**Integrating BirdNET and LiDAR for Breeding Bird Habitat Modelling in a Working Forest**

**Sunny Tseng, Dexter Hodder, and Ken Otter**

# Abstract

# Introduction Understanding the spatial distribution of bird species richness, defined as the number of species occurring across a landscape (cite), is crucial for identifying preferred habitats and guiding conservation priorities (Stattersfield et al., 1998; Garcia, 2006). Traditionally, estimates of richness rely on ground-based bird surveys that provide direct observations of species presence. To extend beyond surveyed locations, models have been developed to predict bird diversity using habitat covariates, thereby linking biodiversity patterns to environmental drivers and supporting conservation and management decisions (cite). In working forests, such models are particularly valuable for identifying priority areas for conservation, informing harvest planning, and evaluating the ecological impacts of forest management strategies (cite).

Obtaining accurate and comprehensive data on bird diversity, however, remains challenging. Ground-based surveys are labor-intensive, time-consuming, and often limited in spatial and temporal coverage, making it difficult to monitor large or remote areas effectively (cite). Moreover, habitat covariates derived from conventional remote sensing methods, such as satellite imagery or aerial photography, generally emphasize horizontal landscape features. Yet vertical habitat structure—such as canopy height, foliage density, and understory complexity—plays a critical role in shaping bird communities (cite). These vertical attributes are often poorly captured in traditional remote sensing products, limiting the accuracy and ecological relevance of bird–habitat models (cite).

Passive acoustic monitoring (PAM) has emerged as a powerful and scalable tool for biodiversity assessment. Autonomous recording units (ARUs) can collect continuous or repeated audio data over large spatial and temporal scales with relatively low human effort (cite). This approach enables detection of vocalizing bird species even in remote or difficult-to-access areas, thereby greatly expanding survey coverage. Numerous studies have demonstrated the effectiveness of acoustic monitoring for estimating species presence (cite), activity patterns (cite), abundance (cite), and community composition (cite).

Light Detection and Ranging (LiDAR) technology provides a complementary perspective by offering detailed three-dimensional information on vegetation structure. Unlike conventional remote sensing, LiDAR captures both horizontal and vertical habitat attributes, making it particularly well-suited for studying bird–habitat relationships. Previous research has demonstrated the utility of LiDAR-derived metrics in modeling bird diversity and habitat suitability across diverse ecosystems. However, the integration of LiDAR with passive acoustic monitoring remains relatively unexplored, despite its potential to improve our understanding of the structural drivers of bird richness.

In this study, we integrate passive acoustic monitoring and LiDAR remote sensing to model bird richness within the John Prince Research Forest, a managed working forest in central British Columbia, Canada. Acoustic data were collected using autonomous recording units (ARUs) during three breeding seasons and processed with the BirdNET algorithm to detect and validate bird vocalizations. LiDAR data provided structural habitat covariates characterizing vertical and horizontal forest complexity. Our objectives were to: (1) compile a comprehensive species list and estimate asymptotic site-level richness; (2) identify groups of species with similar habitat-use patterns; and (3) evaluate the relative importance of LiDAR-derived predictors of bird richness.

Understanding the spatial distribution of richness of a certain taxonomy, defined as determining the number of species across a landscape (cite), is crucial for identifying bird-preferred habitats and guiding conservation priorities (Stattersfield et al., 1998; Garcia, 2006). Traditionally, such richness mapping relies heavily on ground-based bird surveys that provide direct observations of species presence. To extend beyond surveyed locations, models have been developed to extrapolate bird diversity using habitat covariates, enabling the creation of spatially continuous maps that support broad-scale conservation planning (cite). These continuous maps are essential for identifying priority areas for conservation efforts, informing local management decisions, and guiding the selection of zones for logging or harvesting in working forests (cite).

Despite the importance mapping bird richness, obtaining accurate and comprehensive data remains a significant challenge. Ground-based bird surveys are labor-intensive, time-consuming, and often limited in spatial and temporal coverage, making it difficult to monitor large or remote areas effectively (cite). Furthermore, while habitat covariates derived from conventional remote sensing techniques, such as satellite imagery or aerial photography, are widely used to model bird diversity (cite), these data typically emphasize horizontal landscape features. However, vertical habitat structure, such as canopy height, foliage density, and understory complexity, plays a critical role in shaping bird communities (cite). Such vertical information is often lacking or poorly characterized in traditional remote sensing products, limiting the accuracy and ecological relevance of biodiversity models (cite).

Passive acoustic monitoring (PAM) has emerged as a powerful and scalable tool to address some of the challenges in bird biodiversity assessment. Automated recording units (ARUs) enabled continuous or repeated audio data over large spatial and temporal scales with relatively low human effort (cite). This approach allows for the detection of vocalizing bird species even in remote or difficult-to-access areas, greatly expanding survey coverage. Numerous studies have demonstrated the effectiveness of acoustic monitoring for estimating species presence (cite), species activity pattern (cite), abundance (cite), and community composition (cite).

Light Detection and Ranging (LiDAR) technology offers a complementary approach by providing detailed three-dimensional information on vegetation structure. Unlike conventional remote sensing methods that capture mainly horizontal habitat features, LiDAR can characterize both horizontal and vertical aspects of the environment, such as canopy height, foliage density, and vertical complexity. These vertical habitat attributes are especially important for many bird species that utilize different forest strata for foraging, nesting, and shelter. Previous research has successfully used LiDAR-derived metrics to model bird diversity and habitat suitability across a variety of ecosystems and geographic regions. However, the combined application of LiDAR and acoustic data remains relatively unexplored, presenting opportunities to enhance biodiversity mapping and ecological understanding.

In this study, we integrate passive acoustic monitoring and LiDAR remote sensing to model bird biodiversity within the John Prince Research Forest, an active working forest in interior British Columbia, Canada. Acoustic data were collected using 66 autonomous recording units (ARUs) deployed over three breeding seasons, while LiDAR data were acquired during the leaf-on period to generate detailed representations of vertical forest structure. We use these complementary datasets to model the relationship between bird species richness and environmental covariates, with the aim of extrapolating biodiversity patterns across the entire research forest. Specifically, our objectives are to: 1) compile a comprehensive species list and estimate site-specific species richness; 2) identify the most informative LiDAR-derived predictors of bird diversity; and 3) generate a spatially continuous heatmap of breeding bird richness using LiDAR-based extrapolation.

# Methods

## Study area and acoustic recording sites

The study was conducted in the John Prince Research Forest (~150 km2 in area), located in central British Columbia, Canada, within the dry sub-boreal spruce biogeoclimatic zone (**Fig. 1A**). Audio data were collected from 2020 to 2022 throughout the breeding season (early May to late July) between daily dawn chorus (4 am to 7am), using Audio Moths (Hill et al., 2019). ARUs were set to record for 1 minute, followed by 4 minutes pause, throughout the three hours recording period, resulting in 36 one-minute recordings per day per site. Each ARU was placed at least two kilometers apart to minimize spatial correlation. Variability in the number of active ARUs at each site/date (**Fig. 1B**) occurred due to setup logistics (e.g., accessibility of sites during early season) and field/equipment challenges (e.g., battery depletion, firmware issues, disturbances by wildlife). We excluded data from sites which had less than 15 days of ARU data to avoid potential bias due to low surveying effort. This resulted in data from 59 sites, with each site having an average of 106 ARU days across three years, ranging from 32 to 182 ARU days.

## Acoustic data processing and target species



**Fig. 1**. Spatiotemporal distribution of surveying effort. Number of active ARUs during the surveying seasons (A) across the study sites and (B) across the surveying period.

Collected acoustic data were analyzed using the BirdNET Analyzer v2.4 model (*Birdnet-Team/BirdNET-Analyzer*, 2021/2025), utilizing the parameters listed in **Table 1**. The entire dataset, comprising 1.5 terabytes of audio, required approximately 72 consecutive hours of processing. An overall bird community list was generated in four steps: (1) BirdNET detections were initially filtered using a confidence threshold of 0.8. For each detected category, five recording segments with the highest confidence scores were manually reviewed. Sound categories with at least one confirmed vocalization were retained, resulting in 136 sound categories. (2) Non-bird categories, such as Car Engine, Red Squirrel, Wood Frog, and Slender Meadow Katydid, were removed, reducing the list to 129 sound categories. (3) Species not listed in the British Columbia Breeding Bird Atlas (*BC Breeding Bird Atlas*, n.d.), which documents species recorded in the Prince George area since 2008, were excluded, leaving 123 species. (4) Species detected at fewer than two sites or on fewer than two days in June (i.e., main breeding season) were excluded, resulting in a final list of 122 species (**Supplementary Table A**). The overall species list included a diverse range of bird taxa such as raptors, waterfowl, warblers, sparrows, flycatchers, woodpeckers, owls, and other families.

We further refined this overall species list to a target clade of passerine species that have confirmed breeding records in the study region, and which are known to be vocally active during the dawn chorus period. These criteria were used to avoid detection bias given that recordings were scheduled during the dawn chorus of the breeding season. Species-specific BirdNET thresholds were then calculated to retain reliable detections, following the methods of Tseng et al., (2025). For each species, stratified sampling was used to select 360 recording segments, with 20 segments drawn from each 0.05 confidence interval bin (ranging from 0.1 to 1.0). Each segment was manually reviewed, by both listening and spectrogram analysis, to classify detections as true or false positives. For each species, a threshold was selected to achieve a minimum precision of 0.95, indicating that at least 95% of the retained detections were true positives (see Tseng et al., 2025 for the full workflow). The final list of target species included 40 passerines (**Table 2**).

**Table 1**. BirdNET algorithms arguments, with default values and the values used in this study.

| Argument | Default value | Used value |
| --- | --- | --- |
| i | none | -- |
| o | none | -- |
| lat | -1 | same as default |
| lon | -1 | same as default |
| week | -1 | same as default |
| slist | none | same as default |
| sensitivity | 1.0 | same as default |
| min\_conf | 0.1 | same as default |
| overlap | 0 | same as default |
| rtype | table | r |
| threads | 1 | 4 |
| batchsize | 1 | 4 |
| locale | en | same as default |
| sf\_thresh | 0.03 | same as default |
| classifier | none | same as default |
| fmin | 0 | same as default |
| fmax | 15000 | same as default |
| output\_file | none | -- |
| skip\_existing\_results | FALSE | TRUE |

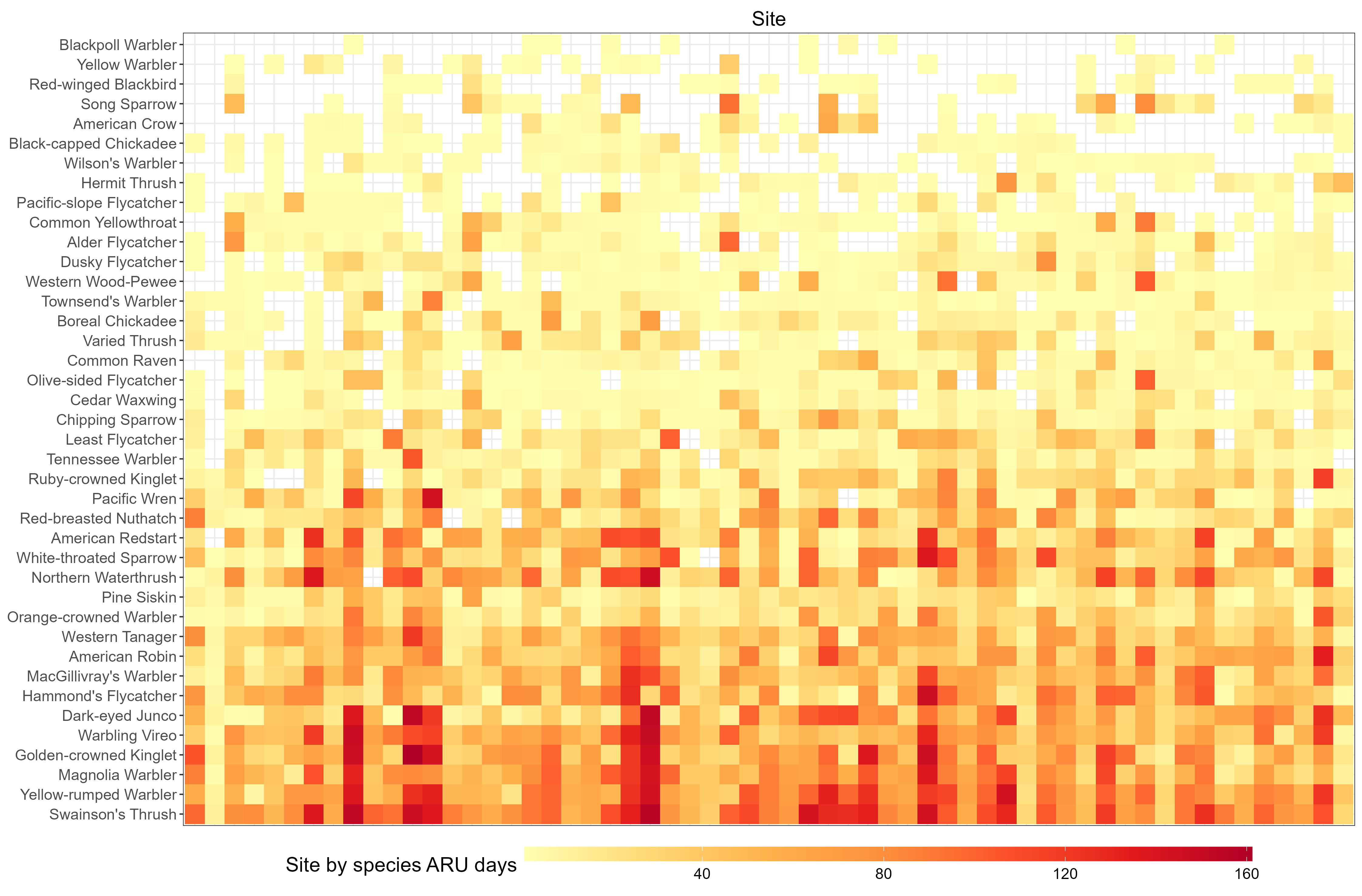
**Table 2**. The 40 breeding dawn chorus passerines selected as target species. For each species, the table reports the total number of BirdNET detections and the number of sites where the species was detected. These values were calculated after applying species-specific BirdNET confidence thresholds that achieved 95% precision, and after excluding sites with fewer than 15 ARU recording days.

| Common name | Scientific name | Threshold (95%) | No. detections | No. sites |
| --- | --- | --- | --- | --- |
| Swainson's Thrush | *Catharus ustulatus* | 0.10 | 327234 | 59 |
| Hammond's Flycatcher | *Empidonax hammondii* | 0.10 | 140580 | 59 |
| Yellow-rumped Warbler | *Setophaga coronata* | 0.10 | 110942 | 59 |
| Warbling Vireo | *Vireo gilvus* | 0.10 | 86199 | 59 |
| Golden-crowned Kinglet | *Regulus satrapa* | 0.10 | 84229 | 59 |
| Magnolia Warbler | *Setophaga magnolia* | 0.10 | 74066 | 59 |
| Dark-eyed Junco | *Junco hyemalis* | 0.10 | 48329 | 59 |
| Western Tanager | *Piranga ludoviciana* | 0.10 | 39421 | 59 |
| American Robin | *Turdus migratorius* | 0.10 | 37644 | 59 |
| MacGillivray's Warbler | *Geothlypis tolmiei* | 0.10 | 30276 | 59 |
| Orange-crowned Warbler | *Leiothlypis celata* | 0.10 | 11907 | 59 |
| Pine Siskin | *Spinus pinus* | 0.10 | 6399 | 59 |
| Northern Waterthrush | *Parkesia noveboracensis* | 0.10 | 145743 | 58 |
| American Redstart | *Setophaga ruticilla* | 0.10 | 74866 | 58 |
| White-throated Sparrow | *Zonotrichia albicollis* | 0.10 | 42408 | 58 |
| Pacific Wren | *Troglodytes pacificus* | 0.10 | 60245 | 57 |
| Red-breasted Nuthatch | *Sitta canadensis* | 0.21 | 15373 | 57 |
| Tennessee Warbler | *Leiothlypis peregrina* | 0.10 | 36095 | 56 |
| Ruby-crowned Kinglet | *Corthylio calendula* | 0.10 | 13718 | 56 |
| Least Flycatcher | *Empidonax minimus* | 0.16 | 36919 | 55 |
| Chipping Sparrow | *Spizella passerina* | 0.10 | 5162 | 54 |
| Cedar Waxwing | *Bombycilla cedrorum* | 0.36 | 1994 | 53 |
| Olive-sided Flycatcher | *Contopus cooperi* | 0.31 | 9201 | 52 |
| Common Raven | *Corvus corax* | 0.10 | 4668 | 52 |
| Varied Thrush | *Ixoreus naevius* | 0.61 | 6693 | 51 |
| Townsend's Warbler | *Setophaga townsendi* | 0.24 | 3039 | 51 |
| Boreal Chickadee | *Poecile hudsonicus* | 0.20 | 2918 | 51 |
| Western Wood-Pewee | *Contopus sordidulus* | 0.10 | 10992 | 50 |
| Dusky Flycatcher | *Empidonax oberholseri* | 0.26 | 3495 | 49 |
| Alder Flycatcher | *Empidonax alnorum* | 0.10 | 21031 | 48 |
| Common Yellowthroat | *Geothlypis trichas* | 0.10 | 4153 | 46 |
| Western Flycatcher | *Empidonax difficilis* | 0.41 | 1306 | 37 |
| Hermit Thrush | *Catharus guttatus* | 0.79 | 3301 | 36 |
| Wilson's Warbler | *Cardellina pusilla* | 0.46 | 603 | 35 |
| Black-capped Chickadee | *Poecile atricapillus* | 0.26 | 606 | 32 |
| American Crow | *Corvus brachyrhynchos* | 0.10 | 1477 | 30 |
| Song Sparrow | *Melospiza melodia* | 0.10 | 3390 | 23 |
| Red-winged Blackbird | *Agelaius phoeniceus* | 0.10 | 204 | 23 |
| Yellow Warbler | *Setophaga petechia* | 0.40 | 394 | 21 |
| Blackpoll Warbler | *Setophaga striata* | 0.85 | 51 | 10 |

## Asymptotic richness detected by ARUs

We analyzed data from 40 target species across 59 sites by calculating, for each site, the number of days on which each species was detected (**Fig. 2**). To characterize similarities in site use among species, we applied k-means clustering to group species with comparable detection patterns across sites. Prior to clustering, we applied data normalization by dividing the number of species ARU days by the total number of ARU days per site, thereby accounting for variation in sampling effort. To determine the optimal number of clusters, we evaluated two approaches: the total within-cluster sum of squares (“elbow method”) using the *factoextra::fviz\_nbclust()* function (cite), and the gap statistic using the *cluster::clusGap()* and *factoextra::fviz\_gap\_stat()* functions (cite). Final clustering was performed using the *kmeans()* function in R.

Direct comparisons of observed species richness across sites (i.e., the number of coloured cells in each column of **Fig. 2**) can be misleading, as survey effort (i.e., number of ARU days) varied among sites. To account for this variation, we used asymptotic richness estimates as a standardized measure of site-level diversity (Chao et al., 2014). We calculated asymptotic richness using the iNEXT (iNterpolation and EXTrapolation) R package (Hsieh et al., 2016), which implements diversity estimation based on Hill numbers. Specifically, we used the *iNEXT::ChaoRichness()* function with incidence data to estimate asymptotic richness and its associated standard errors.



**Fig. 2**. Number of days each species was detected at each site using ARUs. Species in the lower part of the graph, such as American Robin, Dark-eyed Junco, and Golden-crowned Kinglet, were detected across most sites and likely represent generalists. In contrast, species toward the top of the graph, including Yellow Warbler and Red-winged Blackbird, were detected at fewer sites, potentially reflecting more specific habitat preferences.

## LiDAR habitat covariates and model selection

LiDAR data were collected in 2015 during the leaf-off season, with a point density of XYZ. As all sites in the study had not undergone significant disturbance since 2015, the LiDAR should reflect the same relative habitat composition/shifts to allow the use to compare with the audio data. We applied an area-based approach to rasterize the point clouds at a spatial resolution of 10 meters. For modeling purposes, we aggregated habitat covariates within a 100-meter radius around each ARU site to represent the area a bird might be using when detected (cite). A total of 17 LiDAR-derived variables were considered as candidate covariates (**Table 3**). These covariates were classified following the framework outlined in Bakx et al. (2018), which highlights canopy structure and total vegetation metrics as the most commonly associated with bird species richness (see Fig. 4 in Bakx et al., (2019).)

We employed an information-theoretical approach for model selection and inference. Specifically, we fitted models using all possible linear combinations of the LiDAR covariates. Model averaging was then performed to assess variable importance and derive averaged coefficients for prediction. This analysis was conducted using the MuMIn R package (Bartoń, 2025), specifically the *MuMIn::dredge()* function, with the full model being a linear regression model.

**Table 3**. LiDAR-derived habitat covariates used in the analysis, including measures of canopy structure, vegetation density, and ground features.

| Type | Name | Definition |
| --- | --- | --- |
| Topography/  Ground | dem | Average elevation |
| slope | Average slope |
| aspect | Average aspect |
| dist\_wet\_lidar | Shortest distance to the wetland edge |
| d\_LID\_rip\_wet\_str\_le | Distance to LiDAR stream, wetland (polygon), lake edge |
| d\_vri\_polyedge | Distance to VRI polygon stand edge |
| Canopy | cc1\_3 | Average crown closure between 1 – 3 m |
| cc3\_10 | Average crown closure between 3 – 10 m |
| cc10 | Average crown closure above 10m |
| chm | Average canopy height |
| vdi\_95 | Vertical distribution if intercepted canopy elements ranging from 0 and 1 using 95% height |
| Total vegetation | less\_10 | Proportion of polygon with forest age less than 10 years |
| age\_80 | Proportion of polygon with forest age greater than 80 years |
| prop\_decid | Proportion of deciduous tree counts for all trees having DBH > 10cm |
| decid\_dens | Number of deciduous trees per unit area |
| conf\_dens | Number of conifer trees per unit area |
| tree\_dens | Number of total trees per unit area |
| ba\_dens | Sum of basal area of all trees per unit area |

# Result

## Species habitat-use similarity

Based on the species by site distribution (**Fig.2**), k-means clustering grouped species according to their patterns of site use (**Fig.3**). Species within the same cluster exhibited similar distribution patterns across the study sites, which likely reflect shared habitat preferences or frequent co-occurrence. For example, the purple cluster included species such as Yellow-rumped Warbler, Warbling Vireo, and Hammond’s Flycatcher, which were primarily associated with mature coniferous or mixed forest. The orange cluster comprised Ruby-crowned Kinglet, Red-breasted Nuthatch, and Pacific Wren, which were more frequently associated with generalists, both young and matured forests. The green cluster included Chipping Sparrow, Olive-sided Flycatcher, and Tennessee Warbler, which were typically found in early-successional forests with openings and edges and scattered trees (Billerman et al., 2025).

A screenshot of a graph

AI-generated content may be incorrect.

**Fig. 3**. K-means clustering of target species based on site-use patterns, visualized in principal component space. Species within the same cluster exhibit similar distributions across study sites.

## Asymptotic richness estimates

The estimated asymptotic richness for each site ranged from 19.49 to 48.79 species (**Table 4**), with a mean of 36.84 ± 4.48 (SD). Sites were grouped based on survey effort into three categories: low (ARU days < 90), medium (ARU days 90–120), and high (ARU days > 120). The mean difference between estimated asymptotic richness and observed richness was 3.0 species for low effort, 4.6 species for medium effort, and 2.8 species for high effort (Table 4). Species accumulation curves indicated that, regardless of survey effort, the number of species detected reached a plateau at each corresponding effort level (**Fig.4**).

**Table 4**. Estimated asymptotic richness for each site. "ARU days" refers to the number of days the ARU operated at a given site, while "species ARU days" indicates the accumulated number of days a species was detected, summed across multiple species. LCL and UCL are the lower and upper boundary of the estimated asymptotic richness at 95% confidence level.

| Site | ARU days | Species ARU days | Observed richness | Asymptotic richness | No. of species added | LCL | UCL |
| --- | --- | --- | --- | --- | --- | --- | --- |
| N\_15 | 32 | 510 | 34 | 34.24 | + 0.24 | 34.01 | 38.62 |
| N\_25 | 50 | 463 | 31 | 35.41 | + 4.41 | 31.48 | 71.14 |
| N\_11 | 50 | 831 | 34 | 34.16 | + 0.16 | 34.01 | 37.49 |
| 14\_29 | 52 | 606 | 31 | 34.92 | + 3.92 | 31.55 | 59.22 |
| 14\_05 | 59 | 628 | 33 | 38.90 | + 5.9 | 34.05 | 66.10 |
| N\_19 | 65 | 641 | 31 | 48.72 | + 17.72 | 33.49 | 157.04 |
| 14\_02 | 66 | 115 | 19 | 19.49 | + 0.49 | 19.03 | 27.34 |
| N\_23 | 69 | 571 | 31 | 33.22 | + 2.22 | 31.26 | 49.80 |
| 14\_17 | 70 | 753 | 34 | 39.91 | + 5.91 | 35.05 | 67.18 |
| 14\_42 | 71 | 775 | 29 | 29.66 | + 0.66 | 29.06 | 36.72 |
| 14\_23 | 76 | 863 | 37 | 37.99 | + 0.99 | 37.09 | 47.94 |
| 14\_16 | 81 | 1100 | 34 | 34.00 | + 0 | 34.00 | 35.61 |
| N\_05 | 82 | 862 | 31 | 31.16 | + 0.16 | 31.01 | 34.50 |
| 14\_03 | 82 | 990 | 37 | 37.40 | + 0.4 | 37.03 | 42.26 |
| N\_08 | 83 | 453 | 32 | 36.45 | + 4.45 | 32.49 | 72.45 |
| N\_10 | 83 | 1034 | 35 | 41.17 | + 6.17 | 35.96 | 74.68 |
| 14\_35 | 85 | 570 | 31 | 31.10 | + 0.1 | 31.00 | 33.52 |
| 14\_41 | 85 | 959 | 32 | 34.63 | + 2.63 | 32.37 | 50.67 |
| N\_22 | 87 | 926 | 35 | 36.48 | + 1.48 | 35.17 | 47.76 |
| 14\_08 | 89 | 746 | 34 | 35.98 | + 1.98 | 34.27 | 48.23 |
| 14\_30 | 91 | 453 | 31 | 43.12 | + 12.12 | 33.18 | 98.45 |
| 14\_34 | 91 | 1134 | 34 | 34.49 | + 0.49 | 34.04 | 40.19 |
| 14\_21 | 92 | 870 | 36 | 37.48 | + 1.48 | 36.17 | 48.77 |
| N\_16 | 94 | 1012 | 33 | 37.45 | + 4.45 | 33.49 | 73.51 |
| 14\_15 | 98 | 641 | 31 | 35.12 | + 4.12 | 31.67 | 56.54 |
| 14\_19 | 101 | 1145 | 38 | 38.89 | + 0.89 | 38.10 | 46.30 |
| 14\_04 | 102 | 492 | 28 | 48.79 | + 20.79 | 33.02 | 114.06 |
| 14\_18 | 102 | 807 | 30 | 33.96 | + 3.96 | 30.55 | 58.46 |
| 14\_10 | 103 | 865 | 29 | 33.46 | + 4.46 | 29.49 | 69.54 |
| 14\_11 | 104 | 1077 | 34 | 42.91 | + 8.91 | 35.51 | 86.79 |
| 14\_31 | 105 | 1247 | 35 | 35.66 | + 0.66 | 35.06 | 42.75 |
| N\_14 | 105 | 1515 | 35 | 35.99 | + 0.99 | 35.07 | 49.30 |
| N\_21 | 108 | 1068 | 33 | 35.23 | + 2.23 | 33.26 | 51.88 |
| 14\_20 | 109 | 1015 | 35 | 38.96 | + 3.96 | 35.55 | 63.48 |
| 14\_24 | 109 | 1166 | 37 | 41.13 | + 4.13 | 37.67 | 62.56 |
| N\_18 | 111 | 533 | 29 | 31.64 | + 2.64 | 29.37 | 47.72 |
| N\_09 | 111 | 1421 | 36 | 41.95 | + 5.95 | 36.92 | 74.61 |
| 14\_28 | 114 | 985 | 34 | 37.10 | + 3.1 | 34.50 | 53.15 |
| 14\_06 | 115 | 748 | 27 | 31.46 | + 4.46 | 27.49 | 67.58 |
| N\_04 | 115 | 1654 | 35 | 36.98 | + 1.98 | 35.18 | 56.95 |
| N\_12 | 118 | 1645 | 35 | 39.46 | + 4.46 | 35.49 | 75.59 |
| 14\_01 | 121 | 915 | 31 | 38.93 | + 7.93 | 31.98 | 95.30 |
| N\_06 | 121 | 1665 | 36 | 45.92 | + 9.92 | 37.85 | 89.05 |
| 14\_33 | 124 | 1115 | 34 | 34.99 | + 0.99 | 34.09 | 44.99 |
| N\_24 | 137 | 1642 | 37 | 37.25 | + 0.25 | 37.01 | 41.71 |
| N\_13 | 139 | 834 | 35 | 36.49 | + 1.49 | 35.17 | 47.81 |
| 14\_07 | 139 | 1404 | 36 | 36.12 | + 0.12 | 36.01 | 38.91 |
| N\_17 | 139 | 1597 | 34 | 34.50 | + 0.5 | 34.03 | 42.39 |
| 14\_37 | 143 | 1300 | 36 | 37.49 | + 1.49 | 36.17 | 48.81 |
| 14\_40 | 143 | 1455 | 35 | 35.08 | + 0.08 | 35.00 | 37.26 |
| 14\_39 | 145 | 888 | 32 | 38.21 | + 6.21 | 32.97 | 71.87 |
| 14\_38 | 147 | 1393 | 35 | 36.49 | + 1.49 | 35.17 | 47.81 |
| N\_07 | 155 | 1180 | 33 | 34.49 | + 1.49 | 33.17 | 45.82 |
| 14\_25 | 155 | 1841 | 37 | 37.99 | + 0.99 | 37.07 | 51.35 |
| N\_01 | 155 | 1893 | 35 | 35.50 | + 0.5 | 35.03 | 43.40 |
| 14\_09 | 164 | 1990 | 37 | 37.99 | + 0.99 | 37.09 | 48.01 |
| 14\_14 | 167 | 1668 | 33 | 37.47 | + 4.47 | 33.49 | 73.68 |
| 14\_13 | 169 | 1877 | 35 | 42.95 | + 7.95 | 35.98 | 99.45 |
| 14\_26 | 182 | 1934 | 36 | 39.98 | + 3.98 | 36.55 | 64.58 |

A graph of different days

AI-generated content may be incorrect.

**Fig. 4**. Rarefaction curves for sites with different survey efforts: (A) ARU days fewer than 90, (B) ARU days between 90 and 120, and (C) ARU days greater than 120. Each semi-transparent line represents the curve for an individual site.

## Distance to stream, wetland, and lake edge has the highest predictive power for richness

Among the 17 LiDAR covariates, the distance to stream, wetland, and lake edge (d\_lid\_rip\_wet\_str\_le) had the highest sum of model weights, with a value of 0.97 (**Fig.5**). This covariate was also the only one to be significant (α = 0.05) in the averaged model, with a coefficient of -0.039 (p-value = 0.01). The predicted intercept in the average model was 38.86, which was also statistically significant. Other variables, such as site aspect (aspect) and crown closure between 1–3 meters (cc1\_3), had lower sums of model weights (below 0.8) and were not significant (α = 0.05) in the averaged model (**Fig.5**).

A graph of a number of blue and red bars

AI-generated content may be incorrect.

**Fig. 5.** The sum of model weights for the 17 LiDAR covariates, representing the importance of the covariates.

# Discussion

Our large-scale analysis revealed clear differences in habitat preferences among bird species, allowing us to distinguish generalists from specialists (Fig. 2). By integrating acoustic data collected across multiple sites, we were able to identify groups of species with similar habitat-use patterns (Fig. 3), providing insight into species assemblages and potential habitat guilds. These patterns align with expectations based on forest structure and successional stage, highlighting how species co-occurrence reflects shared ecological requirements.

A key limitation of our approach lies in the subset of species included in the analysis. While expanding the study to more species could provide broader ecological insight, it necessitated filtering out species with low BirdNET detection performance to ensure reliable data. Consequently, some members of the avian community were excluded, resulting in a slightly less complete picture of overall community-level patterns. Nevertheless, the remaining target species captured a representative subset of passerines that are vocally active during the breeding season.

When combining asymptotic richness estimates with environmental covariates, we found that proximity to streams, wetlands, and lake edges—derived from LiDAR data—was most strongly associated with richness patterns. Interestingly, the most important predictor was not directly related to canopy structure, suggesting that remotely sensed habitat features reflecting landscape configuration may be sufficient to explain broad-scale biodiversity patterns. These results underscore the importance of riparian and wetland-associated habitats for breeding birds in managed forest landscapes.

One challenge we encountered was how to incorporate the uncertainty associated with asymptotic richness estimates into our models. The confidence limits provided by the iNEXT estimator are not symmetrically distributed, complicating traditional resampling approaches such as bootstrapping. Bayesian modeling frameworks may offer a more flexible approach for integrating these uncertainties into predictive models of biodiversity.

Finally, our study highlights the need for clear guidelines in the use of ARU-derived community data. Rarefaction and asymptotic richness estimates provide defensible approaches to account for variation in survey effort and detection probability. Nevertheless, inherent limitations such as intermittent recordings, potential false negatives, and species-specific variation in detection performance should be considered. Alternative methods, including occupancy modeling or detection rate metrics at appropriate temporal scales, may complement richness-based approaches. Establishing standardized protocols will be critical for ensuring the consistent and effective use of ARU data in biodiversity monitoring and community ecology research.

Paragraph1: Our large-scale analysis revealed clear differences in habitat preferences among bird species, distinguishing generalists from specialists (Fig. 2). In addition, acoustic data collected across multiple sites allowed us to identify groups of species that share similar habitat-use patterns (Fig. 3), providing insight into species assemblages and potential habitat guilds.

Paragraph2: A key limitation of our approach is the restriction on the set of species that could be included in the analysis. Although expanding the study to multiple species provides broader ecological insight, it also necessitates filtering out species with poor BirdNET detection performance. As a result, some members of the community are excluded, leading to a less complete picture of community-level patterns.

Paragraph3: When combining species richness estimates with environmental covariates, we found that distance to water—derived from LiDAR data—was most strongly associated with richness patterns. Interestingly, the final selected covariate was not necessarily specific to LiDAR or any canopy related results, suggesting that remotely sensed information in general may be sufficient to map biodiversity patterns, even at broad spatial scales.

Paragraph4: One challenge we encountered was how to incorporate the uncertainty of asymptotic richness estimates into our models. The lower and upper confidence limits (LCL and UCL) provided are not symmetrically distributed, making it difficult to apply bootstrapping approaches to propagate uncertainty. Bayesian modelling may provide a more flexible framework for integrating this form of uncertainty into biodiversity models.

Paragraph5: Our study highlights the need for general guidelines in the use of ARU-derived community data. Rarefaction curves provide a defensible approach for estimating richness, given the inherent limitations of acoustic detection (e.g., recordings every three seconds, potential false negatives, and species-specific variation in BirdNET performance). Alternative approaches, such as occupancy models (e.g., grouping presence/absence at the daily scale) or detection rate metrics, may also be valuable. Clear methodological guidance will be critical for ensuring consistent and effective use of ARU data in biodiversity and community ecology research.

# Conclusion

# Literature cited

Bakx, T. R. M., Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2019). Use and categorization of Light Detection and Ranging vegetation metrics in avian diversity and species distribution research. *Diversity and Distributions*, *25*(7), 1045–1059. https://doi.org/10.1111/ddi.12915

Bartoń, K. (2025). *MuMIn: Multi-Model Inference* (Version 1.48.11) [Computer software]. https://cran.r-project.org/web/packages/MuMIn/index.html

*BC Breeding Bird Atlas*. (n.d.). Retrieved June 1, 2025, from https://www.birdatlas.bc.ca/

Billerman, S. M., Smith, M. G., Sly, N. D., Medrano, F., Kirwan, G. M., & Keeney, B. K. (2025). *Birds of the World*. Cornell Laboratory of Ornithology. https://doi.org/10.2173/bow

*Birdnet-team/BirdNET-Analyzer*. (2025). [Python]. BirdNET-Team. https://github.com/birdnet-team/BirdNET-Analyzer (Original work published 2021)

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67. https://doi.org/10.1890/13-0133.1

Hill, A. P., Prince, P., Snaddon, J. L., Doncaster, C. P., & Rogers, A. (2019). AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*, *6*, e00073. https://doi.org/10.1016/j.ohx.2019.e00073

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*(12), 1451–1456. https://doi.org/10.1111/2041-210X.12613

Tseng, S., Hodder, D. P., & Otter, K. A. (2025). Setting BirdNET confidence thresholds: Species-specific vs. universal approaches. *Journal of Ornithology*. https://doi.org/10.1007/s10336-025-02260-w

# Supplementary

Table A. The full list of 122 species detected by ARUs at the study site. For each species, the table includes the total number of detections by BirdNET and the number of sites where the species was detected. The values were calculated before applying species-specific threshold, and before dropping any unused sites in the analysis. The species represent a diverse range of taxa, including raptors, waterfowl, warblers, sparrows, flycatchers, woodpeckers, owls, and other families.

| **Family** | **Scientific name** | **Common name** | **No. detections** | **No. sites** |
| --- | --- | --- | --- | --- |
| Accipitridae  (Hawks, Eagles, and Kites) | *Astur atricapillus* | American Goshawk | 4713 | 54 |
| *Buteo jamaicensis* | Red-tailed Hawk | 296 | 50 |
| *Buteo platypterus* | Broad-winged Hawk | 3302 | 57 |
| *Haliaeetus leucocephalus* | Bald Eagle | 398 | 30 |
| *Accipiter striatus* | Sharp-shinned Hawk | 116 | 38 |
| Pandionidae  (Osprey) | *Pandion haliaetus* | Osprey | 238 | 39 |
| Anatida  (Ducks, Geese, and Waterfowl) | *Anas platyrhynchos* | Mallard | 478 | 40 |
| *Branta canadensis* | Canada Goose | 788 | 39 |
| *Mergus merganser* | Common Merganser | 866 | 32 |
| *Bucephala clangula* | Common Goldeneye | 82 | 21 |
| *Anas crecca* | Green-winged Teal | 124 | 25 |
| *Lophodytes cucullatus* | Hooded Merganser | 75 | 13 |
| *Bucephala islandica* | Barrow's Goldeneye | 19 | 7 |
| *Cygnus buccinator* | Trumpeter Swan | 97 | 18 |
| *Spatula discors* | Blue-winged Teal | 24 | 7 |
| *Mareca americana* | American Wigeon | 23 | 4 |
| *Aythya collaris* | Ring-necked Duck | 144 | 14 |
| Trochilidae  (Hummingbirds) | *Selasphorus rufus* | Rufous Hummingbird | 177 | 24 |
| Caprimulgidae (Nightjars and Allies) | *Chordeiles minor* | Common Nighthawk | 469 | 45 |
| Laridae  (Gulls, Terns, and Skimmers) | *Chroicocephalus philadelphia* | Bonaparte's Gull | 31 | 6 |
| Scolopacidae  (Sandpipers and Allies) | *Actitis macularius* | Spotted Sandpiper | 297 | 33 |
| *Gallinago delicata* | Wilson's Snipe | 11401 | 37 |
| *Tringa solitaria* | Solitary Sandpiper | 94 | 23 |
| *Tringa melanoleuca* | Greater Yellowlegs | 994 | 48 |
| Alcedinidae  (Kingfishers) | *Megaceryle alcyon* | Belted Kingfisher | 462 | 51 |
| Falconidae  (Falcons and Caracaras) | *Falco sparverius* | American Kestrel | 143 | 42 |
| *Falco columbarius* | Merlin | 528 | 32 |
| Phasianidae  (Pheasants, Grouse, and Allies) | *Canachites canadensis* | Spruce Grouse | 15906 | 58 |
| *Bonasa umbellus* | Ruffed Grouse | 1000 | 49 |
| *Lagopus lagopus* | Willow Ptarmigan | 62 | 23 |
| Gaviidae  (Loons) | *Gavia immer* | Common Loon | 5530 | 57 |
| Gruidae  (Cranes) | *Antigone canadensis* | Sandhill Crane | 535 | 35 |
| Rallidae  (Rails, Gallinules, and Coots) | *Porzana carolina* | Sora | 3448 | 17 |
| Bombycillidae  (Waxwings) | *Bombycilla cedrorum* | Cedar Waxwing | 5771 | 59 |
| *Bombycilla garrulus* | Bohemian Waxwing | 151 | 40 |
| Cardinalidae  (Cardinals and Allies) | *Piranga ludoviciana* | Western Tanager | 39616 | 61 |
| *Pheucticus ludovicianus* | Rose-breasted Grosbeak | 3278 | 59 |
| Certhiidae  (Treecreepers) | *Certhia americana* | Brown Creeper | 15142 | 61 |
| Cinclidae  (Dippers) | *Cinclus mexicanus* | American Dipper | 1364 | 27 |
| Corvidae  (Crows, Jays, and Magpies) | *Perisoreus canadensis* | Canada Jay | 634 | 55 |
| *Corvus corax* | Common Raven | 4672 | 54 |
| *Corvus brachyrhynchos* | American Crow | 1477 | 30 |
| Fringillidae  (Finches, Euphonias, and Allies) | *Pinicola enucleator* | Pine Grosbeak | 1139 | 61 |
| *Spinus pinus* | Pine Siskin | 6400 | 60 |
| *Loxia leucoptera* | White-winged Crossbill | 6125 | 60 |
| *Coccothraustes vespertinus* | Evening Grosbeak | 552 | 58 |
| *Haemorhous purpureus* | Purple Finch | 10420 | 62 |
| *Leucosticte tephrocotis* | Gray-crowned Rosy-Finch | 1210 | 51 |
| *Haemorhous mexicanus* | House Finch | 4431 | 59 |
| *Loxia curvirostra* | Red Crossbill | 165 | 45 |
| Hirundinidae  (Swallows) | *Tachycineta thalassina* | Violet-green Swallow | 485 | 39 |
| *Tachycineta bicolor* | Tree Swallow | 1094 | 27 |
| Icteridae  (Troupials and Allies) | *Agelaius phoeniceus* | Red-winged Blackbird | 204 | 23 |
| *Molothrus ater* | Brown-headed Cowbird | 87 | 28 |
| *Euphagus carolinus* | Rusty Blackbird | 1395 | 36 |
| Motacillidae  (Wagtails and Pipits) | *Anthus rubescens* | American Pipit | 4368 | 58 |
| Paridae  (Tits, Chickadees, and Titmice) | *Poecile hudsonicus* | Boreal Chickadee | 4430 | 57 |
| *Poecile gambeli* | Mountain Chickadee | 12994 | 61 |
| *Poecile atricapillus* | Black-capped Chickadee | 1615 | 54 |
| Parulidae  (New World Warblers) | *Setophaga magnolia* | Magnolia Warbler | 74084 | 62 |
| *Setophaga coronata* | Yellow-rumped Warbler | 111092 | 62 |
| *Setophaga townsendi* | Townsend's Warbler | 7361 | 61 |
| *Setophaga striata* | Blackpoll Warbler | 4619 | 61 |
| *Geothlypis tolmiei* | MacGillivray's Warbler | 30304 | 61 |
| *Leiothlypis peregrina* | Tennessee Warbler | 36766 | 58 |
| *Setophaga ruticilla* | American Redstart | 74872 | 59 |
| *Seiurus aurocapilla* | Ovenbird | 5286 | 59 |
| *Cardellina pusilla* | Wilson's Warbler | 4738 | 60 |
| *Leiothlypis celata* | Orange-crowned Warbler | 11951 | 60 |
| *Parkesia noveboracensis* | Northern Waterthrush | 146613 | 60 |
| *Setophaga petechia* | Yellow Warbler | 3407 | 42 |
| *Geothlypis trichas* | Common Yellowthroat | 4153 | 46 |
| Passerellidae  (New World Sparrows) | *Spizella passerina* | Chipping Sparrow | 5162 | 54 |
| *Zonotrichia albicollis* | White-throated Sparrow | 42421 | 60 |
| *Zonotrichia leucophrys* | White-crowned Sparrow | 69534 | 62 |
| *Junco hyemalis* | Dark-eyed Junco | 48344 | 62 |
| *Passerella iliaca* | Fox Sparrow | 2861 | 60 |
| *Melospiza lincolnii* | Lincoln's Sparrow | 19790 | 59 |
| *Melospiza melodia* | Song Sparrow | 3390 | 23 |
| *Spizella pallida* | Clay-colored Sparrow | 228 | 39 |
| *Melospiza georgiana* | Swamp Sparrow | 1846 | 56 |
| *Zonotrichia atricapilla* | Golden-crowned Sparrow | 140 | 35 |
| *Pooecetes gramineus* | Vesper Sparrow | 456 | 32 |
| Regulidae  (Kinglets) | *Regulus satrapa* | Golden-crowned Kinglet | 84683 | 61 |
| *Corthylio calendula* | Ruby-crowned Kinglet | 13918 | 59 |
| Sittidae  (Nuthatches) | *Sitta canadensis* | Red-breasted Nuthatch | 23207 | 61 |
| Troglodytidae  (Wrens) | *Troglodytes pacificus* | Pacific Wren | 60989 | 59 |
| Turdidae  (Thrushes and Allies) | *Catharus ustulatus* | Swainson's Thrush | 327431 | 62 |
| *Ixoreus naevius* | Varied Thrush | 32529 | 61 |
| *Catharus guttatus* | Hermit Thrush | 52026 | 61 |
| *Turdus migratorius* | American Robin | 37737 | 62 |
| *Myadestes townsendi* | Townsend's Solitaire | 787 | 55 |
| *Sialia currucoides* | Mountain Bluebird | 1238 | 54 |
| Tyrannidae  (Tyrant Flycatchers) | *Empidonax difficilis* | Western Flycatcher | 12474 | 62 |
| *Empidonax hammondii* | Hammond's Flycatcher | 141949 | 62 |
| *Empidonax minimus* | Least Flycatcher | 44654 | 58 |
| *Empidonax oberholseri* | Dusky Flycatcher | 9356 | 58 |
| *Contopus cooperi* | Olive-sided Flycatcher | 17061 | 61 |
| *Empidonax flaviventris* | Yellow-bellied Flycatcher | 5815 | 61 |
| *Empidonax alnorum* | Alder Flycatcher | 21412 | 49 |
| *Tyrannus tyrannus* | Eastern Kingbird | 48 | 26 |
| *Contopus sordidulus* | Western Wood-Pewee | 10992 | 50 |
| Vireonidae  (Vireos, Shrike-Babblers, and Erpornis) | *Vireo gilvus* | Warbling Vireo | 87002 | 62 |
| *Vireo cassinii* | Cassin's Vireo | 6188 | 55 |
| *Vireo olivaceus* | Red-eyed Vireo | 600 | 35 |
| Ardeidae  (Herons, Egrets, and Bitterns) | *Botaurus lentiginosus* | American Bittern | 359 | 34 |
| Picidae  (Woodpeckers) | *Sphyrapicus varius* | Yellow-bellied Sapsucker | 5261 | 57 |
| *Sphyrapicus nuchalis* | Red-naped Sapsucker | 5182 | 57 |
| *Picoides arcticus* | Black-backed Woodpecker | 2928 | 60 |
| *Picoides dorsalis* | American Three-toed Woodpecker | 6417 | 62 |
| *Dryobates villosus* | Hairy Woodpecker | 1589 | 54 |
| *Dryocopus pileatus* | Pileated Woodpecker | 1910 | 56 |
| *Dryobates pubescens* | Downy Woodpecker | 331 | 49 |
| *Colaptes auratus* | Northern Flicker | 950 | 43 |
| Podicipedidae  (Grebes) | *Podiceps grisegena* | Red-necked Grebe | 26 | 8 |
| Strigidae  (Owls) | *Strix varia* | Barred Owl | 498 | 43 |
| *Bubo virginianus* | Great Horned Owl | 4902 | 57 |
| *Asio otus* | Long-eared Owl | 294 | 40 |
| *Strix nebulosa* | Great Gray Owl | 419 | 36 |
| *Aegolius acadicus* | Northern Saw-whet Owl | 2223 | 54 |
| *Glaucidium gnoma* | Northern Pygmy-Owl | 876 | 48 |
| *Aegolius funereus* | Boreal Owl | 5826 | 44 |