Ecography 37: 137–151, 2014
doi: 10.1111/j.1600-0587.2013.00299.x

Subject Editor: R. Greenberg. Accepted 31 May 2013

Climate and vegetation hierarchically structure patterns of songbird distribution in the Canadian boreal region

S. G. Cumming, D. Stralberg, K. L. Lefevre, P. Sólymos, E. M. Bayne, S. Fang, T. Fontaine, D. Mazerolle, F. K. A. Schmiegelow and S. J. Song

S. G. Cumming and K. L. Lefevre, Dépt des sciences du bois et de la forêt, Pavillon Abitibi-Price, Bureau 2130, Univ. Laval, Québec, QC G1V 0A6, Canada. KL also at: 157 Pearl St. South, Hamilton, ON L8P 3X7, Canada. – D. Stralberg (stralber@ualberta.ca), P. Sólymos and E. M. Bayne, Dept of Biological Sciences, CW 405, Biological Sciences Building, Univ. of Alberta, Edmonton, AB T6G 2E9, Canada. – S. Fang, Alberta Provincial Stroke Strategy, Alberta Health Services, 10044-108 Street, Edmonton, AB T5J IK6, Canada. – T. Fontaine and F. K. A. Schmiegelow, Dept of Renewable Resources, 751 General Services Building, Univ. of Alberta, Edmonton, AB T6G 2H1, Canada. FKAS also at: Northern ENCS Program, Dept of Renewable Resources, Univ. of Alberta, c/o Yukon College, Whitehorse, YT Y1A 5K4, Canada. – D. Mazerolle, Fundy National Park of Canada, PO Box 101, Alma, NB E4H 1B4, Canada. – S. J. Song, Canadian Wildlife Service, Environment Canada, 9250-49th Street, Edmonton, Alberta T6B 1K5, Canada.

Environmental factors controlling the distribution and abundance of boreal avifauna are not fully understood, limiting our ability to predict the consequences of a changing climate and industrial development activities underway. We used a compilation of avian point-count data, collected over 1990–2008 from nearly 36 000 locations, to model the abundance of individual forest songbird species within the Canadian boreal forest. We evaluated 30 vegetation and 101 climatic variables, representing most of the widely-used dimensions of climate space, along with less usual measures of inter-annual variability. Regression tree models allowed us to calculate the relative importance of climate and vegetation variable classes according to avian migration strategy without the need for a priori variable selection or dimension reduction. We tested for hierarchical habitat selection by formulating hypotheses on the locations of variables within the model tree structures. Climate variables explained the majority (77%) of deviance explained over 98 species modelled. As may be expected at high latitudes, we found energy availability (temperature, 65%) to be more important than moisture availability (precipitation, 12%). The contributions of inter- and intra-annual climate variability (28%) were about half that of mean conditions. The relatively large contribution of remotely-sensed vegetation metrics (23%) highlighted the importance of local vegetation heterogeneity controlled by non-climatic factors. The two most important vegetation variables were landcover type and April leaf area index. When selected, these generally occurred in a model's right subtree, consistent with predictions from hierarchical habitat selection theory. When occupying the root node, landcover effectively delineated the historical forest-prairie ecotone, reflecting the current disequilibrium between climate and vegetation due to human land use. Our findings suggest a large potential for avian distributional shifts in response to climate change, but also demonstrate the importance of finer scale vegetation heterogeneity in the spatial distribution of boreal birds.

The boreal forest provides the majority of breeding habitat for many North American passerine birds (Blancher 2003). Boreal habitats in Canada remain relatively intact compared to their counterparts in northern Eurasia (Burton et al. 2006) but industrial development and land conversion have accelerated in recent decades, especially along the southern margins (Hobson et al. 2002, Schneider et al. 2003, Bradshaw et al. 2009). The relatively high rates of climatic change projected at northern latitudes threaten to further alter the distribution and species composition of boreal forests (Ruckstuhl et al. 2008), especially when combined with expected increases in the frequency and severity of fire (Flannigan et al. 2009), drought (Hogg and Bernier 2005), and insect outbreaks (Volney and Fleming 2000). A quantitative understanding of the environmental factors influencing species distribution and abundance is essential for predicting responses to climatic and land-use change (Kerr et al. 2007).

Climate largely determines the terrestrial plant formations of global biomes (Holdridge 1967), which in turn structure broad patterns of species richness of other taxa (Hawkins et al. 2003, Currie et al. 2004, Field et al. 2009). At the level of individual species, the pathways for climatic control are numerous. Temperature can directly limit animal distributions by excluding individuals from regions where physiological tolerances are exceeded (Root 1988, Currie 1991). Climate may also act indirectly through energetic limitations on plant growth and food web development (Brown 1981, Wright 1983). Annual and seasonal variation in temperature, precipitation, and water balance control vegetation form and biomass (Whittaker and Niering 1975, Stephenson 1990), thereby affecting habitat structure for birds (Kissling et al. 2008), as well as distributions of the individual plant species that comprise avian habitats (Lee and Rotenberry 2005, Matthews et al. 2011).

In contrast to the large-scale effects of climate, vegetation composition and structure are well-documented drivers of local bird abundance and community composition (MacArthur 1964, Rotenberry and Wiens 1980, Cody 1981). Vegetation affects the provision of shelter, food and other important aspects of habitat quality at the scales of the territory, patch and landscape (Johnson 1980). Many studies have shown that 0.1–10 ha scale variation in vegetation attributes is well-correlated with distributional patterns of boreal songbirds observed by high-resolution sampling methods, such as point-count surveys (Hobson and Schieck 1999, Drapeau et al. 2000, Hobson and Bayne 2000, Vernier et al. 2008). However, most boreal studies to date have had spatial extents of at most a few thousand square kilometres, which is too small to detect effects of climatic variation.

Departures from climate-driven vegetation patterns may result from local edaphic variability, from successional dynamics associated with natural disturbances such as fire and insect outbreaks, or from human land uses (Burton et al. 2008, Payette et al. 2008, Taylor and Chen 2011). At intermediate or regional scales low rates of disturbance can allow the persistence of 'climate relicts', local vegetation conditions that are uncharacteristic of current climate (Hampe and Jump 2011). The large-scale human conversion of forest lands to agricultural use (Hobson et al. 2002) has resulted in a more dramatic mismatch between actual vegetation patterns and those expected based on climate conditions. This combination of short- and long-term historical factors may lead to species distributions that are not entirely explained by climatic factors.

According to the hierarchical model of habitat selection described by Johnson (1980), climate is viewed as controlling first-order selection, corresponding to the geographic ranges of species or populations, while vegetation influences second-order selection of home-ranges through individual choices made at higher orders of selection (e.g. of foraging sites and food items). By extension, when considering the environmental factors that influence species distribution and abundance, climate conditions are generally thought to be most important at large spatial extents, while the effects of vegetation factors are considered to be primarily local (Forsman and Mönkkönen 2003, Pearson et al. 2004). However, there have been few studies conducted at spatial extents large enough, spatial resolutions of observational data fine enough, and thematic precision of variables high enough to test this model for boreal forest songbirds. In eastern North America, Venier et al. (2004) were unable to distinguish the unique contributions of remotely-sensed landcover from climatic variables in occupancy models of Breeding Bird Atlas data. Using avian occurrence data from the Finnish 10 × 10-km national grid, Luoto et al. (2007) found that adding vegetation variables improved the accuracy of climate-only models at spatial resolutions as coarse as 20 km. However, their study area was small relative to the typical size of avian ranges. In a study of comparable spatial extent, Kissling et al. (2010) showed that adding landcover information significantly improved climate-based models of species richness in Kenya.

As with Kissling et al. (2010) many modelling studies in avian biogeography have used species richness as a response variable. However, a goal of our research is to support the conservation objectives of government agencies responsible for managing individuals and populations of migratory bird species in North America. Therefore, we took a 'deconstructive' species-distribution-modelling approach (Kissling et al. 2010) to disentangle the roles of climate and vegetation as spatial factors influencing the abundance of individual bird species. In this paper we present the first comprehensive analysis of a large collection of avian point-count data, recently assembled from 88 field studies conducted across the Canadian boreal region since the early 1990s (Cumming et al. 2010). The spatial resolution of each survey point in these data was ≈ 100 m. The spatial extent was large enough to encompass large portions of many species' breeding ranges, and marked variation in climatic conditions. We used regression trees (Breiman et al. 1984) to fit hierarchical species distribution models (Guisan and Zimmermann 2000). Model variables included remotely-sensed measures of landcover and productivity, and high-resolution interpolated climate data. The spatial resolution of the observational data was consistent with that of most of the variables.

To characterize the environmental factors of greatest importance to boreal songbirds as a group, we summarized the variable contributions from individual regression tree models across species. Our first objective was to quantify the relative contributions of the classes of climate and vegetation variables, and of various subclasses. Within 'climate' we evaluated monthly temperature and precipitation as well as measures of their within and between-year variability. Within 'vegetation' we distinguished landcover type from remotely-sensed indices of productivity and photosynthetic activity. Our second objective was to determine if the relative importance of variable classes differed among migratory groups. We hypothesized that differing exposures, among migratory groups, to cold boreal climates or to extreme events should influence the relative importance of climate variables. We expected the relative importance of climate to be highest in year-round residents and least in long-distance migrants. Finally, we sought to test the hypothesis that climate and vegetation influences on species distributions are hierarchically structured. Interpreting Johnson's (1980) hierarchy of habitat selection, we hypothesised that climatic factors constrain the areas within which species can occur, while vegetation, or the local details of habitat structure, influence abundance within climatic envelopes (Wiens et al. 1987). We were able to reformulate this as a testable hypothesis on the distribution of the regression tree structures.

Material and methods

Study area and avian survey data

Our study region was the boreal biome in Canada, which extends from the Yukon Territory in the west to

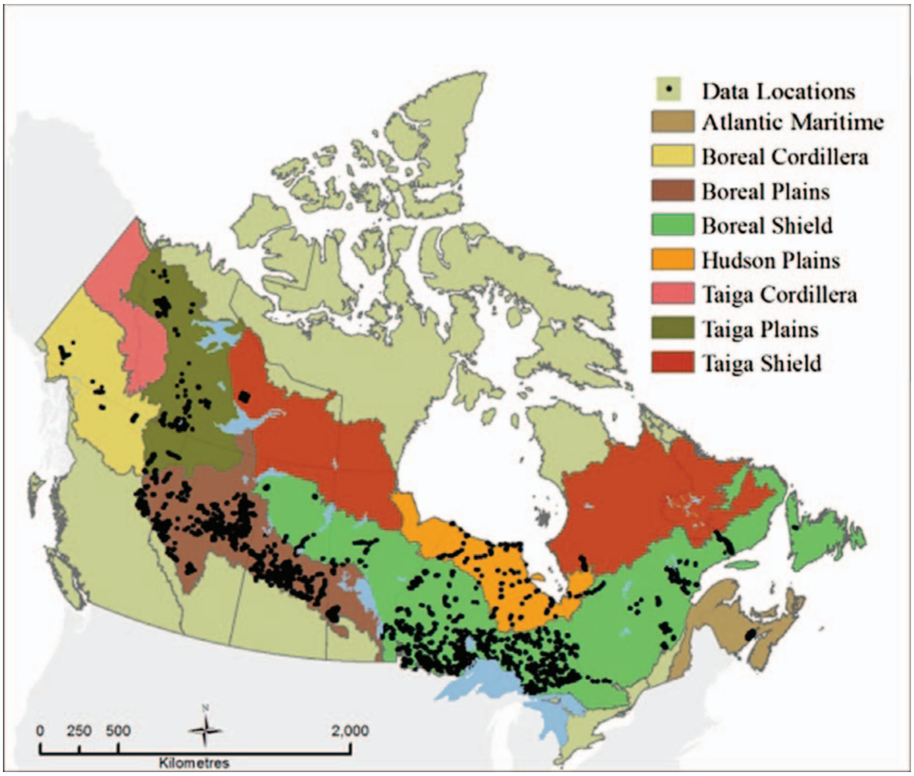

Figure 1. Map of study area with locations of point-count data used in the regression tree models, based on the Boreal Avian Modelling Project database as of February 2009. The dataset represents 53 contributing data partners, 88 research projects, 35 845 locations, and 75 303 point-count surveys.

Newfoundland and Labrador in the east (Fig. 1). Until recently, most of this region was inaccessible by road and sparsely populated; thus road- and volunteer-based surveys such as the Breeding Bird Survey (Sauer et al. 2008) provided very limited coverage. Starting in the early 1990s, academic and government researchers responded to the accelerating pace of industrial development in the region by mounting numerous independent field studies to survey forest songbirds during the breeding season. We analysed data from off-road point-count surveys conducted since 1990, assembled by the Boreal Avian Modelling Project (Cumming et al. 2010). Point counts are a relatively standardised survey method for evaluating habitat requirements of forest songbirds (Ralph et al. 1993), and are widely used in boreal regions (Toms et al. 2006). We used data from 88 distinct projects that collectively surveyed approximately 36 000 unique point-count stations. Spatial sampling effort (Fig. 1) was concentrated in the Boreal Cordillera, Boreal Plains, Boreal Shield, Taiga Plains, Hudson Plains, and Atlantic Maritime ecozones (nomenclature follows Ecological Stratification Working Group 1996). Few studies have been conducted in the more remote Taiga Cordillera and Taiga Shield ecozones.

Environmental variables

We assembled 131 spatial environmental variables (Table 1). Variables were classified hierarchically. The first level distinguished between vegetation and climate variables. Within vegetation, the second level of classification was between landcover and productivity; within climate, the second level distinguished measures of temperature from precipitation. Within these, a third level distinguished between the means and standard deviations of annual values calculated over a 30 yr period.

Vegetation variables

We included descriptions of surface vegetation type (landcover), as well as vegetation productivity—primary productivity, leaf area index, and ‘greenness’ indices (Table 1). Landcover was derived from the 250-m resolution Land Cover Map of Canada 2005 (LCC05); this is on the order of breeding territory sizes (Laurent et al. 2005) and is consistent with the area sampled by point-count surveys. Because some landcover types were absent or uncommon in the boreal region, or poorly represented in the sample, we reclassified the LCC05 legend from 39 to 17 categories (Supplementary material Appendix 1, Table A1). We also included MODIS-based vegetation continuous fields, a multivariate description of the vegetation composition as the proportional coverage of woody plants, herbaceous vegetation and open ground (Table 1). We calculated seven-year means (2000–2006) of gross and net primary productivity (GPP, NPP; Table 1); these measure the amount of photosynthetically active radiation absorbed during the growing season (Zhao et al. 2005). Greenness, or the amount of photosynthesis by terrestrial plants, was measured by

Table 1. Listing and classification of candidate environmental variables for regression tree model development (veg = vegetation; landcov = landcover; product = productivity; clim = climatic; temp = temperature; precip = precipitation; sd = variability). Specific months used varied according to migratory group (Methods). 'Time period' refers to the interval of years over which annual data were used or available. 'Resolution' refers to the approximate spatial precision and scale of measurement for remotely-sensed data; for climate data, it refers to the spatial units to which data were interpolated.

Variable class							
Level 1	Level 2	Level 3	Variable name	Variable description	Time period	Resolution	Source
veg	landcov	NA	LANDCOV	Vegetation class (n = 17)	2005	250 m	MODIS[1]
veg	landcov	NA	BARE	Bare ground proportion (0–1)	2000–2001	500 m	MODIS VCF[2]
veg	landcov	NA	TREE	Tree proportion (0–1)	2000–2001	500 m	MODIS VCF[2]
veg	landcov	NA	HERB	Herbaceous proportion (0–1)	2000–2001	500 m	MODIS VCF[2]
veg	product	NA	LAI*x*	Mean leaf area index (LAI) for month *x*[3]	1998–2005	1 km	SPOT-4[4]
veg	product	NA	GPP	Mean annual gross primary productivity (kg C m^{-2})	2000–2006	1 km	MODIS[5]
veg	product	NA	NPP	Mean annual net primary productivity (kg C m^{-2})	〃	〃	〃
veg	product	NA	NDVI*x*	Mean normalised difference vegetation index (NDVI) from 10-d composite values for month *x*[3]	1998–2005	1 km	SPOT-4[4]
clim	temp	mean	PET	Potential evapotranspiration (cm)[6]	〃	〃	NRCAN[7]
clim	precip	mean	CMI	Climate moisture index (cm)[8]	1971–2000	10 km	〃
clim	temp	mean	GDD > 5	Growing degree days above 5°C	1971–2000	10 km	〃
clim	precip	mean	PCPYR	Annual precipitation (mm)	〃	〃	〃
clim	temp	mean	TAVGYR	Annual mean temperature (°C)	〃	〃	〃
clim	temp	mean	JDAYST	Julian day at start of growing season	〃	〃	〃
clim	temp	mean	JDAYEN	Julian day at end of growing season	〃	〃	〃
clim	temp	mean	GSDAYS	Number of days in growing season	〃	〃	〃
clim	temp	sd	TSEASON	Temperature seasonality (CV)	〃	〃	〃
clim	temp	sd	TMINCOLD	Min temp (°C) of coldest period	〃	〃	〃
clim	temp	mean	TMIN*x*	Mean min temp (°C) for month *x*[3]	1971–2000	Point/10 km	NRCAN[9]
clim	temp	mean	TMAX*x*	Mean max temp (°C) for month *x*[3]	1971–2000	Point/10 km	〃
clim	precip	mean	PCP*x*	Mean precipitation (mm) for month *x*[3]	1971–2000	Point/10 km	〃
clim	temp	mean	TAVG*x*	Mean temperature (°C) for month *x*[3]	1971–2000	Point/10 km	〃
clim	temp	sd	TAVGSD*x*	SD of mean temp (°C) for month *x*[3]	1971–2000	Point/10 km	〃
clim	temp	sd	TMINSD*x*	SD of min temp (°C) for month *x*[3]	1971–2000	Point/10 km	〃
clim	temp	sd	TMAXSD*x*	SD of max temp (°C) for month *x*[3]	1971–2000	Point/10 km	〃
clim	precip	sd	PCPSD*x*	SD of precipitation (mm) for month *x*[3]	1971–2000	Point/10 km	〃

[1] Landcover Map of Canada 2005 (<www.ccrs.nrcan.gc.ca/optical/nlcc_e.php>).
[2] MODIS Vegetation Continuous Fields (<www.glcf.umd.edu/data/vcf>).
[3] x = 3-letter month abbreviation.
[4] Canada 1-km, 10-d, SPOT/VEGETATION composites for growing season 1998–2005 (<http://geogratis.cgdi.gc.ca/>).
[5] Numerical Terradynamic Simulation Group, Univ. of Montana, MODIS GPP/NPP Project (MOD17) (<www.ntsg.umt.edu/>).
[6] Hogg 1997, Penman–Monteith method.
[7] McKenney et al. 2006, <http://cfs.nrcan.gc.ca/projects/3?lang=en_CA>.
[8] Hogg 1997, CMI = cm precipitation – cm potential evaporation per year.
[9] Weather station data interpolated to survey points by Great Lakes Forestry Centre (NRCAN) for model-building; 10-km grids used for prediction.

monthly leaf area index (LAI) and normalised difference vegetation index (NDVI) values. These were obtained from 10-day SPOT-4 imagery products averaged over an eight-year period (1998–2005) (Table 1).

Climate variables

Climate variables were derived from spatially interpolated daily minimum, maximum and mean temperatures and total precipitation (McKenney et al. 2006), summarized annually as monthly means of daily temperatures and monthly totals of daily precipitation for the 30 yr normal period 1971–2000. This period overlapped the sampling dates of most of the avian data, and was also long enough to encompass decadal climate patterns that may affect current distributions. The Julian dates of the beginning and end of the growing season, measures of seasonality and heat sums (McKenney et al. 2006), and indices of potential evapotranspiration and soil moisture availability (Hogg 1997) were also calculated annually over the normal period. We calculated 30-yr means of all the annual statistics (Table 1). We also calculated the 30-yr standard deviations of monthly mean temperatures and total precipitation (Table 1). To minimise spatial error, most climate variables were interpolated to the precise geographic coordinates and elevations of the point-count locations by Natural Resources Canada (Canadian Forest Service, Great Lakes Forestry Centre). For mapping and prediction, we used data interpolated to a 300 arcsecond (roughly 100 km^2) grid. For interpolation methods and other details see McKenney et al. (2006).

Modelling approach

Pairwise correlations among our variables, for example of temperatures in consecutive months, were often high (Cumming and Leblanc 2010). Dimension reduction techniques are sometimes used in such cases to select a small, uncorrelated set of covariates. This simplifies variable selection and, in linear regression models at least, increases the precision of parameter estimates. We did not do this, because our goal was ecological interpretation, not optimal prediction. We wanted to identify groups of variables with high explanatory power across many species, rather than to evaluate a few commonly-used indices chosen a priori. Therefore, we used the model fitting process to identify subsets of variables that were important determinants of species' abundances, in aggregate. The relative importance of different classes of variables was determined by post-hoc analysis of the sample of fitted models, as described below. It was not feasible to develop a priori hypothesis about variable importances for individual species because of the large number evaluated.

Classification and regression trees (CART; Breiman et al. 1984) are widely used to fit species distribution models (Lawler et al. 2006, Prasad et al. 2006), partly because they identify and approximate nonlinear relationships and complex interactions among large sets of predictors (De'ath and Fabricius 2000). CARTs recursively partition the data into a structure like that of a taxonomic key or binary decision tree. A 'leaf' or 'terminal node' represents a relatively homogeneous group of observations. The 'internal nodes' or 'splits' partition a group into two subgroups based on the value of a single variable that minimises a measure of within-group variance. The top-level node, or the first split, is called the 'root'. Each split is said to have a left and a right 'subtree'. By convention, the right subtree has the higher value of the response variable.

We built regression tree models in S-PLUS (< http://spotfire.tibco.com/ >) using the package 'rpart'. We modelled all boreal songbird species that were detected at least at 1% of all stations. The raw data were counts of birds detected at a station over some fixed time period, so they were considered event rates. Accordingly, we used the Poisson splitting method, which is the tree-based analogue of Poisson regression. The response variables were the total counts observed at a station, summed over visits within a year. Samples at the same station in different years were treated as independent. We calculated an exposure term for each visit to account for differences in sampling protocol, and for the effects of Julian date and time of sampling (Supplementary material Appendix 2). The observed counts divided by the exposure equalled the event rate per standard unit effort. This value was the response variable in our models. We used spatial weights to account for clustered sampling (Supplementary material Appendix 2). The minimum size of terminal nodes was set to 200. For each species, we used the default stopping criteria to build an initial tree, and then applied a 10-fold cross-validation procedure to prune the tree back to a parsimonious structure (Venables and Ripley 1994). The pruned trees contained between 2 and 10 terminal nodes, each representing a group of specific variables associated with a mean detection rate, or expected counts. We interpreted the terminal node means as relative abundances rather than densities because the exposure factors did not fully model detection probability.

We used bootstrap sampling (Efron and Tibshirani 1993) to estimate prediction reliability of the terminal nodes (Kuhnert and Mengersen 2003). Because rpart proved unstable in this application we implemented a custom version of its fitting procedure in FORTRAN. For each species' model, we generated 1000 random samples each containing 80% of all locations and refit the model to each sample while holding the tree structure constant. We retained the minimum terminal node size of 200, which sometimes led to further pruning of the tree. The parameter estimates and the mean relative abundances at terminal nodes reported here are the bootstrapped means. Prediction reliabilities are not reported, but see Cumming et al. (2010) for examples.

Canada's boreal avifauna is largely migratory. Only 20% of breeding species are resident in winter while more than 90% of breeding individuals migrate (Schmiegelow and Mönkkönen 2002). The physiological constraints imposed by cold winter conditions are not experienced by most species. More generally, the relevant climate variables for a species depend on when individuals are likely to be present. Accordingly, potential model variables included different subsets of monthly variables, depending on the migratory group (Supplementary material Appendix 3, Table A1), as follows: resident species, all months of the year; short-distance migrants, March–October; long-distance migrants, April–September.

Measuring variable importance

For each model, we calculated the proportion of explained deviance, a measure analogous to the R^2 statistic in linear regression models. We partitioned this value among the splits to give variable-level values, analogous to a partial R^2 (Supplementary material Appendix 3, Table A1). These were weighted by the model-explained deviance and summed over all models or groups of models. This measured variable importance by the proportional contribution of a variable to the total deviance explained by the models. We also calculated frequencies of variable occurrence over all models. Mean variable importance was calculated as the quotient of proportional explained deviance and selection frequency. The importance measures of individual variables were summed by variable class (Table 1) to evaluate the relative importance of climate and vegetation and of their subclasses. This was done for all species combined and separately for each migratory group.

A split on a continuous variable is determined by whether its value is greater than or less than some threshold. The variable has a positive effect when values larger than the threshold are associated with the right subtree and thus a larger mean value of the response. We tested for differing importance of positive and negative effects of the means and standard deviations of monthly temperature and precipitation variables, for all species combined and within migratory groups. As above, importance was measured by the proportion of model deviance explained by a given variable.

Testing hypotheses on tree structure

By definition, the root split explains a plurality of the model deviance. Also by definition, variables within the left and right subtrees generally (although not exclusively) distinguish different levels of low and high mean abundance, respectively. We wanted to determine if different classes of variables (e.g vegetation or climate; Table 1) tended to occur in different parts of the tree (e.g. root node vs right subtrees), and further if any such selection differed between migratory groups. We calculated separate variable selection frequencies for three tree sections: the root node, and left and right subtrees. To evaluate the hierarchy of variables within models, we calculated total selection frequencies by tree section for each variable class, across all species and also for forest-associated species (Supplementary material Appendix 3, Table A1, according to <www.allaboutbirds.org/guide/search>). This was based on the observation that landcover class occurred in the root node only for grassland and wetland birds. For each group of models, we also calculated the total number of nodes in each tree component. The null hypotheses, that class variables were randomly distributed within tree sections, were then tested against multinomial distributions ($\alpha = 0.05$).

Initial inspection of the models led us to formulate the post-hoc hypothesis that positive associations between bird abundance and forested landcover types occur more frequently in the right subtree than in the left subtree or root position. We counted frequencies of positive effects within tree sections for each of the 14 vegetation types and for aggregated forested and non-forested types (Supplementary material Appendix 1, Table A1). We used binomial tests ($\alpha = 0.05$) of significant deviations from random expectations, based on the number of landcover-determined splits in each tree section and the number of vegetation types in each class (13 forested, 3 non-forested).

Results

Regression tree models were constructed for 98 species (Supplementary material Appendix 3). Of the 98 species, 42 were long-distance migrants, 41 were short-distance migrants, and 15 were year-round residents. Model-explained deviance ranged from 15% for chipping sparrow *Spizella passerine* to 80% for Townsend's warbler *Setophaga townsendi*. The species models with the highest explained deviance were generally migratory and had distinct range limits within the study region (Supplementary material Appendix 3, Table A1). Final models ranged in complexity from 2 variables for bobolink *Dolichonyx oryzivorus*, horned lark *Eremophila alpestris*, and Townsend's warbler, to 11 variables for warbling vireo *Vireo gilvus*. The mean number of unique variables per model was 4.9 and the mean number of terminal nodes was 5.2. Explained deviance was not correlated with any of these measures of model complexity.

Across all models, 102 of 131 variables were selected at least once; of these, 13 accounted for 50% of total importance, and 30 variables for 75% (Table 2). Just over half (53) accounted for 90% of total importance. While the variables selected varied widely among species, the individual variables of highest importance across all species were landcover class (LANDCOV), April leaf area index (LAIApr), and mean September temperature (TAVGSep; Table 2). The identity, but not the rank order, of important variables was consistent among migratory groups (Table 2). For long-distance migrants, TAVGSep explained the most variability, followed by mean May minimum temperature (TMINMay), and LAIApr. For short-distance migrants, the top variables were LANDCOV, LAIApr, and TAVGApr; for residents, they were LAIApr, TAVGSep, and LANDCOV.

Variable importance

Climate variables contributed 77% of the total importance across all species—slightly more for long-distance migrants (81%), and slightly less for short-distance migrants (75%) and year-round residents (71%; Fig. 2). Across all species, 65% of the total importance was due to temperature variables, and the greater portion (40%) was explained by mean temperature metrics as compared to 23% for temperature variation. Of the 12% of deviance explained by precipitation variables, 8% was explained by mean vs 4% for variation. Vegetation variables explained the remaining 23%. Of this, productivity metrics explained slightly more (14%) than landcover class (9%; Fig. 2). A vegetation continuous fields variable (% tree cover; Table 1) was selected in only one of 98 models; these variables are not reported further. Although vegetation variables as a whole explained relatively less deviance than climatic factors, the landcover class

Table 2. Cumulative variable importance (up to 0.5) of the most explanatory variables in regression tree models for (A) all species, (B) long-distance migrants, (C) short-distance migrants, and (D) year-round residents. See Table 1 for variable definitions.

(A) All (n = 97)	Cum import	(B) Long-distance (n = 42)	Cum import	(C) Short-distance (n = 41)	Cum import	(D) Residents (n = 15)	Cum import
LANDCOV	0.084	TAVGSep	0.088	LANDCOV	0.114	LAIApr	0.118
LAIApr	0.155	TMINMay	0.150	LAIApr	0.187	TAVGSep	0.229
TAVGSep	0.217	LAIApr	0.205	TAVGApr	0.257	LANDCOV	0.316
TAVGApr	0.258	LANDCOV	0.260	TMINMar	0.318	TMAXSDFeb	0.403
TAVGAug	0.292	PCPSep	0.304	TAVGAug	0.369	TMINFeb	0.464
TMINJul	0.325	TMIN	0.346	PET	0.417	TMAXAug	0.524
TMINMay	0.357	PCPMay	0.387	TMINJul	0.454		
PET	0.385	TMINJul	0.424	TMAXSDMar	0.486		
PCPSep	0.413	NDVISep	0.454	TMAXSDAug	0.514		
TMINMar	0.439	TAVGAug	0.482				
TMAXAug	0.465	PCPSDJun	0.511				
TMINSDJun	0.485						
PCPMay	0.505						

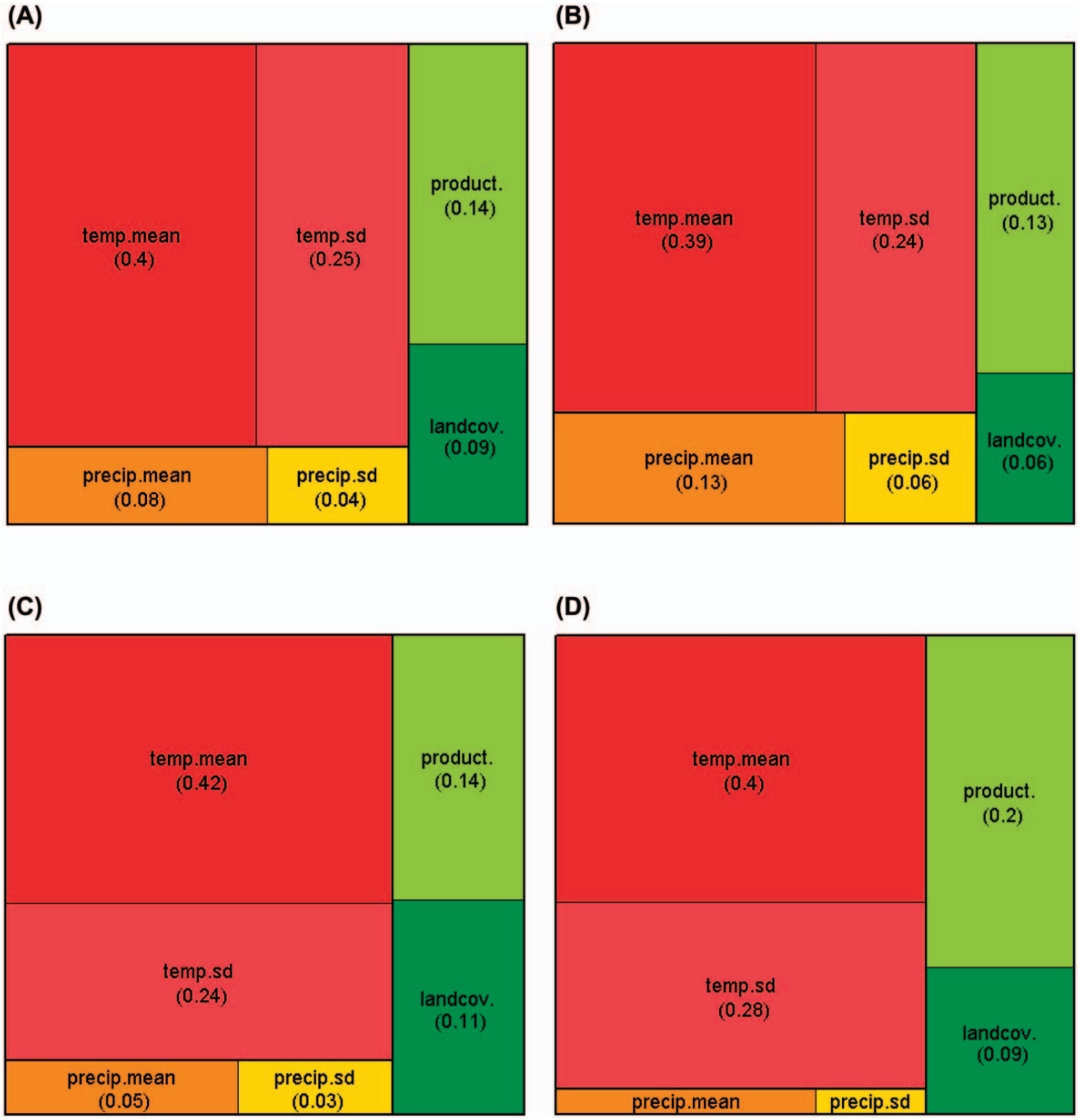

Figure 2. Graphical representation of the relative importance of variable classes (Table 1) over models of (A) all species, (B) long-distance migrants, (C) short-distance migrants, and (D) year-round residents. Variable importance is a measure of the relative contribution of a variable or group of variables to the total deviance explained by a group of models. Climate variables are shown in warm colours; vegetation variables are shown in shades of green. Variable group names refer to pre-defined variable subsets (landcov = landcover; product = productivity; temp = temperature; precip = precipitation; sd = variability) (Table 1).

Table 3. Proportion of total deviance explained by classes of climate variables (Table 1) by positive effects, calculated over all species and by migratory groups (Supplementary material Appendix 3, Table A1).

Species group	precip.mean	precip.sd	temp.mean	temp.sd
All species	0.61	0.70	0.67	0.43
Long-distance migrants	0.56	0.56	0.83	0.35
Short-distance migrants	0.73	1.00	0.50	0.56
Year-round residents	0.76	0.55	0.72	0.25

variable had the largest contribution of any single variable. It was also most frequently selected.

The direction of temperature and precipitation effects on abundance varied across migratory groups (Table 3). Effects of mean temperature, mean precipitation, and precipitation variability were primarily positive across all species (0.67, 0.61 and 0.70 of explained deviance, respectively). Effects of inter-annual temperature variability were primarily negative (1–0.43 = 0.57 of explained deviance; Table 3), particularly during autumn and winter (Supplementary material Appendix 3, Fig. A1). Mean temperature associations for long-distance migrants were mostly positive (0.83 of explained deviance), while effects of temperature variability were primarily negative. Short-distance migrants exhibited more negative associations with mean temperature variables (0.50 of deviance explained, mostly for spring months) compared to other migratory groups, and only positive associations with precipitation variability. Year-round residents demonstrated the most negative responses to temperature variability (1–0.25 = 0.75; Table 3), mostly due to winter and spring variables.

Variable position in trees

Models exhibited highest structural complexity in the right subtree, which contained 50% of all splits, compared to only 30% in the left subtree. Almost all variable classes differed significantly in the rate at which they occurred at the root in comparison to the left and right subtrees (Fig. 3). Variable classes occurring disproportionately at the root were mean temperature and mean precipitation. Temperature variance was underrepresented at root nodes, compared to either subtree. Vegetation productivity occurred disproportionately in left subtrees. Landcover class occurred disproportionately in right subtrees.

Although the landcover class variable occurred primarily in right subtrees, it was present at the root node for seven species, including two long-distance migrants, four short-distance migrants, and one resident, none of which were forest-associated species (Supplementary material Appendix 3, Table A1). For these species, high abundances were associated with non-forested, predominantly agricultural, habitat types.

Multinomial tests revealed that the landcover classes associated with higher mean abundance were significantly related to node location. Positive associations between bird abundance and forested landcover types were significantly more likely to occur in the right subtree than at the root node or in the left subtree, for all migratory groups. When landcover occurred in the right subtree, the distinction between higher and lower abundances was primarily between different types of forested vegetation. When only forest-associated species were analysed, landcover class occurred exclusively in right subtrees (Fig. 3).

Example models for two species

Contrasting examples for the role of vegetation class can be seen in models for the bay-breasted warbler *Setophaga castanea* (Fig. 4) and vesper sparrow *Pooecetes gramineus* (Fig. 5). The bay-breasted warbler is a long-distance migrant known to be associated with outbreaks of spruce budworm *Choristoneura fumiferana* (Venier and Holmes 2011). Based on the root node and map, the northern limits of its Canadian distribution appear to be

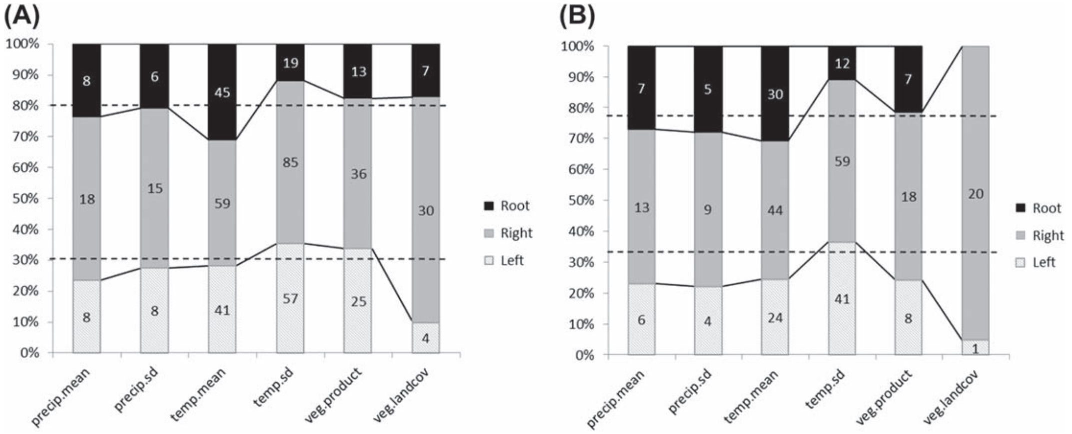

Figure 3. Frequency distributions of variable locations within regression tree models for (A) all species and (B) forest-associated species. Variables are stratified by group (Table 1) and locations are stratified as the root node and the left and right subtrees. Variable counts are standardised by the number of occurrences across all models. Dotted lines represent frequencies expected at random based on the distribution of available nodes. All the frequency distributions differed significantly from the null distribution ($p < 0.05$, based on multinomial tests).

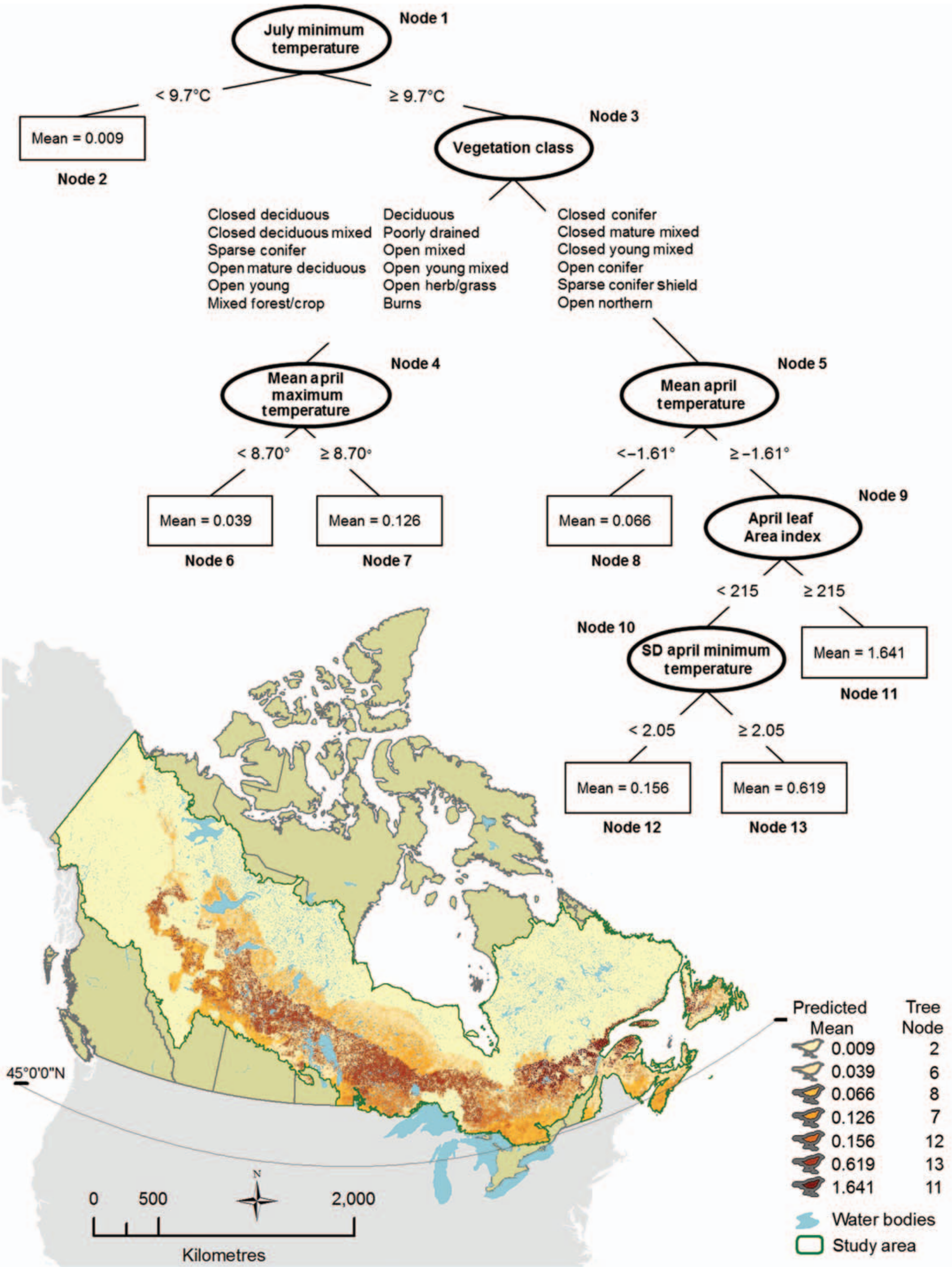

Figure 4. Regression tree model and predictive abundance map for bay-breasted warbler *Setophaga castanea* within the boreal forest region of Canada. See Supplementary material Appendix 3, Table A2 and A3 for additional model details.

associated with a July minimum temperature isocline. This is an example of a positive effect, here of a mean temperature variable. Subject to that constraint, higher abundances are associated with closed mixed forest (young and mature), open coniferous forest (including sparse conifer shield forest), and open northern vegetation types. Within those vegetation types, predicted abundance was highest when mean April temperature was greater than −1.6°C and when April LAI was less than 215 (perhaps an indicator of mixed conifer-hardwood forests), or when the standard deviation of April minimum temperature was greater than 2.0°C.

The vesper sparrow is a short-distance migrant associated with grasslands and open fields (Jones and Cornely 2002). The main split between good and poor habitat (Fig. 5) was vegetation class, with higher abundance associated with mixed forest/cropland. Within ‘good’ habitat, predicted abundance was highest where the standard deviation of

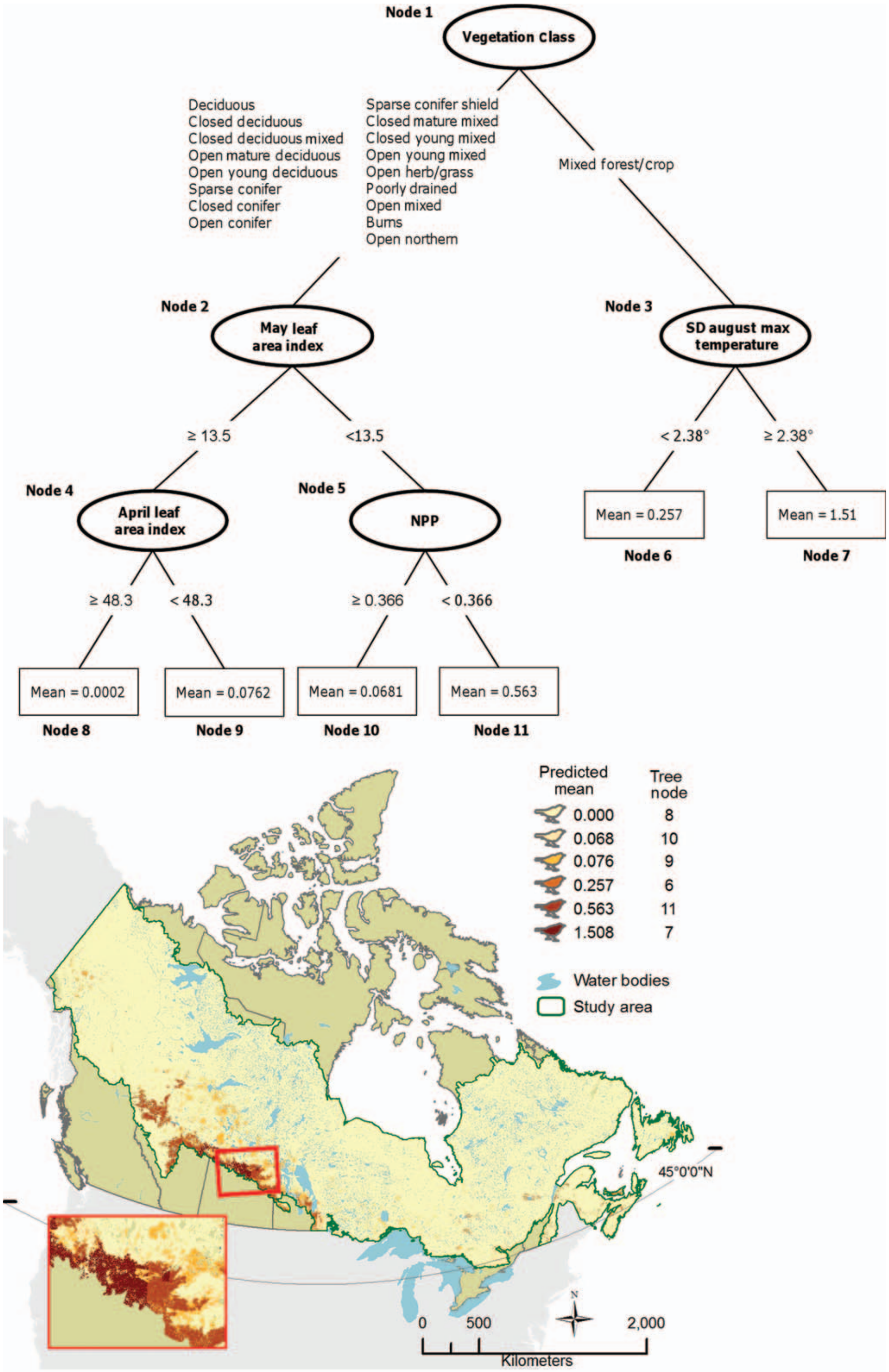

Figure 5. Regression tree model and predictive abundance map for vesper sparrow *Pooecetes gramineus* within the boreal forest region of Canada. The inset shows the distribution of terminal nodes in a region of Saskatchewan where the species is most abundant and highlights the distribution of suitable landcover types along the southern edge of the boreal region, consistent with the conversion of forested to agricultural lands. It also shows the positive effect of the variability of August maximum temperature on abundance of this species, mapped as the vertical split between nodes 6 and 7. See Supplementary material Appendix 3, Table A2 and A3 for additional model details.

August maximum temperature was greater than 2.4°C. Within the left subtree, which describes generally poorer habitat, the highest abundance was found in areas with May LAI less than 13.5 and NPP less than 0.37. We interpret this as a description of the open vegetated habitats that are suitable for this species. The inclusion of LAI and NPP here may compensate for differing classifications of forested and non-forested landcover classes among scenes, or for differences in time of image acquisitions relative to sampling.

Discussion

Climate structures broad-scale avian distributions

Climate variables accounted for the majority (77%) of overall model-explained deviance across all species. Our results support the general finding that avian distributions are well-described by climate characteristics at continental scales (Jiménez-Valverde et al. 2011), perhaps due in part to birds' high mobility and consequent ability to track climate. Indeed, birds are among the taxa most widely documented to have shifted their distributions in response to recent climate change (Parmesan and Yohe 2003, Root et al. 2003).

We found that winter resident species were not more strongly affected by climate variables than were short- and long-distance migrants (70 vs 74 and 81%, respectively). Although low temperatures may ultimately determine the northern distributional limits of resident species via metabolic constraints (Root 1988) the relatively few year-round resident songbirds (15 of 97 species modelled) apparently have wide climatic tolerances (McPherson and Jetz 2007). Indeed, the relationships between mean winter temperatures and relative abundance were generally negative for resident species, indicating this group's tolerance of cold winters. Nonetheless, the northern extent of the boreal forest is the portion for which our data are most limited. Because the range of conditions across the boreal-tundra ecotone is poorly sampled, our ability to determine the absolute importance of climatic factors on winter resident species is limited. Nonetheless, it is notable that the effects of inter-annual variation in temperature were largely negative for this group.

Many migratory bird species have latitudinal range limits that are well-covered by our datasets. This may explain the slightly higher explanatory power of climate variables for these groups, relative to winter residents. In fact, compared to residents, the breeding distributions of migratory species may be more climatically limited due to their affiliations with warmer southern climates during the winter (especially for long-distance migrants). Nevertheless, the relatively small differences in climate variable importance across migratory groups suggests that the explanatory differences between climate and vegetation variables may have more to do with their scales of influence (i.e. with first vs second order selection), and with the degree to which local vegetation characteristics within the boreal forest are not explained by climate.

Vegetation important locally for many species

The effects of local vegetation heterogeneity controlled by factors other than climate were detected by many of our models. Remotely-sensed vegetation metrics contributed 23% of model-explained deviance, across all species. In fact, landcover type and LAI were quite frequently selected as model variables, landcover being among the most important single variables, over all species and within all migratory groups. This should be expected in the boreal forest, where highly distinct vegetation types are broadly distributed across large climatic gradients but with locally-high structural heterogeneity arising from natural disturbances and underlying physiographic variability (Burton et al. 2008). Buermann et al. (2008) suggested that satellite-derived vegetation characteristics may have greatest explanatory power for species with relatively large ranges and moderately specialised habitat associations, especially within areas of relatively homogeneous climate conditions – a characterisation that applies to many boreal-breeding bird species. For example, avian species diversity is generally highest within the mixedwood portions of the boreal region (Hobson and Bayne 2000, Cavard et al. 2011), and the ratio of conifer to hardwood tree species abundance is an important component of habitat suitability for many individual bird species (Drapeau et al. 2000, Vernier et al. 2008, Bayne et al. 2010). As noted, this aspect of vegetation reflects local edaphic factors and disturbance history, not necessarily climate.

Although landcover class was the single most important variable across all species' models, vegetation productivity metrics combined explained a slightly greater portion of total model deviance (14 vs 9%). This is not surprising given the additional information provided by continuous productivity indices, which have the potential to be of greater direct relevance to birds (Bellis et al. 2008). Various indices of vegetation productivity have also been shown to be important positive predictors of bird species richness at sub-continental scales (Ding et al. 2006, Mönkkönen et al. 2006, Honkanen et al. 2010, Elo et al. 2012). For individual species distributions, we found monthly NDVI metrics to be more explanatory than more complex, model-based measures of net and gross primary productivity.

The most explanatory productivity metric in our models, April leaf area index (LAI), has also been found to be useful in predicting tropical bird species distributions (Buermann et al. 2008). In our case the variable appears to represent conifer cover, as leaf-out of boreal deciduous trees does not occur until later in the spring. Although variation in forest cover between needle-bearing and broadleaved species is represented by the LCC05 product, LAI may do a better job of capturing this spectrum, and its seasonality, in a way that is relevant for birds and less subject to classification error. In contrast, the vegetation continuous field (% herbaceous, tree, bare) measures were rarely selected in our models.

Importance of temperature and precipitation gradients

Within the 77% of overall model deviance explained by climate variables, temperature variables explained the largest portion (65% overall), with multi-year means of monthly mean temperatures explaining the majority of that (40% overall). Continental energy gradients, which are primarily a function of latitude and elevation via solar insolation, are well-recognised drivers of species richness patterns (Hawkins et al. 2003, Evans et al. 2008), especially at northern latitudes (Kerr and Packer 1997). Thus, as would be expected, the majority of associations with mean temperature were positive. This was most notable for long-distance migrants, reflecting their southerly affinities (Honkanen et al. 2010). However, nearly one third of temperature relationships were negative (50% for short-distance migrants), suggesting the potential for negative climate-change effects in the future.

In contrast with studies of species richness patterns in Fennoscandia (Honkanen et al. 2010), but consistent with continental-scale analyses of North American bird diversity (Hawkins et al. 2003), we found solar energy (as represented by mean temperature variables) to be more explanatory than vegetation productivity when assessed over multiple species (41 vs 14% of total explained deviance). This suggests that the influence of energy availability on boreal bird distribution and abundance may in fact be explained largely by physiological tolerances (Currie 1991) rather than food availability through trophic cascades (Wright 1983). Alternatively, birds may be affected indirectly by physiological limits to plant species' distribution and abundance. Unfortunately our analysis does not facilitate the separation of these alternative explanations.

The timing within years of associations with important climate factors seemed to vary among migratory groups in accordance with their breeding phenologies. One of the more notable of the frequently selected climatic variables was July minimum temperature. This value has been used to define the limits of the boreal forest (Köppen 1936), so it is perhaps unsurprising that it should also help to explain boreal bird distributions. It remains unclear if this relationship is related to the energetic requirements of trees, of juvenile birds or of their prey. For short-distance migrants, the most important temperature variables were April and August mean temperature; for long-distance migrants the most important were May minimum and September mean temperature. This is consistent with differing constraints on the timing of arrival (Francis and Cooke 1986, Both and Visser 2001) and breeding season length (Benson and Winker 2001) related to migration distance. Furthermore, long-distance migrants are thought to exhibit greater endogenous control of migration timing compared to short-distance migrants, and may therefore be less able to respond to recent changes in early-season climate conditions (Both and Visser 2001, Butler 2003; but see Jonzén et al. 2006).

Precipitation measures and derived metrics of water availability (e.g. actual evapotranspiration) play an important role in the delineation of vegetation types (Stephenson 1990), and have been shown to be positively correlated with the diversity of many taxa, especially plants (Francis and Currie 2003). However, for animal taxa, Hawkins et al. (2003) postulated that water is more important at lower latitudes, while energy is more important at higher latitudes. Our results are consistent with this hypothesis, in that the relative contribution of precipitation variables was low, only 12% over all species. This may be partially explained by the higher spatial variability of precipitation than of temperature and resulting differences in interpolation error (Nalder and Wein 1998). However, we maintain that the general absence of soil moisture deficits in boreal regions (Hogg 1997) is a more compelling reason for the relatively low importance of precipitation. Mean precipitation variables occurred more frequently than expected at the root node. This may reflect distinct longitudinal changes in precipitation regimes that delineate or at least coincide with limits of species distributions, especially for species limited to the eastern or western sides of the continent.

One third of climate effect due to variability measures

The boreal forest biome is a highly variable system (Pastor et al. 1998). One might therefore expect that spatial patterns of variation in climatic factors would be correlated with boreal bird distributions. Indeed, within climatic factors, the second largest deviance component was explained by inter- and intra-annual variability (28% of total = 25% temperature + 3% precipitation; 36% of climate). The relationships were mostly negative, especially for winter resident species, consistent with expected detrimental population effects of climatic variability and extreme events (Parmesan et al. 2000, Stenseth et al. 2002). Negative effects were most prevalent during autumn and winter months, suggesting that many species are most sensitive to inter-annual variability during pre- and post-breeding season periods. It is not clear whether such variability acts directly (e.g. via adult or juvenile survival rates) or indirectly (e.g. via plant growth or insect population size).

Relationships were sometimes positive, most frequently for short-distance migrants and in summer months. An example is vesper sparrow, which had a positive association with the standard deviation of August maximum temperature in the right subtree (Fig. 5). This variable describes a large region of high inter-annual variability corresponding to the grassland portions of south and central Saskatchewan (Cumming and Leblanc 2010, p. 44) that represent the core of the species' distribution in Canada. Similarly, bay-breasted warbler (Fig. 4) exhibited a positive association with the standard deviation of April minimum temperature in the right subtree (Fig. 5), reflecting the high spring variability that characterises its core distribution within the boreal region. Thus, positive associations with inter-annual variability sometimes reflect broad-scale biome delineations. However, we note that overall across species, climatic variability variables occurred disproportionately in left subtrees, which affects the ecological interpretation. For example, a positive effect of high variation within the left subtree may indicate that a 'risky' area was nonetheless better than one totally inhospitable.

The majority of temperature variability effects were due to inter-annual variation (14% of total), but temperature seasonality, or intra-annual variability, accounted for 9% of overall deviance. Its effects were largely negative and limited to long-distance migrants. The boreal biome is characterised by high temperature seasonality, as well as low temperature extremes (Prentice 1990). Thus, the low explanatory power of this class of variables was somewhat unexpected. This may be simply because the class contained only two members (TSEASON and TMINCOLD), albeit important ones. Areas of high seasonality and low minimum temperatures (i.e. northern interior regions) do represent the distributional margins of many warbler species, mostly long-distance migrants that recolonized boreal forests of post-glacial North America from eastern refugia (Weir and Schluter 2004). It is also possible that the extreme temperature conditions associated with high intra-annual variability are associated with reduced insect productivity. Time-series analysis of avian abundance with respect to annual climate data is required to shed light on this matter. New annual climate variables should be specifically calculated from daily data to yield indices designed to test specific ecological hypotheses regarding the relationship between avian and insect phenologies.

Hierarchical structuring by climate and vegetation

The regression tree models that we used, although tending to have lower prediction accuracy than some other methods (Elith et al. 2008), provided intuitive structured depictions of the factors controlling species' abundances. The locations of variables within the tree-based models provided insight into their roles in structuring avian distributions. As demonstrated by the bay-breasted warbler example, mean temperature variables predominantly occupied the root position in the tree, providing the primary split between low and high habitat suitability, corresponding to first order habitat selection (Johnson 1980). Landcover class did appear at the roots of trees for several non-forest species, but in those cases it exclusively represented a positive association with non-forested vegetation types, and likely served as a surrogate for the forest/non-forest boundary. This was the clear interpretation of the model for vesper sparrow, a grassland-associated species not normally found in closed forest. The transition between boreal forest and prairie/parkland biomes or more precisely, between forest and agricultural lands, has shifted progressively northwards since European settlement in the late 19th century, up to the present day (Hobson et al. 2002). The current distribution of forest and agricultural habitats is therefore not determined by current climate. It follows that landcover or other vegetation variables are particularly important when vegetation is in disequilibrium with climate, a condition expected to apply to much of the planet over the next centuries (Solomon 1986, Schneider et al. 2003, Meier et al. 2012).

In most cases, and as predicted, landcover class entered into the model's right subtree, where the discrimination amongst higher abundance levels occurs. There, it tended to identify areas with the highest mean abundance within the coarse climate envelope identified by higher-level splits. This is consistent with Johnson's (1980) second order selection, and supports the generally accepted notion that species distributions are driven primarily by climatic factors at broad continental scales and primarily by vegetation characteristics at finer regional and local scales (Pearson et al. 2004). This finding also reflects the high level of local vegetation heterogeneity that cannot be explained by climatic factors alone. Instead, local vegetation variability is likely to be driven by edaphic, hydrologic, and historic factors that may also be considered in the modelling of avian distributions, especially in the absence of high-quality remotely sensed landcover data.

Conclusion

This paper is the first comprehensive, data-driven synoptic description of avian species distribution and abundance in the Canadian boreal region. Such models are urgently needed in Canada, to establish baseline conditions and inform conservation planning priorities in this era of rapidly expanding industrial development and rapid climate change. The high relative importance of climate variables suggests a large potential for distributional shifts in response to climate change. While many species may have the potential to increase their distributions, many others, especially some of the short-distance migrants, are likely to suffer range contractions, given their often negative associations with temperature variables. Our findings also illustrate the importance of deriving measures of inter-annual variability, rather than relying on long-term averages simply because these are most readily available. However, this may complicate the problem of predicting species responses to changing climate insofar as inter-decadal patterns are not currently well captured by global climate models (Mehta et al. 2010).

The consistent and significant contribution of landcover to the models suggests that local heterogeneity, dictated by edaphic conditions combined with natural and anthropogenic disturbance regimes, will also be a key consideration in evaluating the potential impacts of climate change on boreal forest species. Furthermore, the existing disequilibrium between climate and vegetation in some areas suggests a potential delay in climate-change response. We conclude that the role of vegetation should not be neglected in species models, even at continental extents.

Our findings depended on the interpretation of patterns of variable selection among multiple models, and would not have been possible based on composite models such as of species richness. The specific finding regarding the importance of landcover variables results from testing novel hypotheses formulated on the structure of regression tree models, which as far as we know has not been done before. We believe there is much more to be learned from the analysis of the large suites of species' models that can now be produced for many taxa by machine learning techniques.

Acknowledgements – This publication is a contribution of the Boreal Avian Modelling (BAM) Project, an international research collaboration for the ecology, management and conservation of boreal birds. We acknowledge the BAM Project funding including Environment Canada and U.S. Fish & Wildlife Service, the BAM data partners and Technical Committee members who made this project possible, listed in full at < www.borealbirds.ca/index.php/acknowledgements >. DS was supported additionally by funding from Alberta Ingenuity Fund, Univ. of Alberta, and by a Vanier Canada Graduate Scholarship. We thank Dan McKenney and Pia Papadol of Natural Resources Canada's Great Lakes Forestry Centre for providing climate data custom-interpolated to the point locations of BAM stations. SC and DS contributed equally to this work.

References

Bayne, E. M. et al. 2010. Comparing the predictive capability of forest songbird habitat models based on remotely sensed versus ground-based vegetation information. – Can. J. For. Res. 40: 65–71.

Bellis, L. M. et al. 2008. Modeling habitat suitability for greater rheas based on satellite image texture. – Ecol. Appl. 18: 1956–1966.

Benson, A.-M. and Winker, K. 2001. Timing of breeding range occupancy among high-latitude passerine migrants. – Auk 118: 513–519.

Blancher, P. J. 2003. The importance of Canada's boreal forest to landbirds. – Bird Studies Canada.

Both, C. and Visser, M. E. 2001. Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. – Nature 411: 296–298.

Bradshaw, C. J. A. et al. 2009. Urgent preservation of boreal carbon stocks and biodiversity. – Trends Ecol. Evol. 24: 541–548.

Breiman, L. et al. 1984. Classification and regression trees. – Wadsworth and Brooks/Cole.

Brown, J. H. 1981. Two decades of homage to Santa Rosalia: toward a general theory of diversity. – Am. Zool. 21: 877–888.

Buermann, W. et al. 2008. Predicting species distributions across the Amazonian and Andean regions using remote sensing data. – J. Biogeogr. 35: 1160–1176.

Burton, P. J. et al. 2006. Sustainable management of Canada's boreal forests: progress and prospects. – Ecoscience 13: 234–248.

Burton, P. J. et al. 2008. Large fires as agents of ecological diversity in the North American boreal forest. – Int. J. Wildland Fire 17: 754–767.

Butler, C. J. 2003. The disproportionate effect of global warming on the arrival dates of short-distance migratory birds in North America. – Ibis 145: 484–495.

Cavard, X. et al. 2011. Importance of mixedwoods for biodiversity conservation: evidence for understory plants, songbirds, soil fauna, and ectomycorrhizae in northern forests. – Environ. Rev. 19: 142–161.

Cody, M. L. 1981. Habitat selection in birds: the roles of vegetation structure, competitors, and productivity. – BioScience 31: 107–113.

Cumming, S. G. and Leblanc, M. 2010. The boreal avian modelling project biophysical dataset. Part 2: population and sample distributions. – <www.borealbirds.ca/library/index.php/technical_reports>.

Cumming, S. G. et al. 2010. Toward conservation of Canada's boreal forest avifauna: design and application of ecological models at continental extents. – Avian Conserv. Ecol. 5: 8.

Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. – Am. Nat. 137: 27–49.

Currie, D. J. et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. – Ecol. Lett. 7: 1121–1134.

De'ath, G. and Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. – Ecology 81: 3178–3192.

Ding, T.-S. et al. 2006. Macro-scale bird species richness patterns of the east Asian mainland and islands: energy, area and isolation. – J. Biogeogr. 33: 683–693.

Drapeau, P. et al. 2000. Landscape-scale disturbances and changes in bird communities of boreal mixed-wood forests. – Ecol. Monogr. 70: 423–444.

Ecological Stratification Working Group 1996. A National Ecological Framework for Canada. Report and national map at 1:7,500,000 scale. – Agriculture and Agri-Food Canada, Research Branch <http://sis.agr.gc.ca/cansis/references/1996ew_a.html>.

Efron, B. and Tibshirani, R. J. 1993. An introduction to the bootstrap. – Chapman and Hall.

Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.

Elo, M. et al. 2012. Energy density and its variation in space limit species richness of boreal forest birds. – J. Biogeogr. 39: 1462–1472.

Evans, K. L. et al. 2008. Spatial scale, abundance and the species–energy relationship in British birds. – J. Anim. Ecol. 77: 395–405.

Field, R. et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. – J. Biogeogr. 36: 132–147.

Flannigan, M. D. et al. 2009. Implications of changing climate for global wildland fire. – Int. J. Wildland Fire 18: 483–507.

Forsman, J. T. and Mönkkönen, M. 2003. The role of climate in limiting European resident bird populations. – J. Biogeogr. 30: 55–70.

Francis, A. P. and Currie, D. J. 2003. A globally consistent richness–climate relationship for angiosperms. – Am. Nat. 161: 523–536.

Francis, C. M. and Cooke, F. 1986. Differential timing of spring migration in wood warblers (Parulinae). – Auk 103: 548–556.

Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.

Hampe, A. and Jump, A. S. 2011. Climate relicts: past, present, future. – Annu. Rev. Ecol. Evol. Syst. 42: 313–333.

Hawkins, B. A. et al. 2003. Energy, water, and broad-scale geographic patterns of species richness. – Ecology 84: 3105–3117.

Hobson, K. A. and Schieck, J. 1999. Changes in bird communities in boreal mixedwood forest: harvest and wildfire effects over 30 years. – Ecol. Appl. 9: 849–863.

Hobson, K. A. and Bayne, E. 2000. Breeding bird communities in boreal forests of western Canada: consequences of "unmixing" the mixedwoods. – Condor 102: 759–769.

Hobson, K. A. et al. 2002. Large-scale conversion of forest to agriculture in the boreal plains of Saskatchewan. – Conserv. Biol. 16: 1530–1541.

Hogg, E. H. 1997. Temporal scaling of moisture and the forest-grassland boundary in western Canada. – Agric. For. Meteorol. 84: 115–122.

Hogg, E. H. and Bernier, P. Y. 2005. Climate change impacts on drought-prone forests in western Canada. – For. Chron. 81: 675–682.

Holdridge, L. R. 1967. Life zone ecology. Revised ed. – Tropical Science Center.

Honkanen, M. et al. 2010. Disentangling the effects of area, energy and habitat heterogeneity on boreal forest bird species richness in protected areas. – Global Ecol. Biogeogr. 19: 61–71.

Jiménez-Valverde, A. et al. 2011. Dominant climate influences on North American bird distributions. – Global Ecol. Biogeogr. 20: 114–118.

Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. – Ecology 61: 65–71.

Jones, S. L. and Cornely, J. E. 2002. Vesper sparrow (*Pooecetes gramineus*). – In: Poole, A. (ed.), The birds of North America online. Cornell Lab of Ornithology.

Jonzén, N. et al. 2006. Rapid advance of spring arrival dates in long-distance migratory birds. – Science 312: 1959–1961.

Kerr, J. T. and Packer, L. 1997. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. – Nature 385: 252–254.

Kerr, J. T. et al. 2007. The macroecological contribution to global change solutions. – Science 316: 1581–1584.

Kissling, W. D. et al. 2008. Spatial patterns of woody plant and bird diversity: functional relationships or environmental effects? – Global Ecol. Biogeogr. 17: 327–339.

Kissling, W. D. et al. 2010. Woody plants and the prediction of climate-change impacts on bird diversity. – Phil. Trans. R. Soc. B 365: 2035–2045.

Köppen, W. 1936. Das geographische System der Klimate. – In: Köppen, W. and Geiger, R. (eds), Handbuch der Klimatologie, vol. 1. Gebrüder Bornträger, pp. C1–C44.

Kuhnert, P. M. and Mengersen, K. 2003. Reliability measures for local nodes assessment in classification trees. – J. Comput. Graph. Stat. 12: 398–416.

Laurent, E. J. et al. 2005. Using the spatial and spectral precision of satellite imagery to predict wildlife occurrence patterns. – Remote Sens. Environ. 97: 249–262.

Lawler, J. J. et al. 2006. Predicting climate-induced range shifts: model differences and model reliability. – Global Change Biol. 12: 1568–1584.

Lee, P.-Y. and Rotenberry, J. T. 2005. Relationships between bird species and tree species assemblages in forested habitats of eastern North America. – J. Biogeogr. 32: 1139–1150.

Luoto, M. et al. 2007. The role of land cover in bioclimatic models depends on spatial resolution. – Global Ecol. Biogeogr. 16: 34–42.

MacArthur, R. H. 1964. Environmental factors affecting bird species diversity. – Am. Nat. 98: 387–397.

Matthews, S. N. et al. 2011. Changes in potential habitat of 147 North American breeding bird species in response to redistribution of trees and climate following predicted climate change. – Ecography 34: 933–945.

McKenney, D. W. et al. 2006. The development of 1901–2000 historical monthly climate models for Canada and the United States. – Agric. For. Meteorol. 138: 69–81.

McPherson, J. M. and Jetz, W. 2007. Effects of species' ecology on the accuracy of distribution models. – Ecography 30: 135–151.

Mehta, V. et al. 2010. Decadal climate predictability and prediction: where are we? – Bull. Am. Meteorol. Soc. 92: 637–640.

Meier, E. S. et al. 2012. Climate, competition and connectivity affect future migration and ranges of European trees. – Global Ecol. Biogeogr. 21: 164–178.

Mönkkönen, M. et al. 2006. Energy availability, abundance, energy-use and species richness in forest bird communities: a test of the species–energy theory. – Global Ecol. Biogeogr. 15: 290–302.

Nalder, I. A. and Wein, R. W. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. – Agric. For. Meteorol. 92: 211–225.

Parmesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. – Nature 421: 37–42.

Parmesan, C. et al. 2000. Impacts of extreme weather and climate on terrestrial biota. – Bull. Am. Meteorol. Soc. 81: 443–450.

Pastor, J. et al. 1998. Sustainability and resilience in boreal regions: sources and consequences of variability. – Conserv. Ecol. 2: 16.

Payette, S. et al. 2008. Spatially explicit fire-climate history of the boreal forest-tundra (eastern Canada) over the last 2000 years. – Phil. Trans. R. Soc. B 363: 2299–2314.

Pearson, R. G. et al. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. – Ecography 27: 285–298.

Prasad, A. M. et al. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. – Ecosystems 9: 181–199.

Prentice, K. C. 1990. Bioclimatic distribution of vegetation for general circulation model studies. – J. Geophys. Res. 95: 11811–11830.

Ralph, C. J. et al. 1993. Handbook of field methods for monitoring landbirds. PSW-GTR-144. – Pacific Southwest Research Station.

Root, T. L. 1988. Energy constraints on avian distributions and abundances. – Ecology 69: 330–339.

Root, T. L. et al. 2003. 'Fingerprints' of global warming on wild animals and plants. – Nature 421: 57–59.

Rotenberry, J. T. and Wiens, J. A. 1980. Habitat structure, patchiness, and avian communities in North American steppe vegetation: a multivariate analysis. – Ecology 61: 1228–1250.

Ruckstuhl, K. E. et al. 2008. Introduction. The boreal forest and global change. – Phil. Trans. R. Soc. B 363: 2243–2247.

Sauer, J. R. et al. 2008. The North American breeding bird survey, results and analysis 1966–2007. Version 5.15.2008. – USGS Patuxent Wildlife Research Center.

Schmiegelow, F. K. A. and Mönkkönen, M. 2002. Habitat loss and fragmentation in dynamic landscapes: avian perspectives from the boreal forest. – Ecol. Appl. 12: 375–389.

Schneider, R. J. et al. 2003. Managing the cumulative impacts of land uses in the Western Canadian Sedimentary Basin: a modeling approach. – Conserv. Ecol. 7: 8.

Solomon, A. M. 1986. Transient response of forests to CO_2-induced climate change: simulation modeling experiments in eastern North America. – Oecologia 68: 567–579.

Stenseth, N. C. et al. 2002. Ecological effects of climate fluctuations. – Science 297: 1292–1296.

Stephenson, N. L. 1990. Climatic control of vegetation distribution: the role of the water balance. – Am. Nat. 135: 649–670.

Taylor, A. R. and Chen, H. Y. H. 2011. Multiple successional pathways of boreal forest stands in central Canada. – Ecography 34: 208–219.

Toms, J. D. et al. 2006. Are point counts of boreal songbirds reliable proxies for more intensive abundance estimators? – Auk 438–454.

Venables, W. N. and Ripley, B. D. 1994. Modern applied statistics with S-Plus. – Springer.

Venier, L. and Holmes, S. 2011. Bay-breasted warbler (*Dendroica castanea*). – In: Poole, A. (ed.), The birds of North America online. Cornell Lab of Ornithology.

Venier, L. A. et al. 2004. Climate and satellite-derived land cover for predicting breeding bird distribution in the Great Lakes Basin. – J. Biogeogr. 31: 315–331.

Vernier, P. R. et al. 2008. Generalizability of songbird habitat models in boreal mixedwood forests of Alberta. – Ecol. Model. 211: 191–201.

Volney, W. J. A. and Fleming, R. A. 2000. Climate change and impacts of boreal forest insects. – Agric. Ecosyst. Environ. 82: 283–294.

Weir, J. T. and Schluter, D. 2004. Ice sheets promote speciation in boreal birds. – Phil. Trans. R. Soc. B 271: 1881–1887.

Whittaker, R. H. and Niering, W. A. 1975. Vegetation of the Santa Catalina Mountains, Arizona. V. Biomass, production, and diversity along the elevation gradient. – Ecology 56: 771–790.

Wiens, J. A. et al. 1987. Habitat occupancy patterns of North American shrubsteppe birds: the effects of spatial scale. – Oikos 48: 132–147.

Wright, D. H. 1983. Species–energy theory: an extension of species–area theory. – Oikos 41: 496–506.

Zhao, M. et al. 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. – Remote Sens. Environ. 95: 164–176.

Supplementary material (Appendix ECOG-00299 at <www.oikosoffice.lu.se/appendix>). Appendix 1–3.