Appendix 1

Evaluating trade-offs in spatial versus temporal replication when estimating avian community composition and predicting species distributions

Forest composition varied substantially over the extent of the study area. In southern portions of the study area, upland forest stands are generally dominated by trembling aspen (*Populus tremuloides*), balsam poplar (*P. balsamifera*), white spruce (*Picea glauca*), balsam fir (*Abies balsamea*), and jack pine (*Pinus banksiana*) in varying proportions (Brandt et al. 2013). Further north, coniferous species such as jack pine and black spruce (*Picea mariana*) are more abundant, but become sparse toward the tundra. Across the study area, lowlands mostly consist of a mix of treed wetlands dominated by black spruce or tamarack (*Larix laricinia*), open bogs dominated by ericaceous shrubs, and riparian areas lined with alder (*Alnus* spp.), willow (*Salix* spp.), or birch (*Betula* spp.).

We sampled the avian community using either human observers or autonomous recording units (hereafter ARUs). Following Matsuoka et al. (2014), our standardized point-count methodology specifically included documenting the first minute interval in which each bird was detected and placing birds into the distance interval (0–50, 50–100, ≥100 m) in which they were first detected. These additional data allow correction for components of detection probability following Sólymos et al. (2013). We only conducted surveys during suitable weather conditions from 30 minutes prior to sunrise until four to five hours after sunrise, from the last week of May until the end of the first week in July.

We manually processed one 10-minute and five 3-minute recordings per SSU for recordings made between 2017–2019 in the Saskatchewan/Manitoba study area, and six 3-minute recordings from the Northwest Territories portion of the study area. Potential transcribers were required to take a species identification exam in which they were given 80 recordings of known species identity songs and calls from which all identifying meta-data was removed from the sound files. Transcribers had to achieve a score of ≥ 80% prior to processing data. We used a stratified random draw of available recordings at an SSU to select recordings for processing. We stratified by time of day (early dawn chorus (i.e. 1 hour prior to sunrise until 50 minutes after sunrise), mid dawn chorus (51–150 minutes after sunrise), and late dawn chorus (151–300 minutes after sunrise)) and date of deployment (before or after the median dates available for the site). We generated between four to six randomized samples (depending on availability) for each temporal stratum and selected the first recording draw for analysis; however, this was replaced if there was excessive noise interference (e.g., strong wind, rain) with the next recording in the random selection until an acceptable recording was found. Like the point-count data, ARU recordings were all processed using time of first detection (Farnsworth et al. 2002, Sólymos et al. 2013) sampling using one-minute intervals. Expert listeners estimated counts of individuals using a combination of stereo effect, signal strength, timing of counter-singing events, and occasionally individual variation in song spectral characteristics. Previous work has shown that abundance estimates from ARUs correlate strongly with abundance estimates from point-counts (Van Wilgenburg et al. 2017); however, experimental evidence suggests that abundance estimates tend to be underestimated when birds are very (≥ circa 7 individuals) abundant (Drake et al. 2016). We standardized effort between human point count and ARU surveys by calculating the sum of the counts over the first three one-minute intervals.

We selected survey locations using a spatially balanced hierarchical sampling design previously described by Van Wilgenburg et al. (2020). In brief, sampling was stratified by habitat nested within ecoregions in a sample frame consisting of 5-km wide hexagons encompassing the extent of boreal Canada (as described by Brandt et al. (2013)) plus a buffer extending 100 km into the tundra. The 5-km wide hexagons represent the primary sampling units (PSUs) in our design, within which secondary sampling units (SSUs, or point-count sampling stations) are nested (Van Wilgenburg et al. 2020). PSUs were randomly selected using the Generalized Random Tessellation Stratified sampling algorithm (Stevens and Olsen 2012) using an unequal probability draw weighted by habitat class availability and inversely proportional to access costs (Van Wilgenburg et al. 2020). Secondary sample units were selected from a systematic grid of potential survey locations (spaced 300 m apart) within each PSU using stratified (by habitat classes) random sampling to draw four SSUs. We used the randomly selected SSUs as centroids in clusters of nine SSUs by selecting the eight nearest neighbors around each of the four randomly selected points resulting in four clusters of nine SSUs per PSU. We attempted to sample two clusters of nine per PSU and preferentially attempted to sample the clusters with the highest inclusion probabilities. The additional clusters were oversamples in case safety or accessibility precluded sampling the target clusters (Van Wilgenburg et al. 2020). We began using a precursor to the Boreal Optimal Sampling Strategy (BOSS) design in Saskatchewan during the summer of 2014 and modified the design using the BOSS approach in 2017. The BOSS sampling design was implemented in the Northwest Territories in 2019. We sampled 3251 SSUs distributed amongst 325 PSUs (Figure 1) across Boreal Taiga Plains (123 PSUs; 1050 SSUs), Taiga Shield and Hudson Plains (88 PSUs; 898 SSUs), and Boreal Softwood Shield (123 PSUs; 1243 SSUs) ecozones; note that the number of PSUs tallied across ecozones exceeds the total number of PSUs owing to one PSU containing SSUs in two ecozones. ARUs were used at a total of 166 PSUs across Boreal Taiga Plains (82 PSUs; 686 SSUs), Taiga Shield and Hudson Plains (10 PSUs; 123 SSUs), and Boreal Softwood Shield (74 PSUs; 509 SSUs) ecozones. The remaining PSUs were sampled with human observers.

LITERATURE CITED

Brandt, J. P., M. D. Flannigan, D. G. Maynard, I. D. Thompson, and W. J. A. Volney. 2013. An introduction to Canada’s boreal zone: ecosystem processes, health, sustainability, and environmental issues. Environmental Reviews 21:207–226. https://doi.org/10.1139/er-2013-0040.

Drake, K. L., M. D. Frey, D. Hogan, and R. Hedley. 2016. Using digital recordings and sonogram analysis to obtain counts of yellow rails. Wildlife Society Bulletin 40:346–354. https://doi.org/10.1002/wsb.658.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. The Auk 119:414–425. https://doi.org/10.1093/auk/119.2.414.

Matsuoka, S. M., C. L. Mahon, C. M. Handel, P. Sólymos, E. M. Bayne, P. C. Fontaine, and C. J. Ralph. 2014. Reviving common standards in point-count surveys for broad inference across studies. The Condor: Ornithological Applications 116:599–608. https://doi.org/10.1650/CONDOR-14-108.1.

Sólymos, P., S. M. Matsuoka, E. M. Bayne, S. R. Lele, P. Fontaine, S. G. Cumming, D. Stralberg, F. K. A. Schmiegelow, and S. J. Song. 2013. Calibrating indices of avian density from non-standardized survey data: making the most of a messy situation. Methods in Ecology and Evolution 4:1047–1058. https://doi.org/10.1111/2041-210X.12106.

Stevens, D. L., and A. R. Olsen. 2012. Spatially-balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278. https://doi.org/10.1198/016214504000000250.

Van Wilgenburg, S. L., C. L. Mahon, G. Campbell, L. McLeod, M. Campbell, D. Evans, W. Easton, C. M. Francis, S. Haché, C. S. Machtans, C. Mader, R. F. Pankratz, R. Russell, A. C. Smith, P. Thomas, J. D. Toms, and J. A. Tremblay. 2020. A cost efficient spatially balanced hierarchical sampling design for monitoring boreal birds incorporating access costs and habitat stratification. PLoS ONE 15:e0234494. https://doi.org/10.1371/journal.pone.0234494.

Van Wilgenburg, S. L., P. Sólymos, K. J. Kardynal, and M. D. Frey. 2017. Paired sampling standardizes point count data from humans and acoustic recorders. Avian Conservation and Ecology 12:13. https://doi.org/10.5751/ACE-00975-120113.