**Avian biodiversity mapping using acoustic monitoring and occupancy modelling**

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

# Abstract

# Introduction

Paragraph 1: Importance of biodiversity monitoring in large scale landscapes for evidence-based management.

Mapping biodiversity hotspots, defined as the area with high diversity in species (cite), is crucial for pinpointing bird-preferred habitats and guiding conservation priorities (Stattersfield et al., 1998; Garcia, 2006). This process involves overlaying species distribution maps to identify spatial locations with the highest richness (cite).

Paragraph 2: Current limitation and challenges

However, hotspot mapping requires species distribution maps for all species in the study area, and previous studies have often relied on coarse-resolution distribution maps, such as 5 km grids (Jenkins et al., 2010) or 1.2 km resolution from BirdLife (Merker & Chandler, 2020). While these coarse-resolution hotspots are valuable for large-scale conservation planning, they are less informative for localized management efforts, such as within national or regional parks, where finer resolution is essential.

Paragraph 3: Occupancy modelling with ARU is a statistical approach for estimating species distribution and habitat use. It can deal with non-perfect detection probabilities and providing insights into species occurrence.

Bird distribution maps can be developed using acoustic data combined with occupancy modeling, a type of species distribution model that accounts for imperfect species detection (Campos-Cerqueira & Aide, 2016; Kalan et al., 2015).

Paragraph 4: How combining acoustics monitoring with occupancy modelling can enhance biodiversity assessments by integrating large-scale acoustic data with sophisticated statistical models.

The resolution of these distribution maps is often constrained by the resolution of the environmental covariates, which are typically derived from remote sensing imagery, such as Landsat (30 m) or Sentinel (15 m) data (cite). However, LiDAR, a technology that has gained prominence in recent decades, can provide not only detailed environmental features (~10 m resolution) but also horizontal structural covariates, such as crown coverage and shrub density, which are critical for many songbird species (cite). By integrating acoustic detections with LiDAR-derived environmental covariates through occupancy modeling, it is possible to generate fine-scale species range maps across the landscape. Overlaying these fine-resolution range maps (~10 m) can produce biodiversity a hotspot map that pinpoint areas of highest bird richness, offering valuable insights for conservation within national parks, regional parks, or research forests.

Paragraph 5: Objectives are: 1) provide a biodiversity hotspot map based on acoustic monitoring and occupancy modelling to demonstrate how this approach can reveal critical area for conservation management and illustrate the potential of PAM in large-scale landscape management. And 2) provide a framework for combining the processing of AU data with occupancy modelling.

Using a three-year (2020–2022) dataset from over 60 locations within the John Prince Research Forest (15,000 ha) in interior British Columbia, I will validate detections for all 122 bird species recorded in the forest based on the BirdNET framework developed in my thesis. By integrating acoustic detections with LiDAR-derived environmental covariates through occupancy modeling, I aim to produce a fine-resolution (~10 m) hotspot map, identifying locations with high bird species richness within the John Prince Research Forest. The specific objectives of this research are to: 1) identify key environmental covariates that support high avian occupancy, 2) generate occupancy models and distribution range maps for all 122 bird species detected in the study area, and 3) identify biodiversity hotspots within the research forest by overlaying species distribution maps. This research not only demonstrates the potential of acoustic monitoring but also introduces an innovative approach to mapping avian hotspots at a regional scale, which is applicable to national parks, regional parks, and research forests.

# Methods

## Study area

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 during the breeding season (May to July) between 4am and 7am, using 66 Audio Moths (cite). Recordings were made for 1 minute every 5 minutes. Each ARU was placed at least two kilometers apart to minimize spatial correlation. Variability in the number of active ARUs at each site/date occurred due to setup logistics and field challenges such as battery depletion, firmware issues, or disturbances by wildlife (Fig.1B).



Fig. 1. Number of active ARUs during the surveying seasons.

## Acoustic data processing and target species list

Collected acoustic data were analyzed using the BirdNET Analyzer v2.4 model (cite with GitHub repository), utilizing the Windows Setup option to run the Python module in a local environment (parameters detailed in Table 1). To retain as many detections as possible at the first analysis stage, we set the parameter “min\_conf”, which determines the threshold for ignoring results with confidence below this level, to 0.1. This low threshold was chosen because the optimal confidence threshold varies across species. By keeping as many original detections as possible, we would have the flexibility later to apply species-specific thresholds to filter out false positives. The entire dataset, comprising 1.5 terabytes of audio, required approximately 72 consecutive hours of processing.

The species 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. Categories with at least one confirmed vocalization were retained, resulting in 136 categories. (2) Non-bird categories, such as Car Engine, Red Squirrel, Wood Frog, and Slender Meadow Katydid, were removed, reducing the list to 129 categories. (3) Species not listed in the British Columbia Breeding Bird Atlas (<https://www.birdatlas.bc.ca/>), 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 (main breeding seaason) were excluded, resulting in a final list of 122 species. The final list of 122 species (Table 2) includes a diverse range of taxa such as raptors, waterfowl, warblers, sparrows, flycatchers, woodpeckers, owls, and other families.

For each species, we defined species-specific thresholds for retaining reliable BirdNET detections following the methods recommended by Wood and Kahl (2024) and Tseng et al. (2024). Stratified sampling was used to select 360 recording segments per species, with 20 segments sampled from each 0.5 confidence interval class (ranging from 0.1 to 1.0). Each segment was manually reviewed via listening or spectrogram analysis to classify detections as true or false positives. Logistic regression with a logit link function was applied, modeling BirdNET confidence scores as the predictor and detection accuracy as the response. A threshold achieving a precision of 0.95, indicating that at least 95% of remaining detections were true positives, was then identified for each species. See Tseng et al. 2024 for detailed method. This process was repeated iteratively across all target species.

Table 1. BirdNET algorithms arguments, with default value and the values used in this study. The empty cell in the used value column indicates using the default value.

| 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 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 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 |

## Response variable: detection non-detection matrix

For each species, we prepared the detection-non-detection matrix for occupancy modeling by first applying species-specific confidence thresholds to ensure high precision in the retained detections (Table 2). Detections were grouped into five-day periods, resulting in 18 observation periods per year (6 periods for the month of June, across 3 years of data). For each observation period, we evaluated whether the focal species was present (1), absent (0), or if the ARU was not functioning (NA). Presence was defined as at least one detection of the species within the five-day interval, while absence was indicated by periods without any detections. This process resulted in a binary matrix where rows correspond to sites, columns represent five-day observation periods, and cell values indicate detection (1), non-detection (0), or not available (NA) (Fig. 2). This matrix served as the response variable for subsequent occupancy modeling. We repeated this process for each of the focal species.

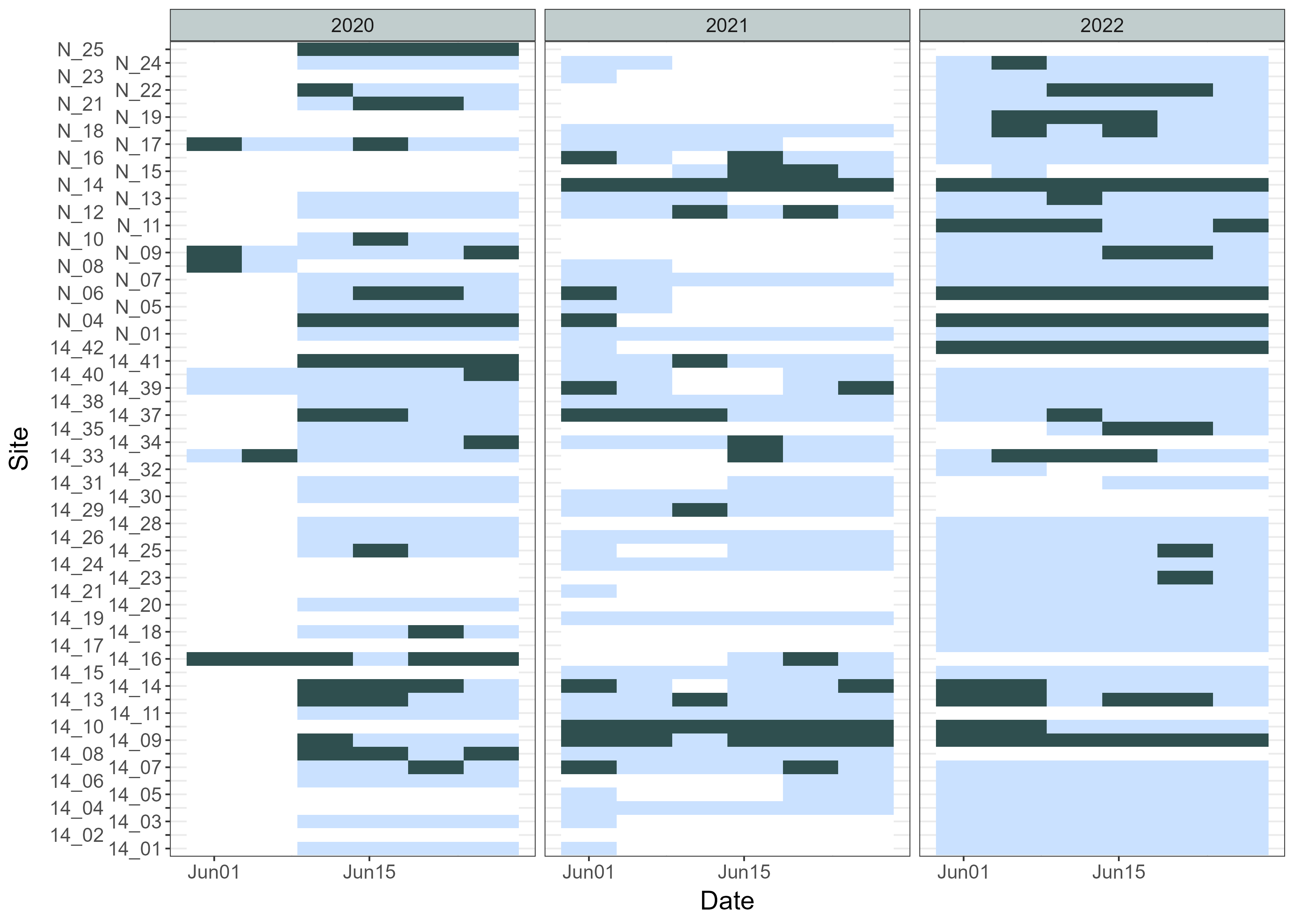


Fig. 2. Detection non-detection matrix for Olive-sided Flycatcher. We can see the presence and absence of the species given a certain site and time period.

## Predictor variables: Lidar covariates

The Lidar data was collected with a point density of XYZ during 2015 leave off season. We used area-based approach to rasterize the point clouds, with a spatial resolution of 10m. We used the canopy and multi-layer metrics, given these LiDAR metrics were shown to be most relevant to bird distribution modelling (Bakx et al., 2019).

Occurrence: cc1\_3m, cc3\_10m, cc10m, chm, wet, decid\_dens, conf\_dens, dist\_wet\_lake\_stream\_lidar, Coarse Woody Debris, CC\_0\_1

Detection: slope, tree\_dens, time of the recording,

## Occupancy modelling

The use of the spOccupancy package – using single or multiple species occupancy modelling

* Guilds such as cavity nesters, warblers, woodpeckers, etc.
* Break up into family group

# Result

Asymptotic richness in sites

Relationship to habitat covariates

Estimation from LiDAR layers

# Discussion

# Conclusion

# Literature cited