bird Species Diversity. *Ecology* 42: 594–598. doi:10.2307/1932254.
- Whittaker RJ, Willis KJ, Field R (2001) Scale and Species Richness: Towards a General, Hierarchical Theory of Species Diversity. *J Biogeogr* 28: 453–470.
- Lefsky MA, Cohen WB, Parker GG, Harding DJ (2002) Lidar Remote Sensing for Ecosystem Studies. *BioScience* 52: 19. doi:10.1641/0006-3568(2002)052[0019:LRSEFS]2.0.CO;2.
- Bergen KM, Gilboy AM, Brown DG (2007) Multi-dimensional vegetation structure in modeling avian habitat. *Ecol Inform* 2: 9–22. doi:10.1016/j.ecoinf.2007.01.001.

Author Contributions

Conceived and designed the experiments: QH RD AS. Performed the experiments: QH. Analyzed the data: QH. Contributed to the writing of the manuscript: QH AS RD SG.

- Goetz S, Steinberg D, Dubayah R, Blair B (2007) Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sens Environ* 108: 254–263. doi:10.1016/j.rse.2006.11.016.
- Imhoff ML, Sisk TD, Milne A, Morgan G, Orr T (1997) Remotely sensed indicators of habitat heterogeneity: Use of synthetic aperture radar in mapping vegetation structure and bird habitat. *Remote Sens Environ* 60: 217–227. doi:10.1016/S0034-4257(96)00116-2.
- Swatantran A, Dubayah R, Goetz S, Hofton M, Betts MG, et al. (2012) Mapping Migratory Bird Prevalence Using Remote Sensing Data Fusion. *PLoS ONE* 7: e28922. doi:10.1371/journal.pone.0028922.
- Culbert PD, Radeloff VC, Flather CH, Kellndorfer JM, Rittenhouse CD, et al. (2013) The Influence of Vertical and Horizontal Habitat Structure on Nationwide Patterns of Avian Biodiversity. *The Auk* 130: 656–665. doi:10.1525/auk.2013.130007.
- Zellweger F, Braunisch V, Baltensweiler A, Bollmann K (2013) Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. *For Ecol Manag* 307: 303–312. doi:10.1016/j.foreco.2013.07.023.
- Kellndorfer J, Walker W, Kirsch K, Fiske G, Bishop J, et al. (2011) National Biomass and Carbon Dataset for the Year 2000. Available: <http://www.whrc.org/mapping/nbcd>.
- Kellndorfer JM, Walker WS, LaPoint E, Kirsch K, Bishop J, et al. (2010) Statistical fusion of lidar, InSAR, and optical remote sensing data for forest stand height characterization: A regional-scale method based on LVIS, SRTM, Landsat ETM+, and ancillary data sets. *J Geophys Res* 115. doi:10.1029/2009JG000997.
- Saatchi SS, Harris NL, Brown S, Lefsky M, Mitchard ETA, et al. (2011) Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc Natl Acad Sci* 108: 9899–9904. doi:10.1073/pnas.1019576108.
- Walker WS, Kellndorfer JM, LaPoint E, Hoppus M, Westfall J (2007) An empirical InSAR-optical fusion approach to mapping vegetation canopy height. *Remote Sens Environ* 109: 482–499. doi:10.1016/j.rse.2007.02.001.
- Baccini A, Goetz SJ, Walker WS, Laporte NT, Sun M, et al. (2012) Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat Clim Change* 2: 182–185. doi:10.1038/nclimate1354.
- Girvetz EH, Greco SE (2007) How to define a patch: a spatial model for hierarchically delineating organism-specific habitat patches. *Landsc Ecol* 22: 1131–1142. doi:10.1007/s10980-007-9104-8.
- Helzer CJ, Jelinski DE (1999) The Relative Importance of Patch Area and Perimeter-Area Ratio to Grassland Breeding Birds. *Ecol Appl* 9: 1448–1458. doi:10.1890/1051-0761(1999)009[1448:TRIOPA]2.0.CO;2.
- Robbins CS, Dawson DK, Dowell BA (1989) *Habitat Area Requirements of Breeding Forest Birds of the Middle Atlantic States*. Wildl Monogr: 3–34.
- Strelke WK, Dickson JG (1980) Effect of Forest Clear-Cut Edge on Breeding Birds in East Texas. *J Wildl Manag* 44: 559–567. doi:10.2307/3808003.
- Broughton RK, Hinsley SA, Bellamy PE, Hill RA, Rothery P (2006) Marsh Tit *Poecile palustris* territories in a British broad-leaved wood. *Ibis* 148: 744–752. doi:10.1111/j.1474-919X.2006.00583.x.
- Hill RA, Hinsley SA, Gaveau DL, Bellamy PE (2004) Predicting habitat quality for Great Tits (*Parus major*) with airborne laser scanning data. *Int J Remote Sens* 25: 4851–4855. doi:10.1080/0143116031000139962.
- Hinsley SA, Hill RA, Fuller RJ, Bellamy PE, Rothery P (2009) Bird species distributions across woodland canopy structure gradients. *Community Ecol* 10: 99–110. doi:10.1556/ComEc.10.2009.1.12.
- Haralick RM, Shanmugam K, Dinstein I (1973) Textural Features for Image Classification. *Syst Man Cybern IEEE Trans On* 3: 610–621. doi:10.1109/TSMC.1973.4309314.
- Coburn CA, Roberts ACB (2004) A multiscale texture analysis procedure for improved forest stand classification. *Int J Remote Sens* 25: 4287–4308. doi:10.1080/0143116042000192367.
- Franklin SE, Hall RJ, Moskal LM, Maudie AJ, Lavigne MB (2000) Incorporating texture into classification of forest species composition from airborne multispectral images. *Int J Remote Sens* 21: 61–79. doi:10.1080/014311600210993.
- St-Louis V, Pidgeon AM, Radeloff VC, Hawbaker TJ, Clayton MK (2006) High-resolution image texture as a predictor of bird species richness. *Remote Sens Environ* 105: 299–312. doi:10.1016/j.rse.2006.07.003.
- St-Louis V, Pidgeon AM, Clayton MK, Locke BA, Bash D, et al. (2009) Satellite image texture and a vegetation index predict avian biodiversity in the Chihuahuan Desert of New Mexico. *Ecography* 32: 468–480. doi:10.1111/j.1600-0587.2008.05512.x.
- Culbert PD, Radeloff VC, St-Louis V, Flather CH, Rittenhouse CD, et al. (2012) Modeling broad-scale patterns of avian species richness across the

- Midwestern United States with measures of satellite image texture. *Remote Sens Environ* 118: 140–150. doi:10.1016/j.rse.2011.11.004.
51. Bailey RG (1995) Description of the Ecoregions of the United States: Miscellaneous Publication.
 52. Robbins CS, Bystrak D, Geissler PH (1986) The breeding bird survey: its first fifteen years, 1965–1979. U.S. Dept. of the Interior, Fish and Wildlife Service. 198 p.
 53. Sauer JR, Hines JE, Fallon JE, Pardieck KL, Ziolkowski DJ Jr, et al. (2011) The North American Breeding Bird Survey, Results and Analysis 1966–2010. USGS Patuxent Wildlife Research Center. Available: <http://www.mbr-pwrc.usgs.gov/bbs/>.
 54. Kendall WL, Peterjohn BG, Sauer JR (1996) First-Time Observer Effects in the North American Breeding Bird Survey. *The Auk* 113: 823–829.
 55. U.S. Geological Survey (2012) List of Species Groupings. Available: <http://www.mbr-pwrc.usgs.gov/bbs/guild/guildst.html>.
 56. Boulmier T, Nichols JD, Hines JE, Sauer JR, Flather CH, et al. (1998) Higher temporal variability of forest breeding bird communities in fragmented landscapes. *Proc Natl Acad Sci U S A* 95: 7497–7501.
 57. Kéry M, Schmid H (2004) Monitoring programs need to take into account imperfect species detectability. *Basic Appl Ecol* 5: 65–73. doi:10.1078/1439-1791-00194.
 58. Vavrek MJ (2011) fossil: palaeoecological and palaeogeographical analysis tools. *Palaeontol Electron* 14: 1T.
 59. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available: <http://www.R-project.org>.
 60. Burnham KP, Overton WS (1978) Estimation of the Size of a Closed Population When Capture Probabilities Vary Among Animals. *Biometrika* 65: 625–633. doi:10.1093/biomet/65.3.625.
 61. Burnham KP, Overton WS (1979) Robust Estimation of Population Size When Capture Probabilities Vary Among Animals. *Ecology* 60: 927–936. doi:10.2307/1936861.
 62. Hines JE, Boulmier T, Nichols JD, Sauer JR, Pollock KH (1999) COMDYN: software to study the dynamics of animal communities using a capture-recapture approach. *Bird Study* 46: 209–217. doi:10.1080/00063659909477247.
 63. Rittenhouse CD, Pidgeon AM, Albright TP, Culbert PD, Clayton MK, et al. (2012) Land-cover change and avian diversity in the conterminous United States. *Conserv Biol J Soc Conserv Biol* 26: 821–829. doi:10.1111/j.1523-1739.2012.01867.x.
 64. Pidgeon AM, Radeloff VC, Flather CH, Lepczyk CA, Clayton MK, et al. (2007) Associations of Forest Bird Species Richness with Housing and Landscape Patterns across the USA. *Ecol Appl* 17: 1989–2010.
 65. Efron B (1987) Better Bootstrap Confidence Intervals. *J Am Stat Assoc* 82: 171–185. doi:10.1080/01621459.1987.10478410.
 66. Breiman L (2001) Random Forests. *Mach Learn* 45: 5–23.
 67. Biau G (2012) Analysis of a Random Forests Model. *J Mach Learn Res* 98888: 1063–1095.
 68. Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2: 18–22.
 69. Hawkins BA, Diniz-Filho JAF, Jaramillo CA, Soeller SA (2007) Climate, niche conservatism, and the global bird diversity gradient. *Am Nat* 170: S16–S27.
 70. Rahbek C, Gotelli NJ, Colwell RK, Entsminger GL, Rangel TFLVB, et al. (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proc R Soc B Biol Sci* 274: 165–174. doi:10.1098/rspb.2006.3700.
 71. Guégan JF, Lek S, Oberdorff T (1998) Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* 391: 382–384. doi:10.1038/34899.
 72. Donovan TM, Flather CH (2002) Relationships among North American Songbird Trends, Habitat Fragmentation, and Landscape Occupancy. *Ecol Appl* 12: 364–374. doi:10.2307/3060948.
 73. Griffiths GH, Lee J (2000) Landscape pattern and species richness: regional scale analysis from remote sensing. *Int J Remote Sens* 21: 2685–2704. doi:10.1080/01431160050110232.
 74. Phillips LB, Hansen AJ, Flather CH (2008) Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production. *Remote Sens Environ* 112: 4381–4392. doi:10.1016/j.rse.2008.08.002.
 75. Hurlbert AH, Haskell JP (2003) The Effect of Energy and Seasonality on Avian Species Richness and Community Composition. *Am Nat* 161: 83–97. doi:10.1086/345459.
 76. Seto KC, Fleishman E, Fay JP, Betrus CJ (2004) Linking spatial patterns of bird and butterfly species richness with Landsat TM derived NDVI. *Int J Remote Sens* 25: 4309–4324. doi:10.1080/0143116042000192358.
 77. Goetz SJ, Sun M, Zolcos S, Hansen A, Dubayah R (2014) The relative importance of climate and vegetation properties on patterns of North American breeding bird species richness. *Environ Res Lett* 9: 034013. doi:10.1088/1748-9326/9/3/034013.
 78. Fagan WE, Cantrell RS, Cosner C (1999) How habitat edges change species interactions. *Am Nat* 153: 165–182. doi:10.1086/303162.
 79. Morgan K, Freedman B (1985) Breeding Bird Communities in a Hardwood Forest Succession in Nova Scotia Canada. *Can Field Nat* 100: 506–519.
 80. North MP, Franklin JF, Carey AB, Forsman ED, Hamer T (1999) Forest Stand Structure of the Northern Spotted Owl's Foraging Habitat. *For Sci* 45: 520–527.
 81. Griffith EH, Sauer JR, Royle JA (2010) Traffic Effects on Bird Counts on North American Breeding Bird Survey Routes. *The Auk* 127: 387–393. doi:10.1525/auk.2009.09056.
 82. Keller CME, Scallan JT (1999) Potential Roadside Biases Due to Habitat Changes along Breeding Bird Survey Routes. *The Condor* 101: 50–57. doi:10.2307/1370445.
 83. Whitehurst AS, Swatantran A, Blair JB, Hofton MA, Dubayah R (2013) Characterization of Canopy Layering in Forested Ecosystems Using Full Waveform Lidar. *Remote Sens* 5: 2014–2036. doi:10.3390/rs5042014.
 84. Hagan JM, Vander Haegen WM, McKinley PS (1996) The Early Development of Forest Fragmentation Effects on Birds. *Conserv Biol* 10: 188–202. doi:10.1046/j.1523-1739.1996.10010188.x.
 85. Cody ML (1974) Competition and the Structure of Bird Communities. Princeton University Press.
 86. Goetz SJ, Steinberg D, Betts MG, Holmes RT, Doran PJ, et al. (2010) Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology* 91: 1569–1576.
 87. Schiewe J (2002) Segmentation of high-resolution remotely sensed data – concepts, applications and problems. Symposium on Geospatial theory, Processing and Applications.
 88. Baatz M, Benz U, Dehghani S, Heynen M, Höltje A, et al. (2003) eCognition user guide. Defin Imaging GmbH.