ECOLOGICAL SOCIETY OF AMERICA

PREPRINT

1 **Exploring the ecological processes driving geographical patterns of breeding bird richness**
2 **in British Columbia Canada**

3 Jessica L. Fitterer[1*], Trisalyn A. Nelson[1], Nicholas C. Coops[2], Michael A. Wulder[3] Nancy A.
4 Mahony[4]

5 [1*] Spatial Pattern Analysis and Research Lab, Department of Geography, University of Victoria,
6 PO Box 3060, Victoria, British Columbia, Canada ,V8W 3R4
7 E-mail: jessyka@uvic.ca
8 Phone: 250-853-3271

9 [1] Spatial Pattern Analysis and Research Lab, Department of Geography, University of Victoria,
10 PO Box 3060, Victoria, British Columbia, Canada, V8W 3R4
11 E-mail: trisalyn@uvic.ca
12 Phone: 250-472-5620

13 [2] Integrated Remote Sensing Studio (IRSS), Department of Forest Resources Management,
14 University of British Columbia, Centre 2424 Main Mall, Vancouver, British Columbia,
15 Canada, V6T 1Z4
16 E-mail: nicholas.coops@ubc.ca
17 Phone: 604-822-6452

18 [3]Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada
19 506 Burnside Road West, Victoria, British Columbia, Canada, V8Z 1M5
20 E-mail: Mike.Wulder@NRCan-RNCan.gc.ca
21 Phone: 250-298-2401

22 [4] Wildlife Research Division, Science and Technology Branch, Environment Canada
23 5421 Robertson Rd., RR 1 Delta, BC V4K 3N2
24 E-mail: nancy.mahony@ec.gc.ca
25 Phone: 604-940-4655

26 *Corresponding author

27 **Abstract**

28   British Columbia's (BC) diverse landscape provides breeding habitat for more than 300

29 avian species and recent development of the BC Breeding Bird Atlas dataset presents key

30 information for exploring the landscape conditions which lead to biological richness. We used

31 the volunteer collected raw breeding bird evidence dataset to analyze the effects of sampling

32 biases on spatial distribution of observed breeding bird species and implemented regression tree

33 analysis (Random Forests) to examine the influence of productivity, ambient energy, and habitat

34 heterogeneity on independently measured breeding bird richness. Results indicated that total

35 breeding species richness is correlated with total survey effort ($\alpha < 0.001$). By stratifying species

36 richness by survey effort, we observed that ambient energy is the top ranked environmental

37 predictor of breeding bird richness across BC which, when used in combination with a number of

38 other environmental variables, explains approximately 40% of the variation in richness. Using

39 our modelled relationships, we predicted breeding bird richness in the areas of BC not presently

40 surveyed between three and six hours. The productive Boreal Plains concentrated around Ft. St.

41 John and Dawson Creek, in the southern portion of the Taiga Plains region, the lowlands of the

42 South and Central Interior, along the Rocky Mountain Trench and the coastal areas of the

43 Georgia Depression are predicted to have the highest categories of breeding richness (35-57

44 unique species). Our results support ongoing species diversity gradient research, which identifies

45 ambient energy as an important factor influencing species diversity distributions in the northern

46 hemisphere. By linking breeding bird richness to environmental data derived from remotely

47 sensed data and systematically collected climate data, we demonstrate the potential to monitor

48 shifts in ambient energy as a surrogate for vertebrate habitat condition affecting population

49 levels. Analyzing the influence of survey effort on species richness metrics we also highlight the

50 need to consider adding attributes to the raw breeding bird dataset to describe observer

51 experience, such as hours or seasons spent surveying, and provide survey dates to allow greater

52 flexibility for removing survey bias. These additions can increase the utility of atlas data for

53 species richness studies useful for conservation planning.

54 **Key words:** species richness, bird atlas, British Columbia, productivity, ambient energy, habitat

55 heterogeneity

56   **1.0 Introduction**

57   British Columbia's (BC) diverse landscape provides breeding habitat for more than 300

58   avian species and the recent development and soon completion of the British Columbia Breeding

59   Bird Atlas data collection presents information for a number of relevant biodiversity assessment

60   questions to be addressed (BC Breeding Bird Atlas, 2012). These may include indentifying

61   where gaps in species observations are, how surveying bias influences estimates of species

62   richness, and what habitat characteristics influence the distribution of breeding bird species. By

63   exploring relationships between bird species richness and environmental variables we can better

64   understand landscape conditions that lead to biological richness and progress toward predicting

65   the response of species to landscape disturbance and change (Orme et al., 2005). Given that

66   avian species are uniquely able to select for spatial and temporal changes in environmental

67   resources birds are considered suitable indicators of habitat condition (Hurlbert and Haskell,

68   2003) and have been examined as potential surrogates for vertebrate species richness (Blair

69   1999; Gregory et al., 2003).Therefore, in establishing the ecological processes governing the

70   spatial distribution of the breeding bird species richness we can provide important information

71   for the preservation of habitat conditions known to support high levels of biological diversity.

72   Key environmental drivers of avian richness at regional to global spatial scales include

73   landscape productivity (Wright, 1983; Currie, 1991; Blackburn and Gaston, 1996; Rahbek and

74   Graves, 2001;Hurlbert and Haskell, 2003; Rensburg et al., 2002; Coops et al., 2009a), ambient

75   energy (Currie, 1991; Lennon et al., 2000; Hawkins et al., 2003a), and habitat heterogeneity

76   (Būhning-Gaese, 1997; Kerr and Packer, 1997; Jetz and Rahbek, 2002; Rahbek et al., 2007).

77   Relationships are hypothesized to occur because productivity, directly and indirectly, limits food

78   and shelter resources (Wright, 1983; Hawkins et al., 2003b; Berry et al., 2007), ambient energy

79   satisfies the thermoregulatory needs of species and influences metabolic rates (Currie, 1991;

80   Brown et al., 2004), and habitat heterogeneity provides niche variety thereby limiting species

81   competition and supporting a diversity of specialist species (Andrén, 1994; Berg, 1997; Farina,

82   1997).

83   Breeding Bird Atlas data offers a unique opportunity to assess spatial effects of these

84   ecological processes on species richness over large areas (e.g., Rensburg et al., 2002, Chown et

85   al., 2003). Species atlases generally rely on a large volunteer base to collect data over substantial

86   landscape extents. The main advantage of the gridded breeding bird atlas is the organizational

87   structure allowing for spatially consistent repeat assessments of bird distributions and abundance

88   (Donald and Fuller, 1998). Assessing the relationship between avian richness and environmental

89   variables requires that environmental data match the spatial extent of gridded atlases (Donald

90   and Fuller, 1998). Remotely sensed imagery and geospatial datasets play a critical role in this

91   aspect by linking avian richness (e.g., Būhning-Gaese, 1997; Atauri and de Lucio, 2001; Hurlbert

92   and Haskell, 2003; Luoto et al., 2004; Coops et al., 2009a, 2009b) to environmental variables

93   from regional (e.g., Luoto et al., 2004) to continental extents (e.g., Coops et al., 2009b).

94   Studies have used satellite derived measures of landscape greenness (surrogate for

95   productivity often measured using Normalized Difference Vegetation Index (NDVI) or fraction

96   of absorbed photosynthetically active radiation (fPAR)) to evaluate the relationship between

97   avian species richness and productivity (Hurlbert and Haskell, 2003; Hawkins, 2004; Evans et

98   al., 2006; Coops et al., 2009a, 2009b; St-Louis et al., 2009). Indirectly measured ambient energy

99   (e.g., evapotranspiration or atmospheric temperature) has been mapped from climate records to

100   display a notable effect on variation in avian richness at regional (Lennon et al., 2000),

101   continental (Currie, 1991), and global scales (Hawkins et al., 2003a). Additionally, within grain

102   landscape heterogeneity represented as the number of land cover types per hectare (Autari and

103   Lucio, 2001; Coops et al., 2009b), elevation variability (Ruggiero and Hawkins, 2008), or texture

104   of NDVI (St-Louis et al., 2009) has been shown to influence the distribution of avian species

105   richness. However, the importance of habitat heterogeneity on the gradient of species richness

106   can vary with study extent and grain size (e.g., Būhning-Gaese, 1997).

107   Understanding the environmental influences on breeding bird richness over BC can

108   provide important information for conservation managers and help scientists move towards

109   monitoring landscape indices as indicators of potential threats (e.g., climate change) to avian

110   breeding habitat which may in turn affect population levels. Thus, an important step toward

111   developing robust habitat models that indicate potential environmental disturbances to bird

112   species is to investigate if the variation in current avian species richness can be predicted with

113   environmental datasets. Data must be collected in a consistent, repeatable manner, over short

114   enough time intervals to signify changes in habitat productivity, ambient energy or habitat

115   configuration. Time-series analysis of high resolution remotely sensed imagery and landscape

116   configuration metrics provide an approach using environmental data to monitor habitat

117   surrogates.

118   The BC Breeding Bird Atlas presents a novel opportunity to assess the spatial effects of

119   habitat patterns on avian species richness. The goal of our study is to explore the relative

120   influence of sampling bias on the observed breeding richness to control survey bias and use the

121   developed richness indices to identify the dominant landscape-scale processes driving the

122   distribution of BC's breeding bird richness. To meet our goal, we explore the association

123   between survey bias and measures of total breeding richness to create species richness indices

124   independent of survey effort. Second, we demonstrate how systematically collected Earth

125   observation datasets and topographically adjusted climate data can represent fundamental

126   landscape-scale processes. Third, we apply non-parametric regression tree analysis to investigate

127   and model relationships between the richness of breeding birds and hypothesized landscape

128   drivers (productivity, ambient energy and landscape heterogeneity). Fourth, we use the modelled

129   relationships relatively unaffected by survey bias to predict breeding richness across the province

130   at uniform survey lengths to provide information for conservation of avian diversity. We

131   conclude by discussing the contributions of work to regional species diversity studies and

132   indirect mapping of the status of species richness through landscape-scale surrogates.

133   **2.0 Methods**

134   **2.1 Study area**

135   British Columbia covers over $940,000 km^2$ with the land mass physiography and climate

136   controlled by the Pacific Ocean to the west, continental air masses in the interior plateaus, and

137   Rocky Mountains to the east (Austin et al., 2008). The province has a highly diverse ecosystem

138   structure comprised of barren and snow covered alpine environments, moderately productive

139   mountainous regions, warm Southern Interior Mountains, moist and productive Taiga and Boreal

140   plains, productive evergreen coastlines and moderately warm and productive interior (Fitterer et

141   al., 2012). The complex landscape structure and ecological processes of BC require predictive

142   breeding bird richness modelling techniques that are able to accommodate non-linear

143   relationships and interaction effects between environmental variables within regions.

144 **2.2 Datasets**

145 **2.2.1 Breeding bird data**

146   Breeding bird species data used in our analysis were collected as part of the BC Breeding

147 Bird Atlas. The BC Breeding Bird Atlas is a five-year volunteer-based project compiling

148 information on the abundance and distribution of all breeding bird species in 10km by 10km

149 gridded areas (British Columbia Breeding Bird Atlas, 2011). Breeding evidence information is

150 collected at different confidence levels ranging from observed during the breeding season (with

151 no breeding evidence) to confirmed (displaying visible breeding activities) (British Columbia

152 Breeding Bird Atlas, 2011). For each observation of a species in a 10km by 10km quadrat, the

153 highest level of breeding confirmation is recorded. These levels indicate the likelihood of

154 breeding with 20 different descriptions available to the bird observer. Nine of the 20 categories

155 are reserved for confirmed breeding activities such as nesting, evidence of bird shells, young

156 present in the nest, and adults carrying food to a nest. Three out of the 20 categories are reserved

157 for observed or possible breeding signs such as sightings during the breeding season and

158 breeding calls. The remaining categories exhibit probable breeding characteristics such as

159 multiple singing males, couple observed in suitable nesting area, among others (British Columbia

160 Breeding Bird Atlas, 2008a). Data used in our study were a compilation of species sightings

161 from 2008-2011. Observer experience ranges from novice to experienced, but no direct attribute

162 of surveying experience is collected. Data provided includes survey effort per sample, observer

163 number, survey number, quadrat identification, and taxonomic information. Data collection is

164 ongoing to 2012 and will result in the collection of new records improving the total survey time

165 in quadrats throughout the province; however, we expect a similar spatial distribution in effort

166 where the hours spent surveying are higher in the accessible and populated regions.

167   **2.2.2 Development of breeding richness indices independent of survey bias**

168   In order to represent breeding bird species richness independent of surveying effort we

169   explored survey intensity (number of surveys conducted per grid cell), survey effort (total hours

170   spent observing) and spatial extent of the surveyed quadrats. We calculated the number of unique

171   species observed within each 10km by 10km grid cell (total breeding richness 2008-2011) and

172   calculated total survey effort (number of hours spent surveying between 2008-2011). We

173   quantified the monotonic association (Spearman's correlation coefficient) between total breeding

174   species richness and total survey effort to understand the spatial dependence. We found a strong

175   positive correlation ($r_s = .816$, $\alpha < .001$) indicating that currently the spatial variance in total

176   survey effort may influence the distribution of observed breeding richness.

177   To develop a suitable correctional method to control sampling bias we investigated the

178   skewness and spatial distribution of the Breeding Bird Atlas data up to 2011. Maximum survey

179   length and total number of surveys per quadrat exhibit positive skewness (3.9 and 3.4,

180   respectively), with the majority of the province having less than six hours of per survey

181   observational effort, and less than three surveys, with a large portion of grid cells having one to

182   two observation visits. The number of surveys per grid cell limited our ability to apply

183   rarefaction estimates, which require random selection of samples per grid cell (Colwell and

184   Coddington, 1994). Instead, we calculated species richness indices within survey effort ranges

185   (i.e., two hour intervals) to provide a relatively uniform sampling distribution because the

186   relationship between survey effort and observed species richness is asymptotic (Colwell and

187   Coddington, 1994).

188   To maximize spatial extent of breeding species information over the areal extent of BC

189   we stratified surveying lengths up to six hours of observation effort (the average maximum

190   survey length per quadrat) as quadrats sampled with total hours or surveys above six hours are

191   generally concentrated in southern urban areas which limits spatial modelling application.

192   Categories of survey effort included one to two hours, three to four hours and five to six hours.

193   Once stratified into two-hour interval survey lengths the quadrats had relatively few completed

194   surveys (average 1.6, 1.3 and 1.2 samples respectively); therefore, the individual survey with the

195   maximum amount of unique species sighted was used to represent breeding bird species richness

196   within each survey stratification category and 10km x 10km grid cell.

197   **2.2.3 Landscape indices background and modelling**

198   Productivity both as a permanent and seasonal resource is documented to influence the

199   spatial distribution of avian species richness (Hurlbert and Haskell, 2004; Hawkins, 2004; Coops

200   et al., 2009b). For instance, it is understood that resident avian richness is limited by the lowest

201   production period and migratory species by the seasonal production pulse (Hurlbert and Haskell,

202   2003; Hawkins, 2004), while the distribution of total avian species richness responds to

203   cumulative landscape production representing both seasonality and maximum production levels

204   (Coops et al., 2009b). The relationship between productivity and organisms is understood to

205   exhibit increased interaction between organisms when productivity is high and consequently

206   disturbance is low (see Southwood, 1988). Further, highly productivity regions are understood to

207   provide food (metabolic need), shelter and nesting resources for migratory and resident species,

208   with dispersive species following seasonal changes in vegetation productivity (Berry et al.,

209   2007). To assess the influence of production on avian richness distribution, remotely derived

210   landscape indices of productivity, surrogates of growth and seasonality, have been developed

211   over broad scales for both annual (Berry et al., 2007; Coops et al., 2009a, 2009b) and seasonal

212   (Hurlbert and Haskell, 2003) time scales. However, annual indices of productivity have been

213 hypothesized to obscure landscape relationships leaving one to consider the time scale in which

214 bird distribution and abundance data were collected (Hurlbert and Haskell, 2003; Hawkins,

215 2004).

216   We created productivity indicators from 8-day maximums of Moderate Resolution

217 Imaging Spectroradiometer measured fraction of absorbed photosynthetically active radiation

218 (fPAR) by the plant canopy (MCD15A2) spanning February to September, the same temporal

219 extent as the breeding bird season (British Columbia Breeding Bird Atlas, 2008b).

220 Photosynthetically active radiation spans the visible electromagnetic spectrum from 400 to

221 700nm and is absorbed by the Chlorophyll a & b, Carotene, and Xanthophylls pigments in plants

222 for photosynthesis. Contrastingly, vegetation markedly reduces its absorption of energy between

223 700 and 1300 nm (red edge to near infrared region). These reflectance properties are measured as

224 the proportion between the radiation received by the surface and the returned reflectance to the

225 remote sensor. MODIS fPAR values range from 0%, signifying barren land (i.e., rock or snow

226 cover), to close to 100% for dense vegetation cover (Coops et al., 2008) and can be used as input

227 parameters for gross primary productivity models because vegetation growth is related to the rate

228 at which vegetation absorbs visible energy (Berry et al., 2007). High productivity is inferred

229 when fPAR remains high over a growing season and vice versa for low productivity. Monthly

230 maximums of fPAR were created from the 8-day composite datasets to minimize the effect of

231 cloud cover on the remotely sensed imagery. Subsequent to monthly aggregation, seasonal

232 (February-September) indices from 2008-2011 were derived. Indices included seasonal

233 minimum vegetation cover, maximum vegetation cover, cumulative sum, and coefficient of

234 variation representing vegetation seasonality. Productivity indices were then average over the

235 2008-2011 breeding bird dataset collection period and re-sampled to 10km by 10km resolution

236   using a neighbourhood average to match the spatial resolution at which the avian breeding

237   evidence was collected. Water bodies were masked to avoid skewed productivity values.

238   Ambient energy, measured in the form of water-energy variables or atmospheric

239   temperature, is documented to influence the variation in vertebrate species richness (e.g., Currie,

240   1991, Lennon et al., 2000; Hawkins et al., 2003a) particularly in the terrestrial areas of Northern

241   Hemisphere where thermoregulatory needs of species must be met (Currie,1991; Hawkins et al.

242   2003a). The relative importance of ambient energy is displayed at both regional (Lennon et al.,

243   2000) and global extents (Hawkins et al., 2003a). For example, summer air temperature is a

244   documented predictor of avian species richness through the geographic extent of the United

245   Kingdom (Lennon et al. 2000) and water-energy variables, such as, potential evapotranspiration

246   account for a significant variation in species diversity in the Northern Hemisphere (Hawkins et

247   al. 2003a). Static climate indicators have been shown to influence the gradient of avian species

248   richness in addition to air temperature and moisture availability, such that at regional scales

249   average elevation influences avian species richness, with species favouring lower elevations

250   (Farina, 1997). Elevation range, an indirect measure of climate variation (Ruggiero and Hawkins,

251   2008), has also been identified to enhance the prediction of avian richness (Davies et al., 2007).

252   We produced estimates of ambient energy using MODIS 8-day maximum 1km land

253   surface temperature data (MYA11A2), also aggregated to an approximate monthly maximum.

254   MODIS land surface temperatures (units Kelvin) were calculated from daily conditions

255   represented in the thermal infrared bands during clear sky conditions (Wan et al., 2004). Land

256   surface temperature (LST) is interpreted as the temperature of the canopy top in dense vegetation

257   covers, or soil surface in barren areas (Wan et al., 2004). Subsequent to monthly aggregation,

258   seasonal (February-September) indices from 2008-2011 were derived. Indices include annual

259   seasonal minimum, maximum and range in seasonal LST (February to September) which were

260   averaged over the 2008-2011 time period and re-sampled to 10km by 10km regions using a

261   neighbourhood average.

262   In addition, to land surface temperature, topographically adjusted climate record data

263   were used to represent atmospheric temperature and moisture availability. Using a re-sampled 1

264   km digital elevation model created from the Canadian Digital Elevation product (available from

265   www.geobase.ca) we developed indices representing mean annual temperature, mean summer

266   precipitation, mean annual precipitation, annual climate moisture deficit and annual reference

267   atmospheric evaporative demand over the province of BC from interpolated climate data using

268   the ClimateWNA program for 2008 and 2009. Mean annual temperature, mean annual

269   precipitation, and mean summer precipitation (May to September) are directly calculated from

270   observed monthly weather station data and interpolated across the province (Wang et al., 2012).

271   Both mean annual and summer precipitation were included in our models because the mean

272   summer precipitation did not extend the temporal length of the breeding season for all species in

273   BC (British Columbia Breeding Bird Atlas, 2008b). Reference atmospheric evaporative demand

274   (E_{ref}) and climate moisture deficit (CMD) were derived using Hargreaves equations (see Wang et

275   al., 2012 for equation details). Reference atmospheric evaporative demand is the amount of

276   moisture lost to atmospheric evaporation and climate moisture deficit is interpreted as a drought

277   variable, calculated as the summation of the monthly differences between E_{ref} and precipitation

278   (Wang et al., 2012). These 1km climate indices were re-sampled from 1km spatial resolution to

279   10km by 10km using a neighbourhood average. Additionally, to indirectly represent BC's

280   climate we re-sampled 25 m Canadian Digital Elevation Product to 10 km by 10km using a

281 neighbourhood average and calculated the within pixel range in elevation as a surrogate of

282 climate variation.

283   The strength of habitat heterogeneity as a predictor of the distribution of avian species

284 richness often depends on the scale of analysis. For example, in the Mediterranean region,

285 habitat heterogeneity has a stronger relationship with species richness than ambient energy

286 (Autari and Lucio, 2001). Similarly, regional predictive models of richness in Chihuahuan Desert

287 of New Mexico were improved by a combination of productivity and habitat heterogeneity

288 measures (St. Louis et al., 2009). Conversely, at broad spatial scales, Ruggiero and Hawkins

289 (2008) concluded that the climatic gradients influence species richness and must be controlled

290 for before the importance of habitat heterogeneity can be realized. Therefore, we postulate that

291 habitat heterogeneity may be a more influential indicator of avian distributions at regional scales

292 or in warmer regions where thermoregulatory needs of species are already satisfied (see Kerr and

293 Packer, 1997).

294   To evaluate the influence of habitat heterogeneity on BC's breeding birds we modelled

295 niche variety using 25m spatial resolution elevation data provided by the Canadian Digital

296 Elevation Product and 2009 500m MODIS Terra and Aqua (Version 005, University of

297 Maryland) land cover (MCD12Q1). Terrain heterogeneity was estimated as the coefficient of

298 variation in elevation within each 10km by 10km grid cell over BC. The 2009 500m MODIS

299 Terra and Aqua (Version 005, University of Maryland) land cover (MCD12Q1) was used to

300 calculate the number of different land cover types (land cover complexity) in each 10km by

301 10km grid cell and also provided land cover information to calculate the queens case percentage

302 of like adjacencies of the nearest-neighbour re-sampled 1km land cover pixels, aggregated to a

303 10km by 10km using a neighbourhood maximum.

304   **2.3 Random Forest modelling using survey length stratified species richness indices**

305   Given the correlation between total survey effort and total breeding bird species richness;

306   maximum species richness per survey and quadrat stratified by one to two hours, three to four

307   hours, and five to six hours of observational effort were used when characterizing the

308   relationships between breeding bird species richness and the landscape indices, representing

309   productivity, ambient energy and landscape heterogeneity. Relationships were modelled using

310   regression trees created in Random Forests R statistical package (described in Breiman, 2001).

311   Regression trees repeatedly partition the selected environmental indices using a hierarchical

312   structure and binary splits based on singular environmental variables that best explain the

313   variation in our breeding bird species richness indices (De'ath and Fabricius, 2000). The "best"

314   split is determined by iteratively examining the landscape indices splits and selecting the

315   environmental variance (split) that maximizes the heterogeneity between the resulting breeding

316   richness groups (ANOVA) (Prasad et al., 2006). Regression trees provide a flexible and robust

317   method for ecological studies when the relationships between dependent and independent

318   variables are non-linear or exhibit interaction effects (De'ath and Fabricius, 2000; Prasad et al.,

319   2006).

320   We used the Random Forests algorithm to grow 999 regression trees from bootstrapped

321   samples of the original data frame for each survey effort stratification (one to two hours, three to

322   four hours and five to six hours). In addition to the bootstrap sample, the Random Forests

323   algorithm employs a second randomization step where, at each node within the singular tree

324   variables are randomly chosen to determine the best binary split to explain the variation in

325   breeding bird richness (Prasad et al., 2006). Our models randomly selected six environmental

326   variables to choose the best split at each node. Trees are aggregated by averaging the singular

327   tree prediction results to produce the final prediction (Prasad et al., 2006).

328   Using an ensemble of trees has been shown to improve prediction capabilities by

329   reducing the severity of the boundaries of prediction created by singular trees, which can yield a

330   reduction in prediction variance (Prasad et al., 2006). Additionally, the second randomization

331   step is utilized to decrease the correlation between trees in the forest and reduces the effects of

332   redundant environmental features, thus improving model error rate (see Archer and Kimes,

333   2008). In addition, by allowing variables that might have been obscured by dominant

334   environmental predictors to be selected we introduce diversity into the trees that can display

335   interaction effects between variables that may otherwise have been concealed (Strobl et al.,

336   2008). To analyze the residual effect of survey effort on breeding bird species richness we

337   created two Random Forests models for each stratification; one model including and the other

338   excluding survey effort.

339   To assess which environmental variables influenced the distribution of breeding bird

340   species richness at the 10km by 10km scale we produced variable importance tables for each

341   model. The first variable importance is calculated using the out-of-bag data (data not included in

342   the bootstrap sample). Because when 999 trees are grown the error rate is generalized (Prasad et

343   al., 2006). The percentage increase in the mean square error is calculated by randomizing each

344   environmental variable before prediction and comparing the predicted species richness levels to

345   the out-of-bag species richness levels (Breiman, 2001). The increase is then calculated as the

346   difference between the average mean square error before and after randomization for that

347   variable (Prasad et al., 2006; Archer and Kimes, 2008). If the environmental variable does not

348   provide a significant amount of predictive power, when it is randomized the increase in the mean

349 square error value remains low (Archer and Kimes, 2008). The second importance variable

350 signifies the increase in node purity, which is the average reduction in the sum of the squares

351 within the split species richness groups achieved by all splits on the specified environmental

352 variable (Prasad et al., 2006).

353 **2.4 Evaluation of predictive modelling**

354 Using the modelled relationships derived from the Random Forests we predicted

355 breeding bird species richness within each observed grid cell at each survey effort stratification

356 to assess the differences between predicted and observed levels of breeding bird species richness.

357 We mapped the prediction results of the three to four hour and five to six hour stratification

358 excluding the survey effort covariate to visually assess spatial patterns of error, and reported the

359 root mean, minimum and maximum error between observed and predicted breeding bird richness

360 and test if the residuals are normally distributed. The one to two hours model was excluded from

361 mapping and model performance evaluation because survey effort remained to have a notable

362 influence on the distribution of breeding bird richness (see Section 3.1). Therefore, we used the

363 three to four and five to six hour models excluding survey effort because the breeding richness

364 stratifications are relatively unaffected by the two hour surveying interval (see Section 3.1). We

365 also calculated the coefficients of determination to analyze the model's predictive power.

366 **2.5 Prediction of independently measured breeding richness**

367 We applied the three to four hour and five to six hour modelled relationships, excluding

368 survey effort covariate, to predict breeding bird richness in the areas not presently surveyed

369 between three to six hours by the BC Breeding Bird Atlas volunteers. To analyze the similarity

370 between breeding richness predictions we compared the three to four model predictions to the

371   five to six hour predictive results and characterize the spatial distributions of the breeding

372   richness predictions using the environmental variables most often selected to reduce the sum of

373   the squares between the breeding richness partitions. Additionally, we described the spatial

374   location of breeding richness groups using the Ministry of Environment's Ecoprovince

375   classification of British Columbia (accessed from BC Government Open Data Licence at

376   www.data.gov.bc.ca on May 2nd, 2012, Figure 1) and highlighted predicted high richness

377   locations as areas where additional field surveying would be warranted to confirm distributions

378   at similar or higher levels of survey effort.

2.6 Observation bias

380   In our final analysis, we assessed observer bias, which is known to obscure the analysis

381   of data collected during volunteer avian breeding surveys (e.g., Sauer et al., 1994; Link and

382   Sauer, 1998; Fitzpatrick et al., 2009). We assessed observational bias on a pixel by pixel basis

383   determining how species richness changes with survey effort. Where the range in stratified

384   species richness was large between the stratifications and spatially variable (e.g., southern

385   Vancouver Island, Southern Interior and Ft. St. John area) we created line graphs with survey

386   effort intervals (one to two hours, three to four hours, five to six hours) on the x-axis and

387   maximum species richness (stratified by survey effort) on the y-axis. Deviations outside of a

388   positive linear or asymptotic relationship between additional survey effort and species richness

389   are understood as observer bias. The bias can be attributed to a multitude of affects since

390   breeding richness in each cell is represented by one survey (the maximum observed richness).

391   Examples of the causes of bias may include observation date within the breeding season,

392   observer experience and location within quadrat. We overlaid the range in breeding richness with

393   our predictive errors in the three to four hour model and five to six hour model to explore if a

394   connection exists between maximum residuals and observer bias.

395   **3.0 Results**

396   **3.1 Influence of survey effort on species richness prediction**

397   Random Forests variable importance indicators (e.g., Tables 1 and 2) represent the

398   variables explaining greater than 10% of breeding richness distribution. The impact of survey

399   effort is evident in the one to two hour model, as survey effort increased the prediction

400   inaccuracy (increased the mean square error) by 61.21% when randomized and was the top

401   ranked node purity variable. Conversely, our three to four hour and five to six hour models were

402   relatively unaffected by survey effort with the covariate increasing the mean square error when

403   their values were randomized and used to predict breeding richness by just 7% and 3.71%

404   respectively. As such, we focussed subsequent analysis on the three to four hour and five to six

405   hour models excluding the survey effort covariate (Tables 1 and 2).

406   **3.2 Drivers of breeding richness**

407   The three to four hour and five to six hour models are similar in their selection of

408   environmental variables, though variation in ranking was expected given the different spatial

409   extents and frequencies of the datasets (n= 817, n = 492, respectively). However, both models

410   showed that at the spatial resolution (10km by 10km) and extent of our analysis, moisture levels,

411   temperature, and elevation were consistently selected as important predictors of breeding bird

412   richness (Tables 1 and 2). Additionally, the variations in land cover types and percentage of like

413   adjacencies between the dominant land cover types had the lowest ranks for increasing

414   prediction inaccuracy when their values were randomized and used for prediction. When

19

415 randomized in the three to four-hour model land cover heterogeneity increased prediction

416 inaccuracy by 3.72% and the percentage of like adjacencies by 5.81%. Similarly, in the five to

417 six hour model, land cover heterogeneity increased the mean square error of prediction accuracy

418 by 3.16% and the percentage of like adjacencies by 5.16%. These results indicate that land cover

419 heterogeneity indices are the least useful predictors of breeding bird richness at our grain of

420 analysis and study area.

421   Our three to four hour model exhibited that climate moisture deficit is top ranked variable

422 for increasing both mean square error and node purity (Table 1 and 2). Subsequent to climate

423 moisture deficit, average elevation, mean summer precipitation, average fPAR, mean annual

424 precipitation, and mean annual temperature are top ranking variables for increasing the mean

425 square error in breeding richness predictions when randomized, meaning that these six variables

426 contributed the most to the accurate prediction of breeding richness (Table 1). The node purity

427 indicates that for all developed trees climate moisture deficit and mean summer precipitation

428 were selected most often to reduce the sum of squares in the breeding richness partitions (Table

429 1), demonstrating that moisture and temperature influence the spatial distribution of breeding

430 richness.

431   The spatial extent of our five to six hour observed maximum breeding richness is

432 relatively small given there are fewer surveys (Figure 2, see observed breeding richness map).

433 Despite the reduction in the number of quadrats surveyed, the top ranking environmental

434 variables for predicting breeding richness are similar to our three to four hour model. The top

435 ranking environmental variables increasing the mean square error when randomized include

436 average elevation, average seasonal land surface temperature, mean annual precipitation, mean

437 summer precipitation, mean annual temperature and average fPAR. The two variables that differ

438   in the top six ranks for the three to four hour and five to six hour models are climate moisture

439   deficit and average seasonal land surface temperature. However, these indicators are akin to

440   displaying the variance between warm and cool areas of BC. Similarly, between the two models,

441   the node purity results indicate that mean summer precipitation and annual evaporative demand

442   are most often selected to reduce the sum of squares in the breeding richness partitions (Table 2).

443   These results are consistent with our three to four hour model using moisture and the interaction

444   between air temperature and moisture to explain the variation in the distribution of breeding

445   richness.

446   **3.3 Model performance**

447   Visualizing the observed (1 to 82 and 1 to 80 species) and predicted (8 to 60 and 11 to 64

448   species) ranges of species richness for both the three to four hour and five to six hour models

449   respectively (Figure 2), we conclude that ambient energy variables are explaining a dominant

450   proportion of the distribution in breeding richness. The calculated coefficients of determination

451   from the Random Forests models excluding survey effort indicate that environmental variables

452   explain 43.99% (three to four hour model) and 41.89% (five to six hour model) of the variation

453   in breeding bird richness.

454   Evaluating the residuals, our three to four hour model had a root mean square error of six

455   species with maximum over-prediction of 16 species and under-prediction of 24 species. Our

456   five to six hour model performed similarly with a root mean square error of seven species,

457   maximum over-prediction of 21 species and under-prediction of 22 species. Both models display

458   normal distributions in their prediction errors (One-Sample Kolmogorov-Smirnov Test, α .90 and

459   .95 respectively). Despite the range in the residuals, the model predictions have similar spatial

460   patterns between the observed and predicted breeding bird richness distributions (Figure 2).

461　These results provide confidence that the environmental variables are predicting the spatial

462　distribution of breeding richness.

3.4 Breeding bird richness predictions

464　　　　Using our evaluated three to four hour, and five to six hour modelled relationships we

465　predicted breeding bird species richness in the regions of BC not represented by survey lengths

466　between 3 to 6 hours. Both stratifications forecast the highest breeding richness (30 to 50 unique

467　species and 34 to 57 unique species, respectively) to be located in the low-lands of the South and

468　Central Interior, Boreal Plains concentrated around Ft. St. John and Dawson Creek, in the

469　southern portion of the Taiga Plains, the low-land coastal areas along the Strait of Georgia,

470　southern portion of Vancouver Island, and along the valley in the Southern Interior Mountains

471　(Figure 3). Conversely, the lowest species richness levels are predicted to occur in our highest

472　elevations in the Northern Boreal Mountains region, coastal mountains and Southern Interior

473　Mountains (Figure 3). Overall, the highest breeding bird richness levels are in the warmer, low-

474　land areas with higher evaporative demand and moisture deficits (Figure 4).

475　　　　The spatial distributions are similar between models because the primary variables used

476　to split breeding richness into heterogeneous groups are influenced by ambient energy. For

477　example, both modelled relationships use summer precipitation and the water-energy variables

478　(represented by climate moisture deficit and reference evaporative demand) for their primary and

479　secondary most used splitting indices (Table 2). The third and fourth ranked variables for

480　reducing the sum of the squares within breeding partitions are average elevation and indicators of

481　temperatures represented by evaporative demand and average seasonal land surface temperature

482　respectively (Table 2). Similarly, the third and fourth ranking variables for reducing the sum of

483　squares are mean annual temperature and elevation range (Table 2). From, the modeled

484 relationships and the predicted richness levels we can conclude that the distribution of bird

485 breeding richness is dependent on moisture and temperature with the highest breeding richness

486 located in the Boreal and Taiga Plains and the South and Central Interior. These regions are

487 characterized by our indices to have high temperatures, low summer and annual precipitation

488 levels, low range in elevation and generally lower elevation than the rest of the mountainous

489 province of BC (Figure 4).

490   For the majority of the province the prediction differences are within five unique species

491 of each other (Figure 3). Generally, the different predictions are along the north and south coast

492 of BC where the terrain and ecosystems are most complex (Fitterer et al., 2012). Given that the

493 ecosystem dynamics within the 10km by 10km grain could encompass many different habitat

494 characteristics such as coastline, rain forests and urban areas we would expect breeding richness

495 to be harder to consistently measure and model. The most pronounced differences in breeding

496 richness (11 to 25 unique species) are located along the Georgia Strait, west coast of Vancouver

497 Island and east coastline of the Queen Charlotte Islands (Figure 3). Two factors may be

498 influencing these results. First, the three to four hour model exhibits a more even spatial

499 distribution of samples across Vancouver Island and coast. Second, as the survey effort increases

500 to the five to six hour stratification of breeding richness also increases (Figure 2). Therefore,

501 throughout the Coast and Mountains Ecoprovince both survey length and the amount of cells

502 surveyed contributes to predictive differences. Similarly, the predictive differences in the

503 southern Taiga Plains could be attributed to lack of surveys and inconsistent observations of

504 richness (see observed ranges in Figure 2), generally, the three to four hour model has more

505 surveys to support the breeding richness prediction in the Taiga Plains. Observational timing and

506 experience as highlighted in (Figure 5) could cause predictive differences in the South and

507 Central Interior, particularly the negative differences where the observed data used to create the

508 modelled relationships are less stable (i.e., breeding richness between stratifications does not

509 follow a positive relationship with survey effort). Regardless of the differences, the South and

510 Central Interior are modelled as hot spots of breeding richness and the similarity in spatial

511 distribution of breeding richness between the two predictions supersedes the differences.

512 **4.0 Discussion**

513 **4.1 Effect of sampling bias on species richness indices**

514 Evaluating raw breeding bird richness and total breeding richness survey effort we

515 highlight an association between total number of unique species observed and surveying effort in

516 the 2008-2011 BC Breeding Bird Atlas dataset, meaning the spatial distributions of raw species

517 richness hot spots may signify observer density and total survey effort rather than breeding

518 richness. For shorter sampling times (one to two hours) small increases in the amount of time

519 spent in the cell positively affected the number of species sighted per survey. We observed that

520 when survey periods are longer, three to four hours or five to six hours breeding richness

521 distributions are indicative of observation experience, observational timing and bird habitat

522 selection rather than survey effort variation from cell to cell.

523 The important influence of survey effort at intervals shorter than two hours indicates that

524 BC's common species are sighted in shorter survey times and surveys generated from longer

525 time intervals are more appropriate building relationships within environmental indicators. To

526 improve atlas data for richness modelling it would be useful to implement a minimum

527 observation period (e.g., longer than three hours). We also support previous research by stressing

528 that sampling extent is equally important as survey frequency and length when modeling and

529   predicting breeding richness to provide confidence in the observed breeding richness (Robertson

530   et al., 2010).

531   Given that species richness indices have become an important information resource for

532   designing reserve networks to protect the greatest number of species, conducting biological

533   monitoring through species surrogates and providing information for adaptive species

534   management where total species richness is unknown (Pearman and Weber, 2007), it is

535   necessary for researchers to uncover sampling bias before analyzing the spatial distribution of

536   species richness and linking these distributions to environmental factors to ensure relationships

537   accurately represent species-habitat interactions. Observational bias can skew the level of species

538   richness observed per survey (e.g., Figure 5), obscuring the positive trajectory expected between

539   species richness and survey effort before an asymptote is reached. Within BC Breeding Bird

540   Atlas data, we have identified observational biases, which may have affected the consistency of

541   the relationships formed between BC's breeding bird richness and environmental variables

542   within our stratified models.

543   For example, if we visually compare the highest error classes in our three to four hour

544   and five to six hour models (Figure 2) with the range in the species richness levels between all

545   stratifications (Figure 5), we find that the maximum over and under predictions overlay with the

546   highest ranges in the observed breeding richness. Assessing the breeding richness on a pixel

547   based assessment of how breeding richness changes with survey effort using examples extracted

548   from the southern Vancouver Island, Okanagan region and Peace River area (Figure 3.5) it

549   appears that observational effect could be distorting our ability to accurately predicted species

550   richness. Therefore, confounding factors such as observation timing, experience and location of

551   the observer within the 10km by 10km quadrat may influence the breeding richness levels

552   because the relationship between maximum breeding richness observed and survey effort does

553   not follow positive trend. Until more data are collected to control observation bias and enable

554   more advanced correctional methods, we found that stratifying data by survey effort served to

555   explore environmental variables associated with species richness. To mitigate observation

556   uncertainties when using singular surveys for analysis for environmental species modelling it

557   would be beneficial to include the date and time of observation and judgement on their personal

558   observational skills (e.g., number of seasons spent surveying) in raw breeding evidence datasets.

559   The addition of these attributes will allow researchers to stratify by observation levels and

560   similar observation time periods, potentially reducing observational bias when more surveys are

561   available.

562   **4.2 Important indicators of breeding bird richness**

563   Despite the variability in the spatial coverage of our stratified models we found that the

564   interaction between moisture and temperature modeled as climate moisture deficit and reference

565   evaporative demand were the most important (primary and secondary most selected splitting

566   variables) indicators for explaining the variation in the distribution of breeding bird richness

567   (Table 2). Higher levels of breeding bird richness were found in low-lands, valleys and warmer

568   areas of BC; areas which have the highest evaporation of moisture (Figure 4). Examples of areas

569   with the highest level of breeding bird richness include the heart of the Boreal Plains region,

570   southern Vancouver Island and the South and Central Interior.

571   Similar to our primary splits of breeding bird richness levels we found that available

572   energy continues to play an important role in predicting the richness of breeding bird as

573   elevation, precipitation, water-energy variables and temperature are within the top ranks for

574   increasing the inaccuracy of our models when randomized. Additionally, average seasonal fPAR

575   represents the importance of productivity on breeding habitat selection and is linked with

576   available energy, as vegetation is dependent on water and sunlight for growth (Hawkins et al.,

577   2003a). With these findings, we highlight consistent patterns of ambient energy variables in the

578   top six ranking predicative environmental variables emphasizing the importance of climate

579   moisture deficit, reference evaporative demand, precipitation, elevation and average seasonal

580   land surface temperature for explaining breeding richness distributions across BC. We conclude

581   that for the observed extents ambient energy (or rather water and temperature) control the

582   distribution of breeding bird richness in BC. Our results are consistent with Currie (1991) study

583   highlighting that broad-scale patterns of species richness are associated with climate and H-

584   Acevedo and Currie (2003) research concluding that summer bird richness co-varies with annual

585   temperature and precipitation. The results are also consistent with metabolic theory of ecology

586   where it is theorized that areas with high productivity and temperatures lead to faster metabolism

587   and shorter generation times resulting in increased species diversity (Brown et al., 2004).

588   Contrary to the theory that niche variety increases species richness (Andrén, 1994; Berg

589   1997) we observed landscape heterogeneity indices ranked low in predicting breeding bird

590   richness. Given our spatial extent and resolution, modelled heterogeneity did not increase

591   performance. However, quite possibly heterogeneity may influence BC's breeding bird richness

592   at smaller spatial extents where the energy gradient is uniform, or at different grid cell

593   resolutions. For example, we were unable to study the influence vegetation structural

594   complexity, species diversity or flowering properties know to affect the distribution of birds by

595   providing food, shelter and nesting resources (e.g., Neave et al., 1996) because of the coarse

596   spatial resolution of the Bird Atlas Data within which we are unable to represent observation

597   scale landscape details, which also provide interesting insights into the habitat selection of avian

598   species. Our study was conducted using a spatial resolution of 10km by 10km and heterogeneity

599   metrics have been observed to have a weakened relationship with species richness as grid cell

600   resolutions increase from 4km to 16km (Būhning-Gaese, 1997). Additionally, our results support

601   Kerr and Packer's (1997) indicating that in the northern latitudes available energy is a more

602   important indicator of species richness than habitat heterogeneity, suggesting that breeding

603   habitats must first meet the thermoregulatory needs of species.

604   The spatial distribution of the error terms and residual ranges show that our models

605   explain a portion of the variation in breeding bird richness over the province. We theorize that

606   landscape processes operating at a finer scale than the 10km by 10km spatial grain of our study

607   also influence breeding richness. Substantial variations in the landscape composition and

608   structure are expected within a 10km by 10km area as BC has a dynamic habitat structure,

609   exhibiting distinct ecosystem characteristics at smaller spatial resolutions (e.g., Fitterer et al.,

610   2012). Following theory described by the modifiable unit areal problem (MAUP), the variance of

611   the environmental and landscape characteristics will be relatively high within each cell and low

612   between cells (Wiens, 1989; Jelinski and Wu 1996). The decrease in modelled landscape

613   variance may be masking landscape structure and processes occurring within grain resolution

614   (Wiens, 1989; Jelinski and Wu 1996). Thus, larger grain sizes will affect the emergence of

615   relationships between species richness and environmental variables, which function at smaller

616   scales. To improve the ability of scientists to link and validate environmental processes with BC

617   breeding bird data, it would advantageous for volunteers to record simple habitat descriptions

618   whenever possible (such as heavily forested, grassland, etc.) with species observations and

619   record the approximate distance travelled within a gridded location as habitat characteristics can

620   change within a 10km by 10km area. However, to meet these goals volunteers would need to be

621   provided with resources to estimate distance traveled such as a Global Positioning Systems

622   found in cellular devises (e.g., Wiehe et al., 2008).

623   **4.3 Predictions of breeding bird richness independent of sampling bias**

624   Our predictions of BC's breeding bird richness in the cells not currently represented by

625   three to six hours of survey effort display similar trends between models increasing our

626   confidence in the extrapolation of our predictions. The similar spatial trends occur despite the

627   limited range of species richness values surveyed within the three to four and five to six hour

628   stratifications and dissimilar spatial extent between the two models influencing the relationship

629   between our environmental indicators and species richness (see Figures 2 and 3). Analyzing the

630   distribution of our predicted breeding bird richness, we note that high breeding bird richness

631   levels are located in the warmer, more inhabitable regions of BC presenting a long-standing issue

632   of provisioning land for species conservation where humans and industry thrive (Freemark,

633   2006). Our study supports existing research (e.g., Freemark, 2006) indicating BC's avian

634   richness peaks in the south, particularly the south central areas of the province and southern

635   Vancouver Island, where parks and protected lands are limited (Freemark, 2006). Our mapped

636   predictions of breeding bird richness could be used to highlight potential gaps in the protection

637   of biological diversity if overlaid with a parks and protected areas map.

638   However, extrapolating predictions outside of the spatial range of your model introduces

639   inaccuracies; therefore, as a caveat we note the breeding richness levels predicted in the Northern

640   Boreal Mountains, coastal mountains and northern coast were supported by a limited number of

641   surveys. While gridded atlases seek to optimize the spatial extent of data collected (Donald and

642   Fuller, 2010), human access to observation areas still plays a major role in data collection and is

643   evident in the spatial distribution of the BC Breeding Atlas surveyed grid cells. When additional

644   resources are available for data collection, efforts should focus on gathering breeding

645   information outside the ranges easily accessible by roadways to support our knowledge of the

646   selection of remote areas as breeding habitat. Similarly, the areas of BC predicted to have high

647   species richness, such as South and Central Interior regions, Boreal and Taiga Plains, the coastal

648   areas along the Strait of Georgia, southern portion of Vancouver Island, and valley in the

649   Southern Interior Mountains, are ideal targets for future field surveying.

650   Breeding bird atlas data are a unique data source for studying broad scale environmental

651   interactions between vertebrates and habitat compositions and structures. For example, atlas data

652   have been successful used to assess the effects of afforestation of invasive tree species on

653   grassland avifauna communities in Mpumalanga Province, South Africa (Allan et al., 1997).

654   Abundance data have been integrated with atlas data to study the large-scale movements of birds

655   across Australia (Griffioen and Clarke, 2002) and developing important breeding bird areas used

656   to discuss the creation of species protected areas in North England were identified through north

657   England atlas data (Brown et al., 1995). Additionally, gridded bird data displays considerable

658   promise for linking environmental data derived from remotely sensed imagery to vertebrate

659   diversity (e.g., Coops et al., 2009b). These studies, and our own, provide insights into the

660   potential information breeding bird atlas data can supply for monitoring threats to vertebrate

661   habitat and guiding biodiversity conservation planning for the allocation of protected lands or

662   areas of concern for strategic species abundance and diversity monitoring.

663   For example, the emergence of a strong connection between BC's breeding richness and

664   available energy (temperature and water indicators) demonstrates the potential to monitor

665   changes in regional climate as an indicator of habitat condition for supporting biodiversity

666   (Nagendra, 2001; Duro, 2007; Gillespie et al., 2008). By linking vertebrate richness with

667   available energy indicators automatically collected by weather station data and remotely sensed

668   imagery we highlight a data source appropriate for indicating threats to breeding species

669   diversity levels once surveying is complete making efficient use of limited conservation

670   resources given the amount of money, time and organization needed for collecting species

671   information (Franklin, 1993). However, we should note that the explained relationships between

672   breeding bird species richness and habitat conditions might also be affected by environmental

673   conditions, such as climate change in the migration areas and wintering habitats of BC's

674   migratory breeding birds, which requires analysis of environmental conditions in more than one

675   region to quantify and predict.

676   **5.0 Conclusion**

677   Our research expands upon ecological studies examining the significance of food

678   resources (productivity), thermoregulatory needs (ambient energy) and niche habitat

679   (heterogeneity) on vertebrate habitat selection. From our models, we have concluded that

680   precipitation levels and climate drive the distribution of breeding bird richness across the

681   province of BC. Our findings support a large body of research that has found water-energy to be

682   the fundamental driver of species diversity (see Hawkins et al., 2003a) theorized by some to

683   occur because of the increase in metabolic rates in warmer regions leading to greater speciation

684   (see Brown et al., 2004 for discussion).

685   By presenting suitable indices for modelling environmental factors and by utilizing a

686   flexible predictive model able to accommodate non-linear relationships, interaction affects and

687   predictor correlations we present transferable methodology for species richness modelling. We

688   believe model performance could be improved as more breeding bird data are collected. Notably,

689   we model and analyze the predictive power of environmental indicators derived from freely

690   available data to encourage the extension of our findings to be used as a stepping-stone for

691   landscape monitoring as a surrogate for vertebrate habitat conditions affecting population levels.

692   The benefit of making atlas data freely available throughout project time period will

693   encourage scientists to undertake analysis to uncover potential data bias or quality issues that can

694   be corrected before data collection is complete to optimize data collection efforts. Fostering a

695   connection between atlas organizers and data analyzers could enhance the breadth of atlas use for

696   habitat modelling and strategic biological monitoring in areas of concern and provide observers

697   with an understanding of the variety of scientific uses atlas data hold. This will require

698   engagement between data analyzers, who can provide knowledge based research products, and

699   volunteer data collectors, who can communicate the feasibility of additional data collection.

700   Ideally a few simple additions to attribute data collection could increase the flexibility in how

701   survey data can be utilized by scientists. Examples of these attributes include consistently

702   recording survey time and date, approximate area covered within the quadrat, simple descriptions

703   of breeding habitat types, and observer experience such as number of seasons spent surveying.

704   Future field work should target areas of inconsistent breeding richness forecasts (Coast and

705   Mountains, Taiga Plains regions) and areas not currently surveyed but expected to support high

706   levels of species richness (Central and South Interior, productive Boreal Plains and valley of

707   Southern Interior Mountains). To reach these goals, resources are needed for data collection in

708   by academics of agencies in the inaccessible regions of BC to augment the rich data source

709   volunteers have contributed. Undoubtedly, the richness of gridded bird data makes it a unique

710   data sources for scientist to investigate the large-scale relationships between birds and

711   environmental factors and would not be possible without the generous efforts of the volunteers.

712   **Acknowledgements**

720 **Literature Cited**

721 Allan, D. G., Harrison, J. A., Navarro, R., Van Wilgen, B. W. and Thompson, M. W. 1997. The

722   impact of commercial afforestation on bird populations in Mpumalanga Province, South

723   Africa — Insights from bird-atlas data. Biological Conservation 79:173-185.

724 Andrén, H. 1994. Effects of Habitat Fragmentation on Birds and Mammals in landscapes with

725   different proportions of suitable habitat: A Review. Oikos 71:355-366.

726 Archer, K. J., and Kimes, R. V. 2008. Empirical characterization of random forest variable

727   importance measures. Computational Statistics and Data Analysis 52: 2249-2260.

728 Atauri, J. A., and de Lucio, J. V. 2001. The role of landscape structure in species richness

729   distribution of birds, amphibians, reptiles and lepidopterans in Mediterranean

730   landscapes.Landscape Ecology 16:147-159.

731 Austin, M. A., Buffett, D. A., Nicolson, D. J., Scudder, G. G. E., and Stevens, V. 2008. Taking

732   nature's pulse: the status of biodiversity in British Columbia. www.biodievrsitybc.org

733 Berg, Å. 1997. Diversity and abundance of birds in relation to forest fragmentation, habitat

734   quality and heterogeneity. Bird Study 44:355-366.

735 Berry, S., MacKey, B., and Brown, T. 2007. Potential applications of remotely sensed vegetation

736   greenness to habitat analysis and the conservation of dispersive fauna. Pacific Conservation

737   Biology 13:120-127.

738   Blackburn, T. M., and Gaston, K. J. 1996. Spatial patterns in the species richness of birds in the

739   New World. Ecography 19:369-376.

740   Blair, R.B.1999. Birds and butterflies along an urban gradient: surrogate Taxa for assessing

741   biodiversity? Ecological Applications 9:164-170.

742   Breiman, L. E. O. 2001. Random Forests. Machine Learning 45:5-32.

743   British Columbia Breeding Bird Atlas. 2012. About the atlas.

744   http://www.birdatlas.bc.ca/english/index.jsp.

745   British Columbia Breeding Bird Atlas. 2011. Data accessed from Nature Counts, a node of the

746   Avian Knowledge Network, Bird Studies Canada. http://www.naturecounts.ca.

747   British Columbia Breeding Bird Atlas. 2008*a*. Coding sheets.

748   http://www.birdatlas.bc.ca/bcdata/codes.jsp?lang=enandpg=breeding.

749   British Columbia Breeding Bird Atlas. 2008*b*. Species breeding dates.

750   http://www.birdatlas.bc.ca/download/brdates.pdf.

751   Brown, A. F., Stillman, R. A., and Gibbons, D. W. 1995. Use of breeding bird atlas data to

752   identify important bird areas   : a northern England case study Use of breeding bird atlas

753   data to identify important bird areas   : a northern England case study. Bird Study 42:132-

754   143.

755   Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, B. 2004. Toward a

756   Metabolic Theory of Ecology. Ecology 85: 1771-1789.

757   Būhning-Gaese, K. 1997. Determinants of avian species richness at different spatial scales.

758   Journal of Biogeography 24:49-60.

759   Chown, Steven L., van Rensburg, B. J., Gaston, K. J., Rodrigues, A. S. L., and Van Jaarsveld, A.

760   S. 2003. Energy, species richness, and human population size: conservation implications at

761   a national scale. Ecological Applications 13: 1233-1241.

762   Clarke, K., Lewis, M., and Ostendorf, B. 2011. Additive partitioning of rarefaction curves:

763   removing the influence of sampling on species-diversity in vegetation surveys. Ecological

764   Indicators 11:132-139.

765   Colwell, R. K., and Coddington, J. A. 1994. Estimating terrestrial biodiversity through

766   extrapolation. Philosophical transactions of the Royal Society of London. Series B,

767   Biological Sciences 345:101-118.

768   Coops, N. C., Waring, R. H., Wulder, M. A., Pidgeon, A. M., and Radeloff, V. C. 2009*a*. Bird

769   diversity: a predictable function of satellite-derived estimates of seasonal variation in

770   canopy light absorbance across the United States. Journal of Biogeography 36:905-918.

771   Coops, N.C., Wulder, M.A., and Iwanicka, D. 2009*b*. Exploring the relative importance of

772   satellite-derived descriptors of production, topography and land cover for predicting

773   breeding bird species richness over Ontario, Canada. Remote Sensing of Environment

774   113:668-679.

775   Coops, N.C., Wulder, M. A., Duro, D. C., Han, T., and Berry, S. 2008. The development of a

776   Canadian dynamic habitat index using multi-temporal satellite estimates of canopy light

777   absorbance. Ecological Indicators 8):754-766.

778   Currie, D. J. 1991. Energy and large scale patterns of animal and plant species richness.

779   American Naturalist 137:27-49.

780   Davies, R. G., Orme, C. D. L., Storch, D., Olson, V. A., Thomas, G. H., Ross, S. G., Ding, T.S.,

781   Rasmussen, P.C., Bennett, P.M., Owens, I.P.F., Blackburn, T.M., and Gaston, K. J. 2007.

782   Topography, Energy and the Global Distribution of Bird Species Richness. Proceedings:

783   Biological Sciences 274:1189-1197.

784   De'ath, G., and Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple

785   technique for ecological data analysis. Ecology 81:3178-3192.

786   Donald, P. F., and Fuller, R. J. 1998. Ornithological atlas data: a review of uses and limitations.

787   Bird Study 45:129-145.

788   Duro, D. C., Coops, N. C., Wulder, M. A., and Han, T. 2007. Development of a large area

789   biodiversity monitoring system driven by remote sensing. Progress in Physical Geography

790   31:235-260.

791   Evans, K. L. 2006. Abundance, species richness and energy availability in the North American

792   avifauna. Global Ecology and Biogeography 15:372-385.

793   Farina, A. 1997. Landscape structure and breeding bird distribution in a sub-Mediterranean agro-

794   ecosystem. Landscape Ecology 12:365-378.

795   Fitterer, J. L., Nelson, T. A., Coops, N. C., and Wulder, M. A. 2012. Modelling the ecosystem

796   indicators of British Columbia using Earth observation data and terrain indices. Ecological

797   Indicators 20:151-162.

798   Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M., and Elkinton, J. S. 2009. Observer bias and the

799   detection of low-density populations. Ecological Applications 19:1673-1679.

800   Franklin, J. F. 1993. Preserving biodiversity: species, ecosystems, or landscapes? Ecological

801   Applications, 3:202-205.

802   Freemark, K. E., Meyers, M., White, D., Warman, L. D., Kiester, A. R., and Lumban-tobing, P.

803   2006. Species richness and biodiversity conservation priorities in British Columbia, Canada.

804   Canadian Journal of Zoology 84:20-31.

805   Gillespie, T. W., Foody, G. M., Rocchini, D., Giorgi, A. P., and Saatchi, S. 2008. Measuring and

806   modelling biodiversity from space. Progress in Physical Geography 32:203-221.

807   Gregory, R. D., Noble, D., Field, R., Marchant, J., Raven, M., and Gibbons, D. W. 2003.Using

808   birds as indicators of biodiversity. Ornis Hungaria 12-13:11-24.

809   Griffieon, P. A., and Clarke, Michael, F. 2002. Large-scale bird-movement patterns evident in

810   eastern Australian atlas data. Emu 102:99-125.

811   H-Acevedo, D., and Currie, D. J. 2003. Does climate determine broad-scale patterns of species

812   richness? A test of the causal link by natural experiment. Global Ecology and Biogeography

813   12:461-473.

814   Hawkins, Bradford A, Field, R., Cornell, H. V., Currie, D. J., Guégan, J.-F., Kaufman, D. M.,

815   Kerr, J. T., et al. 2003*a*. Energy, water, and broad-scale geographic patterns of species

816   richness. Ecology 84:3105-3117.

817   Hawkins, B A, Porter, E. E., and Felizola Diniz-Filho, J. A. 2003*b*. Productivity and history as

818   predictors of the latitudinal diversity gradient of terrestrial birds. Ecology 84:1608-1623.

819   Hawkins, B. A. 2004. Summer vegetation, deglaciation and the anomalous bird diversity

820   gradient in eastern North America. Global Ecology and Biogeography 13:321-325.

821   Hurlbert, A. H., and Haskell, J. P. 2003. The effect of energy and seasonality on avian species

822   richness and community composition. The American naturalist 161:83-97.

823   Jelinski, D. E., and Wu, J. 1996. The modifiable areal unit problem and implications for

824   landscape ecology. Landscape Ecology 11:129-140.

825   Jetz, W., and Rahbeck, C. 2002. Geographic range size and determinants of avian species

826   richness. Science 297:1548-1551.

827   Kerr, J.T., and Packer, L. 1997. Habitat heterogeneity as a determinant of mammal species

828   richness in high-energy regions. Nature 385:252-254.

829   Lennon, J. J., Greenwood, J. J. D., and Turner, J. R. G. 2000. Bird Diversity and Environmental

830   Gradients in Britain: A test of the species-energy hypothesis. Journal of Animal Ecology

831   69:581-598.

832   Link, W. A., and Sauer, J. R. 1998. Estimating population change from count data : application

833   to the north american breeding bird survey. Ecological Applications 8:258-268.

834   Luoto, M., Virkkala, R., Heikkinen, R. K., and Rainio, K. 2004. Predicting bird species richness

835   using remote sensing in boreal agricultural-forest mosaics. Ecological Applications

836   14:1946-1962.

837   Nagendra, H. 2001. Using remote sensing to assess biodiversity. International Journal of Remote

838   Sensing 22:2377–2400.

839   Neave, H. M., Cunningham, R. B., Norton, T. W. and Nix, H. A. 1996. Biological inventory for

840   conservation evaluation III. Relationships between birds, vegetation and environmental

841   attributes in southern Australia. Forest Ecology and Management 85:197-218.

842   Orme, C. D. L., Davies, R. G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V. A., Webster, A.

843   J., Ding, T., Rasmussen, P. C., Ridgely, R. S., Stattersfield, A. J., Bennett, P. M.,

844   Blackburn, T. M., Gaston, K. J., and Owens, I. P. F. 2005. Global hotspots of species

845   richness are not congruent with endemism or threat. Nature 436:1016-1019.

846   Pearman, P. B., and Weber, D. 2007. Common species determine richness patterns in

847   biodiversity indicator taxa. Biological Conservation 138:109-119.

848   Prasad, A., Iverson, L., and Liaw, A. 2006. Newer classification and regression tree techniques:

849   bagging and Random Forests for ecological prediction. Ecosystems 9:181-199.

850   Rahbek, C., and Graves, G. R. 2001. Multi-scale assessment of patterns of avian species

851   richness. Proceedings of the National Academy of Sciences of the USA 98:4534-4539.

852   Rahbek, Carsten, Gotelli, N. J., Colwell, R. K., Entsminger, G. L., Rangel, T. F. L. V. B., and

853   Graves, G. R. 2007. Predicting continental-scale patterns of bird species richness with

854   spatially explicit models. Proceedings Biological sciences / The Royal Society 274:165-74.

855   Rensburg, B. J., Chown, S. L., and Gaston, K. J. 2002. Species richness, environmental

856   correlates, and spatial scale: a test using South African birds. American Naturalist 159:566-

857   577.

858   Robertson, M. P., Cumming, G. S., and Erasmus, B. F. N. 2010. Getting the most out of atlas

859   data. Diversity 16:363-375.

860   Ruggiero, A., and Hawkins, B. A. 2008. Why do mountains support so many species of birds?

861   Ecography 31:306-315.

862   Sauer, J. R., Peterjohn, B. G., and Link, W. A. 1994. Observer differences in the North American

863   Breeding Bird Survey. The Auk 111:50-62.

864   Southwood, T. R. E. (2012). Tactics, strategies and templates. Oikos 52: 3-18.

865   St-Louis, V., Pidgeon, A. M., Clayton, M. K., Locke, B. A., Bash, D., and Radeloff, V. C. 2009.

866   Satellite image texture and a vegetation index predict avian biodiversity in the Chihuahuan

867   Desert of New Mexico. Ecography 32:468-480.

868   Strobl, C., Malley, J., and Tutz, G. 2009. An introduction to recursive partitioning: rationale,

869   application, and characteristics of classification and regression trees, bagging, and random

870   forests. Psychological methods 14:323-48.

871   Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.L. 2004. Quality assessment and validation of the

872   MODIS global land surface temperature. International Journal of Remote Sensing 25:261-

873   274.

874   Wang, Tongli, Hamann, A., Spittlehouse, D. L., and Murdock, T. Q. 2012. ClimateWNA- High-

875   Resolution Spatial Climate Data for Western North America. Journal of Applied

876   Meteorology and Climatology 51:16-29.

877   Wiehe, S. E., Carroll, A. E., Liu, G. C., Haberkorn, K. L., Hoch, S. C., Wilson, J. S., and

878   Fortenberry, J. D. 2008. Using GPS-enabled cell phones to track the travel patterns of

879   adolescents. International journal of health geographics 7:22-33.

880   Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385-397.

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

882   506.

883   Yuhong,T. Zhang,Y., Knyazikhin, Y., Myneni, R., Glassy, J., Dedieu, G., and Running, S.

884   2000. Prototyping of MODIS LAI and FPAR algorithm with LASUR and LANDSAT

885   data. IEEE Transactions on Geoscience and Remote Sensing 38:2387-2401.

886 Table 1 Top six variable performances in predicting breeding bird richness for each model. The

887 percentage increase in the mean square error is the calculated average prediction error rate if the

888 covariate of interest is randomized and used to predict breeding richness. Large increases in the

889 mean square error indicate the variable is important for accurately predicting breeding richness.

890 Differences between the models variable selections are italicized.

Top Ranked Predictor Performance in our Breeding Bird Richness Models			
3 to 4 hr Model		5 to 6 hr Model	
Percent Increase in Mean Square Error			
Climate moisture deficit	20.57	Average elevation	18.08
Average elevation	20.11	*Average land surface temperature*	14.76
Mean summer precipitation	18.27	Mean annual precipitation	14.36
Average fPAR	17.50	Mean summer precipitation	13.24
Mean annual precipitation	16.5	Mean annual temperature	13.17
Mean annual temperature	15.79	Average fPAR	12.47

891   Table 2 Top six predictor variable performances for reducing the sum of the squares within the breeding

892   bird richness partitions. Larger node purity values indicate variables selected most often to predict the

893   distribution in breeding bird richness. Differences between the models variable selections are italicized.

Top Ranked Predictor Performance in our Breeding Bird Richness Models			
3 to 4 hr Model		5 to 6 hr Model	
Increase in Node Purity			
Climate moisture deficit	17080	Mean summer precipitation	10702
Mean summer precipitation	14605	Evaporative demand	10528
Evaporative demand	14557	Average elevation	10247
Average elevation	13252	*Average land surface temperature*	8738
Mean annual precipitation	12370	Mean annual precipitation	8953
Elevation range	11518	Elevation range	8080

894   Figure 1 Ecoprovince Ecosystem Classification of British Columbia, data accessed through the

895   BC Government Open Data Licence at www.data.gov.bc.ca (last accessed May 2nd, 2012).

896   Figure 2 Random Forests model performance, three to four and five to six hour breeding richness

897   stratification models. In order from top to bottom the maps display observed, predicted and the

898   difference between observed and predicted distributions in breeding bird species richness for the

899   survey stratifications.

900   Figure 3 Maps of predicted distributions (categorized using Jenks Natural Breaks) of breeding

901   bird richness regions of BC not presently surveyed between 3 to 6 hours. These maps predict

902   richness based on uniform three to four hour and five to six hour survey effort in each 10km by

903   10km quadrat across BC. The final map, model difference, is the breeding richness prediction

904   difference between the five to six hour and three to four hour models.

905   Figure 4 Selection of the top ranked splitting predictors of breeding bird richness. The variables

906   represented are a climate moisture deficit, reference evaporative demand, average elevation,

907   mean summer precipitation, average land surface temperature and elevation range.

908   Figure 5 Display of the range in the observed breeding richness over the two-hour intervals and

909   graphs depicting the relationship between breeding richness and survey effort within the range.

910   These results, highlight observational bias in the stratified (one to two hour, three to four hour,

911   five to six hour) maximum observed species richness as we would expect the graphs to display

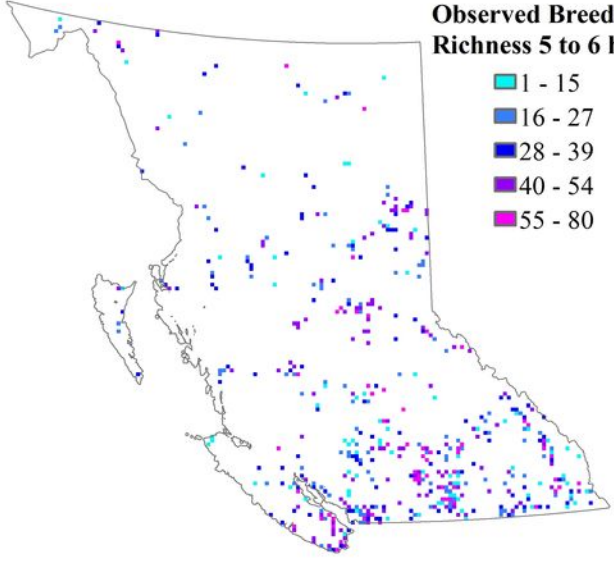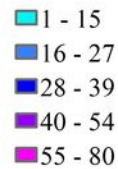912   positive trend as survey effort (hours) increase.
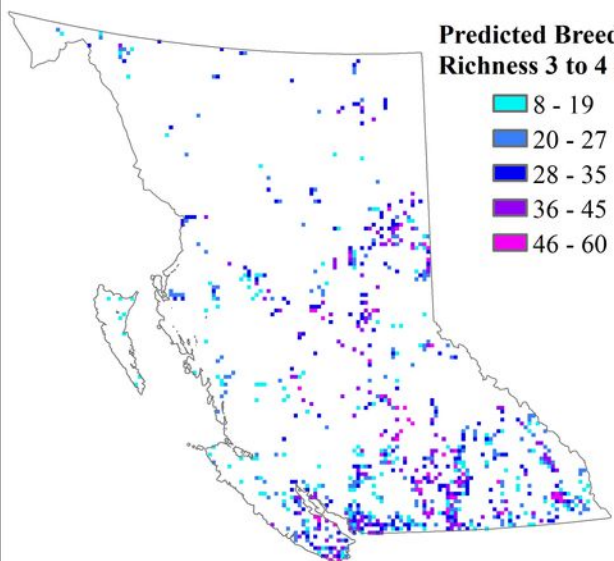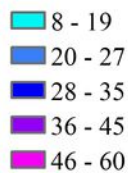
Ecoprovinces of British Columbia
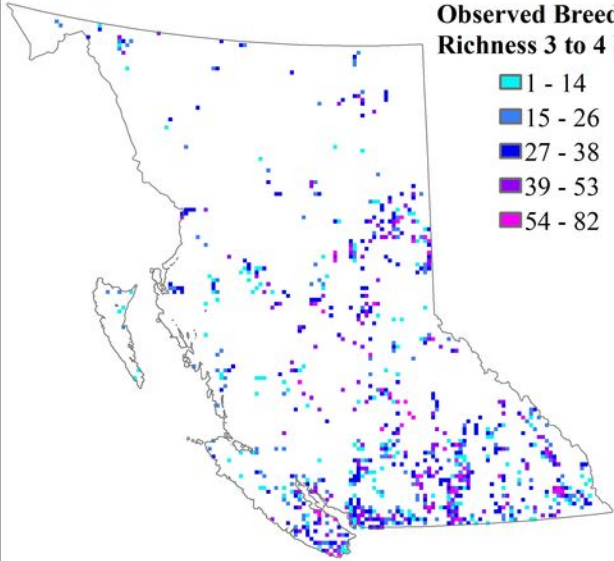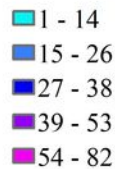
- Northern Boreal Mountains
- Taiga Plains
- Boreal Plains
- Sub-Boreal Interior
- Coast and Mountains
- Central Interior
- Southern Interior Mountains
- Southern Interior
- Georgia Depression

Observed Breeding Richness 3 to 4 hrs
- 1 - 14
- 15 - 26
- 27 - 38
- 39 - 53
- 54 - 82

Observed Breeding Richness 5 to 6 hrs
- 1 - 15
- 16 - 27
- 28 - 39
- 40 - 54
- 55 - 80

Predicted Breeding Richness 3 to 4 hrs
- 8 - 19
- 20 - 27
- 28 - 35
- 36 - 45
- 46 - 60

Predicted Breeding Richness 5 to 6 hrs
- 11 - 20
- 21 - 28
- 29 - 38
- 39 - 49
- 50 - 64
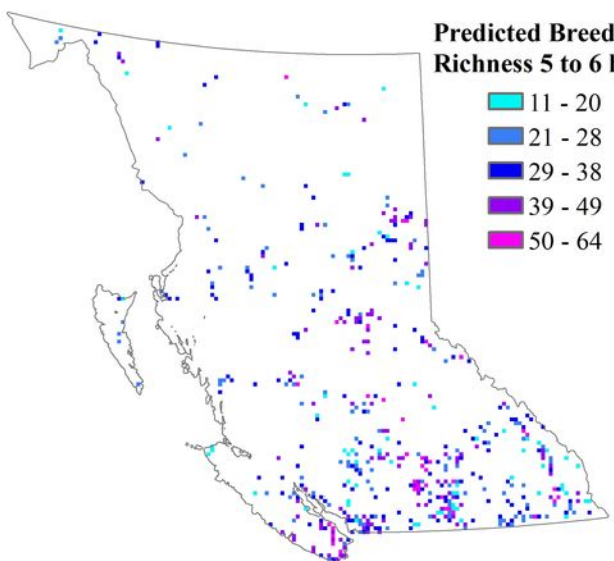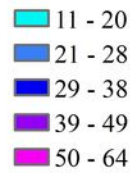
Prediction Error 3 to 4 hr Model
- -16 - -10
- -9 - -5
- -4 - 5
- 6 - 10
- 11 - 24
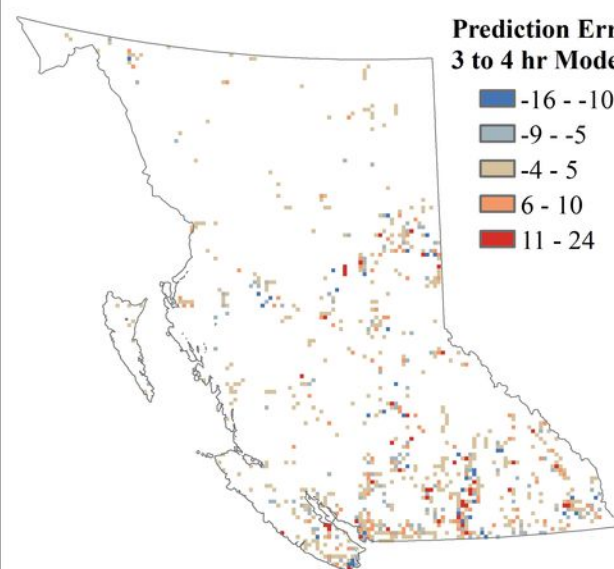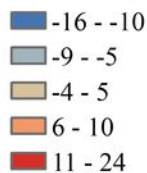
Prediction Error 5 to 6 hr Model
- -21 - -10
- -9 - -5
- -4 - 5
- 6 - 10
- 11 - 22
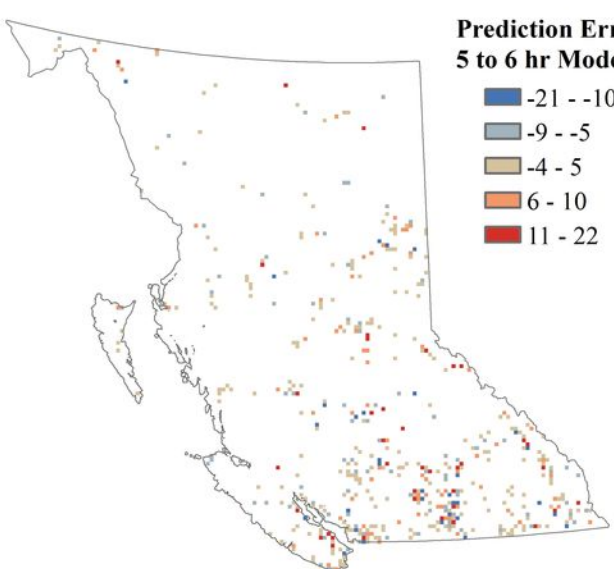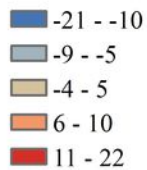
3-4 hr Breeding Bird Species Richness Prediction

- 13 - 21
- 22 - 25
- 26 - 29
- 30 - 34
- 35 - 50

5-6 hr Breeding Bird Species Richness Prediction

- 15 - 25
- 26 - 29
- 30 - 33
- 34 - 38
- 39 - 57

Model Difference

- -15 - -11
- -10 - -6
- -5 - 5
- 6 - 10
- 11 - 25

3-4 hr Pred.

5-6 hr Pred.

CMD

E_{ref}

Avg. Elev.

M. S. Precip.

Avg. LST
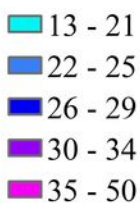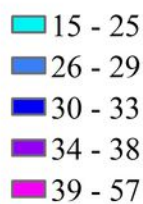
Elev. Range
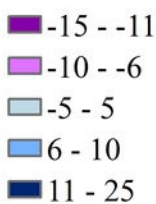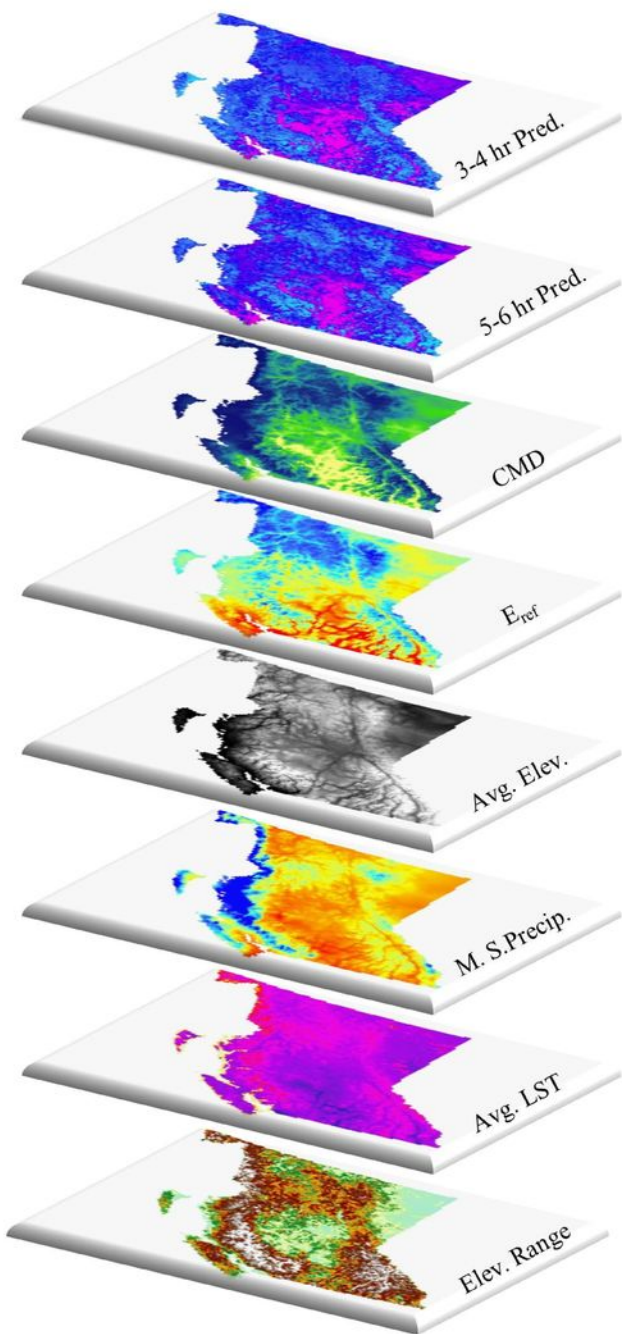
Range in Observed Breeding Bird Richness

- 1 - 12
- 13 - 24
- 25 - 40
- 41 - 77