Avoiding 'uninformed opportunism' by understanding the value of biodiversity feature and cost data in conservation prioritization

Abstract:	Although many decision-support tools have been developed to identify approaches likely to maximize the benefit – or return – on conservation investments, outcomes do not necessarily improve with better data. We evaluated the extent to which the relative value of data describing land cost and biodiversity features depended upon the structure and variability of data layers, species attributes, conservation targets, and selectivity thresholds for habitat suitability for two contrasting groups of birds (forest and human-associated) in the Pacific Northwest of North America. We analyzed distributions of focal bird species and tax-assessed land values using integer linear programming and a Marxan-like prioritization model to identify sets of 1-ha planning units that met conservation targets for the lowest possible cost (hereafter termed 'efficiency'). By comparing different prioritization scenarios, we found that the relative value of biodiversity data increased as planning efforts became more ambitious and/or selective (i.e., higher conservation targets or acceptable threshold of habitat suitability), or when species disproportionately occupied costly sites (i.e., 'cost-correlated species' whose distributions were positively related to land value). Incorporating data on land cost tended to improve efficiency of solutions, but did so at a diminishing rate as the relative variability of biodiversity to cost data increased. However, to the extent that improvements in the precision and resolution of biodiversity data capture more spatial variability in ecological features, their value in prioritization exercises should also increase. Our results indicate that the most variable cost or biodiversity feature data layer tended to drive solutions, which provides some guidance about prioritizing data acquisition in financially-constrained planning efforts. Overall, our work underscores the need to carefully consider conservation targets and thresholds across different planning contexts due to the influence of spatial variation in biodiversity feature and cost layers on the optimized solutions to alternative planning scenarios.

SCHOLARONE™
Manuscripts

1

1 **Abstract**

2 Although many decision-support tools have been developed to identify approaches likely to maximize the

3 benefit – or return – on conservation investments, outcomes do not necessarily improve with better data.

4 We evaluated the extent to which the relative value of data describing land cost and biodiversity features

5 depended upon the structure and variability of data layers, species attributes, conservation targets, and

6 selectivity thresholds for habitat suitability for two contrasting groups of birds (forest and human-

7 associated) in the Pacific Northwest of North America. We analyzed distributions of focal bird species

8 and tax-assessed land values using integer linear programming and a Marxan-like prioritization model to

9 identify sets of 1-ha planning units that met conservation targets for the lowest possible cost (hereafter

10 termed 'efficiency'). By comparing different prioritization scenarios, we found that the relative value of

11 biodiversity data increased as planning efforts became more ambitious and/or selective (i.e., higher

12 conservation targets or acceptable threshold of habitat suitability), or when species disproportionately

13 occupied costly sites (i.e., 'cost-correlated species' whose distributions were positively related to land

14 value). Incorporating data on land cost tended to improve efficiency of solutions, but did so at a

15 diminishing rate as the relative variability of biodiversity to cost data increased. However, to the extent

16 that improvements in the precision and resolution of biodiversity data capture more spatial variability in

17 ecological features, their value in prioritization exercises should also increase. Our results indicate that

18 the most variable cost or biodiversity feature data layer tended to drive solutions, which provides some

19 guidance about prioritizing data acquisition in financially-constrained planning efforts. Overall, our work

20 underscores the need to carefully consider conservation targets and thresholds across different planning

21 contexts due to the influence of spatial variation in biodiversity feature and cost layers on the optimized

22 solutions to alternative planning scenarios.

23

24 **Keywords:** Conservation planning, conservation prioritization, linear programming, optimization,

25 biodiversity feature data

26   **Introduction**

27   Conservation decision-makers often respond to constraints on human effort, funding, or data

28   quality by favoring the acquisition of land with low transaction, management, or opportunity costs. Yet in

29   the absence of a robust decision framework, such decisions can result in protecting land with limited

30   ecological value, a process dubbed 'uniformed opportunism' (Pressey et al. 1993). Uniformed

31   opportunism can both reduce the efficiency of conservation actions and cultivate the misperception that

32   adequate protection can be achieved solely by meeting area-based targets (Pressey and Bottrill 2008,

33   Bennett and Arcese 2013, Schuster and Arcese 2013). Decision-support tools developed to address these

34   and other issues in conservation have applied the design concepts of ecological complementarity and

35   irreplaceability to more efficiently conserve mapped biodiversity 'features' while simultaneously

36   minimizing various 'cost' metrics (Pressey et al. 1993, Gonzales et al. 2008, Ball et al. 2009). Systematic

37   spatial planning tools have since employed integer linear-programming (ILP) to identify cost-effective

38   solutions to highly complex problems in conservation planning, based on large numbers of mapped

39   features at a wide range of temporal and spatial scales (e.g., Beyer et al. 2016, Wilson et al. 2018,

40   Schuster et al. 2018). A recent survey indicated that of spatial prioritizations intended for

41   implementation, almost all (96%) included species data, but far fewer that included spatial data on land

42   value (24%) or implementation costs (33%; Sinclair et al. 2018). However, few empirical studies have

43   examined how variation in the costs of land acquisition and biodiversity feature data can best be used to

44   avoid uninformed opportunism in conservation prioritization (Pressey et al. 1993, Knight and Cowling

45   2007, Pressey and Bottrill 2008, Arponen et al. 2010).

46   Theory suggests that uniformed opportunism becomes more likely as spatial variation in the costs

47   of conservation actions exceed spatial variation in the ecological benefits or other features of interest, and

48   empirical studies suggest this situation is common and sometimes extreme (Balmford et al. 2003, Ferraro

49   2003, Naidoo et al. 2006, Polasky 2008, Bode et al. 2008, but see Perhans et al. 2008). A corollary of this

50   theory is that as the spatial variation of one feature layer becomes large relative to others, the more

3

51   variable layer increasingly drives solutions (Naidoo et al. 2006). However, despite the potential influence

52   of spatial variation in biodiversity feature or cost data on the solutions obtained, empirical tests of these

53   theoretical expectations predictions are scarce (Arponen 2012, Armsworth et al. 2017). In particular, few

54   studies quantify the contribution of biodiversity feature data on the efficiency of optimized solutions or

55   identify conditions under which 'informed opportunism' in area-based conservation plans is most likely

56   to be achieved.

57   In this paper, we estimate the value of biodiversity feature and land cost data on the efficiency of

58   systematic conservation plans to protect focal birds of the Pacific Northwest of North America.

59   Specifically, we examined how the relative value of cost and biodiversity data varied with (1) data

60   structure and variability, (2) species attributes, (3) conservation targets, (4) and decision rules regarding

61   acceptable levels of habitat suitability (Table 1). Because our study aimed to elucidate general principles

62   underlying efficient conservation planning, rather than to identify a portfolio for real-world

63   implementation, we focused our examination and findings using two contrasting groups of birds – those

64   associated either with forest or human settlements.

65   **Methods**

66   **Study area**

67   We focused on a 27,250 km^2 portion of the Georgia Basin, Puget Trough and Willamette Valley

68   of the Pacific Northwest region spanning the US and Canada (Fig. 1), corresponding to the climate

69   envelope indicative of the Coastal Douglas-fir (CDF) Biogeoclimatic zone in southwestern British

70   Columbia (Meidinger and Pojar 1991). Land cover in the region is diverse, with approximately 57% of

71   the land in forest, 8% as savanna or grassland, 5% in cropland, and 10% being urban or built.

72   **Data Layers**

73   *Biodiversity data.* Our prioritizations were run with eBird data, which is a citizen-science effort that has

74   produced the largest and most rapidly growing biodiversity database in the world (Hochachka et al. 2012,

4

75   Sullivan et al. 2014). From the 2013 eBird Reference Dataset (http://ebird.org/ebird/data/download) we

76   used a total of 12081 checklists in our study area, then filtered these checklists to retain only those <1.5

77   hours in duration, <5 km travelled, and a maximum of 10 visits to a given location (unpublished R code;

78   Hochachka, pers. com.). Sampling locations <100 m apart were collapsed to one location, yielding 5470

79   checklists from 2160 locations, visited from 1-10 times and 2.53 times on average. Following Schuster et

80   al. (2014, 2017) we used a combination of quantitative models and expert elicitation to identify which

81   species were associated either with forest habitat or with human-dominated habitat, such as built or

82   residential land (Supplemental Material methods, Supplementary Table 1).

83   *Cadastral layer and land cost.* We incorporated spatial heterogeneity in land cost (Ando et al. 1998,

84   Polasky et al. 2001, Ferraro 2003, Naidoo et al. 2006) in our plan by using cadastral data and 2012 land

85   value assessments from the Integrated Cadastral Information Society of BC, resulting in 193,623

86   polygons for BC (Schuster et al. 2014). Cadastral data, including tax assessment land values from

87   Washington State came from the University of Washington's Washington State Parcel Database

88   (https://depts.washington.edu/wagis/projects/parcels/; Version: StatewideParcels_v2012n_e9.2_r1.3; Date

89   accessed: 2015/04/30), as well as San Juan County Parcel Data with separate signed user agreement. The

90   combined cadastral layer included 1.92M polygons. Cadastral data, including tax assessment land values

91   from Oregon State had to be sourced from individual counties, which included Benton, Clackamas,

92   Columbia, Douglas, Lane, Linn, Marion, Multnomah, Polk, Washington and Yamhill. The combined

93   cadastral layer for Oregon included 605,425 polygons.

Conservation prioritization

95   To assess the importance of biodiversity data, we compared prioritizations using both cost and

96   biodiversity data to prioritizations using only cost. In both cases, the goal was to identify a set of planning

97   units that captured a given percentage of each species' total occupancy across the entire study region.

98   When prioritizing sites with biodiversity data, we modeled the 'minimum set problem' in conservation

99   planning wherein the goal is to minimize the cost of the solution whilst ensuring that all conservation

5

100   targets are met. This objective is similar to that used in Marxan and detailed in Rodrigues et al. (2000).

101   As such, we used a Marxan-like approach to find the minimum set of planning units that met the given

102   occupancy targets ranging from 5-100% (in 5% increments) for the lowest possible cost. When

103   prioritizing sites without biodiversity data ('uninformed opportunism'), we used a C-rank approach

104   (Supplementary Material Appendix A), whereby sites were selected from the cheapest to most expensive

105   until occupancy targets were satisfied for all species. To explore the influence of constraints on habitat

106   quality, we obtained optimal solutions to our spatial planning problem using three, progressively

107   conservative thresholds for identifying suitable habitat (e.g., $p(occ) \geq 25\%$, $\geq 50\%$, or $\geq 75\%$ likely to

108   occupy a site); doing so was achieved simply by excluding sites with estimated occupancy probability

109   less than the threshold indicated. For all scenarios, we used 1 km^2 planning units, generated by

110   aggregating the species and cost data to this coarser resolution from the original 1-ha cells.

111   The relative value of cost data was assessed by comparing prioritizations generated with both cost

112   and biodiversity feature data, to prioritizations based only on the latter. The value of biodiversity feature

113   data was estimated similarly, by comparing the cost of scenarios that included biodiversity data to those

114   based only on cost (i.e., uninformed opportunism, C-rank). In both cases, we solved the Marxan-like

115   prioritization problem for occupancy targets ranging from 5-100%, in 5% increments, while maintaining a

116   occupancy threshold $\geq 75\%$ to ensure that only high quality habitat was selected. When using cost data

117   we selected the cheapest set of planning units that met the occupancy targets; without cost data, we

118   selected the smallest number of 1 km^2 planning units that met habitat area and quality targets.

119   The above prioritizations were repeated for the 10 forest and 10 human-associated species to

120   explore the consquences of spatial variation in cost, under the expectation that the more variable layer

121   would be disproportionately influential on the prioritized solution. All prioritizations were run using the

122   the prioritizr package (Hanson et al. 2018) in R (R Core Team 2018).

123

124   **Relative variation in costs and benefits**

125   We explored how the relative variation in biodiversity and cost data drove prioritization solutions

126   by examining scenarios in which the coefficient of variation (CV) of the biodiversity data was 2, 4, 8, or

127   16 times the CV of the cost data. To do so, we added a fixed quantity to the cost of each planning unit,

128   which increased the mean cost without altering the standard deviation, thereby decreasing the CV. This

129   quanitity (Δ_{cost}) was chosen based on the following formula:

130
$$\Delta_{cost} = \frac{CV_{relative} \cdot SD_{cost}}{CV_{benefit}} - \mu_{cost}$$

131   where SD is the standard devition, μ is the mean, $CV = SD/\mu$ is the coefficient of variation, $CV_{relative} =$

132    $CV_{benefit}/CV_{cost}$, and Δ_{cost} is the amount added to the cost layer to achieve the desired relative CV of 2,

133   4, 8, or 16. Throughout this process, the benefit CV was held constant and measured as the average CV of

134   the species occupancy layers. We then performed all of the prioritizations described above for each of the

135   relative CV values. In each case (with and without cost data; with and without biodiversity data), we

136   produced cost-benefit curves illustrating the cost, as a percentage of the total cost of the entire study

137   region, to achieve a given occupancy target. More efficient solutions are depicted with steeper cost-

138   benefit curves and reach a higher occupancy target for lower cost. As such, we used the area under the

139   cost-benefit curves as a metric of the efficiency of prioritization approaches across all occupancy targets.

140

141   **Results**

142   Land cost and biodiversity feature data varied widely across our study area for both focal species

143   group. Planning unit costs varied over 8 orders of magnitude, from \$744 to 44.1 billion dollars per km^2

144   (mean = \$78 ± 565 million; CV = 7.25). The coefficients of variation in species occupancy probability

145   predictions ranged from 0.407 to 1.415 (Supplementary Material Table 1). On average, the predicted

146   occurrence of human-associated species was positively related to land cost ($r_{cost} = 0.083 ± 0.094$; mean ±

147   standard deviation), whereas forest species occurrence declined with land cost ($r_{cost} = -0.066 ± 0.053$;

148   mean ± standard deviation; Supplementary Material Table 1).

149   Contrary to the assumption that biodiversity feature data reliably enhances the efficiency of

150   spatially-optimized conservation plans, we found that the relative value of cost and biodiversity data

151   varied by context. First, the value of biodiversity data and efficiency of solutions increased as planning

152   efforts adopted more ambitious conservation targets, and/or became more restrictive by raising the

153   threshold for occupancy, or habitat suitability (Fig. 2, 3). Second, although incorporating land cost in

154   prioritizations tended to make scenarios more cost-effective, efficiency gains declined as the relative

155   variability of biodiversity feature to land cost data increased (Fig. 4, Supplementary Material Fig. 1).

156   Third, we observed that biodiversity data tended to drive solutions more so when spatial variation in

157   biodiversity feature data was high, relative to spatial variation in cost data (Fig. 5, Supplementary

158   Material Fig. 2). These relationships support our expectation that the most variable data layer was likely

159   to be most influential of optimized solutions.

160   The influence of biodiversity feature and land cost data on solutions also differed among focal

161   species as a consequence of underlying correlations between species occurrence and land cost. For

162   example, human-associated birds were much more likely to occupy land that varied greatly in cost than

163   did species relying on mature forest. Although human-associated species are not often targeted for

164   conservation, there are many instances where species of conservation concern are likely to occur in high-

165   cost landscapes (e.g., Coastal California Gnatcatcher, *Polioptila californica californica*). Prioritizations

166   for such 'cost-correlated' species were most efficient when both land cost (Fig. 4) and biodiversity feature

167   data (Fig. 5) were incorporated. In contrast, gains in efficiency achieved by including land cost and/or

168   biodiversity feature data were more modest for mature forest species, whose predicted occurrence was not

169   strongly correlated with variation in land cost.

170

171   **Discussion**

172   Biodiversity feature and land cost data are frequently used to prioritize portfolios of sites

173   potentially capable of achieving conservation goals at the lowest land cost. We estimated the relative

8

174   influence of biodiversity feature and land cost data empirically and illustrated the effect of spatial

175   variation in cost and biological data, on solutions to conservation scenarios over a wide range of targets

176   and thresholds for habitat suitability. Despite some contextual effects, four rules-of-thumb emerged from

177   our analyses of these effects.

178   First, we found that including land costs in spatial prioritizations led to more efficient solutions in

179   almost all cases. Consideration of land or opportunity cost has been widely shown to improve cost-

180   efficiency of biodiversity conservation (Dilkina et al. 2017, Manhaes et al. 2018) and/or reduces negative

181   impacts on extractive and recreational sectors (Mehri et al. 2017).   The value of cost data was similarly

182   demonstrated in a review of global conservation decisions for seven taxonomic groups, for which

183   biodiversity data were typically less influential than socioeconomic concerns (Bode et al. 2008). Yet

184   despite the fact that a vast majority of conservation professionals favorably regarded the inclusion of cost-

185   effectiveness in planning exercises, most consider cost to be less important than other program design

186   elements (Grand et al. 2017) and, hence, seldom include cost as part of return-on-investment evaluations

187   (Murdoch et al. 2007). Indeed, a recent survey of individuals conducting spatial prioritizations showed

188   that only one-quarter to one-third of prioritizations incorporated land value or cost of implementation

189   (Sinclair et al. 2018), suggesting a potential disconnect between motivation and practice in optimization

190   exercises. One barrier to including cost may be the highly variable and aggregated ways that costs are

191   estimated and/or reported (Cook et al. 2017).

192   Second, biodiversity feature data became more influential of scenario outcomes as conservation

193   targets became more ambitious (e.g., scenarios protecting 75% vs. 25% of suitable habitat; Fig. 5). This

194   finding is interesting because conservation targets vary widely in practice; for example, the Convention

195   on Biological Diversity aims to protect 17% of terrestrial ecosystems, whereas the Nature Needs Half

196   movement aims to conserve 50% of 846 ecoregions globally (e.g., natureneedshalf.org). Still higher

197   targets may be applied to species of particular concern to conservation, such as endemic, range-restricted,

198   or critically endangered species.

9

199   Third, the value of biodiversity feature data tended to increase with thresholds used to identify

200   suitable habitat (e.g., probability of occupancy $\geq$ 75% vs. 25%; Fig. 5), underscoring the potential

201   influence of precision in maps used to set thresholds for suitable habitat. For example, uniform range

202   maps (e.g., International Union for the Conservation of Nature (IUCN) Citation, BirdLife International

203   Citation) are widely used in conservation prioritization, but may contribute little spatial variance when

204   used as biodiversity feature data. In contrast, improvements to uniform, expert-elicited, and other course-

205   scale map products are occurring rapidly as citizen-science data are used to enhance existing and create

206   new map products based on multi-species assemblages (e.g., Carroll et al. 2018, Stralberg et al. 2018).

207   Fourth, our most general finding was that the value of biodiversity feature or land cost data

208   depended on its relative variability ($CV_{relative}$) and relationship to each other, and on the extent to which

209   species occurrence patterns were correlated with spatial variation in land cost. As variability in land cost

210   increased relative to variability in biodiversity data, cost increasingly drove solutions and vice versa – a

211   finding that is consistent with Ferraro (2003) and Naidoo et al. (2006). Land cost had particularly strong

212   effects prioritization scenarios targeting 'cost-correlated species', i.e., species whose probability of

213   occurrence increased in areas with high mean and variance in land cost. These effects appeared as

214   comparatively larger efficiency gains in human-associated (positively correlated to cost) than forest-

215   associated birds (weakly negatively correlated to cost). Alternatively, when biodiversity features and

216   costs were negatively correlated in space – as was the case for forest birds in our study, cost had less

217   influence relative to biodiversity data alone. Other empirical studies have also found cost data to be more

218   variable than biodiversity feature data, often by several orders of magnitude (Balmford et al. 2003,

219   Ferraro 2003, Naidoo et al. 2006, Polasky 2008, Bode et al. 2008). Perhans et al. (2008) reported that

220   ecological data tended to be more variable than cost data when selecting among parcels of similar type

221   and value. Taken together, these results and our own suggest that spatial variation in feature data can be

222   used to anticipate its influence on optimized solutions to complex planning problems and, potentially, to

223   evaluate the marginal value of 'better' data given the additional costs or effort required to collect it.

10

224   Spatial prioritizations are increasingly used to guide conservation and a recent survey showed that

225   74% of prioritizations intended for implementation produced action on-the-ground (Sinclair et al. 2018).

226   Because prioritization exercises inform real-world decisions, understanding the manner in which solutions

227   are influenced by the types of data layers included is imperative. We showed that incorporating cost data

228   dramatically improve the efficiency of conservation planning solutions, particularly when biodiversity

229   feature and cost data were positively correlated in space (e.g., when target species occurrence increased

230   with land cost), and when spatial variation in cost exceeded spatial variation in benefits. We further

231   showed that biodiversity feature data became more influential of planning solutions and efficiency as

232   conservation targets and/or the minimum thresholds of habitat suitability were increased, especially in

233   cost-correlated species. One challenge potentially arising for planners is that, in practice, spatial variation

234   in land cost often exceeds variation in biodiversity feature data, especially in areas dominated by humans

235   (Balmford et al. 2003, Polasky 2008), and it is often much easier to estimate than biodiversity feature or

236   benefit data (Naidoo & Adamowicz 2006). Consequently, there may be cases where the marginal value

237   of additional or more precise biodiversity feature data has little or no effect on optimized solutions. It is

238   also the case that access to or the affordability of cost data varies regionally and can be very hard to

239   estimate, such as when tenure is uncertain or contested. Nevertheless, we suggest that considering

240   correlations between cost and benefit data and variability in them should help decision-makers prioritize

241   investments in data acquisition and refinement when attempting to maximize efficiency in spatial

242   prioritizations of land for conservation.

243   **References**

244   Akaike, H. 1974. A new look at the statistical model identification. – IEEE Trans. Automatic Control

245   19:716–723.

246   Ando, A. Camm, J. Polasky, S. and Solow, A. 1998. Species Distributions, Land Values, and Efficient

247   Conservation. – Science 279:2126–2128.

11

248   Armsworth, P. R. Jackson, H. B. Cho, S. Clark, M. Fargione, J. E. Iacona, G. D. Kim, T. Larson, E. R.

249   Minney, T. and Sutton, N. A. 2017. Factoring economic costs into conservation planning may not

250   improve agreement over priorities for protection. – Nature Comm.s 8:2253.

251   Arponen, A. 2012. Prioritizing species for conservation planning. – Biodiv. and Conserv. 21(4):875-893.

252   Arponen, A. Cabeza, M. Eklund, J. Kujala, H. and Lehtomaki, J. 2010. Costs of Integrating Economics

253   and Conservation Planning. – Conserv. Biol. 24(5):1198-1204.

254   Ball, I. R. Possingham, H. P. Watts, M. 2009. Marxan and relatives: software for spatial conservation

255   prioritisation. – Spatial conservation prioritisation: Quantitative methods and computational

256   tools:185–195.

257   Balmford, A. Gaston, K. Blyth, S. James, A. and Kapos, V. 2003. Global variation in terrestrial

258   conservation costs, conservation benefits, and unmet conservation needs. – Proc. Nat. Acad. Sci.

259   USA 100(3):1046-1050.

260   Bennett, J. and Arcese, P. 2013. Human influence and classic island biogeographic predictors of rare

261   species occurrence. – Conserv. Biol. 27: 417-421

262   Beyer, H. L. Dujardin, Y. Watts, M. E. and Possingham, H. P. 2016. Solving conservation planning

263   problems with integer linear programming. – Ecol. Modelling 328:14–22.

264   Beyer, H. L. 2012. Geospatial Modelling Environment (Version 0.7.2.1). (software). – URL:

265   http://www.spatialecology.com/gme.

266   Bode, M. Wilson, K. A. Brooks, T. M. Turner, W. R. Mittermeier, R. A. McBride, M. F. Underwood, E.

267   C. and Possingham, H. P. 2008. Cost-effective global conservation spending is robust to taxonomic

268   group. – Proc. Nat. Acad. Sci. USA 105(17):6498-6501.

12

269   Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical

270   information-theoretic approach. – Springer Verlag, New York, NY.

271   Carroll C. Parks, S.A. Dobrowski, S.Z. Roberts, D.R. 2018. Climatic, topographic, and anthropogenic

272   factors determine connectivity between current and future climate analogs in North America. –

273   Global Change Biology First published: 02 July 2018 https://doi.org/10.1111/gcb.14373.

274   Cook, C. N. Pullin, A. S. Sutherland, W. J. Stewart, G. B. and Carrasco, L. R. 2017. Considering cost

275   alongside the effectiveness of management in evidence based conservation: A systematic reporting

276   protocol. – Biol. Conserv. 209:508-516.

277   De Wan, A, Sullivan, P. J. Lembo, A. J. Smith, C. R. Maerz, J. C. Lassoie, J. P. Richmond, M. E. 2009.

278   Using occupancy models of forest breeding birds to prioritize conservation planning. – Biol.

279   Conserv. 142:982–991.

280   Dilkina, B. Houtman, R. Gomes, C. P. Montgomery, C. A. McKelvey, K. S. Kendall, K. Graves, T. A.

281   Bernstein, R. and Schwartz M. K.. 2017. Trade-offs and efficiencies in optimal budget-constrained

282   multispecies corridor networks. – Conserv. Biol. 31(1):192-202.

283   ESRI. 2012. ArcGIS 10.1 – Economic and Social Research Institute Inc., Redlands, CA.

284   http://www.esri.com/.

285   Ferraro, P. 2003. Assigning priority to environmental policy interventions in a heterogeneous world. – J.

286   Policy Anal. Manage. 22(1):27-43.

287   Fiske, I. J. and Chandler, R. B.. 2011. unmarked : An R Package for Fitting Hierarchical Models of

288   Wildlife Occurrence and Abundance. – J. Stat. Software 43:128–129.

13

289   Gonzales, E. K. Arcese, P. Schulz, R. and Bunnell, F.L. 2003. Strategic reserve design in the central

290   coast of British Columbia: integrating ecological and industrial goals. – Can. J. For. Res. 33: 2129–

291   2140.

292   Grand, L. Messer, K. D. and Allen,W,III. 2017. Understanding and Overcoming the Barriers for Cost-

293   effective Conservation. – Ecol. Econ. 138:139-144.

294   Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat

295   models. – Ecol. Lett. 8:993–1009.

296   Hansen, M. C. Potapov, P. V. Moore, R. Hancher, M. Turubanova, S. A. Tyukavina, A. Thau, D.

297   Stehman, S. V. Goetz, S. J. Loveland, T. R. Kommareddy, A. Egorov, A. Chini, L. Justice, C.O. and

298   Townshend, J.R.G. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. –

299   Science 342:850–853.

300   Hanson, J.O. Schuster, R. Morrell, N. Strimas-Mackey, M.,Watts, M. E. Arcese, P. Bennett, J. and

301   Possingham, H.P. 2018. prioritizr: Systematic Conservation Prioritization in R. – R package version

302   3.0.4.   https://CRAN.R-project.org/package=prioritizr

303   Hochachka, W. M. Fink, D. Hutchinson, R.A. Sheldon, D. Wong, K. and Kelling, S. 2012. Data-intensive

304   science applied to broad-scale citizen science. – Trends Ecol. & Evol. 27:130–137.

305   Homer, C. G. Dewitz, J. A. Yang, L. Jin, S. Danielson, P. Xian, G. Coulston, J. Herold, N. D. Wickham,

306   J. D. Megown, K. 2015. Completion of the 2011 National Land Cover Database for the

307   conterminous United States—Representing a decade of land cover change information. – Phot. Eng.

308   Remote Sens. 81:345–354.

309   Jewell, K. J. Arcese, P. and Gergel, S. 2007. Robust predictions of species distribution: Spatial habitat

310   models for a brood parasite. – Biol. Conserv. 140:259–272.

14

311   Knight, A. T. and Cowling, R. M. 2007. Embracing opportunism in the selection of priority conservation

312   areas. – Conserv. Biol. 21(4):1124-1126.

313   Lawler, J. J. and Edwards, T. C. 2006. A Variance-Decomposition Approach to Investigating Multiscale

314   Habitat Associations. – Condor 108:47–58.

315   Mackenzie, D. I. Nichols, J. D. Lachman, G. B. Droege, S. J. Royle, J. A. and Langtimm, C. A. 2002.

316   Estimating site occupancy rates when detection probabilities are less than one. – Ecology 83:2248–

317   2255.

318   Manhaes, A.P. Loyola, R. Mazzochini, G.G. Ganade, G. Oliveira-Filho, A.T. and Carvalho, A.R. 2018.

319   Low-cost strategies for protecting ecosystem services and biodiversity. – Biol. Conserv. 217:187-

320   194.

321   Meidinger, D. and Pojar, J. 1991. Ecosystems of British Columbia. – British Columbia Ministry of

322   Forests, Victoria, BC.

323   MES. 2008. Terrestrial Ecosystem Mapping of the Coastal Douglas-Fir Biogeoclimatic Zone. – Madrone

324   Environmental Services LTD., Duncan, BC. Mandrone Environmental Services LTD., Duncan, BC.

325   Murdoch, W. Polasky, S. Wilson, K.A. Possingham, H.P. Kareiva, P. and Shaw, R. 2007. Maximizing

326   return on investment in conservation. – Biol. Conserv. 139(3-4):375-388.

327   Naidoo, R. and Adamowicz, W.L. 2006. Modeling opportunity costs of conservation in transitional

328   landscapes. – Conserv. Biol. 20:490-500.

329   Naidoo, R. Balmford, A. Ferraro, P. J. Polasky, S. Ricketts, T. H. and Rouget, M. 2006. Integrating

330   economic costs into conservation planning. – Trends Ecol. & Evol. 21(12):681-687.

15

331   Perhans, K. Kindstrand, C. Boman, M. Djupstrom, L. B. Gustafsson, L. Mattsson, L. Schroeder, L.

332   Weslien, J.M. and Wikberg, S. 2008. Conservation Goals and the Relative Importance of Costs and

333   Benefits in Reserve Selection. – Conserv. Biol. 22(5):1331-1339.

334   Polasky, S. Camm, J. D. and Garber-Yonts, B. 2001. Selecting Biological Reserves Cost-Effectively: An

335   Application to Terrestrial Vertebrate Conservation in Oregon. – Land Econ. 77:68–78.

336   Polasky, S. 2008. Why conservation planning needs socioeconomic data. – Proc. Nat. Acad. Sci. USA

337   105(18):6505-6506.

338   Pressey, R. Humphries, C. C. Margules, C. Vanewright, R. and Williams, P. 1993. Beyond Opportunism

339   - Key Principles for Systematic Reserve Selection. – Trends Ecol. & Evol. 8(4):124-128.

340   Pressey, R. L. and Bottrill, M. C. 2008. Opportunism, Threats, and the Evolution of Systematic

341   Conservation Planning. – Conserv. Biol. 22(5):1340-1345.

342   R Development Core Team. 2012. R: A language and environment for statistical computing 2.15.2,

343   http://www.r-project.org. – R Foundation for Statistical Computing, Vienna, Austria.

344   R Core Team. 2018. R: A language and environment for statistical computing. – R Foundation for

345   Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

346   Schuster, R. and Arcese, P. 2013. Using bird species community occurrence to prioritize forests for old

347   growth restoration. – Ecography 36:499–507.

348   Schuster, R. Wilson, S. Rodewald, A.D. Arcese, P. Fink, D. Auer, T. and Bennett, J. 2018. Optimizing

349   Conservation of Migratory Species over Their Full Annual Cycle in the Western Hemisphere. –

350   bioRxiv. https://doi.org/10.1101/268805.

351   Sinclair, S.P. Milner-Gulland, E.J. Smith, R.J. McIntosh, E.J. Possingham, H.P. Vercammen, A. Knight,

352   A.T. 2018. The use, and usefulness, of spatial conservation prioritizations. – Conserv. Letters

353   2018: e12459.

354   Stralberg D. Carroll, C. Pedlar, J.H. Wilsey, C.B. McKenney, D.W., Nielsen, S.E., Fortin, M-J. 2018.

355   Macrorefugia for North American trees and songbirds: Climatic limiting factors and multi-scale

356   topographic influences. – Glob. Ecol.and Biogeogr. 27: 690-703.

357   Sullivan, Aycrigg, Barry, Bonney, Bruns, Cooper, Dhondt, Dietrich, Farnsworth, Fink, Gerbracht,

358   Gomes, Hochachka, Iliff, Lagoze, La Sorte, Merrifield, Phillips, Reynolds, Rodewald, Rosenberg,

359   Trautmann, Wiggins, Winkler, Wong, Wood, Yu, Kelling. 2014. The eBird enterprise: an

360   integrated approach to the development and application of citizen science. – Biol. Conserv. 169:31-

361   40.

362   Wang, T. Hamann, A. Spittlehouse, D. and Carroll, C. 2016. Locally downscaled and spatially

363   customizable climate data for historical and future periods for North America. – PLoS One.

364   https://doi.org/10.1371/journal.pone.0156720

365   Wilson S. Schuster, R. Rodewald, A.D. Bennett, J., Smith, A. and Arcese, P.. 2018. Prioritize diversity or

366   declining species? Trade-offs and synergies in spatial planning for the conservation of migratory

367   birds. bioRxiv, 429019. –

368   White, G. C. and Burnham, K. P.1999. Program MARK: survival estimation from populations of marked

369   animals. – Bird Study 46:120–139.

17

370   Table 1. Range of values or conditions evaluated across prioritization scenarios.

Attribute	Values
Relative variability of data	CV of biodiversity data were 2, 4, 8, or 16 times that of cost data
Conservation target	0-100% of populations protected
Habitat suitability (occupancy threshold)	A species has a 25%, 50%, or 75% probability of occurrence
Species group	Forest vs. human-associated birds
Land cost	Incorporated vs not incorporated
Biodiversity data	Included vs. not included

371
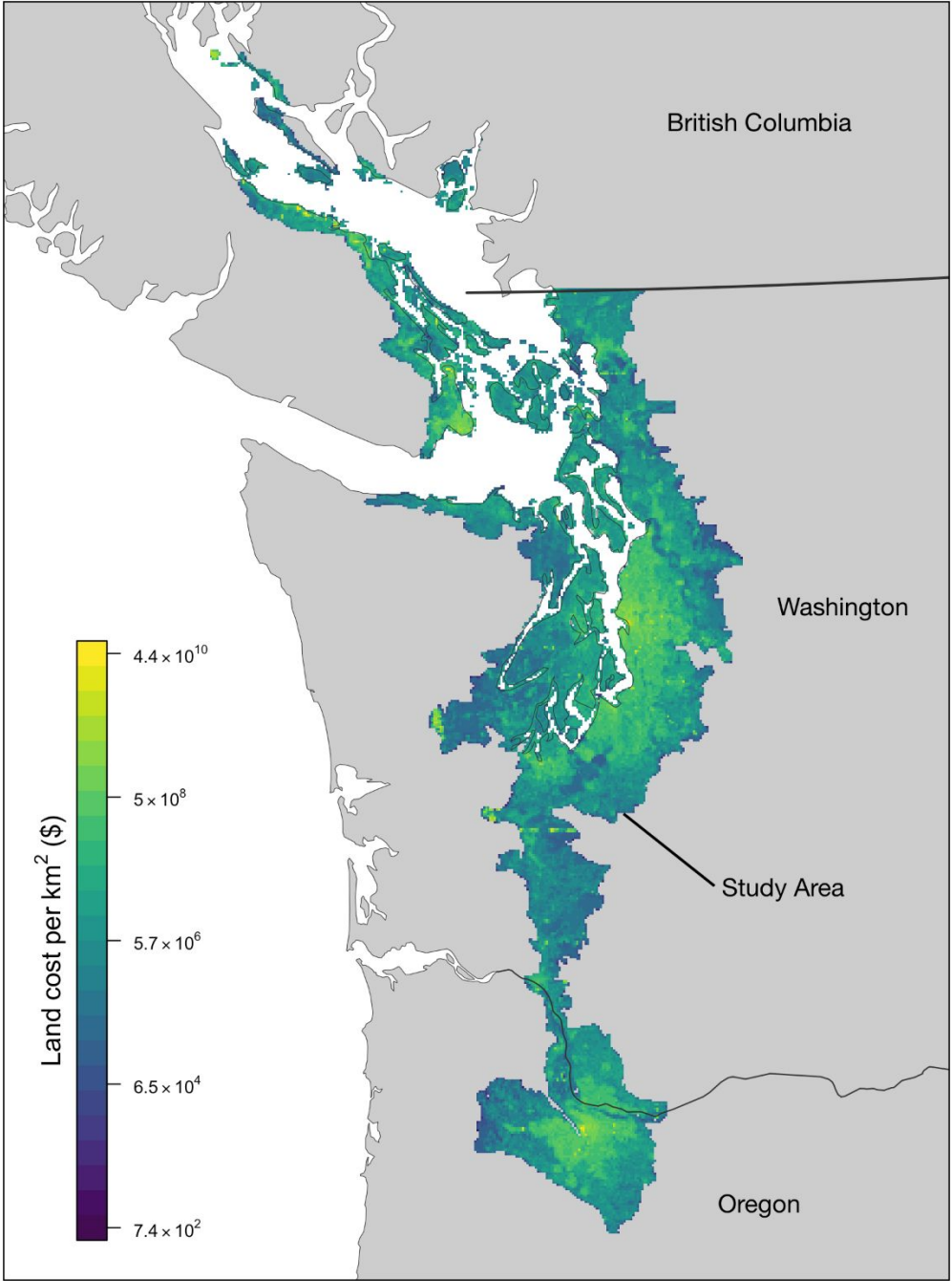
Figure 1. The focal area for this study showing the cost of 1 km^2 planning units.
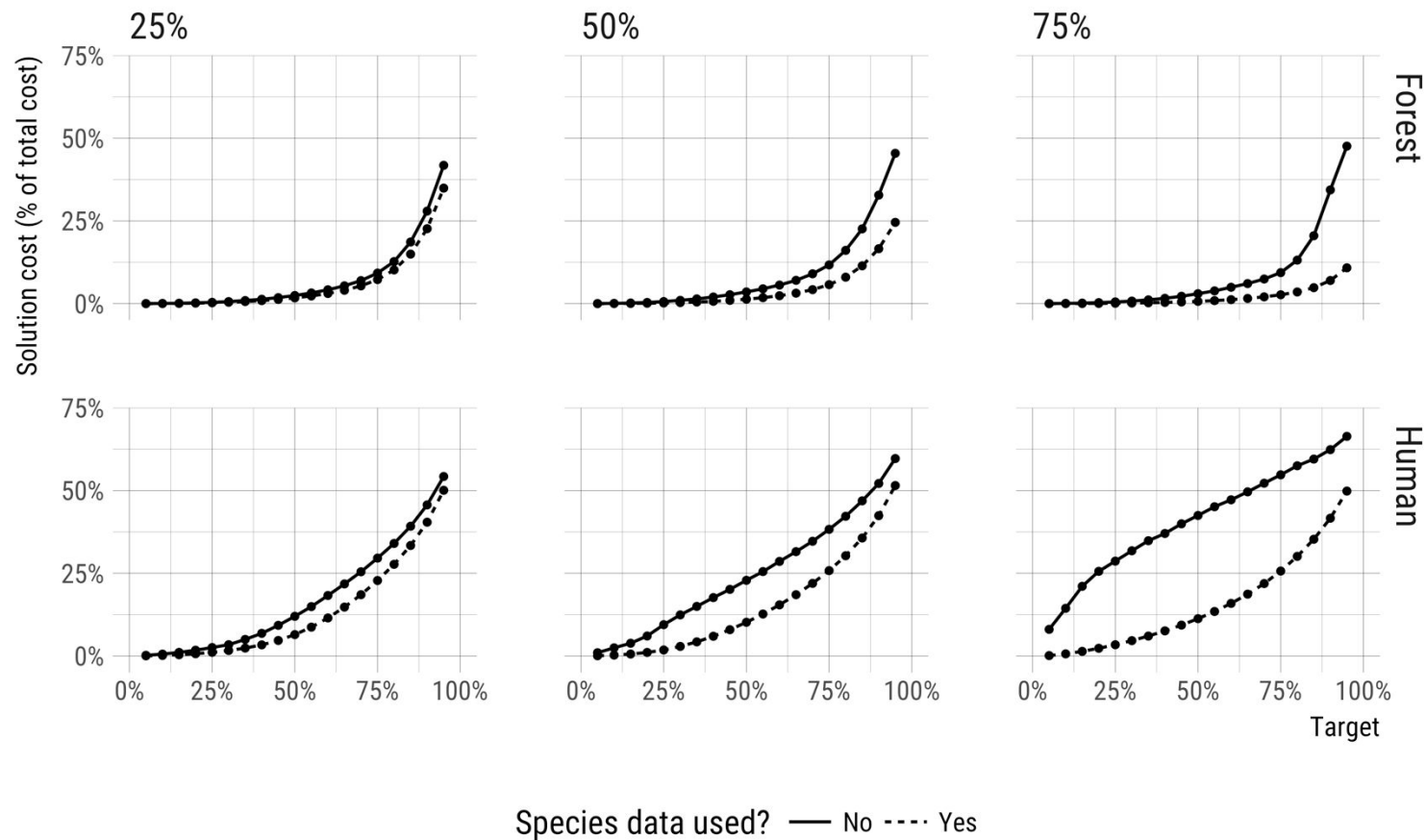
Figure 2. Using biodiversity feature data in conservation prioritization (dashed line) improved the efficiency of meeting conservation targets as compared to using parcel cost alone (solid line). Cost-only solutions were derived purchasing land from least to most expensive until targets were met (C-rank prioritization; biodiversity data only used to determine when targets were meet; see Methods). Only parcels meeting the indicated occupancy threshold (25%, 50%, or 75%) were used to ensure the selection of parcels where species were very likely to occur.
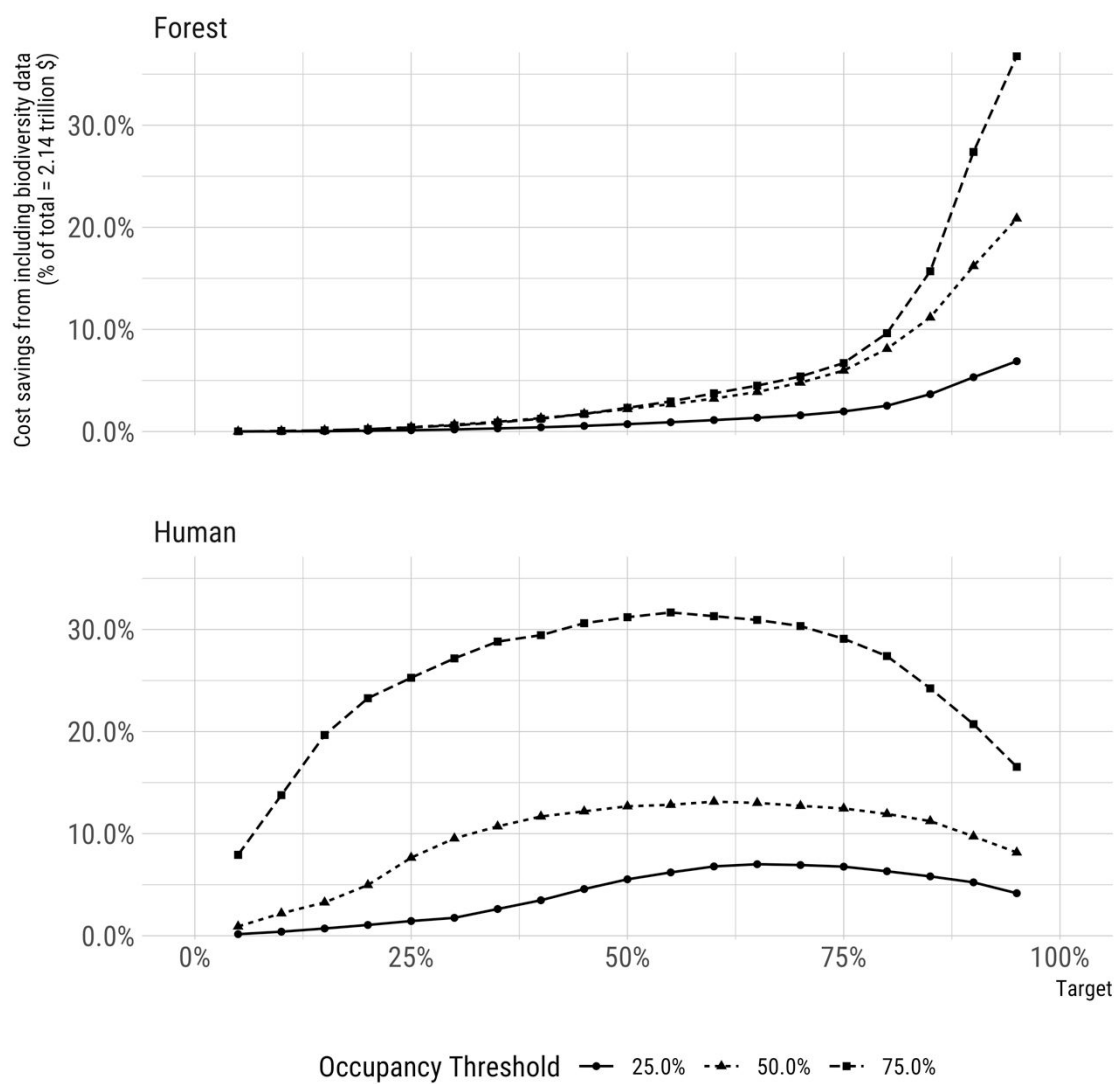
Figure 3. Cost savings varied widely across conservation targets and occupancy thresholds when including or excluding biodiversity feature data in Marxan-like prioritizations for forest and human-associated species. Restricting the prioritization to only select higher quality habitat (i.e. increasing the occupancy threshold), led to greater cost savings from including biodiversity data. Similarly, higher occupancy targets also led to an increase in the cost savings from including biodiversity data.
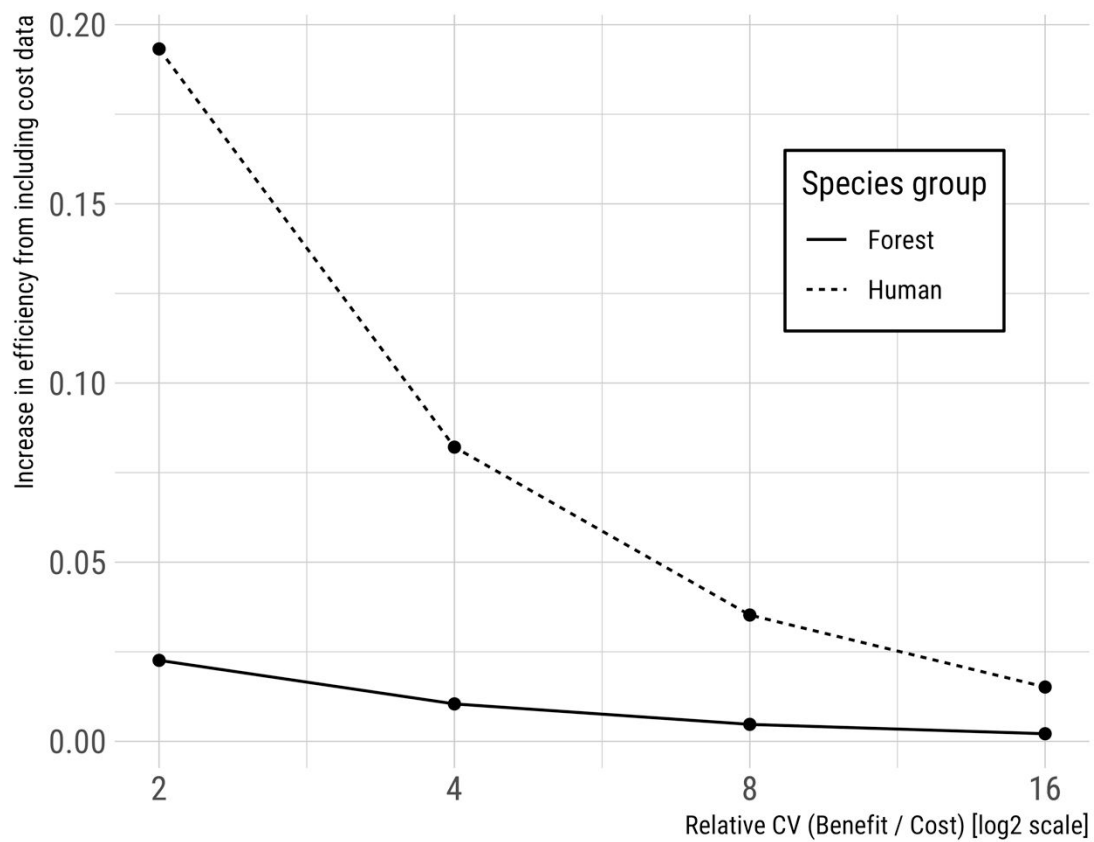
Figure 4. Fractional gain in efficiency when using both cost and biodiversity data, as compared to biodiversity data alone, declined as the relative variability of costs decreased. Human-associated species (dashed line) experienced a greater gain in efficiency from incorporating cost data than forest-associated species (solid line).
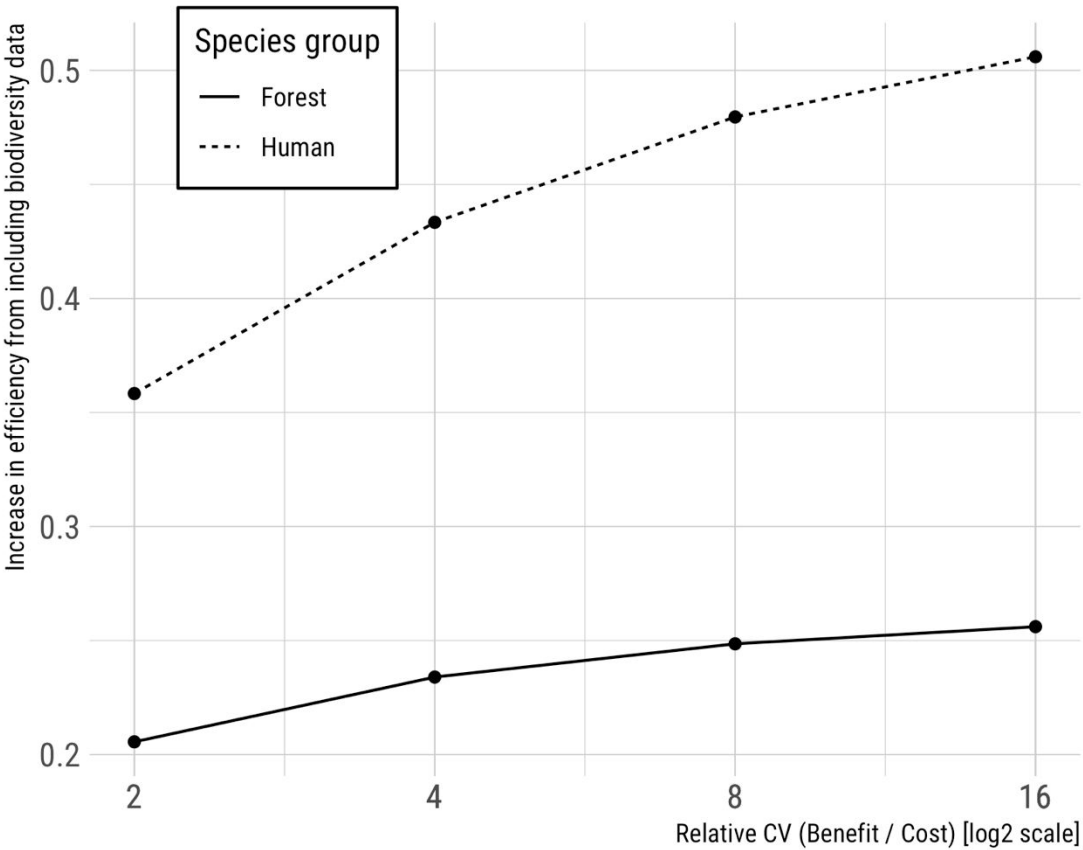
Figure 5. Fractional gain in efficiency from including biodiversity data in addition to cost data, compared to cost data alone, increased as the relative variability of biodiversity data increased. Human-associated species (dashed line) experienced a greater gain in efficiency from incorporating species data than forest-associated species (solid line).

23

Supplementary Information

Landscape covariates

Because birds are well-known to respond to many fine and coarse scale habitat features (Lawler and Edwards 2006), we developed covariate descriptors of landscape condition and context using coarse (1km) and fine (100m) scale features to advance early work conducted at coarse scales alone (De Wan et al. 2009). For modelling species detection and occurrence, we chose candidate predictors based on their proven ability to predict species occurrence at site and landscape levels in similar exercises or regions (Guisan and Thuiller 2005, Jewell et al. 2007, Schuster & Arcese 2013). All covariate names appear in Table S2 and were derived from the following sources: (i) Terrain Resource Information Management (TRIM, http://archive.ilmb.gov.bc.ca/crgb/pba/trim/specs/specs20.pdf), (ii) Terrestrial Ecosystem Mapping (TEM) of the CDF Zone (MES 2008), (iii) National Hydrography Dataset (http://nhd.usgs.gov/data.html), (iv) National Land Cover Database 2011 (Homer et al. 2015), (v) Global Forest Change (Hansen et al. 2013), (vi) OpenStreetMap data extracts (http://download.geofabrik.de/, accessed 2015-09-17), and (vii) Species distribution maps for 21 tree species in the region (Wang et al. 2016, Schuster and Arcese, unpublished data). Our dataset comprised 37 predictor covariates of site and landscape condition, derived at each of 2160 avian point count locations. All covariates were created using Geospatial Modeling Environment (Beyer 2012) in conjunction with ArcGIS 10.1 (ESRI 2012) and R v. 2.15.2 (R Development Core Team 2012). Due to their widely varying scales, all covariates were standardized about their mean value, to ensure that importance was not driven by measurement scale (White and Burnham 1999).

Occupancy and detection models

We assumed no variation in site occupancy across sampling occasions to minimize model complexity, thus assumed a closed population for all species (Mackenzie et al. 2002). The R package unmarked v. 0.9-9 (Fiske and Chandler 2011) provided the framework for all species models, which

necessarily include two parts: occupancy and detection (Mackenzie et al. 2002). To estimate detectability, we used seven observation specific covariates (observer ID, Julian date, time of day, count type, effort hours, distance travelled, number of observers). For each of 73 bird species we fitted detectability models (without parameterizing occupancy) using a machine learning algorithm we developed to work with the unmarked package (Supplementary Material Appendix A, modified from Schuster and Arcese, 2013) and then ranked each by AIC (Akaike 1974). We then ranked all candidate models by AIC and averaged those with ΔAIC ≤ 7 from the top ranked one (Burnham and Anderson 2002). The covariates from the averaged model were retained for further analysis as fixed parts of further models. Next we used the same machine learning algorithm from above to model occupancy using the 37 predictor covariates introduced earlier. We again ranked all candidate models by AIC and averaged those with ΔAIC ≤ 7 from the top ranked one (Burnham and Anderson 2002) to create our final averaged model for each of the 73 focal species.

Expert elicitation and habitat associations

We asked 11 professional ornithologists with >5 years of local experience to rank the likelihood of 47 species in 10 focal habitat types using photographic and text descriptions of herbaceous, shrub, woodland, wetland, four forest types (pole, young, mature and old), and 2 human-dominated habitats (rural, urban). Experts ranked species by association (low = -1, medium = 0 or highly associated = 1) with each of 10 focal habitat types in CDF habitats. We then averaged ranks to describe habitat associations for each bird and habitat type and created two community association scores indicative of Human-dominated and Old Forest habitat associations that were standardized between 0 and 1 by dividing through the maximum value possible, where:

$$OldForest = \frac{-2*Herb - 1*Shrub - 0.5*Pole + 0.5*YFor + 1*MFor + 2*OFor}{7}$$

$$Human = \frac{2*URB + 1*RUR}{3}$$

We selected two groups to illustrate contrasting scenarios – one in which species were positively associated with land cost (human-associated species) and one in which species were negatively or uncorrelated with land cost (forest species). For the Human-associated species, we selected the 10 species with the highest Human community association score: European Starling (*Sturnus vulgaris*), House Sparrow (*Passer domesticus*), Rock Dove (*Columba livia*), American Crow (*Corvus brachyrhynchos*), House Finch (*Haemorhous mexicanus*), European Collared Dove (*Streptopelia decaocto*), Anna's Hummingbird (*Calypte anna*), American bushtit (*Psaltriparus minimus*), Barred Owl (*Strix varia*), and Stellar's Jay (*Cyanocitta stelleri*). For the species associated with forest, we selected the 10 species with the highest Old Forest community association score: Pileated Woodpecker (*Hylatomus pileatus*), Brown Creeper (*Certhia americana*), Red-breasted Nuthatch (*Sitta canadensis*), Hairy Woodpecker (*Leuconotopicus villosus*), Red Crossbill (*Loxia curvirostra*), Pacific Slope Flycatcher (*Empidonax difficilis*), Pacific Wren (*Troglodytes pacificus*), Varied Thrush (*Ixoreus naevius*), Townsend's Warbler (*Setophaga townsendi*), and Golden-crowned Kinglet (*Regulus satrapa*).

Supplementary Table 1. Cell-wise coefficient of variation (CV) and correlation with land cost (Pearson's correlation coefficient; r_{cost}) of each species' occupancy probability layer.

Forest species			
Common Name	**Scientific Name**	**CV**	**r_{cost}**
Hairy Woodpecker	*Picoides villosus*	0.751	-0.104
Pileated Woodpecker	*Dryocopus pileatus*	0.579	-0.094
Pacific-slope Flycatcher	*Empidonax difficilis*	0.587	-0.110
Red-breasted Nuthatch	*Sitta canadensis*	0.427	-0.047
Brown Creeper	*Certhia americana*	0.522	-0.081
Pacific Wren	*Troglodytes pacificus*	0.468	-0.127
Golden-crowned Kinglet	*Regulus satrapa*	0.466	-0.033
Varied Thrush	*Ixoreus naevius*	0.545	-0.057
Red Crossbill	*Loxia curvirostra*	0.474	0.057
Townsend's Warbler	*Setophaga townsendi*	0.883	-0.061
All forest (mean ± sd)		**0.570 ± 0.143**	**-0.066 ± 0.053**
Human-associated species			
Rock Pigeon	*Columba livia*	1.047	0.110
Eurasian Collared-Dove	*Streptopelia decaocto*	0.990	-0.034
Anna's Hummingbird	*Calypte anna*	0.811	0.161
Barred Owl	*Strix varia*	1.415	-0.039
Steller's Jay	*Cyanocitta stelleri*	0.407	-0.065
American Crow	*Corvus brachyrhynchos*	0.419	0.106
Bushtit	*Psaltriparus minimus*	0.922	0.176
European Starling	*Sturnus vulgaris*	0.702	0.086
House Finch	*Haemorhous mexicanus*	0.789	0.160
House Sparrow	*Passer domesticus*	0.993	0.167
All human (mean ± sd)		**0.850 ± 0.300**	**0.083 ± 0.094**
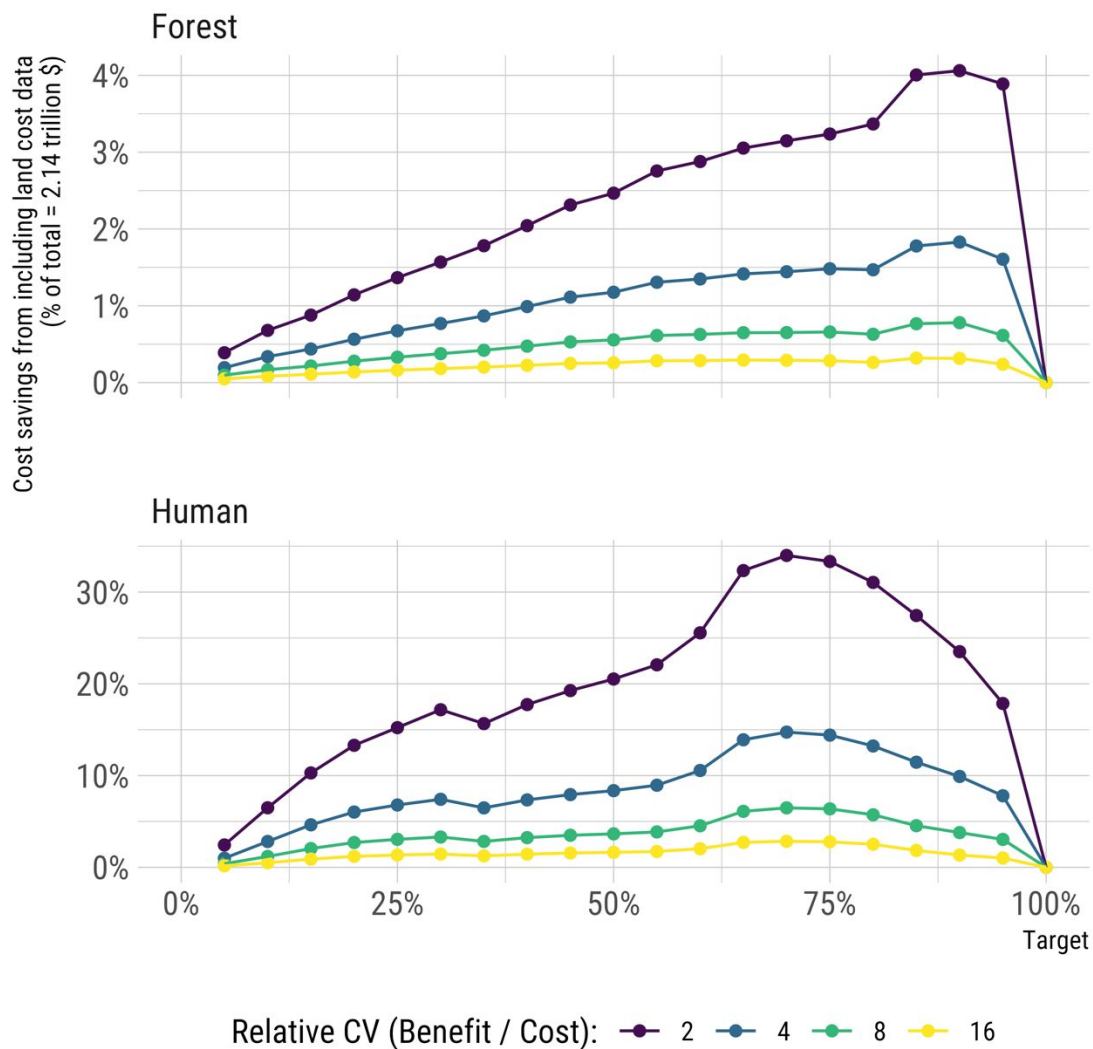
Supplementary Figure 1. The savings associated with using cost and biodiversity data, compared to using only biodiversity data, when prioritizing land across different representation target levels. These savings increased as the coefficient of variation of cost data increase relative to biodiversity data (decreasing relative CV). All prioritizations were performed using an occupancy threshold of 75%. Note that the relative CV is shown on a logarithm scale to highlight differences in color between the lines.
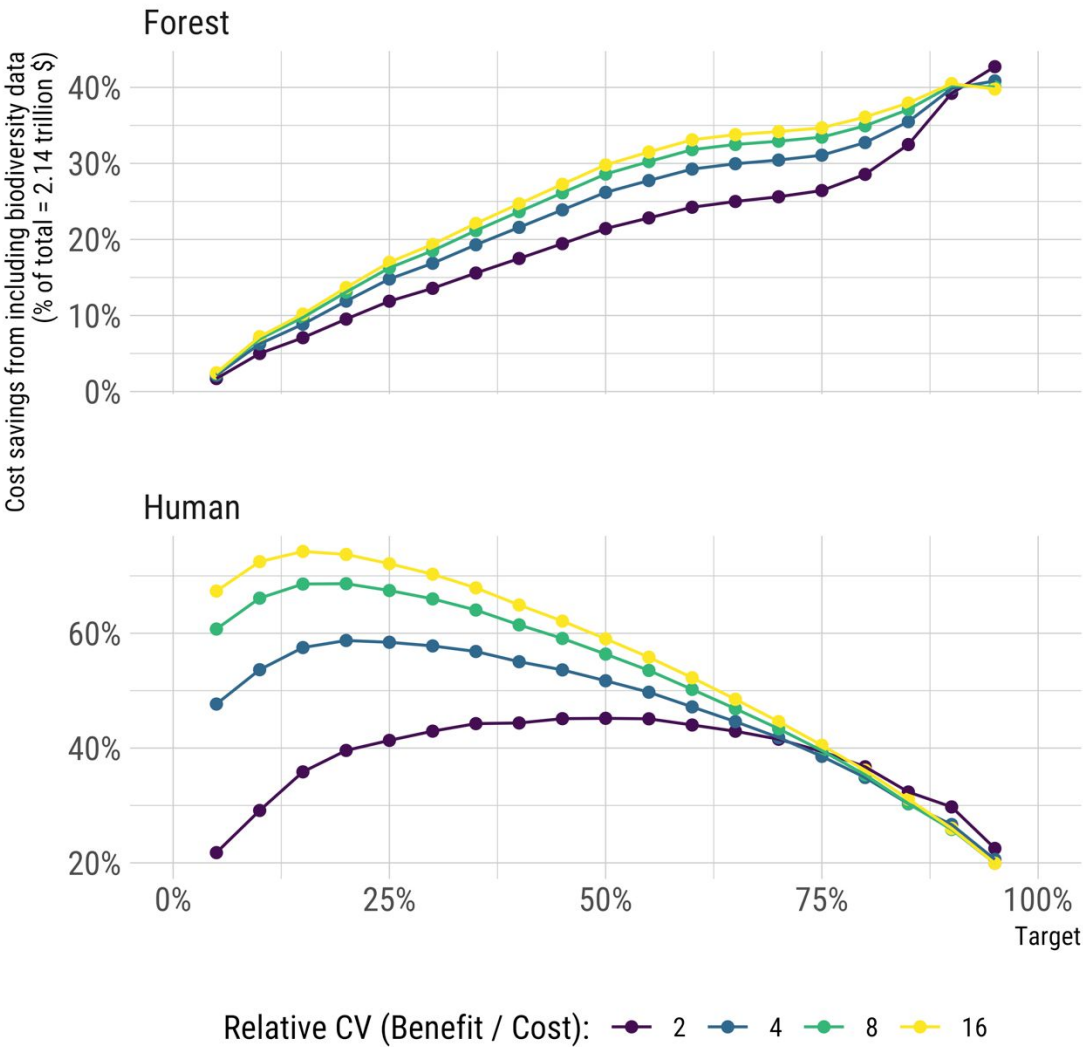
Supplementary Figure 2. The savings associated with using biodiversity and cost data, compared to using only cost data, when prioritizing land across different representation target levels. These savings increased as the coefficient of variation of savings data increase relative to cost data. All prioritizations were performed using an occupancy threshold of 75%. Note that the relative CV is shown on a logarithm scale to highlight differences in color between the lines.

Appendix A. R code of the function developed to fit occupancy models using a machine learning approach and to run the C-rank prioritization approach.

```
f.AICc.occu.sig <- function(start.model, blocks, max.iter = NULL, detocc = 1,

                            AICcut = 1, p.crit = 0.15, flag_se = FALSE, um.frame = "",

                            print.log = TRUE, ...) {

  # f.AICc.occu.sig: a function for "stepwise" regression using occupancy models of package
unmarked

  # start.model: initial model e.g. occu(~1~1, UMF)

  # detocc: if set to 1 (default) runs the function on the occupancy side; 2 does the
detectability

  # foreback: if TRUE covariates are allowed to be dropped; if FALSE only adding covariates
allowed

  # Parts based on Forward.lmer by Rense Nieuwenhuis (http://www.rensenieuwenhuis.nl/r-sessions-
32/)

  # with some additions by for Nick Isaac

  # Author: Richard Schuster (mail@richard-schuster.com

  # 08 October 2015


  modlst <- c(start.model)

  x <- 2

  if (detocc == 1) {

    coeff <- length(start.model@estimates@estimates$state@estimates)

  } else {

    coeff <- length(start.model@estimates@estimates$det@estimates)

  }


  best <- FALSE

  model.basis <- start.model

  keep <- list(start.model)

  AICmin <- AIClst <- AICc(start.model)

  # cutoff for when to exclude values

  cutoff <- 20


  # critical z value alhpa = 0.3 (1.036433); 0.15 (1.439531); 0.1 (1.644853); 0.05 (1.959964)

  zc <- qnorm(1 - (p.crit / 2))



  # Maximum number of iterations cannot exceed number of blocks, but this is also the default

  if (is.null(max.iter) | max.iter > length(blocks)) max.iter <- length(blocks)
```

```
# Setting up the outer loop
for (ii in 1:max.iter) {
  models <- list()
  coeff <- coeff + 1
  cnt <- 1
  for (jj in 1:length(keep)) {
    # Iteratively updating the model with addition of one block of variable(s)
    for (kk in 1:length(blocks)) {

      # check if blocks[kk] is already in the model, if so go to next kk
      if (detocc == 1) {
        if (blocks[kk] %in% names(keep[[jj]]@estimates@estimates$state@estimates)) next
      } else {
        if (blocks[kk] %in% names(keep[[jj]]@estimates@estimates$det@estimates)) next
      }



      if (detocc == 1) {
        form <- as.formula(paste("~. ~. + ", blocks[kk]))
      } else {
        form <- as.formula(paste("~. + ", blocks[kk], "~. "))
      }


      if (class(dummy <- try(update(keep[[jj]], form, data = um.frame))) == "unmarkedFitOccu")
{
        flag <- 0

        # check if model converged
        if (dummy@opt$converge != 0) flag <- 1

        # check if there is any NAN's in the SE's of the occupancy side
        if (flag == 0 && detocc == 1) {
          if (any(diag(vcov(dummy@estimates@estimates$state)) < 0) ||
            any(sqrt(diag(vcov(dummy@estimates@estimates$state))) > cutoff)) {
            flag <- 1
          }
        }
```

```
    # check if there is any NAN's in the SE's of the detection side
    if (flag == 0 && detocc == 2) {
      if (any(diag(vcov(dummy@estimates@estimates$det)) < 0) ||
        any(sqrt(diag(vcov(dummy@estimates@estimates$det))) > cutoff)) {
        flag <- 1
      }
    }


    # check for repeat models
    if (flag == 0) {
      for (bb in 1:length(AIClst)) {
        if (round(AICc(dummy), digits = 6) == round(AIClst[bb], digits = 6)) {
          flag <- 1
          break
        }
      }
    }


    # Remove models where z < zc or SE = 0 for a beta > 0
    if (flag == 0 && flag_se == TRUE) {
      if (detocc == 1) {
        if (any(abs(dummy@estimates@estimates$state@estimates[-1] /
          sqrt(diag(vcov(dummy@estimates@estimates$state)))[-1]) <
          zc)) {
          flag <- 1
        }
      } else {
        if (any(abs(dummy@estimates@estimates$det@estimates[-1] /
          sqrt(diag(vcov(dummy@estimates@estimates$det)))[-1]) <
          zc)) {
          flag <- 1
        }
      }
    }
}
else {
  flag <- 1
}
```

```
    # add dummy model to the model list if it passes all previous tests

    if (flag == 0) {

      models[[cnt]] <- dummy

      modlst[[x]] <- models[[cnt]]

      AIClst <- c(AIClst, AICc(models[[cnt]]))

      x <- x + 1

      cnt <- cnt + 1

    }

  }

}


if (length(LL <- unlist(lapply(models, function(x) {

  AICc(x)

}))) == 0) {

  break

}

keep <- list()

k <- 1

cont <- 0

# check for improvement in AIC, if none stop loop

for (mm in order(LL, decreasing = FALSE)) {

  if (LL[mm] < AICmin + AICcut) {

    if (detocc == 1) {

      if (length(models[[mm]]@estimates@estimates$state@estimates) == coeff) {

        keep[[k]] <- models[[mm]]

        k <- k + 1

        if (LL[mm] < AICmin) {

          AICmin <- LL[mm]

          cont <- 1

        }

      }

    }

    else {

      if (length(models[[mm]]@estimates@estimates$det@estimates) == coeff) {

        keep[[k]] <- models[[mm]]

        k <- k + 1

        if (LL[mm] < AICmin) {
```

```
                AICmin <- LL[mm]

                cont <- 1

            }

          }

        }

      }

      else {

        break

      }

    }

    rm(models)

    gc()

    if (length(keep) == 0) break

  }

  ## Create Model List

  fitlst <- fitList(fits = modlst)

  modsel <- modSel(fitlst, nullmod = NULL)


  ## Return the gathered output

  return(list(model = model.basis, modlst = modlst, fitlst = fitlst, modsel = modsel))

}


###################################################
###################################################
##### Post-processing functions           #####
###################################################
###################################################


occu.subset <- function(model, cutoff = 2) {

  # subsetting the above function output so it can be used with

  # model.avg from package MuMIn

  # 2015-08-04

  mdlst <- model

  f.occ.aic <- unlist(lapply(mdlst, function(x) AIC(x)))

  min.aic <- min(f.occ.aic)

  f.occ.delta <- f.occ.aic - min.aic

  return(mdlst[f.occ.delta < cutoff])

}
```

```
occu.rem.non.sig <- function(model, p.crit = 0.15) {

  # critical z value alhpa = 0.3 (1.036433); 0.15 (1.439531); 0.1 (1.644853); 0.05 (1.959964)

  zc <- qnorm(1 - (p.crit / 2))


  model <- model$modlst


  sig <- unlist(lapply(model, function(x) all(abs(x@estimates@estimates$state@estimates[-1] /
      sqrt(diag(vcov(x@estimates@estimates$state)))[-1]) > zc)))


  return(model[sig])
}


##################################################
##################################################
##### Functions from package MuMIn         #####
##################################################
##################################################
.coefarr.avg <-
  function(cfarr, weight, revised.var, full, alpha) {
    weight <- weight / sum(weight)
    nCoef <- dim(cfarr)[3L]
    if (full) {
      nas <- is.na(cfarr[, 1L, ]) & is.na(cfarr[, 2L, ])
      cfarr[, 1L, ][nas] <- cfarr[, 2L, ][nas] <- 0
      # cfarr[, 1L:2L, ][is.na(cfarr[, 1L:2L, ])] <- 0
      if (!all(is.na(cfarr[, 3L, ]))) {
        cfarr[, 3L, ][is.na(cfarr[, 3L, ])] <- Inf
      }
    }


    avgcoef <- array(
      dim = c(nCoef, 5L),
      dimnames = list(dimnames(cfarr)[[3L]], c(
        "Estimate",
        "Std. Error", "Adjusted SE", "Lower CI", "Upper CI"
      ))
    )
```

35

```r
    for (i in seq_len(nCoef))
      avgcoef[i, ] <- par.avg(cfarr[, 1L, i], cfarr[, 2L, i], weight,
        df = cfarr[, 3L, i], alpha = alpha, revised.var = revised.var
      )


    avgcoef[is.nan(avgcoef)] <- NA
    return(avgcoef)
  }



.makecoefmat <- function(cf) {
  no.ase <- all(is.na(cf[, 3L]))
  z <- abs(cf[, 1L] / cf[, if (no.ase) 2L else 3L])
  pval <- 2 * pnorm(z, lower.tail = FALSE)
  cbind(cf[, if (no.ase) 1L:2L else 1L:3L],
    `z value` = z, `Pr(>|z|)` = zapsmall(pval)
  )
}


add.coefmat <- function(object) {
  is.arm <- ncol(object$msTable) == 6L && (colnames(object$msTable)[6L] == "ARM weight")

  weight <- object$msTable[, if (is.arm) 6L else 5L]

  object$coefmat <- .makecoefmat(.coefarr.avg(
    object$coefArray, weight,
    attr(object, "revised.var"), TRUE, 0.05
  ))
  return(object)
}
```

36

R code for implementing C-rank prioritization.

```
#' Select sites via a C-Rank prioritization
#'
#' C-Rank prioritization is a greedy prioritization algorithm that selects site
#' starting at the cheapest and proceeding in order of increasing cost, until a
#' specified target level of representation is met for all of the input features.
#'
#' @param cost A RasterLayer object of the cost to protect each planning unit.
#' @param features A RasterStack object of biodiviersity features, with each
#'   feature typically being the occupancy or abundance of a given species.
#' @param target A value between 0-1 specifying the percent of the total
#'   biodiveristy value that must be met for all features.
#'
#' @return A logical vector specifying whether each planning unit has been
#'   selected or not.
prioritize_crank <- function(cost, features, target) {
  # convert to vectors and matrices, removing NA cells
  cost_v <- cost[]
  included_pu <- which(!is.na(cost_v))
  cost_v <- cost_v[included_pu]
  features_m <- t(features[][included_pu, ])
  species <- names(features)

  # order to choose PUs in
  cost_order <- order(cost_v)

  # decision variable
  x <- rep(FALSE, length(cost_v))

  # total amount
  total_amount <- as.numeric(features_m %*% rep(TRUE, length(cost_v)))

  # watched
  spp_cost <- setNames(rep(NA_real_, length(species)), species)
  spp_n <- setNames(rep(NA_integer_, length(species)), species)

  for(i in seq_along(cost_order)) {
```

```
    x[cost_order[i]] <- TRUE

    # check representation

    pct_rep <- as.numeric(features_m %*% x) / total_amount

    target_met <- pct_rep > target

    # keep track of cost associated with meeting each target

    if (any(target_met & is.na(spp_cost))) {

      cost_x <- sum(cost_v[x], na.rm = TRUE)

      spp_cost[target_met] <- dplyr::coalesce(spp_cost[target_met], cost_x)

      spp_n[target_met] <- dplyr::coalesce(spp_n[target_met], i)

    }

    # if all targets met, finish

    if (all(!is.na(spp_cost))) {

      break()

    }

  }

  # prepare results

  x_out <- rep(FALSE, raster::ncell(cost))

  x_out[included_pu] <- x

  x_out

}
```