Remote Sensing of Environment 119 (2012) 35–42

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Performance of habitat suitability models for the endangered black-capped vireo built with remotely-sensed data

Chad B. Wilsey [a,*], Joshua J. Lawler [a], David A. Cimprich [b]

[a] School of Forest Resources, University of Washington, Seattle, WA, 98195-2100, United States
[b] Fort Hood Military Reservation, Bldg 1939, Rod & Gun Club Loop @ 53rd, Killeen, TX 76544-5028, United States

ARTICLE INFO

Article history:
Received 19 June 2011
Received in revised form 29 October 2011
Accepted 3 December 2011
Available online 14 January 2012

Keywords:
Vegetation
Aerial imagery
Aerial altimetry
LiDAR
cforest
Variable importance
Black-capped vireo

ABSTRACT

Identifying suitable habitat is critical to endangered species management and recovery. However, this basic task is often complicated by the rarity of the species in question and the limited availability of high quality environmental data. The endangered Black-capped Vireo (*Vireo atricapilla*) occupies early seral shrub communities generated by fire. The ability to predict vegetation structure is particularly important for mapping vireo habitat because this species occupies transitional, non-equilibrium vegetation types. We use presence data for territorial male vireos collected throughout the Fort Hood Military Reservation, Texas, to construct habitat suitability models using vegetation type (mapped from aerial imagery), soil data, and laser altimetry (LiDAR)-derived measures of vegetation structure. LiDAR produces a three-dimensional, high-resolution representation of vegetation structure across broad spatial scales. We built models that incorporated LIDAR outputs as well as the other habitat predictors using a non-parametric machine-learning algorithm, cforest. Models built using a single predictor class (vegetation structure or type or soil) performed similarly across 25 bootstrapped training and test datasets (median accuracies 76%, 74%, and 79%, respectively). Models incorporating two predictor classes performed better (80–81%) and only slightly worse than the full model (82%). Furthermore, vegetation type and soil data were more important predictors of habitat suitability than structural measures in the full model. Predictive maps suggest that models incorporating vegetation type and soil would be most useful for habitat conservation and management applications on Fort Hood. The addition of LiDAR-derived variables to the model further distinguishes suitable habitats from potentially suitable but presently overgrown areas. In the absence of detailed vegetation data, models based on structural measures performed well when combined with soil data. This could be useful in a broader regional context in which vireos occupy a greater variety of vegetation types with a common structure.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Vegetation structure (i.e. physiognomy) and species composition are important determinants of habitat suitability for birds (Hutto, 1985; James, 1971; Rotenberry, 1985) as well as predictors of overall avian diversity (MacArthur & MacArthur, 1961). Field-based sampling methods for quantifying vegetation structure and composition are widely used (Barber & Martin, 1997), but may be inadequate for characterizing vegetation across broad spatial extents. Vegetation plot surveys may be combined with air photo interpretation to scale-up results into thematic maps of landscapes (Bergen et al., 2009; Newton et al., 2009). These high-resolution vegetation maps provide a detailed snapshot of vegetation composition and configuration, but may require months to years in fieldwork and image classification. Now widely available, remote-sensing methods such as laser altimetry, also known as LiDAR (light detection and ranging), allow for continuous estimates of vegetation structure (e.g., canopy height and percent cover) across large areas (Lefsky et al., 2002; Vierling et al., 2008). LiDAR datasets can be collected in days or weeks and processed in months, making them attractive for many conservation and management applications. If physiognomy is equivalent or superior to vegetation composition as a predictor of habitat suitability, then rapid habitat assessments using LiDAR-derived datasets may be preferable to field-based methods.

LiDAR is an active remote sensing technology for mapping three dimensional surfaces. A LiDAR device emits laser pulses at regular intervals as it is flown above the target surface. The LiDAR sensor then records the time it takes for the pulses to return. Recording of multiple returns and time delays from millions of laser pulses gives an estimate of surface complexity in three dimensions. Sensors can either record discrete returns or large-footprint waveforms (Lefsky et al., 2002). Both are collected by sensors affixed to an aircraft and can be used for high resolution (0.25–5 m) mapping of terrain and vegetation (Bergen et al., 2009; Lefsky et al., 2002). Terrestrial-based

* Corresponding author. Tel.: +1 206 543 5772.
E-mail addresses: cbwilsey@uw.edu (C.B. Wilsey), jlawler@uw.edu (J.J. Lawler), david.a.cimprich2.civ@mail.mil (D.A. Cimprich).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2011.12.002

LiDAR, in which data are collected from a fixed location, is used in forestry applications and may become appropriate for detailed assessment of micro-scale vegetation structure in the future (Michel et al., 2008).

The potential applications of LiDAR-derived datasets to modeling species-habitat relationships are broad and currently underdeveloped (Bergen et al., 2009; Lefsky et al., 2002; Vierling et al., 2008). To date, LiDAR-derived estimates of vegetation structure have been applied to a range of terrestrial systems, including conifer forests and alpine meadows (Graf et al., 2009; Müller et al., 2010), deciduous woodlands (Goetz et al., 2010; Hinsley et al., 2006), mixed conifer/deciduous forests (Clawges et al., 2008; Goetz et al., 2007), riparian forests (Seavy et al., 2009), forest understory (Martinuzzi et al., 2009), shrub and rangelands (Streutker & Glenn, 2006), and agricultural or mixed-use landscapes (Bradbury et al., 2005; Nelson et al., 2005). Individual-species distributions (Goetz et al., 2010; Graf et al., 2009; Nelson et al., 2005), species diversity (Goetz et al., 2007; Müller et al., 2010; Seavy et al., 2009), and habitat quality (Goetz et al., 2010; Hinsley et al., 2006)have all been modeled using LiDAR-derived measures. One study of avian diversity suggested that LiDAR-derived structural measures performed better than vegetation composition in predicting site diversity (Müller et al., 2010). Again, where this is the case, assessment using LiDAR-derived estimates of vegetation physiognomy may be more time and cost efficient than field-based measurements.

We used LiDAR data to define and map potential habitat for the black-capped vireo (*Vireo atricapilla*) on the Fort Hood Military Reservation, Texas. The black-capped vireo breeds in shrubland habitats historically maintained by fire and shallow soils and persisting 5–30 years (Graber, 1961). These transitional habitats are time-consuming to monitor and map, but initial efforts suggest they may be identifiable with LiDAR surveys (Leyva et al., 2002). Model-based habitat mapping using remotely-sensed data could save considerable effort and aid in the designation and management of critical vireo habitats. Furthermore, the black-capped vireo is endangered and an accurate model of habitat suitability would be useful for designating critical habitat. We constructed and compared predictive habitat models using all combinations and subsets of LiDAR-derived data on vegetation structure, field-derived data on vegetation composition, and soils data. Models were fitted using a machine-learning algorithm and permutation analyses were used to assess variable importance.

2. Methods

2.1. Study area

Fort Hood, an 87,890-hectare military installation in north-central Texas (97°44′W, 31°12′N), supports the largest known population of black-capped vireos under a single management agency (Cimprich & Kostecke, 2006). Fort Hood is located at the intersection of the Edwards Plateau and the Crosstimbers and Southern Tallgrass Prairie ecoregions. Topography includes numerous flat-topped mesas with steep gullies and mesic bottomlands. The installation is covered by woodlands and upland forest (47%), grasslands (34%), and small amounts of shrubland and riparian forest (4% and 3%, respectively).

2.2. Study species

The black-capped vireo is a nearctic-neotropical migrant passerine that breeds in north central Mexico, Texas, and central Oklahoma and winters in west central Mexico (Graber, 1961). In Texas, the vireo nests in shrub thickets comprised primarily of short, deciduous shrub and tree species arranged in clumps on the landscape (Grzybowski et al., 1994). These irregularly shaped thickets are 1–3 m tall, provide 30–60% woody vegetative cover, and are separated by grasslands or rock pavement (Bailey & Thompson, 2007). Shrublands have been observed to persist on Fort Hood in areas with shallow soils or where fires ignited by artillery are common (Pekins, 2006). Otherwise, they represent an early stage in forest succession that will transition into taller shrubs and trees (Cimprich & Kostecke, 2006).

2.3. Data sources

An assessment of vireo presence throughout Fort Hood was conducted in 2002 and 2003. All areas identified from aerial photos as potential vireo habitat were visited during the breeding season (Cimprich & Kostecke, 2006). Areas were searched on foot by walking randomly through potentially suitable vegetation. Song playbacks and multiple visits were used to increase detections. Locations of all calling male vireos were marked with geographic positioning system (GPS) receivers.

We used LiDAR data to generate four variables for modeling vireo habitat (Table 1). LiDAR data were acquired along 42 flight-lines over Fort Hood during a four-day period (Table 2). Raw LiDAR data were processed by the contractor using DASHMap software (Optech Inc., 2006) to create geo-corrected LiDAR point clouds that were used to construct a digital elevation model (DEM). The DEM was quality checked by the contractor against 950 test points to estimate a 95% confidence interval for the DEM of $\pm$0.47 m (Optimal Geomatics, 2009).

We generated maps of canopy height and two measures of woody cover across Fort Hood using the original geo-corrected LiDAR point cloud files (LAS file format) and the GridMetrics function in the FUSION software package (McGaughey, 2009). Estimated height was the mean height of all first returns in a grid cell excluding points with heights less than 1 m to avoid calculating grassland heights (as in Seavy et al., 2009). The first cover estimate was the percentage of first returns between 1 and 30 m out of all first returns and provides an estimate of total woody cover. The second

Table 1
Response and predictor variables and data sources used in constructing black-capped vireo habitat models for Fort Hood, TX.

Variable	Description	Original format	Reference
VIREO	All areas identified as potential vireo habitat were visited in 2002 and 2003 and locations of observed vireos recorded.	Point shapefile	Cimprich & Kostecke, 2006
HEIGHT	Mean height of woody vegetation. Calculated using LiDAR return data from surveys conducted on Fort Hood in March 2009.	LAS files and LiDAR-derived DEM	Optimal Geomatics, 2009
COVER	Percent cover of woody vegetation. Percent of LiDAR returns measuring between 1 and 30 m in height.	LAS files and LiDAR-derived DEM	Optimal Geomatics, 2009
COVER2	Percent cover of woody vegetation less than 3 m tall. Percent of LiDAR returns measuring between 1 and 3 m in height.	LAS files and LiDAR-derived DEM	Optimal Geomatics, 2009
EDGE	Edge density. Total edge length within a grid cell divided by the area of the cell. Edges delineated from the 3 m resolution HEIGHT grid.	HEIGHT grid	
VEG	Manual delineation and classification of vegetation type based on aerial imagery and vegetation sampling data. Sixteen types identified, e.g., shin oak shrubland, shin oak-juniper woodland, riparian woodland, and grassland.	Polygon shapefile	Reemts & Teague, 2007
SOIL	Soil depth to a restrictive layer extracted from the Soil Survey Geographic Database (SSURGO) produced by the US Department of Agriculture Natural Resource Conservation Service	Polygon shapefile	USDA-NRCS, 2007

Table 2
LiDAR specifications.

Collection dates	21–25 March 2009
Aircraft	Cessna Turbo 206 Stationair
Sensor	Optech ALTM S/N 04SEN155
Laser pulse density	~1 point/m^2
Laser pulse rate	50 kHz
Flying height	2200 m
Scan angle from nadir	±21°
Horizontal accuracy	1.42 m
Vertical accuracy	1.47 m

cover estimate was the percentage of first returns between one and three meters out of all first returns. This is a proxy measure for the percent cover of low-level vegetation used by nesting vireos. In all calculations, developed areas and points above 30 m (likely from man-made structures) were excluded. Grid cells with no woody vegetation were assigned heights and cover values of zero.

One challenge of using high resolution data is identifying the appropriate sample frame or grid-cell size (Seavy et al., 2009; Smith et al., 2008). Initial applications of LiDAR-derived measures to habitat suitability modeling for birds suggest that these measures should be summarized at different spatial scales for different species (Seavy et al., 2009). Therefore, height and cover data were summarized in raster grids at six spatial resolutions (10, 25, 50, 100, 250, and 500 m) in ArcGIS 9.3 (ESRI, 2009).

We converted a 3-m resolution map of vegetation height into a binary presence/absence map of woody-vegetation and used it to calculate edge density. Edge density, the total distance of all edges in a grid cell divided by its area (Neel et al., 2004), was calculated in a GIS at the six spatial resolutions by summing the total edge distance from the 3-m resolution map within the larger areas. Edge density is known to show non-linear responses to patch size and degree of aggregation, but the machine-learning algorithm used to build the models has the ability to capture this non-linearity (Neel et al., 2004). Spearman's rank correlations were also calculated among all LiDAR-derived measures to inform interpretation of variable importance measures. A map of vegetation types across Fort Hood was generated in 2007 through manual delineation of vegetation patches from 0.35-m resolution digital orthophotos taken during 2004 (Reemts & Teague, 2007). Patch classification was based on detailed vegetation surveys at 193, 400-m^2 classification plots distributed throughout Fort Hood. These were analyzed in a cluster analysis identifying 70 vegetation classes. These classifications and 358 validation points were used to generate a photo-interpretation guide that incorporated information on vegetation, soils, topography, and known disturbance history for patch delineation and classification. The resulting vegetation map includes 7693 polygons classified into the 70 vegetation classes. We merged the 70 classes into 16 vegetation types representing shrub and woodland habitats important to vireos (see Table 1 and Supplementary Fig. A1 for examples). The minimum delineated patch size was ~0.5 ha. We converted polygon data into grids at the six spatial resolutions, assigning classes based on the center point of each grid cell to preserve spatial pattern in the polygon dataset (Bian & Butler, 1999).

Shrublands suitable to vireos on Fort Hood are known to occur in areas with shallow soils and rock pavement (Bailey & Thompson, 2007). Therefore, a map of soil depth (cm) to a restrictive layer was extracted from the US Department of Agriculture Natural Resource Conservation Service Soil Survey Geographic (SSURGO) database for Bell and Coryell Counties, Texas (USDA-NRCS, 2007) using the Soil Data Viewer 5.2 tool (USDA-NRCS, 2008) for ArcMap 9.2. These maps are based on soil surveys conducted in Bell and Coryell counties in 1977 and 1985, respectively. The soil map includes 1233 polygons with soil depths ranging from 28 to 201 cm. The minimum delineated patch size was also ~0.5 ha, and we again converted this dataset into raster grids at the six spatial resolutions. We tested for independence between classifications of vegetation type and soil depth with contingency table analysis (X^2). We also converted LiDAR-derived measures into categorical variables (five bins were used with bin width depending on variable range) and compared these to classifications of vegetation type and soil depth to assess independence.

Mapping polygon data to fine resolution raster grids can both improve polygon representation and introduce error. Improvements occur where a smaller grid-cell size more accurately depicts a patch's perimeter. Conversely errors are introduced where smaller patches of potential habitat or non-habitat are obscured by the original, coarser resolution data. We chose to grid vegetation and soil data to 10, 25, and 50-m resolution grids because the gains of better representing patch perimeters likely outweighed the potential for introducing errors. Because the vegetation classification included savanna vegetation types that described grassy areas with scattered individual trees and shrubs, it is likely that most of the very small patches of shrubs were described in these vegetation types. Nonetheless, by gridding the data at finer resolutions, we introduced error into the vegetation and soils data.

Vireo presence data were converted into raster grids at the six resolutions by designating as a presence any grid cell containing a presence location, and designating remaining cells as pseudo-absences.

2.4. Habitat suitability models

Habitat suitability models were calculated with cforest (Strobl et al., 2007), a variation of the Random Forest Predictors (RF) classification algorithm (Breiman, 2001). RF is a non-parametric machine-learning tool notable for (1) requiring no a priori assumptions about the relationship between predictor and response variables, (2) the ability to model nonlinear relationships and interactions among variables, and (3) high prediction accuracy (Prasad et al., 2006). In direct comparisons, RF performed similarly or better than parametric methods such as linear discriminant analysis, generalized linear regression, classification trees, artificial neural networks, and generalized additive models (Cutler et al., 2007; Lawler et al., 2006; Prasad et al., 2006).

The `cforest` function (CF) of the 'party' package version 0.9-995 in R (Hothorn et al., 2010) is not as widely applied as RF. CF is based on a conditional inference statistical framework (Strobl et al., 2007) that results in less bias towards predictor variables with many potential cut points (Strobl et al., 2007). CF is recommended when predictor variables include both categorical and continuous variables and when continuous variables differ in range (Strobl et al., 2007).

We set aside one third of grid cells in a test dataset to assess accuracy and to calculate additional test statistics. The number of randomly selected grid cells used for calibration purposes varied with grid cell size, ranging from 1615 cells for the 500-m resolution model to 6600 cells for the 10-m resolution model. Prevalence, the ratio of presence to absence grid cells, has been shown to impact calculations of some performance metrics (Allouche et al., 2006). We therefore maintained a ratio of one presence to two pseudo-absences in all test and training datasets. Preliminary analyses suggested that performance was sensitive to the composition of the training and test datasets. Therefore, we randomly generated 25 training and test datasets at each spatial scale and calculated performance metrics for each dataset (as in Santos et al., 2010). This approach, along with the use of a machine-learning algorithm as opposed to a standard statistical method such as logistic regression, precluded the use of information-theoretic methods for model selection and comparison. Therefore, we used boxplots to summarize performance metrics across the 25 datasets.

We first constructed models using only each data class, i.e. LiDAR-derived structural measures, the vegetation composition map, or the soil depth map, summarized across the six spatial scales. For each data class we identified the best-performing resolution. We then constructed structure + composition, structure + soil depth and composition + soil depth models as well as the full model (structure + composition + soil depth).

2.5. Model assessment

We calculated several measures of model fit because there are many such measures, each with its own advantages and limitations (Allouche et al., 2006; Fitzgerald & Lees, 1994; Lobo et al., 2008; Manel et al., 2001; Stehman, 1997). CF models output the probability of presence for each grid cell in the test dataset. We generated a receiver operating characteristic (ROC) curve (ROC v1.16 package Carey, 2007) for each CF model (Hanley & McNeil, 1982) and calculated the area under the ROC (the area is commonly referred to as the AUC) to assess model fit (Fielding & Bell, 1997; Hanley & McNeil, 1982). The ROC curve was also used to select a threshold value for classifying presences from predicted probabilities that minimized the difference between the rates of omission and commission error (Jiménez-Valverde & Lobo, 2007). We calculated Cohen's κ statistic (vcd package Meyer et al., 2010) for agreement between two raters (Fitzgerald & Lees, 1994) based on this threshold. The κ statistic provides a measure of accuracy adjusted for prevalence as well as the relative performance of the classifier across classes (Fitzgerald & Lees, 1994; Manel et al., 2001). In addition, we reported the overall accuracy of the classification. Accuracy estimates were also used to assess whether models performed better than a null model in which presence and absences were randomly assigned to the test data with the same proportions as prevalence. An empirical distribution of null model performance was generated from 1000 simulations of the null model and used to estimate a probability for the observed accuracy of each CF model.

Variable importance was estimated as in RF: calculating the mean decrease in model accuracy based on cross-validation of the training data resulting from the random permutation of each predictor variable within each classification tree (Breiman, 2001; Hothorn et al., 2010). All habitat modeling algorithms and statistical analyses were conducted using the R statistical package v2.6.1 (R Development Core Team, 2010).

3. Results

3.1. Single predictor class models

Habitat models constructed with LiDAR-derived measures of vegetation structure performed well, correctly predicting between 68% and 78% of observations in the 25 test datasets (Fig. 1). The relationship between model performance and resolution was curvilinear with models at intermediate resolutions (25- and 50-m) performing best. Models at the 25- and 50-m resolution correctly predicted a median (across test datasets) of 76% of observations. Other measures of model performance were similar. AUC scores for LiDAR models ranged from 0.75 to 0.86 and κ values ranged from 0.33 to 0.54.

The relative importance of the four structural variables differed across resolutions in the LiDAR models. In the top-performing 25-m resolution model, relative variable importance (listed in declining order) was: edge density (EDGE), vegetation height (HEIGHT), percent cover of vegetation 1–3 m tall (COVER2), and total percent woody cover (COVER, Fig. 2). COVER2 was most important in the 10-m resolution model, whereas EDGE was most important in coarser resolution models. Cover measures were positively correlated (Spearman's rank $\rho = 0.91$, p-value < 0.001) and correlated with height (COVER: $\rho = 0.85$, COVER2: $\rho = 0.64$; p-values < 0.001) at 25-m resolution. No other LiDAR-derived measures were correlated by more than $|\rho| = 0.5$ at 25-m resolution.

Model performance increased monotonically with resolution for the vegetation composition and soil depth models. The best performing vegetation structure and soil depth models correctly predicted a median 74% and 79% of observations (Fig. 1, results not shown for all resolutions). Median AUC and κ values were 0.85 and 0.46 for the vegetation composition model and 0.84 and 0.54 for the soil depth model. Threshold dependent accuracy and κ values for both the vegetation composition and soil models were sensitive to the minimum distance threshold used, performing worse and better, respectively, than suggested by AUC scores. All models performed better than a null model that randomly assigned presences and absences according to prevalence (p-value < 0.001).

Classifications of vegetation type were not independent of soil depth in all training datasets (X^2 tests, df ~ 272, p-values < 0.0001). Vegetation structure measures (HEIGHT, COVER, COVER2, EDGE) were also dependent on soil depth (X^2 tests, df = 72, p-values < 0.0001) and vegetation type (X^2 tests, df = 68, p-values < 0.0001) classifications.

3.2. Multi-predictor class models

Combining data classes (structure + soil, structure + composition, and composition + soil) improved median accuracies in all cases (Fig. 1). Median accuracies were 81%, 80%, and 80%, respectively. Combining data classes reduced sensitivities to the minimum distance threshold for the accuracy and κ measures, but these were still lower than expected given the AUC values for the composition + soil model. Threshold-independent AUC values showed even less discrepancy among models. The vegetation structure and soil model performed slightly better across all performance metrics. The full model, including vegetation structure, vegetation composition and soil depth measures, correctly predicted a median 82% of all observations. Vegetation type was consistently the most important predictor in the full model, followed by soil depth, edge density and height, and cover measures (Fig. 2). Again, all models performed better than a null model randomly assigning presences and absences according to prevalence (p-value < 0.001).

4. Discussion

The curvilinear response of LiDAR model performance measures to changes in data resolution (Fig. 1) suggests that vireos are best associated with habitat structure summarized with 25- to 50-m resolution grid cells (0.05–0.25 ha). Similar sample areas were used by Grzybowski et al. (1994) and Bailey and Thompson (2007) to calculate cover and height (0.04 ha) and percent cover and edge density (0.2 ha) measures, respectively, in their vireo habitat suitability models. This is now the third analysis of vireo habitat in which these measures and spatial scales have proven important predictors of vireo habitat suitability. The 10-m grid cells may be too small to capture the patchy combination of grass and shrubs associated with vireos, resulting in poor performance of the finer resolution models. Aggregation to coarser resolutions introduces errors (Bian & Butler, 1999) that likely decrease model performance.

The relative importance of structural variables in distinguishing habitat suitability in our models differed from a comparison of nesting and non-nesting vegetation on Fort Hood (Bailey & Thompson, 2007). We found edge density, height and percent cover in the 1–3 m height class were more important than woody cover in distinguishing presence and pseudo-absence locations (Fig. 2), whereas Bailey and Thompson (2007) found woody cover to have a greater impact than edge density in 0.2-ha resolution models. Anecdotal observation by D. Cimprich has found that vireos occupy sites across a range of total woody cover provided there is sufficient foliage close to the ground. Our cover measures were correlated with each other and height, and were dependent on classifications of vegetation type and soil depth. Therefore, edge density likely has a higher

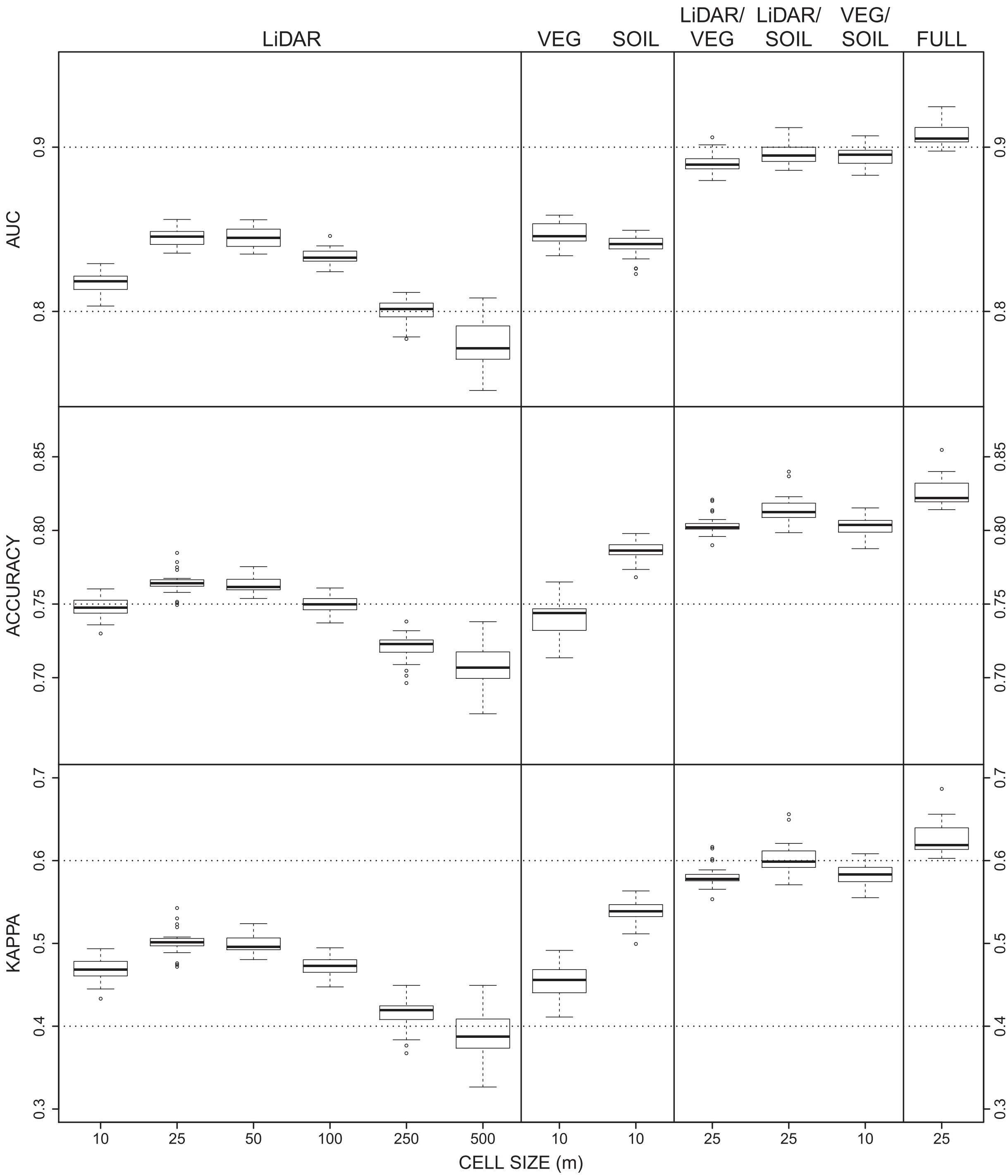

Fig. 1. Threshold-independent (AUC) and threshold-dependent (overall accuracy and Cohen's κ) measures of model performance for (1) models constructed with only LiDAR-derived measures of vegetation structure at six spatial resolutions (LiDAR), or maps of vegetation type (VEG) or soil depth (SOIL); (2) models combining pairs of predictor classes; and (3) a full model incorporating all three data classes (FULL). Performance measures for the LiDAR-based models are reported for a range of spatial resolutions but measures for VEG, SOIL, and combined models are reported for the best-performing spatial resolution (either 10 or 25 m). Dashed lines mark commonly referenced bounds for assessing model fit for AUC (Swets, 1988) and κ (Landis & Koch, 1977) as well as an arbitrary minimum acceptable threshold of 75% accuracy.

relative importance in our models because it provides unique information about vegetation structure. The exception to this is the 10-m resolution model where the small grid-cell size likely reduces the information contained in the edge density metric because of insufficient coverage. The relative importance of total woody cover and cover in the 1–3 m height class are reversed in the full model because they are both dependent on the vegetation type and soil depth classifications which are absent from the model with only structural variables.

All predictor classes contributed to a full model that outperformed all other models. Variable importance rankings support the accepted protocol on Fort Hood for identifying suitable habitat in the field: searching first for appropriate vegetation composition and then for appropriate structure. Because LiDAR cannot distinguish among plant species, importance of vegetation composition ranks higher than structural measures. For example, an area with appropriate structure composed mainly of elbowbush (*Forestiera pubescens var.*

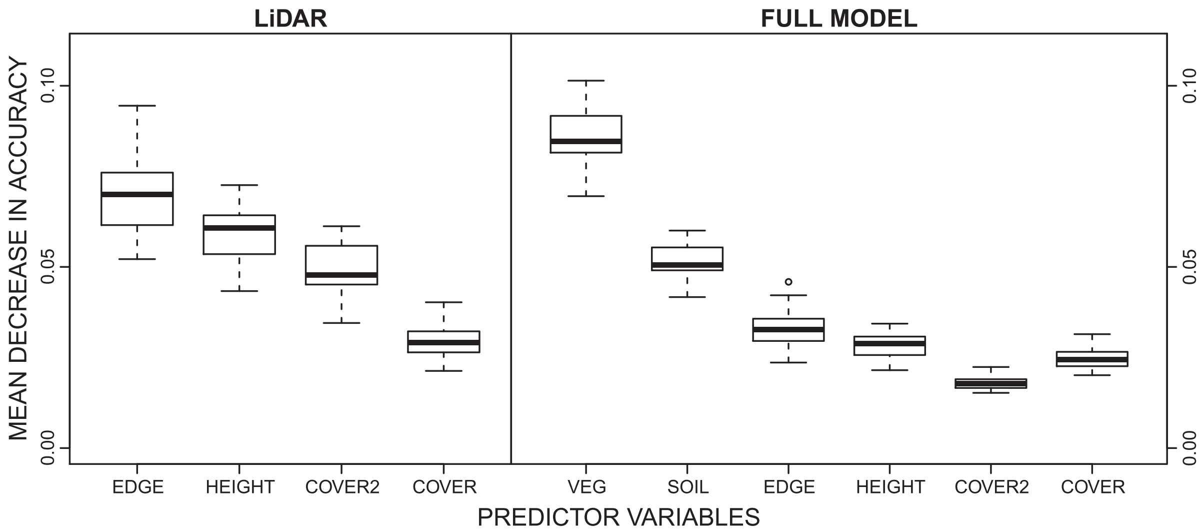

Fig. 2. Mean decrease in accuracy when data for each predictor variable are permutated as a measure of variable importance in the 25-m resolution LiDAR model and FULL model incorporating all predictor classes. Descriptions of variables and their names are provided in Table 1.

pubescens), flame-leaf sumac (*Rhus lanceolata*), or juniper (*Juniperus* spp.) would have few to no vireos present, whereas an equivalent area with Texas redbud (*Cercis canadensis var. texensis*), Texas red oak (*Quercus buckleyi*), or Texas ash (*Fraxinus texensis*) would have a moderate number of vireos, and an area with Shin oak (*Quercus sinuata var. breviloba*) would have many vireos. Vegetation classifications and soil depths are more important than structural features in the full model because vegetation type implies a certain structure. For example, stands of Shin oak co-dominant with juniper are more often taller, denser woodlands, whereas Shin oak stands are more often shrublands. Vegetation type is also dependent on soil depth. Structural attributes provide a final filter determining habitat suitability because some suitable vegetation is overgrown. Correlations between LiDAR-derived measures of vegetation structure and measures derived from aerial photos and field surveys have been observed elsewhere as well (Müller et al., 2009).

Maps predicting habitat suitability across Fort Hood based on models generated from one of the 25 datasets also support the variable importance rankings (Supplementary Figs. A2 and A3). Predictions based on either vegetation type or soil include areas that are suitable or potentially suitable. Vegetation composition identifies all of the best habitat, but also large patches of habitat, particularly in southeastern Fort Hood, where few vireos have been documented. This area is dominated by suitable oak species, but is too tall and overgrown for vireos. Predictions based on soil depth highlight areas with shallow soils that either are currently shrublands or have the potential to be shrublands in the future. The vegetation-structure model identifies suitable habitat structure throughout Fort Hood, including some regions where very few vireos were observed. These are regions with suitable structure, but dominated by juniper or shrub species avoided by vireos (Bailey & Thompson, 2007; Grzybowski et al., 1994). A map based on vegetation composition is more suitable for conservation action because it identifies areas of current and potential habitat, whereas a model based on structure is more of a snapshot that lacks context. This may be peculiar to the vireo because it prefers early seral communities. A species preferring mature communities may be better served by a model of vegetation structure.

The best models performed well in both threshold-independent (AUC) and threshold-dependent (κ and overall accuracy) measures of model fit. Median AUC values for our models were equal to or higher than reported values in other avian habitat suitability modeling efforts using LiDAR-derived measures with (Graf et al., 2009) and without (Seavy et al., 2009) measures of vegetation composition. Median κ values were higher than values reported for a LiDAR-derived model of grouse habitat (Graf et al., 2009). Differences in study design may impact these performance metrics, so is suffices to say that our models performed similarly to other published attempts. Median accuracies were above 75% for a majority of models. This level of accuracy would be acceptable for most habitat management applications. However, managers may choose to select a more conservative threshold (one that minimizes commission error) than we used for designation of critical habitat.

This study effectively applied non-parametric machine-learning methods to model habitat suitability. The conditional inference statistical framework of CF demonstrates less bias when building classification trees with both categorical and continuous predictor variables (Strobl et al., 2007). The use of 25 randomly generated test and training datasets provided information on the uncertainty surrounding estimates of model performance. This proved to be a useful method for comparing model performance among machine-learning algorithms.

The age and asynchrony of the data sources used likely introduced errors into our models. Vegetation growth over time, disturbance-induced changes in vegetation structure and composition, and annual variation in the presence/absence of vireos throughout Fort Hood would explain much of this error. In spite of this, models performed well, suggesting that broader patterns of vegetation structure, composition, and vireo occupancy remain consistent through time. Vireo habitat may not be as dynamic as previously thought, which has positive implications for the species as a whole. That said, successfully modeling vireo occupancy does not imply that density patterns have not changed over time. This model as well as others employing remotely sensed data should be validated in the field with current presence/absence or other data before being applied in a management context. Furthermore, model performance would likely be lower if applied to a new location.

Inaccuracies in LiDAR-derived height estimates have been observed in shrub-steppe vegetation (Sankey & Bond, 2011; Streutker & Glenn, 2006). However, in both studies inaccuracies did not translate into an inability to classifying vegetation structural types based on height. The same occurred in this study. The vertical inaccuracy of our LiDAR dataset was relatively high (1.47 m), but heights differed between modeled suitable and unsuitable vireo habitats. The use of edge density, calculated from a binary map of woody vegetative cover derived from height data, likely decreased the impact of height measurement errors. In general, vertical inaccuracies in the LiDAR dataset probably decreased the importance of canopy height as a predictor of vireo presence, but did not preclude high overall model performance.

4.1. Conclusions

The black-capped vireo is an endangered species with a narrow breeding distribution and habitat requirements restricted by vegetation species composition and structural attributes. We found that remotely sensed data products can be used effectively to scale-up predictions of habitat suitability. Our results suggest that in central Texas, vegetation composition is a more important predictor of vireo occupancy than structural attributes, but that structure can serve to either improve predictions as a filter for vegetation or to predict habitat in the absence of vegetation data when combined with a second data source such as soil depth. Field studies indicate that vireo habitat preferences outside of the Fort Hood region include other shrub and tree species, but often with the same shrubland structure (Grzybowski et al., 1994; Wilkins et al., 2006). Therefore, LiDAR-derived measures of vegetation structure may be a suitable method for predicting range-wide habitat suitability when broad-scale LiDAR data becomes available. Non-parametric machine-learning methods, such as cforest, are an effective tool for modeling habitat when using correlated measures.

In the ongoing discussion of whether physiognomy or species composition is a stronger predictor of avian species diversity (MacArthur & MacArthur, 1961; Müller et al., 2010), Müller et al. (2010) recently demonstrated that LiDAR-derived structural measures were more effective than field-based measures of vegetation species composition at predicting avian species composition. For an individual species, preference for physiognomy or species composition is a natural history trait, and does not necessarily fit into a larger pattern reflecting overall species diversity. Both vegetation composition and structural variables contributed to a model of breeding habitat suitability for the black-throated blue warbler (*Dendroica caerulescens*, Goetz et al., 2010). *V. atricapilla* is an interesting case study because in the area around Fort Hood it is associated with several deciduous woody species, e.g. *Quercus sinuate var. breviloba*, but only as early seral shrublands (Bailey & Thompson, 2007; Graber, 1961; Grzybowski et al., 1994). Although structure does serve as a final filter, our results suggest that species composition can be used to predict vireo occupancy with high accuracy and is potentially more meaningful for vireo conservation because it identifies current and potential habitats. Areas that have the right mix of species could be managed in such a way as to produce the appropriate vegetation structure. Incorporation of other species-specific factors, such as soil characteristics, further improve model performance and can be combined with structural data in the absence of vegetation composition data. Current estimates of the extent of *V. atricapilla* habitat across Texas are imprecise (Wilkins et al., 2006). Our findings suggest that LiDAR-derived measures of vegetation structure can effectively identify these ephemeral habitats.

Supplementary materials related to this article can be found online at doi:10.1016/j.rse.2011.12.002.

Acknowledgements

This research was completed as part of the Strategic Environmental Research and Development Program (SERDP) RC-1541 project. The authors would also like to acknowledge the following individuals for their involvement: Charlotte Reemts, Richard Kostecke, Marion Noble, Robert McGaughey, and Betsy Bancroft.

References

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*, 1223–1232.

Bailey, J. W., & Thompson, F. R. (2007). Multiscale nest-site selection by black-capped vireos. *Journal of Wildlife Management*, *71*, 828–836.

Barber, D. R., & Martin, T. E. (1997). Influence of alternate host densities on brown-headed cowbird parasitism rates in black-capped vireos. *Condor*, *99*, 595–604.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., et al. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research-Biogeosciences*, *114*.

Bian, L., & Butler, R. (1999). Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data. *Photogrammetric Engineering and Remote Sensing*, *65*, 73–84.

Bradbury, R. B., Hill, R. A., Mason, D. C., Hinsley, S. A., Wilson, J. D., Balzter, H., et al. (2005). Modelling relationships between birds and vegetation structure using airborne LiDAR data: A review with case studies from agricultural and woodland environments. *Ibis*, *147*, 443–452.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Carey, V. (2007). *ROC: utilities for ROC, with uarray focus v1.16.0: Bioconductor.* http://www.bioconductor.org

Cimprich, D. A., & Kostecke, R. M. (2006). Distribution of the black-capped vireo at Fort Hood, Texas. *The Southwestern Naturalist*, *51*, 99–102.

Clawges, R., Vierling, K., Vierling, L., & Rowell, E. (2008). The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment*, *112*, 2064–2073.

Cutler, D. R., Edwards, T. C., Jr., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., et al. (2007). Random forests for classification in ecology. *Ecology*, *88*, 2783–2792.

ESRI (2009). *ArcGIS 9.3.1.* Redlands, CA: ESRI.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*, 38–49.

Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of Environment*, *47*, 362–368.

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 Sensing of Environment*, *108*, 254–263.

Goetz, S. J., Steinberg, D., Betts, M. G., Holmes, R. T., Doran, P. J., Dubayah, R., et al. (2010). Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology*, *91*, 1569–1576.

Graber, J. W. (1961). Distribution, habitat requirements, and life history of the black-capped vireo (*Vireo atricapilla*). *Ecological Monographs*, *31*, 313–336.

Graf, R. F., Mathys, L., & Bollmann, K. (2009). Habitat assessment for forest dwelling species using LiDAR remote sensing: Capercaillie in the Alps. *Forest Ecology and Management*, *257*, 160–167.

Grzybowski, J. A., Tazik, D. J., & Schnell, G. D. (1994). Regional analysis of black-capped vireo breeding habitats. *Condor*, *96*, 512–544.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hinsley, S. A., Hill, R. A., Bellamy, P. E., & Balzter, H. (2006). The application of lidar in woodland bird ecology: Climate, canopy structure, and habitat quality. *Photogrammetric Engineering and Remote Sensing*, *72*, 1399–1406.

Hothorn, T., Hornik, K., Strobl, C., & Zeileis, A. (2010). *Package 'party' version 0.9-995 in R.* http://cran.r-project.org/web/packages/party/party.pdf [accessed 6 July 2010]

Hutto, R. L. (1985). Habitat selection by nonbreeding, migratory land birds. In M. Cody (Ed.), *Habitat selection in birds* (pp. 455–476). Orlando, FL: Academic Press.

James, F. C. (1971). Ordinations of habitat relationships among breeding birds. *Wilson Bulletin*, *83*, 215–236.

Jiménez-Valverde, A., & Lobo, J. M. (2007). Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecologia*, *31*, 361–369.

Landis, J. R., & Koch, G. G. (1977). Measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lawler, J. L., White, D., Neilson, R. P., & Blaustein, A. R. (2006). Predicting climate-induced range shifts: Model differences and model reliability. *Global Change Biology*, *12*, 1568–1584.

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*, 19–30.

Leyva, R. I., Henry, R. J., Graham, L. A., & Hill, J. M. (2002). Use of LiDAR to determine vegetation vertical distribution in areas of potential black-capped vireo habitat at Fort Hood, Texas. *Endangered species monitoring and management at Fort Hood, Texas: 2002 annual report*. Fort Hood, Texas, USA: The Nature Conservancy, Fort Hood Project.

Lobo, J. M., Jimenez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*, 145–151.

MacArthur, R., & MacArthur, J. W. (1961). On bird species-diversity. *Ecology*, *42*, 594–598.

Manel, S., Williams, H. C., & Ormerod, S. J. (2001). Evaluating presence–absence models in ecology: The need to account for prevalence. *Journal of Applied Ecology*, *38*, 921–931.

Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., et al. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, *113*, 2533–2546.

McGaughey, R. J. (2009). *FUSION/LDV: Software for LIDAR data analysis and visualization.* Seattle, WA: US Department of Agriculture Forest Service, Pacific Northwest Research Station, University of Washington.

Meyer, D., Seileis, A., & Hornik, K. (2010). *vcd: Visualizing Categorical Data 1.2-8.* http://cran.r-project.org/web/packages/

Michel, P., Jenkins, J., Mason, N., Dickinson, K. J. M., & Jamieson, I. G. (2008). Assessing the ecological application of lasergrammetric techniques to measure fine-scale vegetation structure. *Ecological Informatics*, *3*, 309–320.

Müller, J., Moning, C., Bassler, C., Heurich, M., & Brandl, R. (2009). Using airborne laser scanning to model potential abundance and assemblages of forest passerines. *Basic and Applied Ecology*, *10*, 671–681.

Müller, J., Stadler, J., & Brandl, R. (2010). Composition versus physiognomy of vegetation as predictors of bird assemblages: The role of lidar. *Remote Sensing of Environment*, *114*, 490–495.

Neel, M. C., McGarigal, K., & Cushman, S. A. (2004). Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landscape Ecology*, *19*, 435–455.

Nelson, R., Keller, C., & Ratnaswamy, M. (2005). Locating and estimating the extent of Delmarva fox squirrel habitat using an airborne LiDAR profiler. *Remote Sensing of Environment*, *96*, 292–301.

Newton, A. C., Hill, R. A., Echeverria, C., Golicher, D., Benayas, J. M. R., Cayuela, L., et al. (2009). Remote sensing and the future of landscape ecology. *Progress in Physical Geography*, *33*, 528–546.

Optech Inc. (2006). *DASHMap Data Processing Software.* : Optech Inc..

Optimal Geomatics (2009). *Fort Hood LiDAR Survey*. Huntsville, AL: Optimal Geomatics, Inc..

Pekins, C. E. (2006). Armored military training and endangered species restrictions at Fort Hood, Texas. *Federal Facilities Environmental Journal*, *2*, 37–50.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181–199.

R Development Core Team (2010). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing3-900051-07-0http://www.R-project.org

Reemts, C. M., & Teague, J. (2007). Vegetation classification and mapping of Fort Hood. *Endangered species monitoring and management at Fort Hood, Texas: 2007 annual report*. Fort Hood, Texas, USA: The Nature Conservancy, Fort Hood Project.

Rotenberry, J. T. (1985). The role of habitat in avian community composition — Physiognomy or floristics. *Oecologia*, *67*, 213–217.

Sankey, T. T., & Bond, P. (2011). LiDAR-based classification of sagebrush community types. *Rangeland Ecology & Management*, *64*, 92–98.

Santos, M. J., Greenberg, J. A., & Ustin, S. L. (2010). Using hyperspectral remote sensing to detect and quantify southeastern pine senescence effects in red-cockaded woodpecker (*Picoides borealis*) habitat. *Remote Sensing of Environment*, *114*, 1242–1250.

Seavy, N. E., Viers, J. H., & Wood, J. K. (2009). Riparian bird response to vegetation structure: A multiscale analysis using LiDAR measurements of canopy height. *Ecological Applications*, *19*, 1848–1857.

Smith, K. M., Keeton, W. S., Donovan, T. M., & Mitchell, B. (2008). Stand-level forest structure and avian habitat: Scale dependencies in predicting occurrence in a heterogeneous forest. *Forest Science*, *54*, 36–46.

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, *62*, 77–89.

Streutker, D. R., & Glenn, N. F. (2006). LiDAR measurement of sagebrush steppe vegetation heights. *Remote Sensing of Environment*, *102*, 135–145.

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*, 25.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285–1293.

USDA-NRCS (2007). *US Department of Agriculture National Resources Conservation Service Soil Survey Geographic (SSURGO) Database.* : US Department of Agriculturehttp://SoilDataMart.nrcs.usda.gov/ [accessed 30 July 2008]

USDA-NRCS (2008). *US Department of Agriculture National Resources Conservation Service Soil Soil Data Viewer 5.2.* : US Department of Agriculturehttp://soilddataviewer.nrcs.usda.gov [accessed 11 January 2010]

Vierling, K., Vierling, L., Gould, W., Martinuzzi, S., & Clawges, R. (2008). Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, *6*, 90–98.

Wilkins, N., Powell, R. A., Conkey, A. T., & Snelgrove, A. G. (2006). *Population status and threat analysis for the Black-capped Vireo.* College Station, TX: Department of Wildlife and Fisheries Sciences, Texas A&M University.