Evaluating the Effectiveness of Biodiversity Surrogates for Conservation Planning in the Boreal Region of Canada

Target journal : Ecosphere

September 13, 2019

Authors: order TBD

Pierre R. Vernier1\* ([pierre.vernier@gmail.com](mailto:pierre.vernier@gmail.com))

Marc Edwards1 (medwards219@gmail.com)

Kimberly A. Lisgo1 (klisgo@ualberta.ca)

Fiona K. A. Schmiegelow1 (fschmieg@gmail.com)

Diana Stralberg1,2 (stralber@ualberta.ca)

Steve Cumming3 (stevec.boreal@gmail.com)

1Department of Renewable Resources, University of Alberta, 751 General Services Building, Edmonton, Alberta T6G 2H1 Canada

2Department of Biological Sciences, University of Alberta, CW 405, Biological Sciences Building, Edmonton, Alberta T6G 2E9 Canada

3Département des sciences du bois et de la foret, Université Laval, Québec, Québec G1V 0A6 Canada

\*Corresponding author: Pierre R. Vernier: [pierre.vernier@gmail.com](mailto:pierre.vernier@gmail.com)

**Abstract (max 350 words)**

**Key Words (max 12 words):** Biodiversity surrogates, effectiveness, conservation planning, boreal region.

# Introduction

Ecological reserves that are designed using modern conservation planning methods should be representative of regional biodiversity (Margules and Pressey 2000). However, comprehensive biodiversity data are generally not available, so conservation practitioners need to use biodiversity surrogates when assessing representation (Caro and O’Doherty 1999, Wiens et al. 2008, Hortal et al. 2009, Urbina-Casanova et al. 2016). The use of surrogates assumes that areas that are representative for them will also be representative for regional biodiversity. Surrogates can be taxonomic, where species data are used as indicators for wider biodiversity, or environmental, where physical and biological data are used (Grantham et al 2010). Environmental surrogates can be discrete classified layers such as physiographic features or land types (e.g. Lombard et al. 2003, Oliver et al. 2004), continuous layers such as climatic variables (e.g. Sarkar et al. 2005, Hanson et al. 2017), or a combination of variables in multivariate space (e.g. Beier and Albuquerque 2015, Albuquerque and Beier 2018). Environmental surrogates have the advantage of being easily mapped over wide areas and therefore provide a low-cost alternative for large scale conservation planning. It is generally assumed that representing a variety of environmental conditions will also represent a range of biodiversity values, however tests of environmental surrogates have produced varying results (e.g. Oliver et al. 2004, Sarker et al. 2005, Grantham et al. 2010, Mellin et al. 2011, Beier et al. 2015, Engelbrecht et al. 2016), as have tests for biodiversity surrogates in general (e.g. Rodrigues and Brooks 2007). Grantham et al (2010) suggest that this variability is due to differences in surrogates used, study regions, measures of surrogate effectiveness, and biodiversity features that the surrogates are intended to represent.

Evaluations of biodiversity surrogates range in their level of integration with conservation planning. Pattern-based tests directly measure the spatial relationship between the surrogate and test features on the landscape (Grantham et al. 2010, Andrefouet et al. 2012). This approach can be used to test, for example, the statistical significance of correlations between surrogates and test features (e.g. Oliver et al. 2004, Su et al. 2004), but does not directly evaluate the effectiveness of surrogates for conservation planning. Selection-based techniques generally involve the selection of conservation areas using the surrogates, then evaluate the representation of the test features in the selected areas (Grantham et al. 2010, Andrefaut et al. 2012). According to Rodrigues and Brooks (2007), this approach focuses on evaluating the extent to which areas selected for surrogates capture the test features.

Methods for conservation planning in the boreal forest region of Canada are increasingly accounting for ecosystem dynamics (Leroux et al. 2007, Leroux and Rayfield 2014). For example, the concept of minimum dynamic reserves (MDR) has been suggested for the construction of ecological benchmarks – large, intact reserves designed to maintain large-scale natural disturbances and other ecological processes (Arcese and Sinclair 1997, Wiersma 2005, Schmiegelow et al. 2014). However, environmental surrogates in the boreal region are under-represented in the current reserve network (Andrew et al. 2014) and expansions to the network are needed if Canada is to meet targets set by the Convention on Biological Diversity (SCBD 2010). Expansions to the reserve network should be science-based (Coristine et al. 2018), and will require the use of biodiversity surrogates, but little has been done to test the effectiveness of surrogates in this region (but see Warman et al. 2004, Powers et al. 2013). Remotely sensed environmental surrogates are increasingly available for the Canadian boreal (e.g. Powers et al. 2013) and recent advances in the availability of species models in the boreal region (Stralberg et al. 2015) make it possible to test the effectiveness of surrogates to represent a range of species across the entire Canadian boreal region. This presents an opportunity to integrate surrogacy evaluations into existing conservation planning frameworks such as those using the National Ecological Framework for Canada (Ecological Stratification Working Group 1996). Environmental surrogates are especially useful in the boreal region where species data are scarce (Schmiegelow et al. 2012).

There's a certain amount of confusion in the use of the terms representativeness and representation. According to Kukkala and Moilanen (2013), "representation can be interpreted as the extent of occurrence of a particular species or other biodiversity feature within a specific area", whereas "representativeness can be seen as a broader concept indicating whether a reserve network represents the full variety of biodiversity at all levels of organization."

In this study we focus on representativeness by evaluating if networks comprised of ecological benchmarks that are representative of four environmental surrogates are also representative of several test features, specifically boreal caribou, songbird species, and waterfowl guilds. For each test feature, we evaluate the hypothesis that there is a positive linear relationship between networks that are representative of the species or groups of species and those that are representative of the surrogates. Rather than constructing and selecting networks using subjective area-based targets, we use watershed catchments as building blocks to create benchmark networks based on criteria of size, intactness and hydrological connectivity. We also address representation with a case study of an ecoregion in north-central Canada that identifies benchmark networks that maximize representation (i.e. extent of occurrence) of test species.

# Methods

## Study area

Our study region is the Canadian boreal forest (Brandt 2013), stratified into ecoregions using the National Ecological Framework for Canada (Ecological Stratification Working Group 1996; Figure 1). Ecoregions are stratified into areas that are characterized by distinctive regional ecological factors, including climate, physiography, vegetation, soil, water, and fauna. A total of 94 ecoregions occur in the boreal region (Table 1) and, among those, benchmark networks were developed for 72 of them, of which 52 contained networks that were considered representative of their ecoregion based on criteria defined in the following paragraphs. Our focus in this study is on the 52 ecoregions with representative benchmark networks.

**Figure 1 & Table 1**

## Biodiversity surrogates

In this paper we define surrogates as a set of environmental features that we expect, when represented in a reserve network, to conserve broader biodiversity. The surrogates serve as indicators of large-scale environmental variation of ecological patterns and processes that are assumed to influence biodiversity at coarse spatial scales. We used four environmental surrogates (Figure 2):

1. Gross primary productivity (GPP) is an indicator of the amount of carbon absorbed by photosynthesis, measured in kg per meter squared per year (C kg/m2/yr). It is calculated as an average of 15 years of data (2000-2014) from the MODIS satellite and has a resolution of 1 km2 (Zhao and Running 2010);
2. Climate moisture index (CMI) is a measure of the water deficit (or surplus) in soil, calculated as the yearly average precipitation minus yearly potential evapotranspiration expressed in units of cm of water per year (Hogg 1997). We used data from Hamann et al. 2013 with resolutions of 1-km2 for western North America and 4-km2 for North America, resampling the latter to 1-km2.
3. Land cover (LCC) is a categorical descriptor of vegetation on the landscape and was measured using the 2010 Land Cover of North America based on 250m MODIS satellite imagery, containing 19 cover types (CEC 2013).
4. Lake edge density (LED) is an indicator of the density of riparian habitat, measured in km/km2 within 100 km2 circular units. This dataset was created using the 1:1,000,000 lakes and wide rivers coverage from the Atlas of Canada’s National Scale Frameworks Hydrology (v6.0, NRCAN 2009);

**Figure 2**

## Test features

We refer to the biodiversity features we are trying to represent as test features or test species. We used predictive maps of caribou habitat and bird densities as test features to evaluate the effectiveness of the four environmental surrogates. For caribou, we used a 1-km2 resolution raster map created from a national resource selection function (RSF) model for boreal woodland caribou in Canada. The map was produced by Environment Canada (2011) and quantifies the predicted probability of habitat selection relative to its availability. The national-scale RSF model was developed using animal location data collected by provincial and territorial wildlife departments. Data points from 581 radio-collared caribou distributed among the 27 caribou ranges from 2000 to 2010 were used.

For birds, we used a set of raster maps representing the predicted densities of 80 species of boreal birds[[1]](#footnote-1). The maps were developed using boosted regression trees (BRT) to model the relationship between each species and several biophysical, climate, and ecological variables that were available consistently across the boreal region (Stralberg et al. 2015). We used map-based predictions rather than the underlying survey data to evaluate test features because, unlike the original data, the maps use bird-habitat relationships to estimate density across the entire region of interest. They are also robust to differences in the variety of field survey methods used across the boreal (Solymos et al. 2013). Moreover, there is a paucity of survey data in many boreal reserve networks constructed using conservation planning methods due to their design characteristics i.e., they are selected to be large and free from human influences implying poor accessibility and high survey cost.

We evaluated 11 individual songbird species and 3 waterfowl guilds (Table 2, Appendix S1) that were selected as species of interest in two related projects in the boreal region of Canada and Alaska. The species were selected in consultation with partner organizations and external experts based on several criteria including threat status, vulnerability to future landscape change, geographic distribution, biology, social/cultural importance, and relavance to forestry. These species all reside within both Canada and Alaska and are representative of broad habitat types that occur across the boreal (Table 2).

We also evaluated three composite metrics, the first comprised of all available bird species (“AllBirds”), the second comprised of 53 species associated with forested habitats (“ForestBirds”), and the third comprised of all waterfowl species (“AllWaterfowl”). For AllBirds, we calculated the composite metric by summing all individual songbird densities to create a map showing overall density for the 80 species. Likewise, we summed individual waterfowl densities to create a map showing overall density for the 17 species.

## Benchmarks networks

For each ecoregion, we generated a set of benchmark areas that were: 1) large enough to sustain the expected maximum fire size in the ecoregion according to Leroux et al.’s (2007) minimum dynamic reserve (MDR) approach, 2) hydrologically connected through the use of catchments as building blocks (Vernier and Lisgo 2011 or Lisgo et al. *in prep*), and 3) largely intact as measured using GFWC’s (2014) Human Access dataset. Benchmark areas were constructed using the Benchmark Builder software (BEACONs 2017) by assembling catchments until the MDR-based target for the ecoregion was reached. MDR values were based on a boreal-wide regionalization of fire regimes and vegetation detailed in Lisgo et al. (*in prep*) and represent the minimum reserve size needed to maintain all habitat types under the regional fire regime. Since Builder constructs benchmarks to be hydrologically connected, and since the ecological framework of Canada does not line up with hydrological boundaries, we constructed benchmarks within the ecoregion plus its intersecting fundamental drainage areas (FDA) delineated using the Atlas of Canada’s National Scale Frameworks Hydrology (NRCAN 2009). All benchmark construction began with seed (headwater) catchments located within the ecoregion, but benchmarks were allowed to grow outside of the ecoregion to maintain hydrologic connectivity and meet size requirements.

We combined benchmarks into networks to ensure that at least some of the networks in each ecoregion were representative of the four biodiversity surrogates. We first determined how many benchmarks were required in each ecoregion to construct a representative network by assessing the ability of networks to represent their respective ecoregions for the four surrogates. Once the number of required benchmarks per network was determined, we constructed all possible combinations of network, or if all combinations was too computationally intensive we constructed a random subset of 10,000 networks. We then removed any highly overlapping network to minimize redundancy in the network sample. This procedure resulted in the creation of a range benchmark networks, from highly representative to minimally representative. We then selected a subset of networks which we considered to be most representative and least representative (non-representative) using the following criteria: representative networks included those networks for which all four surrogates had a dissimilarity value ≤ 0.2 while dissimilarity values for surrogates in non-representative networks were all > 0.2. Additional details on the methods used to construct and select networks are provided in the Supplementary Information.

## Relationship between species and surrogates

Our main objective was to evaluate the relationship between surrogates' representation based on CMI, GPP, LED, and land cover, and species' representation based on boreal caribou habitat, songbird densities, and waterfowl guild densities at the ecoregion-level. For each network in each ecoregion, we calculated dissimilarity metrics for the four surrogates (Figure 2) to quantify how well the networks represented the distribution of surrogates in their ecoregion. The dissimilarity metrics were calculated by comparing the surrogates’ distribution in each network to its distribution in its ecoregion. For the three continuous surrogates, GPP, CMI and LED, we used the Kolmogorov-Smirnov (KS) statistic. For LCC, a categorical surrogate, we used the Bray-Curtis statistic. Both measures range from 0 to 1 and indicate increasing dissimilarity.

We repeated the dissimilarity calculations for the test species. For each individual songbird, we only included ecoregions whose density map intersected its known range boundary[[2]](#footnote-2). This was done to avoid including ecoregions where a species was essentially non-existent but predicted to be present by the BRT models, which tend to predict very low densities rather than absolute zeroes. We calculated the KS statistic for each network and each species by comparing the distribution of density values within the network to its distribution within the ecoregion. Consequently, the KS statistic measures the representativeness of a particular species in a network compared to the ecoregion it is embedded in. We also calculated the representativeness values for several groups of birds (Table 2): all birds (“AllBirds”), forest-associated birds (“ForestBirds”), all waterfowl (“AllWaterfowl”), cavity-nesting waterfowl (“CavityNesters”), ground nesting waterfowl (“GroundNesters”), and overwater nesting waterfowl (“OverwaterNesters”). We did this by summing the density values for species belonging to each group in each ecoregion and subsequently calculating the KS statistic as described above. In using the KS statistic, our focus was on evaluating whether networks were effective at representing the regional variation in habitat quality for the test species.

We developed multiple linear regression models to evaluate the relationship between the representativeness of each test species and the representativeness of four environmental surrogates within benchmark networks at the ecoregion level. We first selected, for each ecoregion with benchmark networks within an ecoregion, the lesser of all representative or non-representative networks (Table 1). We then select up to 10 times as many networks from the other group. For example, for ecoregion 215 which includes 23 representative and 2310 non-representative networks, we selected all 23 representative networks and a random subset of 231 non-representative networks. In most ecoregions, this resulted in all networks being included in the models. We developed regression models for each ecoregion: one for all birds and forest birds, one each for the waterfowl guilds, one for caribou, and one each for the 11 songbird species. For each model for each ecoregion, we related the KS value of the test species (SPP\_KS) to the linear combination of KS and BC values for the predictor variables (CMI\_KS, GPP\_KS, LED\_KS, LCC\_BC):

SPP\_KS = CMI\_KS + GPP\_KS + LED\_KS + LCC\_BC

For each species’ model, we recorded the coefficient values of each predictor variable (surrogate) along with the model R-squared (R2) as a measure of the strength of the relationship and the root mean squared error (RMSE) as a measure of the uncertainty in the mean response. We also report the absolute value of the *t*-statistic for each model parameter as an estimate of the contribution of each variable to the model i.e., variable importance (Kuhn and Johnson 2013). All analyses were carried out using R v. 3.6.1 (R Development Core Team 2019).

We need a small section describing the additional step(s) to create the regression models based on a subset of networks i.e., those that are representative and non-representative based on our criteria.

In the Supplementary Information we provide a case study and example code for running the evaluation of representativeness. We also include an additional analysis of representation which allows for a comparison of the two approaches. For this latter analysis, we evaluated representation using the area-adjusted proportion of the species population or core habitat within the network. This is measured as the ratio of the proportion of population (or core habitat) within a network to the proportion of the ecoregion that is in the network. A ratio near 1 indicates that the network contains a representative proportion of the population, based on its relative area to the ecoregion. Values below 1 would indicate a lower potential proportion of the population. This approach is in contrast to the representativeness analysis which evaluates the ability of surrogates to capture the full distribution of test features for each of the test species.

# Results

Across all ecozones, there was a moderate to strong positive linear relationship between the representativeness of test species and the representativeness of the four surrogates (Table 2, Figs 4 & 5). This relationship was generally strongest and more consistent across ecozones for All Birds and Forest Birds in comparison to individual species. For All Birds and Forest Birds, the four surrogates explained 51-66% and 46-63% of the variation in KS values, respectively. For both groups of species, the relationship was strongest in ecozones 9 and 11, and weakest in ecozone 4. Model uncertainty, as measured by RMSE, was low across all ecozones, ranging between 0.04-0.07 for All Birds and 0.04-0.08 for Forest Birds. The direction of the coefficients for All Birds and Forest Birds was generally positive for all surrogates in all ecozones. Based on the absolute value of the *t*-statistic, CMI was the most important surrogate in 7 of 9 ecozones for All Birds and 8 of 9 ecozones for Forest Birds. The effect size of CMI ranged from 0.23-0.43 for All Birds and 0.21-0.49 for Forest Birds. In only 2 cases (both for All Birds) were coefficients negative, but in both cases the values were too close to zero to be meaningful.

**Table 2 and Figure 3**

Among individual bird species, the strength and uncertainty of the relationship varied by species and ecozone with no evident patterns for either species or ecozones. In addition, the direction and significance of the coefficients for individual bird species varied across ecozones more than they did for All Birds and Forest Birds. Three of the 11 individual bird species had sufficient densities to be evaluated across all ecozones: Boreal Chickadee, Swainson’s Thrush, and White-winged Crossbill. In contrast, Canada Warbler and Blackburnian Warbler could only be evaluated in 3 and 4 ecozones, respectively. On average, the relationship between surrogate and species representativeness was strongest for Swainson’s Thrush (R2 = 0.40) following by Boreal Chickadee (R2 = 0.36) and Brown Creeper (R2 = 0.33). The strength of the average Caribou model was 0.30 but this was mainly due to one ecozone with a very strong relationship (R2 = 0.72). The strongest bird relationships occurred for Blackburnian Warbler in ecozone 6B (R2 = 0.71), and Black-throated Green Warbler in ecozone 6B (R2 = 0.66). However, these same species had much weaker relationships in the other ecozones. Pine Grosbeak and Canada Warbler were the most consistent species across ecozones with the strength of the relationships ranging from 0.20-0.35 and 0.15-0.40, respectively.

Across all individual bird species, CMI was the most important predictor in 57% of the models (54 of 95 models), followed by GPP (21%) and LCC (15%). Olive-sided Flycatcher was the only exception, with LCC being most important in 3 ecozones while the other surrogates were most important in 2 ecozones each. Conversely, LED was the most important surrogate in only 7 models. For all models (test features x ecozones) the direction of the coefficient of the most important surrogate was always positive. With few exceptions (coefficient hovering around zero), CMI was the only surrogate that had a consistently positive effect across all test features. There were some exceptions, for example the direction of the coefficient for LED for OSFL and LCC for BRCR was more often negative than positive. The direction of the coefficient of CMI was generally only positive, with a few exceptions where it was near or just below 0. However, in those cases, the coefficient was very close to 0 and had a very low t-statistic value.

Caribou models were developed for 5 of the 9 ecozones where sufficient data was available. The relationship between representativeness of habitat and surrogates was also positive and linear for Caribou but generally weak (R2 = 0.11-0.26) with the exception of ecozone 9 (R2 = 0.72). The two most important surrogates were LCC (3 ecozones) and GPP (2 ecozones). In both cases, the direction of the coefficients varied greatly, being negative in some ecozones and positive in others. The effect sizes of the surrogates also ranged widely, from -0.03-1.17 for LCC and -0.31-0.53 for GPP. Caribou was the only test species for which CMI was never the most important variable.

We developed a Shiny app to allow readers/conservation planners in the boreal to explore the results of the analysis and identify species/ecoregion combinations that are adequaly or well represented by benchmark networks selected using surrogates (and other criteria).

# Discussion

The use of biodiversity surrogates is a key part of the conservation planning process and, specifically, the identification of important conservation areas. Surrogates are intended to represent the biodiversity of a region including its species, biotic communities, habitats and ecosystems (Margules and Pressey 2000, Caro 2010). Currently, there is little consensus about which surrogates are most effective and this largely depends on ecoregional context, conservation objectives, and the amount of information available on individual species or environmental features and their relationship to the regions' biodiversity (Grantham et al 2010). The boreal region of Canada is a large and relatively intact landscape with limited access and geographically restricted field-based information on the distribution and abundance of species. Consequently, the use of environmental indicators and, in particular, those that can be mapped across the region, provide the most logical choices as indicators of the regions' biodiversity. We have selected four such indicators and the purpose of this study was to evaluate those indicators in terms of their effectiveness at representing a subset of the regions' biodiversity, specifically boreal caribou and songbirds. Because of the largely intact nature of our study region, our focus was on assessing representativeness as opposed to maximizing representation of selected species (*sensu* Kukkala and Moilanen 2013). Essentially, we wanted to evaluate whether our surrogates were effective at identifying conservation areas that also represented the range of habitat conditions and species densities that occur in the boreal region. The use of specific map-based surrogates that are representative of other elements of biodiversity would greatly facilitate large-scale conservation planning in remote areas with many information gaps such as the boreal region of Canada.

In this study, we used a reserve selection approach to evaluate if networks comprised of ecological benchmarks that are representative of four environmental surrogates are also representative of boreal caribou and songbirds. In general, we found that effectiveness varied by species and ecozones. However, groups of species (All Birds and Forest Birds) performed more consistently than individual species across ecozones in terms of strength, uncertainty (less variation in the strength of the relationship), effect sizes and direction. Among the four surrogates, Climate Moisture Index was found to be the most important for groups of species as well as most individual species. Together, the results lend some support for environmental indicators, and Climate Moisture Index in particular, as being effective surrogates that are representative of biodiversity in the boreal region of Canada.

The next two paragraphs are mostly repeating the results. Summarize and roll in to paragraph 2.

We used two metrics to represent broad groups of boreal bird species, All Birds which included all species with a minimum density measured for each ecoregion and Forest Birds, a subset of All Birds that included up to 53 forest associated species per ecozone (Stralberg et al. 2015). Both metrics revealed similar patterns of representativeness across ecozones. The strength of the relationships with the surrogates was moderate to strong and consistent across ecozones, varying relatively little and with no evident geographic pattern. Likewise, model uncertainty was relatively low and consistent among all ecozones. One surrogate, Climate Moisture Index, was consistently found to be the most important variable in most ecozone models in comparison to the other surrogates. In contrast, Lake Edge Density most often had a minor effect on the models and was never the most important surrogate for any of the models. The direction of the coefficients was generally positive with all surrogates with few exceptions, and where there was an exception, the effect size was very close to zero and thus not significant. Overall, the consistency of results for these two metrics is encouraging and provides some evidence for the effectiveness of the four surrogates in the boreal region.

Among individual bird species, the results were more equivocal and there was much variation between species and across ecozones. There were few evident patterns related to average species density at the ecozone level with the exception of Swainson’s Thrush, which had the highest density among test species in each ecozone and also had the strongest relationship with the surrogates averaged across all ecozones. At the ecozone level, Blackpoll Warbler and Black-throated Green Warbler had the strongest relationships with the surrogates, both within ecozone 6B in Newfoundland. There was no geographic pattern in the relationships i.e., east to west or north to south. Several unmeasured factors may explain some of the variation in the relationships including topography, landscape composition and structure, land use, and natural and anthropogenic disturbance history.

Our choice of using the Kolmogorov–Smirnov (KS) statistic as a measure of representativeness was based on using a simple approach that uses the full range of habitat quality of species density for our analysis. Other surrogacy tests seem to focus on species richness or meeting abundance targets with the assumption that high environmental diversity will lead to high biodiversity by creating lots of habitat types. In our study, we are prioritizing benchmarks that have the same distribution of habitat quality of species density as the ecoregion. The focus is thus on maintaining representativeness of relatively intact landscapes rather than maximizing representation for selected species. In our view, both approaches are valid, with the former approach beneficial in large relatively intact landscapes and the latter approach being complementary but more suited to targeting reserves for species at risk. We demonstrated both approaches in a case study for one of the ecoregions comprising ecozone 6A in the central boreal region (Supp Info). ME: I think we could discuss the differences in results a bit here. Why are our representative networks not the best at representation? Except for boreal chickadee? Maybe BOCH has a skewed habitat distribution such that the ecoregion has lots of good habitat. I could see that causing a convergence in the best networks for both methods. ME: This is a different type of surrogacy and we should frame it carefully. Is there literature we can cite supporting reserves where habitat distributions are representative of the ecoregion? Once we have justified this, we can state that using habitat models is more detailed than simple species richness, and less subjective than target-based approaches to effectiveness.

Based on table in step 8 of the case study, it seems representation and representativeness can be achieved for some species but not others in a single network. We should discuss this in discussion.

Our approach offers a realistic test of surrogacy because we do not construct the benchmarks using the surrogates. Ecological benchmarks are built based on best practices for ensuring intactness, size and hydrologic connectivity. This approach provides us with a large suite of benchmarks to use as test cases for surrogacy. In addition to testing the relationship within our test units (benchmarks), we also carry out a systematic conservation planning exercise by constructing benchmark networks and evaluating the surrogates across a full range of representative to non-representative networks. Our results suggest that networks of ecological benchmarks that are selected for high levels of surrogate representation, will also be effective at representing groups of species (All Birds and Forest Birds) and to a lesser extent individual species. The approach lends itself well to the design of an effective conservation network that is representative of biodiversity in the boreal region of Canada.

Our approach has some limitations which could be addressed in future research. We evaluated effectiveness using current predicted species distributions. This was done because of the gaps in the distribution of sampling sites across the boreal. Our study, like most that evaluate surrogates, is not directly comparable because of suggest that this variability is due to differences in surrogates used, study regions, measures of surrogate effectiveness, and biodiversity features that the surrogates are intended to represent (Grantham et al. 2010). However, future research should consider using actual bird survey data to evaluate effectiveness of benchmark networks, at least for those regions of the boreal where there is sufficient point data. In addition, it would be of interest to use predicted future distributions of habitats and species densities to evaluate effectiveness of benchmark networks under a range of alternative climate change scenarios.

# Acknowledgements

# Supporting Information

Additional Supporting Information may be found online at: <https://github.com/prvernier/surrogates>

* 01\_species\_names.md
* 02\_test\_features.md
* 03\_creating\_networks.md
* 04\_evaluating\_surrogates.Rmd (04\_evaluating\_surrogates.html)

## Detailed model summaries. Summary of ecozone-level regression models relating species KS to surrogates for All Birds, Forest Birds, Caribou, and 11 individual bird species. Ecoregions and Nets indicates the number of ecoregions and networks, respectively, that were used in each ecozone/test species’ model. For each of the four surrogates, we provide the coefficients and *t*-statistic as an estimate of the contribution of each variable to the model; bold coefficients indicate the most important surrogate for a particular species/ecozone model.

# Literature Cited

Andrew, M.E., M.A. Wulder, and J.A. Cardille. 2014. Protected areas in boreal Canada: a baseline and considerations for the continued development of a representative and effective reserve network. Environmental Reviews 22:135-160.

Arcese, P., and A.R.E. Sinclair. 1997. The role of protected areas as ecological baselines. Journal of Wildlife Management 61:587-602.

Albuquerque, F., and P. Beier. 2018. Improving the use of environmental diversity as a surrogate for

species representation. Ecology and Evolution 8:852-858.

Andrefouet, S., M.A. Hamel, and M. Dalleau. 2012. Distinction between effective pattern-based

and selection-based biodiversity surrogates is essential: caveats for managers. Marine Ecology Progress Series 452:287-295.

Barker, N. K. S., S. G. Cumming, and M. Darveau. 2014. Models to predict the distribution and abundance of breeding ducks in Canada. Avian Conservation and Ecology 9(2): 7.

http://dx.doi.org/10.5751/ACE-00699-090207

BEACONs. 2017. Manual Benchmark Builder Version 3.3.15. BEACONs Project, University of Alberta, Edmonton. AB.

Beier, P., and F. Albuquerque 2015. Environmental diversity as a surrogate for species representation. Conservation Biology 29: 1401-1410.

Beier, P., P. Sutcliffe, J. Hjort, D.P. Faith, R.L. Pressey, and F. Alburqurque. 2015. A review of selection-based tests of abiotic surrogates for species representation. Conservation Biology 29(3):668-679.

Brandt J. P., M.D. Flannigan, D.G. Maynard, and I. D. Thompson. 2013. An introduction to Canada’s boreal zone: ecosystem processes, health, sustainability, and environmental issues. Environmental Reviews 226:207–226.

Caro, T.M. and G. O’Doherty. 1999. On the use of surrogate species in conservation biology. Conservation Biology 13(4):805-814.

CEC. 2013. 2010 Land Cover of North America at 250 meters, Edition 1.0. Commission for Environmental Cooperation, Montréal, QC.

Coristine, L.E., A.L. Jacob, R. Schuster, S.P. Otto, N.E. Baron, N.J. Bennett, S.J. Bittick, C. Dey, B. Favaro, A. Ford, L. Nowlan, D. Orihel, W.J. Palen, J.L. Polfus, D.S. Shiffman, O. Venter, and S. Woodley. 2018. Informing Canada’s commitment to biodiversity conservation: A science-based framework to help guide protected areas designation through Target 1 and beyond. FACETS 3 :531-562.

Ecological Stratification Working Group. 1996. A National Ecological Framework for Canada. Ottawa/Hull, ON, Canada: Agriculture and Agri-Food Canada, Research Branch, Centre for Land and Biological Resources Research and Environment Canada, State of Environment Directorate.

Engelbrecht, I., M. Robertson, M. Stolz, and J.W. Joubert. 2016. Biological Conservation 197:171-179.

Environment Canada. 2011. Scientific Assessment to Inform the Identification of Critical Habitat for Woodland Caribou (*Rangifer tarandus caribou*), Boreal Population, in Canada: 2011 update. Ottawa, Ontario, Canada. 102 pp. plus appendices.

GFWC. 2014. Human Access of Canada's Landscapes. Global Forest Watch Canada. https://databasin.org/datasets/0c54d369b225471ea7e9f7999ce94cc0

Grantham, H. S., R.L. Pressey, J. Wells, J., and A. J. Beattie. 2010. Effectiveness of biodiversity surrogates for conservation planning: different measures of effectiveness generate a kaleidoscope of variation. PLoS ONE 5(7):e11430.

Hamann A, Wang TL, Spittlehouse DL, Murdock TQ. 2013. A comprehensive, high‐resolution database of historical and projected climate surfaces for Western North America. Bulletin of the American Meteorological Society, 94, 1307–1309.

Hanson, J.O., J.R. Rhodes, C. Riginos, and R.A. Fuller. 2017. Environmental and geographic variables are

effective surrogates for genetic variation in conservation planning. PNAS 114(48):12755-12760.

Hogg, E. H. 1997. Temporal scaling of moisture and the forest – grassland boundary in western Canada. Agriculture and Forest Meteorology 84:115‐122.

Hortal, J., M.B. Araujo, and J. M. Lobo. 2009. Testing the effectiveness of discrete and continuous environmental diversity as a surrogate for species diversity. Ecological Indicators 9(1):138-149.

Kukkala, A. S., and A. Moilanen. 2013. Core concepts of spatial prioritisation in systematic conservation planning. Biological Reviews 88:443-464. http://dx.doi.org/10.1111/brv.12008

Kuhn, M., and K. Johnson. 2013. Applied Predictive Modeling. Springer

Leroux, S.J., and B. Rayfield. 2014. Methods and tools for addressing natural disturbance dynamics in conservation planning for wilderness areas. Diversity and Distributions 20:258–271.

Leroux, S. J., F.K. Schmiegelow, R.B. Lessard, and S. G. Cumming. 2007. Minimum dynamic reserves: a framework for determining reserve size in ecosystems structured by large disturbances. Biological Conservation 138(3):464-473.

Lisgo, K.A., F.K.A. Schmiegelow, S.J. Leroux, and S.G. Cumming. In prep. Benchmarks across the Boreal: Designing large protected areas in one of the World’s remaining wilderness regions.

Lombard, A.T., R.M. Cowling, R.L., Pressey, and A.G. Rebelo. 2003. Effectiveness of land classes as surrogates for species in conservation planning for the Cape Floristic Region. Biological Conservation 112:45-62.

Margules, C.R., and R.L. Pressey. 2000. Systematic conservation planning. Nature 405(6783):243-253.

Mellin, C., S. Delean, J. Caley, G. Edgar, M. Meekan, R. Pitcher, R. Przeslawski, R., A. Williams, and C. Bradshaw. 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: A global meta-analysis. PLoS ONE 6(6):e20141.

NRCan. 2009. Atlas of Canada 1,000,000 National Frameworks Data, Hydrology Version 6.0: A practical guide to the datasets. Natural Resources Canada, Ottawa, ON. <ftp://ftp.geogratis.gc.ca/pub/nrcan_rncan/archive/vector/framework_cadre/drainage_areas/1M_HYDRO_GUIDE_EN_2009.pdf>

Oliver, I., A. Holmes, J.M. Dangerfield, M. Gillings, A.J. Pik, D.R. Britton, M. Holley, M.E. Montgomery, M. Raison, V. Logan, R.L. Pressey, and A.J. Beatie. 2004. Land systems as surrogates for biodiversity in conservation planning. Ecological Applications 14(2):485-503.

Powers, R.P., N.C. Coops, J.L. Morgan, W.A. Wulder, T.A. Nelson, C.R. Drever, and S.G. Cumming. 2012. A remote sensing approach to biodiversity assessment and regionalization of the Canadian boreal forest. Progress in Physical Geography 37(1):36-62.

R Development Core Team. 2019. R: a language and environment for statistical computing. R Foundation

for Statistical Computing, Vienna, Austria.

Rodrigues, A. S. L., and T. M. Brooks. 2007. Shortcuts for Biodiversity Conservation Planning: The Effectiveness of Surrogates. Annual Review of Ecology, Evolution, and Systematics 38(1):713-737.

Sarkar, S., J. Justus, T. Fuller, C. Kelley, J. Garson, and M. Mayfield. 2005. Effectiveness of Environmental Surrogates for the Selection of Conservation Area Networks. Conservation Biology 19(3):815-825.

Schmiegelow, F. et al. 2012. Predictive tools for the monitoring and assessment of boreal birds in Canada, 2009-2012. Annual Report to Environment Canada by the Boreal Avian Modelling Project.

Schmiegelow, F.K.A., S.G. Cumming, K.A. Lisgo, K.A., S.J. Leroux, and M.A. Krawchuk. 2014. Catalyzing Large Landscape Conservation in Canada’s Boreal Systems: The BEACONs Project Experience, in: Lewitt, J.N. (Ed.), Conservation Catalysts: The Academy as Nature’s Agent. Lincoln Institute of Land Policy, Cambridge Massachusetts, pp. 97–122.

Secretariat of the convention on biological diversity. 2010. COP-10 decision X/2. https://www.cbd.int/decision/cop/?id=12268

Stralberg, D., S.M. Matsuoka, A. Hamann, E.M. Bayne, P. Solymos, F.K.A. Schmiegelow, X. Wang, S.G. Cumming and S.J. Song. 2014. Projecting boreal bird responses to climate change: the signal exceeds the noise. Ecological Applications, 25, 52‐69.

Stralberg, D., S.M. Matsuoka, A. Hamann, E.M. Bayne, P. Sólymos, F.K.A. Schmiegelow, X. Wang, S.G. Cumming, and S.J. Song. 2015. Projecting boreal bird responses to climate change: the signal exceeds the noise. Ecological Applications 25(10):52-69.

Su, J.C., D.M. Debinski, M.E. Jakubauskas, and K. Kindscher. 2004. Beyond species richness: community similarity as a measure of cross-taxon congruence for coarse-filter conservation. Conservation Biology 18(1):167-173.

Urbina-Casanova, R., F. Leubert, P. Pliscoff, and R. A. Scherson. 2016. Assessing floristic representativeness in the protected areas national system of Chile: are vegetation types a good surrogate for plant species? Environmental Conservation 43:1-9.

Vernier, P. and Lisgo, K., 2011. Creating catchments for the boreal forest of Canada: A semi-automated procedure using ArcGIS and Arc Hydro Tools. Available at: <http://www.beaconsproject.ca/assets/beacon/PDF/catchments_manual.pdf>

Warman, L.D., M. Forsyth, A.R.E. Sinclair, K. Freemark, H.D. Moore, T.W. Barrett, R.L. Pressey, and D. White. 2004. Species distributions, surrogacy, and important conservation regions in Canada. Ecology Letters 7:374-379.

Wiersma, Y.F. 2005. Environmental benchmarks vs. ecological benchmarks for assessment and monitoring in Canada: Is there a difference? Environmental Monitoring and Assessment 100:1-9.

Wiens, J. A., G. D. Hayward, R. S. Holthausen, and M. J. Wisdom. 2008. Using surrogate species and groups for conservation planning and management. Bioscience 58:241–252.

Zhao, M., and S. W. Running. 2010. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. Science 329:940-943.

**Table 1.** Range of representative and non-representative benchmark networks within the ecozones of the boreal region of Canada. Criteria used to differentiate representative and non-representative networks are described in the methods section. Networks were also filtered to reduce spatial overlap. The actual number of ecoregions and networks used varied by species based on the extent of their ranges. See Supplementary Information for actual number of networks in each ecoregion.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ecozone** | **Number of ecoregions** | **Ecoregions with networks** | **Range of representative networks** | **Range of non-representative networks** |
| 4 | 15 | 9 | 6 - 470 | 80 - 2277 |
| 5 | 9 | 9 | 1 - 142 | 129 - 882 |
| 6A | 14 | 10 | 2 - 152 | 8 - 781 |
| 6B | 5 | 1 | 8 | 145 |
| 9 | 6 | 3 | 1 - 31 | 6 - 49 |
| 11 | 6 | 5 | 17 - 2384 | 3 - 433 |
| 12 | 11 | 10 | 31 - 1262 | 3 - 1665 |
| 14 | 3 | 3 | 12 - 309 | 9 - 990 |
| 15 | 3 | 2 | 23 - 56 | 234 - 2310 |

Table 2. Common and scientific names for test features (species) used in the evaluation of surrogates. See Appendix 1 for codes, common names, and scientific names for all species included in the composite indicators.

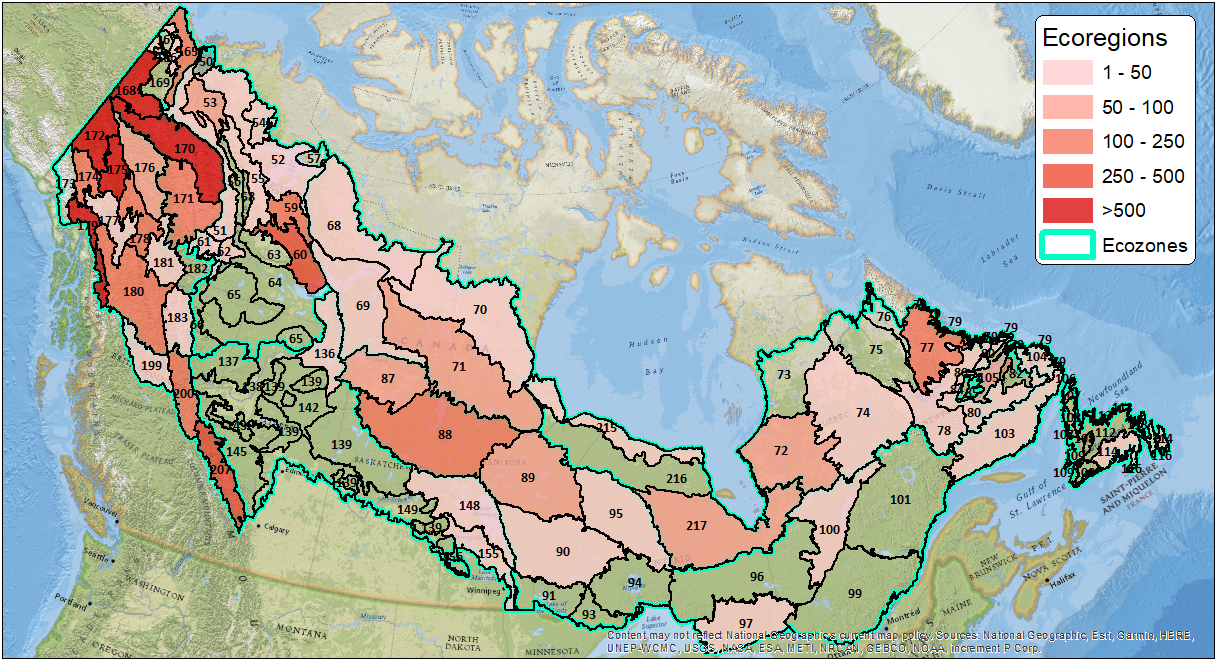
|  |  |  |  |
| --- | --- | --- | --- |
| Group | Code | Common name | Latin name |
| Ungulates | Caribou | Boreal caribou | *Rangifer tarandus* |
| Songbirds | BLBW | Blackburnian Warbler | *Setophaga fusca* |
|  | BOCH | Boreal Chickadee | *Poecile hudsonicus* |
|  | BRCR | Brown Creeper | *Certhia americana* |
|  | BTNW | Black-throated Green Warbler | *Setophaga virens* |
|  | CAWA | Canada Warbler | *Cardellina canadensis* |
|  | CMWA | Cape May Warbler | *Setophaga tigrina* |
|  | OSFL | Olive-sided Flycatcher | *Contopus cooperi* |
|  | PIGR | Pine Grosbeak | *Pinicola enucleator* |
|  | RUBL | Red-winged Blackbird | *Agelaius phoeniceus* |
|  | SWTH | Swainson's Thrush | *Catharus ustulatus* |
|  | WWCR | White-winged Crossbill | *Loxia leucoptera* |
|  | AllBirds | All birds | 80 modelled species |
|  | ForestBirds | All forest birds | 53 modelled species |
| Waterfowl | AllWaterfowl | All waterfowl | 23 species |
|  | CavityNesters | Cavity nesting waterfowl | 6 species |
|  | GroundNesters | Ground nesting waterfowl | 11 species |
|  | OverwaterNesters | Overwater nesting waterfowl | 6 species |

**Table 3.** Summary of ecoregion-level regression models relating species KS to surrogates for All Birds, Forest Birds, Caribou, and 11 individual bird species. Ecoregions and Nets indicates the number of ecoregions and networks, respectively, that were used in each ecozone/test species’ model. For each of the four surrogates, we provide the coefficients and *t*-statistic as an estimate of the contribution of each variable to the model; bold coefficients indicate the most important surrogate for a particular species/ecozone model.

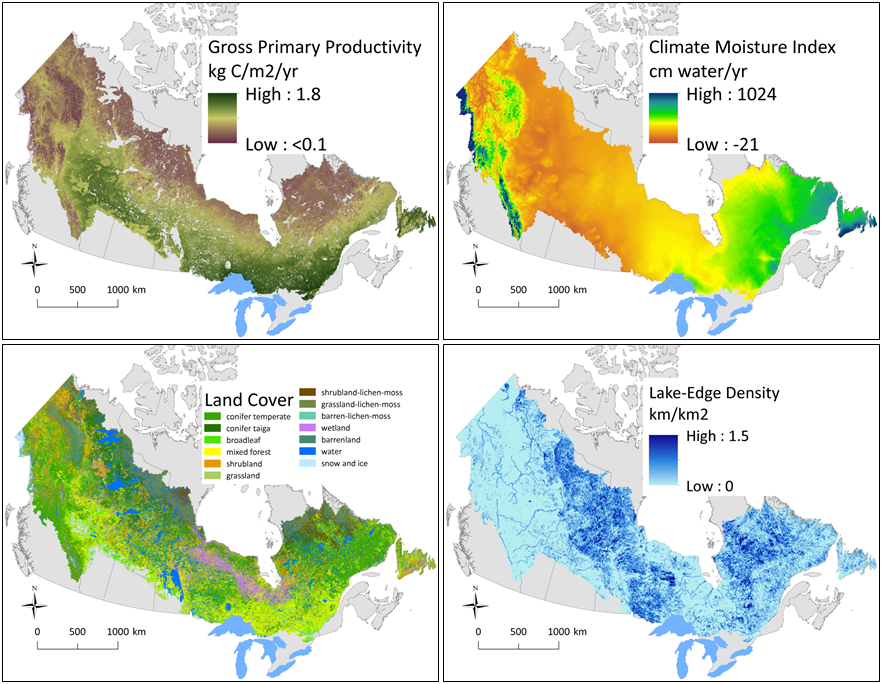
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Species | N | CMI | GPP | LED | LCC | R2 | RMSE |
| AllBirds | 52 | 0.14  (-0.81-1.08) | 0.29  (-1.29-2.71) | 0.02  (-2.58-2.08) | 0.13  (-2.19-1.84) | 0.44  (0.06-0.96) | 0.08  (0.02-0.16) |
| ForestBirds | 52 | 0.11  (-1.25-0.81) | 0.33  (-1.80-2.04) | 0.03  (-1.72-2.41) | 0.18  (-2.69-1.37) | 0.46  (0.05-0.95) | 0.09  (0.01-0.15) |
| allWaterfowl | 52 | 0.14  (-1.40-1.53) | 0.06  (-0.95-1.77) | 0.31  (-0.43-1.67) | 0.08  (-1.30-1.73) | 0.47  (0.00-0.99) | 0.08  (0.02-0.15) |
| CavityNesters | 52 | 0.10  (-0.86-2.00) | -0.10  (-1.92-1.27) | 0.28  (-2.22-1.33) | 0.40  (-0.60-1.72) | 0.47  (0.02-1.00) | 0.08  (0.01-0.14) |
| GroundNesters | 52 | 0.15  (-0.57-1.95) | 0.23  (-1.36-1.84) | 0.13  (-0.71-1.87) | 0.07  (-1.64-2.42) | 0.44  (0.02-1.00) | 0.08  (0.01-0.15) |
| OverwaterNesters | 52 | 0.15  (-1.11-1.42) | 0.14  (-1.40-1.77) | 0.19  (-1.06-1.42) | 0.00  (-3.05-1.71) | 0.45  (0.02-0.99) | 0.07  (0.01-0.14) |
| blbw | 8 | 0.42  (0.03-0.99) | 0.44  (-1.73-1.18) | -0.13  (-1.84-0.43) | 0.06  (-2.03-1.29) | 0.75  (0.43-0.97) | 0.06  (0.01-0.08) |
| boch | 52 | 0.22  (-0.84-1.60) | 0.23  (-0.96-1.93) | 0.12  (-0.89-1.87) | 0.19  (-2.16-1.21) | 0.50  (0.11-1.00) | 0.09  (0.01-0.15) |
| brcr | 23 | 0.28  (-0.40-1.29) | 0.07  (-2.14-1.52) | -0.03  (-0.77-0.78) | 0.34  (-0.96-1.55) | 0.52  (0.02-0.97) | 0.08  (0.03-0.13) |
| btnw | 17 | 0.22  (-0.37-0.89) | 0.59  (-0.71-2.21) | -0.15  (-2.53-0.77) | 0.20  (-1.56-1.50) | 0.62  (0.20-1.00) | 0.07  (0.01-0.12) |
| cawa | 11 | 0.52  (-0.11-1.97) | 0.59  (-0.98-1.56) | 0.05  (-1.50-0.56) | -0.19  (-1.78-0.58) | 0.70  (0.28-0.98) | 0.07  (0.02-0.13) |
| cmwa | 17 | 0.22  (-0.39-1.12) | 0.38  (-1.23-1.30) | 0.03  (-0.98-0.64) | 0.21  (-0.43-0.93) | 0.58  (0.03-0.97) | 0.08  (0.01-0.14) |
| osfl | 42 | 0.18  (-1.07-1.12) | 0.29  (-2.63-3.14) | 0.15  (-1.07-1.95) | -0.07  (-3.53-2.23) | 0.46  (0.03-1.00) | 0.08  (0.01-0.15) |
| pigr | 52 | 0.27  (-0.86-1.18) | 0.32  (-0.86-1.60) | -0.01  (-1.29-1.15) | 0.08  (-1.53-1.72) | 0.54  (0.06-0.99) | 0.07  (0.01-0.16) |
| rubl | 52 | 0.18  (-0.97-1.06) | 0.31  (-1.14-2.67) | 0.03  (-0.64-0.83) | 0.11  (-1.90-2.17) | 0.47  (0.03-0.99) | 0.08  (0.01-0.16) |
| swth | 49 | 0.31  (-1.15-1.26) | 0.16  (-2.29-2.17) | -0.03  (-2.22-1.27) | 0.52  (-1.27-4.59) | 0.60  (0.09-0.99) | 0.08  (0.01-0.15) |
| wwcr | 52 | 0.23  (-1.05-1.26) | 0.34  (-0.95-2.32) | -0.02  (-1.07-1.45) | 0.06  (-2.77-1.25) | 0.45  (0.03-0.99) | 0.09  (0.01-0.15) |

**Table 3.** The number of times CMI, GPP, LED, and LCC was selected as the most important variable, based on the absolute value of the t-statistic in each test feature / ecozone model, for songbirds and waterfowl test features. See Appendix 1 for results for each test feature/ecozone model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group** | **Test feature** | **CMI** | **GPP** | **LED** | **LCC** |
| Songbirds | Blackburnian Warbler | 3 | 3 | 0 | 2 |
|  | Boreal Chickadee | 20 | 12 | 7 | 13 |
|  | Brown Creeper | 9 | 6 | 1 | 7 |
|  | Black-throated Green Warbler | 5 | 7 | 0 | 5 |
|  | Canada Warbler | 4 | 5 | 1 | 1 |
|  | Cape May Warbler | 4 | 8 | 2 | 3 |
|  | Olive-sided Flycatcher | 13 | 20 | 5 | 4 |
|  | Pine Grosbeak | 21 | 14 | 3 | 14 |
|  | Rusty Blackbird | 15 | 18 | 7 | 12 |
|  | Swainson’s Thrush | 22 | 12 | 2 | 13 |
|  | White-winged Crossbill | 20 | 10 | 9 | 13 |
|  | AllBirds group | 13 | 21 | 3 | 15 |
|  | ForestBirds group | 15 | 16 | 5 | 16 |
|  | Sub-total | 164 | 152 | 45 | 118 |
| Waterfowl | AllWaterfowl guild | 17 | 12 | 14 | 9 |
|  | CavityNesters guild | 9 | 10 | 20 | 13 |
|  | GroundNesters guild | 14 | 16 | 12 | 10 |
|  | OverwaterNesters guild | 17 | 15 | 13 | 7 |
|  | Sub-total | 57 | 53 | 59 | 39 |
| Ungulates | Caribou |  |  |  |  |
| All species | Total | 221 | 205 | 104 | 157 |



**Figure 1.** The boreal region of Canada was stratified into ecoregions (black outlines) using the National Ecological Framework for Canada (Marshall et al. 1999).. The number of networks meeting size and intactness criteria by ecoregion are indicated by the value in the centre of the polygon as well as the shading of the polygon i.e., increasing shades of grey indicates higher number of networks.



|  |  |
| --- | --- |
|  |  |

**Figure 2.** Top four maps show the distribution of the four environmental surrogates in the boreal region. From top left to bottom right: Gross Primary Productivity (GPP), Climate Moisture Index (CMI), North America Land Cover 2005 (LCC), and Lake-edge Density (LED). The bottom two maps show the predicted density of Canada Warbler and habitat suitability of boreal caribou.

1. Data and report are available on Zenodo (<https://zenodo.org/search?page=1&size=20&q=Stralberg>). [↑](#footnote-ref-1)
2. <https://www.allaboutbirds.org/> [↑](#footnote-ref-2)