Protecting Biodiversity (in All Its Complexity): New Models and Methods

Laura J. Pollock, Louise M.J. O'connor, Karel Mokany, Dan F. Rosauer, Matthew V. Talluto, Wilfried Thuiller

HAL Id: hal-03493786
https://hal.science/hal-03493786v1

Submitted on 21 Nov 2022

1 **Protecting biodiversity (in all its complexity): new models and methods**

2

3 Laura J. Pollock[1,2*], Louise M. J. O'Connor[2], Karel Mokany[3], Dan F. Rosauer[4], Matthew V.

4 Talluto[5, 6] and Wilfried Thuiller[2*]

5

6

7 [1]Department of Biology, McGill University, 1205 Dr. Penfield, Montréal, Québec H3A 1B1,

8 Canada.

9 [2]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Laboratoire d'Écologie

10 Alpine, F-38000 Grenoble, France

11 [3]CSIRO, PO Box 1700, Canberra, Australian Capital Territory 2601, Australia

12 [4]Research School of Biology, Australian National University, Acton, Australian Capital

13 Territory, 2601, Australia.

14 [5]Department of Ecohydrology, Leibniz Institute for Freshwater Ecology and Inland Fisheries

15 Müggelseedamm 310, 12587 Berlin, Germany.

16 [6]Department of Ecology, University of Innsbruck, Innrain 52, AT-6020 Innsbruck, Austria

17

18 *Corresponding Author: L. Pollock, lajosy@gmail.com, +33 6 1041 3058, @lajosy0

19

20 **Keywords**

21 Conservation prioritization, multi-trophic diversity, species distribution models, phylogenetic

22 diversity, functional diversity

23

24

25 **Abstract**

26 We are facing a biodiversity crisis at the same time as we are acquiring an unprecedented view of the

27 world's biodiversity. Vast new datasets (e.g. species distributions, traits, phylogenies, and interaction

28 networks) hold knowledge to better comprehend the depths of biodiversity change, reliably anticipate

29 these changes, and inform conservation actions. To harness this information for conservation, we need to

30 integrate the largely independent fields of biodiversity modelling and conservation. We highlight new

31 developments in each respective field, early examples of how they are being brought together, and ideas

32 for a future synthesis so conservation decisions can be made with a fuller awareness of the biodiversity at

33 stake.

34

35 **Main Text**

36

37 *We need greater diversity in biodiversity conservation*

38 Aside from a small fraction of well-known species with established conservation status, most of

39 the world's biodiversity is poorly integrated into conservation efforts. Even for iconic clades

40 such as birds and mammals, only a subset of species are considered in conservation despite a

41 continuously refined understanding of the ecology and evolution of these clades[1,2]. Lesser

42 known taxa are nearly always overlooked[3] even those most at risk of extinction[4]. First, we

43 need to expand the focus of biodiversity in conservation programs to consider more of the tree of

44 life, even taxa without names[5,6], especially as increasing evidence suggests 'invisible'

45 diversity is important to ecosystem functioning [7]. Second, we must consider the role of species

46 in ecosystems[8]. Biodiversity is not simply the mere sum of species, but includes complex

47 interaction networks with fluxes of energy and nutrients[9,10]. Even conservation efforts for

48 focal charismatic species would benefit from considering the other species enabling their

49 existence. Biodiversity losses impact all facets of biodiversity (species, genetic diversity,

50 **functional diversity**, **phylogenetic diversity**, see Glossary) and spatial extents ranging from

51 local assemblages to the global pool of biodiversity. Considering these dimensions is essential to

52 prevent widespread biodiversity loss.

53

54 ***Conservation evaluation often falls short for under-represented biodiversity***

55 *Simplified metrics likely miss many types of biodiversity*

56 Despite scientific recognition of the importance of biodiversity across multiple facets, most

57 planning and on-the-ground conservation still focus on a simplified representation of biodiversity

58 such as the amount or quality of habitat. Habitat-based metrics are used more often in

59 conservation than in ecology[11], and are often the basis of conservation offsetting, wetland

60 mitigation, conservation easements, conservation and management of listed species (alongside

61 population abundances, connectivity, and threat information when available), and are used at a

62 regional extent in coordinated reserve designs (e.g. Natura 2000,

63 https://ec.europa.eu/environment/nature/natura2000/index_en.htm). Reliance on habitat metrics

64 can overlook important biodiversity that falls outside of commonly held notions of high-quality

65 habitat[12], and these metrics do not necessarily represent population persistence[11]. While the

66 use of habitat metrics certainly has follow-on benefits to some diversity, its surrogacy value for

67 even well-known species is unclear[13], and unknown for many taxa and for different

68 **biodiversity facets**.

69

70   In cases where multiple species are considered together in applied conservation, they are most

71   often represented as species richness (the count of species in areas)[11]. The use of species

72   richness cannot address the known biases in extinction risk towards certain clades (e.g.

73   amphibians[14]) and functional types (e.g. top predators). While species richness and other

74   species-based metrics can be an effective surrogate for phylogenetic diversity[10] or functional

75   diversity[15], its surrogacy value depends on the conservation scenario[16] and tends to decrease

76   in scenarios where resources are limited (i.e. few sites available for protection)[17]. We have

77   very limited evidence for surrogacy value of different biodiversity facets outside of iconic animal

78   groups (e.g. does aboveground animal richness relate to soil microbial diversity?). Finally, how

79   species interact within ecosystems is almost never considered in applied conservation despite

80   links to the important conservation goals of ecosystem functioning and resilience[9].

81

82   *Independent conservation efforts fail to protect the world's biodiversity*

83   Safeguarding multiple facets of biodiversity—especially for the purposes of preventing

84   extinction, protecting our biological heritage, and retaining future options for society—requires

85   an understanding of the global biodiversity pool, arguably more than for other conservation aims

86   (e.g. managing local ecosystems or protecting ecosystem services)[18]. But conservation actions

87   are typically local or regional, which can lead to unintended outcomes for global

88   biodiversity[19]. First, species are declining globally[20], but not necessarily locally[21–23],

89   leading to a homogenization of biodiversity where different areas have increasingly similar

90   biodiversity[20]. Second, conservation priorities established for local species assemblages[16] or

91   even for species at a national scale[19] differ from global biodiversity priorities[16,19]. Lastly,

92   the emphasis on protecting large patches and connected habitats (e.g. species corridor

93   initiatives), which is based on research on metapopulations and conservation genetics, risks

94   overlooking important biodiversity (e.g. endemic species, highly functional communities) that

95   occur in small and isolated habitat patches [24]. A more spatially-coordinated approach is needed

96   that recognizes how local conservation efforts impact the regional and global biodiversity pool.

97

98   *The potential to bring together biodiversity modelling and conservation*

99   The fields of ecology and conservation planning already consider the spatial extents necessary

100   for this coordination, but they do so in slightly different ways. In systematic conservation

101   planning (SCP), biodiversity features (e.g. species distributions) are considered in planning units.

102   Planning units are compared across the entire area of interest using the principles of spatial

103   **complementarity** and **irreplaceability**. In ecology, a similar hierarchy can be found in the

104   delineation of $\boldsymbol{\alpha}$**-**(within-site or within-ecosystem diversity), $\boldsymbol{\beta}$**-**(differentiation of sites in a

105   region; e.g. compositional turnover), and $\boldsymbol{\gamma}$**-diversity** (diversity of the entire region). We refer to

106    $\boldsymbol{\gamma}$**-diversity** as the total set of biodiversity in the area of interest, which could have a regional or

107   global extent, to better align with conservation planning. Both fields are progressing

108   independently toward a multi-scale, multi-facetted view of biodiversity (Figure 1). Now is the

109   time to integrate these fields to rapidly expand the amount of biodiversity data that reaches

110   conservation practice (Figure 2).

111

112   **Advances in biodiversity models**

113   While steady progress has been made on addressing the major biodiversity shortfalls with

114   increased data collection (e.g. citizen science[25], environmental DNA) and global initiatives to

115   compile data on species occurrences (e.g. GEO-BON initiative; geobon.org, GBIF; gbif.org),

116   traits (e.g. TRY;www.try-db.org), genetic and phylogenetic data (e.g. GenBank), and species

117 interaction data (e.g. GloBI; globalbioticinteractions.org), major shortfalls remain[26].

118 Biodiversity models are essential tools to rapidly overcome these shortfalls. We consider

119 **biodiversity models** to be any model that makes predictions for biodiversity (including species

120 occurrences or abundances, traits, phylogenetic placement, or species interactions). Here, we

121 highlight two model classes: (1) **imputation models** for missing biodiversity data and (2) **spatial**

122 **biodiversity models** for predicting taxa or other biodiversity facets for unsampled locations,

123 which often serve as an input for conservation planning (Figure 1).

124

125 *Imputation models*

126 Imputation models are needed to fill in the many gaps in biodiversity datasets. For example, they

127 can generate missing trait values for taxa[27], which can be estimated from phylogenies,

128 assuming traits have a general pattern of conserved evolution[28] or through multivariate

129 imputation that does not rely on phylogenies[29]. Increasingly sophisticated approaches are also

130 being used to generate phylogenies when some gene sequences are missing[30], and to combine

131 multiple datasets into aggregated phylogenies either based on the phylogenies themselves ('super

132 trees' or preferably from a combined gene matrix[31]. There is also growing interest in

133 predicting interactions between species based on their trait attributes (e.g. predator-prey trait

134 matching[32]) and spatial co-occurrence data[33]. Predicting species interactions across many

135 species or at multiple locations is difficult with severely limited data, but can be addressed by

136 combining empirical data and expert knowledge[34].

137

138 While imputation models are greatly expanding the coverage of biodiversity datasets, challenges

139 remain. Data collection is systematically geographically and taxonomically biased, which could

140 bias predictions, especially for rare taxa. We are increasingly relying on the same datasets for

141 different types of analyses, so there is also the danger of circularity if those analyses are later

142 combined (e.g. using the phylogeny to predict trait values and later combining phylogeny and

143 traits in spatial models). We need more robust tests for the impact of these biases, and to ensure

144 that uncertainty is propagated from the original data sources to all subsequent analyses. A partial

145 solution is to integrate methods and create a transparent and documented pipeline of analyses

146 (Figure 2), but we still must address the underlying causes of bias and use this knowledge to

147 guide future collection.

148

149
150 *Spatial Biodiversity models*
151
152 The Wallacean shortfall (inadequate knowledge of geographic distributions) greatly hinders

153 progress in identifying conservation priorities[26]. This shortfall is particularly problematic at

154 local scales—it is easier to define a regional or national set of species than to pinpoint the precise

155 distributions of species at fine spatial scales[26]. Biodiversity models are essential tools to

156 address this shortfall. In many cases, they can provide continuous coverage of species

157 distributions at a resolution relevant for conservation purposes[35]. The most common examples

158 are single species distribution models (i-SDMs; Fig. I), which have been and will continue to be

159 invaluable tools for conservation applications [35]. Advanced models can additionally

160 accommodate multiple facets (e.g. traits and phylogenies) and integrate elements of α- and β-

161 diversity (Box 1). A notable advance is the ability to combine a top-down approach focused on

162 an entire species assemblage with a bottom-up approach focused on single species (**'sideways'**

163 **biodiversity models**; Box 1, Box 2). One such model, Joint Species Distribution Models (j-

164 SDMs, reviewed in [36]), offers potential improvements by modeling all species simultaneously

165 [37,38], enabling conditional predictions [39–41], and has a flexible hierarchy amenable to

166 adding other information such as traits[37], phylogenies[42], geographic space[43], temporal

167 data[44], and detection bias[45], which impact multifaceted biodiversity estimates[46]. Models

168 that can account for species interactions[47] are an exciting arena for future developments.

169

170 A key advantage of these advanced biodiversity models is the ability to process highly

171 multidimensional datasets (e.g. many taxa, sites, and environmental variables). Combining

172 multiple, large biodiversity datasets can quickly result in a computationally intractable problem.

173 Solutions include dimensionality reduction in a Bayesian framework[36,48] or via machine

174 learning[49]. However, what these approaches gain in flexibility may come at a cost in terms of

175 generality, and the relationships derived between explanatory variables and diversity metrics are

176 often empirical with little to no grounding in ecological theory. An important future research

177 avenue will be to rigorously test new biodiversity models to evaluate their capacity for aiding

178 ecological inference and for predicting to new situations (e.g. novel climates or altered

179 ecosystems).

180

181 **Advances in conservation evaluation of multifaceted biodiversity**

182 While the methods described above could offer a richer representation of biodiversity and

183 improved predictive abilities, identifying conservation solutions requires more than just

184 biodiversity predictions. These predictions must be translated to meaningful outputs that can be

185 used to meet defined objectives (e.g. protected area targets) or provide forecasts for decision

186 makers. In systematic conservation planning (SCP)[50], priorities are established that efficiently

187 meet goals for representation of biodiversity under a range of spatial or other constraints (e.g. in

188 a **spatial prioritization**). Targets can be set on diversity (e.g. 10% of each species range

189 protected[17]) or approaches without targets such as algorithms based on **weighted**

190 **endemism**[16,19]. While i-SDMs are commonly used as inputs to spatial prioritization, more

191 advanced biodiversity models are rarely used (see Box 2 for exceptions). Multiple biodiversity

192 facets are also rarely considered in conservation despite decades of recognized conservation

193 benefit of *e.g.* phylogenetic diversity[51]. However, in recent years methodological

194 advances[52–54] have enabled the use of phylogenies and functional diversity [16,53] in

195 conservation planning studies at regional[6,52] and global[16,17] extents, and initiatives such as

196 EDGE (evolutionarily distinct and globally endangered; edgeofexistence.org) are promoting the

197 conservation of multiple biodiversity facets. Species interactions are even more rarely

198 considered, although emerging examples show how interactions can be used to evaluate

199 ecological network collapse[55], and set spatial priorities[56] and management scenarios that

200 account for predators and their prey[57]

201

202 **The way forward: integrating biodiversity models and conservation evaluation**

203 Modeling has a critical, yet under-appreciated role in conservation planning and practice,

204 especially with respect to setting and achieving global conservation targets (Aichi targets for

205 threatened species, habitat loss, and protected areas) in the Strategic Plan for Biodiversity 2011-

206 2020 and the post-2020 global biodiversity framework[58]. Models that predict the distribution

207 of species in space and time are crucial to these global initiatives, and i-SDMs are already widely

208 used in conservation[35]. We believe the more advanced biodiversity models have vast potential

209 for representing not just species but entire communities, clades, or functional roles (see early

210 examples in Box 2), but they first must be better-integrated into conservation efforts.

211

212   The advantages of this integration are many. First, a combined approach retains the unique

213   strengths of each individual approach. Conservation planning would benefit from having access

214   to the most advanced statistical machinery including model-calibration, model validation, and

215   the propagation of uncertainty. Such models can also incorporate more flexible hierarchical

216   structures that account for biases in sampling biodiversity, different sources of inputs (e.g.

217   museum location data with community-based eDNA sampling), different extents (α- and β-

218   diversity), and different resolutions of input data. They can take advantage of new data types

219   (e.g. tracking data[59]) as they become available. Conversely, the field of biodiversity modelling

220   would benefit from tailoring predictions for their use in tangible conservation scenarios and

221   targets.

222

223   Second, an integrated approach reduces any redundancies that arise from using similar datasets

224   and similar approaches in an unsynchronized manner. The potential for redundancy becomes

225   greater the larger and the more complex the representation of biodiversity becomes. Redundancy

226   could be reduced on two fronts—the raw data inputs and the methodological similarities (e.g.

227   combining modelling and conservation optimization algorithms as discussed below).

228

229   Finally, this synthesis would make the pipeline of data-implementation more efficient and likely

230   to result in a better representation of under-represented biodiversity in conservation. Recent calls

231   to streamline analytical pipelines in ecology[60–62] point to increased efficiency [63] of readily

232   available technological solutions, such as user-friendly interfaces (e.g. R Shiny, Dash).

233   Importantly, the integration of conservation goals, targets, and scenarios into modelling would

234   require stakeholder input into the modelling decisions[64], which could lead to increased

235 coordination, trust and uptake of these models, decreasing the 'black box' problem of

236 biodiversity models[60]. It would engage biodiversity scientists beyond academia to understand

237 the needs of stakeholders and provide more tailored conservation guidance.

238

239 **Steps forward**

240 We recommend three related areas of focus to speed integration of biodiversity models and

241 conservation: (1) adapting biodiversity models to accommodate conservation goals and

242 scenarios, (2) combining biodiversity models and conservation optimization algorithms, and (3)

243 re-evaluating model validation in light of conservation goals and scenarios. Recent examples

244 show this integration is indeed possible (Box 2).

245

246 In many cases, statistical methods already exist to adapt biodiversity models for use in

247 conservation. For example, management actions (or management history) could be included as

248 model predictors, which could then be used to make predictions for different future actions.

249 More complex methods of scenario-building are also possible through altering various model

250 parameters [65,66](Box 2) or even adding sub-models of ecological processes (e.g. population

251 viability). Previously collected 'prior' information (information on species threat or response to

252 management action) can be combined with newly collected data to refine model estimates[67].

253 The challenge will be to find the best data and models for the conservation question at hand[68],

254 requiring close collaboration between modelers and conservation practitioners[65].

255

256 Combining conservation optimization algorithms with biodiversity models is another important

257 step (Figure 2) that will enable relevant ecological processes to be directly considered in the

258 conservation alternatives[65]. This integration will require some method developments, but will

259 benefit from the similarities in data types (Figure 1), especially when definitions are

260 synchronized (e.g. α-diversity is equivalent to diversity in planning units, β-diversity to

261 complementarity[69], γ-diversity to the total set of biodiversity). In particular, β-diversity could

262 be used in ecological models with scenarios (Box 2) or within optimization algorithms,

263 streamlining methods and opening the possibility of evaluating conservation value of

264 biodiversity not represented in a single-species approach (e.g. ecological interaction

265 networks[70]).

266

267 Finally, model evaluation—while always a critical step in biodiversity modelling—is especially

268 important with conservation scenarios that are often extrapolations to new situations (e.g.

269 different environmental conditions, alternative restoration scenarios, poorly known taxa). Model

270 evaluation will help address questions such as which type of model to use. For some

271 conservation applications, a highly flexible, accurate model is likely appropriate (e.g. predicting

272 the current distribution of a threatened species). For others, it might be preferable to have models

273 more closely aligned with ecological theory (e.g. when predicting to novel environments such as

274 climate change) or when the conservation question depends on a particular variable (e.g.

275 properties of a connectivity corridor). While there are examples that evaluate the ability of

276 advanced biodiversity models to predict rare species[71] and their interactions[72], very few

277 studies have tested how well models extrapolate to new conditions with independent datasets, or

278 whether they predict realistic conservation outcomes. Much more research is needed to

279 understand how biodiversity models will perform in an applied setting.

280
281

282 *Concluding Remarks*

283 We urgently need to understand how biodiversity change impacts critical ecosystem functioning

284 and prunes the tree of life. We propose an integration of advanced biodiversity models with

285 conservation goals, targets, assessments, and practice. This integration would improve the ability

286 to rapidly evaluate biodiversity data, make predictions and recommend conservation action for

287 taxa, communities, and ecosystems. The backbone of this integration already exists with recent

288 developments in biodiversity modelling and conservation. Early examples show this integration

289 is possible, could be extended to a number of other conservation efforts and used to address

290 unresolved questions of how to best protect the world's biodiversity (see Outstanding

291 Questions).

292

298

299

300 **Glossary Box**

301

302 **Alpha (α) diversity:** quantifies local diversity, usually species richness or effective number of

303 species (Shannon exponent)

304

305   **Beta (*β*) diversity:** quantifies the dissimilarity or turnover between locations in terms of

306   composition. Can be pairwise between locations, global (average dissimilarity across the region)

307   or the distinctiveness of a location compared to all others in the region.

308

309   **Biodiversity facets:** categories of biodiversity that describe relevant taxonomic or ecological

310   information to support biodiversity evaluation, and that apply to and can be (relatively) easily

311   measured for all or most taxa. Examples include taxonomic diversity, phylogenetic diversity,

312   genetic diversity, trait/functional diversity and network diversity.

313

314   **Biodiversity models:** statistical or process-based models that are used to make inferences and

315   predictions about the effect of the environment on biodiversity, accounting for ecological

316   processes explicitly or implicitly. Biodiversity can be represented in units ranging from

317   individuals to entire communities and additionally represent attributes of those taxa (e.g.

318   abundances, functional traits, phylogenetic position, threat).

319

320   **Complementarity (conservation planning):** quantifies the difference between locations in

321   terms of species or features represented within them (two locations are fully complementary if

322   their pairwise β-diversity is equal to 1).

323

324   **Complementarity (ecology):** two species are complementary if they fill different roles in an

325   ecosystem or use resources differently.

326

327   **Functional diversity (FD):** the diversity of functional forms in a species set (or community)

328   measured by a variety of metrics that use dendrograms or representations in multidimensional

329   space.

330

331   **Gamma (γ) diversity:** quantifies the total diversity across locations for an entire area of interest,

332   which could be at regional or global scale depending on the study.

333

334   **Imputation model:** a model that estimates values for missing data points in biodiversity datasets

335   (e.g. trait values for a particular taxa).

336

337   **Irreplaceability:** the extent to which a location or species is distinct from all others (opposite of

338   redundancy). For example, a location where an endemic species occurs is irreplaceable; an

339   irreplaceable species has a unique position in its community (in terms of function or interactions,

340   i.e. distinctive) or in the phylogeny.

341

342   **Macro-ecological model**: a biodiversity model that uses a top-down approach to model α- or ß-

343   diversity directly rather than modeling the distributions of the component taxa.

344

345   **Phylogenetic diversity (PD):** The evolutionary diversity represented by sets of taxa, with the

346   most common metric (Faith's PD) being the branch length of the minimum spanning tree

347   connecting a set of species on a phylogeny[51] (sets of species can be from a single area or

348   multiple areas combined).

349

350 **'Sideways' biodiversity models:** models that predict the distribution of biodiversity with a

351 combination of bottom-up (i.e. single-taxon predictions) and top-down (i.e. models of the

352 properties of an assemblage or community of taxa) approaches.

353

354 **Spatial biodiversity models:** biodiversity models that are explicitly spatial, in which inferences

355 and predictions are made for biodiversity in certain locations (e.g. planning units or grid cells).

356

357 **Spatial prioritization:** A form of systematic conservation planning (SCP) that selects a set of

358 areas that maximizes conservation value given other constraints (e.g. cost, protected areas,

359 feasibility).

360

361 **Weighted endemism:** the ratio between a species local range of occurrence and its total range.

362 Note: this is different than the definition of endemism which is the extent to which a species

363 range is restricted to a particular location.

364

365

366 **Figure 1.** How biodiversity facets, different scales of biodiversity and conservation concerns

367 factor into biodiversity modelling and conservation evaluation. Most modelling and evaluation

368 involve particular target species rather than other facets of diversity (e.g. evolutionary history,

369 species interactions). Most biodiversity models focus on ecological processes rather than

370 explicitly considering conservation concerns, although many models are built for threatened

371 species or species of conservation interest resulting in a mostly unidirectional flow of

372 information from models to evaluation. In the case of spatial conservation planning, both models

373 and evaluation use similar data inputs and spatial organization of biodiversity. Both have local

374 scale (or planning units) and a regional or global scale. In particular, the use of metrics or

375 algorithms to differentiate sites from one another (beta diversity in modelling, complementarity

376 in conservation planning) are treated similarly in modelling and planning.

377

378 **Figure 2.** The steps and information flow in a typical data-implementation pipeline (left) and a

379 streamlined version (right) that combines models and evaluation in one step. We expect that

380 more biodiversity information would make it through to the application stage (orange arrows)

381 with a unified rather than disconnected method of evaluating biodiversity from the data

382 products to their use in conservation applications, and that the increasing transparency and

383 additional communication between biodiversity scientists and stakeholders is required by the

384 combined approach.

385

386 **Figure I.** (Embedded in Box 1). The dimensions of recent biodiversity models. Models can have

387 different types of response objectives: within-site models, between-site models and models of

388 individual species. These response objectives fall along two axes: (1) top-down approaches

389 where α-diversity or other diversity metrics are the focus versus bottom-up where individual

390 species are the focus (vertical axis), and (2) whether the focus is on predicting biodiversity

391 within sites or between sites (horizontal axis). DynamicFOAM: Dynamic Framework for

392 Occurrence Allocation in Metacommunities, GDM: Generalized Dissimilarity Modelling, MEM:

393 macro-ecological model, i-SDM: individual or single-species distribution model, s-SDM:

394 stacked species distribution model, j-SDM: joint species distribution model, SESAM: Spatially

395 explicit species assemblage modelling.

396

397 **Box 1. Modelling biodiversity from the top-down, bottom-up and sideways**

398   Biodiversity models can be defined by their 'response objectives' and internal structure

399   (particularly the response variable—what is being predicted). These response objectives can be

400   represented with axes that describe whether the objective is a top-down versus bottom-up or a

401   within- versus between-site depiction of biodiversity (Fig. I). Current modelling approaches can

402   be mapped onto this space, with potentially the most information rich approaches (grey circle in

403   Fig. I) cross-linking biodiversity facets (traits and phylogenies) while jointly predicting

404   community composition (e.g. species abundance or occurrence), within-site community

405   diversity, and between-site turnover.

406

407   In the top-down approach, known as a **macro-ecological model** (MEM), α-diversity is itself is

408   the response (e.g. species richness or phylogenetic diversity) and the input data is a site by

409   diversity matrix. β-diversity models (e.g. GDM; Generalized Dissimilarity Modelling[73] or

410   MBM; multi-facetted biodiversity models[49]) estimate turnover in species or phylogenetic

411   diversity[74] between sites as a function of differences in environmental and geographic

412   distances.

413

414   Bottom-up approaches, where individual species are the response variables (single-species

415   distribution model, i-SDM), are by far the most common type of biodiversity model used. Note

416   that α-diversity can also be derived from aggregating models of individual species together in a

417   stacked species distribution model (s-SDM). This feature is important for conservation planning,

418   because both the species pool of a set of sites and diversity metrics for entire species

419   assemblages can be calculated (e.g. endemism-based metrics). However, resulting diversity

420 metrics are simply the sum of the parts—there is no information on species associations or

421 community-level attributes.

422 'Sideways' methods are methods that combine top-down and bottom-up approaches: Joint

423 Species Distribution Models (j-SDMs), which are the multi-species version of i-SDMs that

424 model effects of individual species and communities[39,40], and can include shared responses

425 between species and can include other information such as phylogeny and traits in a hierarchical

426 framework, and spatial and temporal dynamics[66]; Dynamic Framework for Occurrence

427 Allocation in Metacommunities (DynamicFOAM)[75], which balances richness and turnover to

428 predict occurrences for all species, SESAM: Spatially explicit species assemblage modelling[76],

429 which adjusts combined s-SDM predictions based on those of macro-ecological models.

430

431 **Box 2. Hybrid approaches for integrating advanced biodiversity models and conservation**

432 **decisions**

433 We have emphasized the need for models that join different spatial and taxonomic dimensions of

434 biodiversity that can be combined with conservation evaluation. Here are examples from two

435 classes of advanced biodiversity models (Box 1) that directly address applied conservation

436 questions. There is much scope for building upon and extending these early examples.

437
438   *1. Macro-ecological models*

439 For taxa for which there is no abundance or time-series data, models based on α- and β-diversity

440 (Box1) can also be used to directly design, predict, and compare conservation scenarios. For

441 example, ß-diversity models can summarise the expected effects on future biodiversity of

442 changes to habitat condition, extent or management, as change in effective habitat area[77]. With

443 sophisticated models of environmental or land-use change for example, new protected areas,

444   habitat loss[78] or degradation and climate change[79], scenarios can be planned for the most

445   beneficial effect on biodiversity even at global extents[77], and can also be used with

446   phylogenetic[74] or functional[49] ß-diversity.

447   *2. 'Sideways' biodiversity models*

448   Multi-species models that contain a hierarchy for a species-level and community (or ecosystem)-

449   level (Box 1) are widely relevant to conservation applications because they can simultaneously

450   estimate ecological processes relevant for individual species and for larger-scale attributes (e.g.

451   ecosystem properties). Perhaps the best example of this approach has been for evaluating

452   management strategies in fisheries, in which models of fish species (or sets of species) are

453   combined with ecosystem level attributes[64]. In this way, important ecological and ecosystem

454   constraints can be considered (e.g. population dynamics in response to habitat or predators)

455   along with management scenarios (harvest controls)[57]. Recent versions allow for evaluation of

456   harvest impacts, survey designs, and protection scenarios for finescale habitats in a changing

457   climate[66]. While these more elaborate designs are only possible for a subset of well-studied

458   taxa, this will be increasingly possible with ever-improving sampling and imputation models.

459

460   **References**

461   1   Jetz, W. *et al.* (2012) The global diversity of birds in space and time. *Nature* 491, 444–448

462   2   Mazel, F. *et al.* (2017) Global patterns of β-diversity along the phylogenetic time-scale: The

463   role of climate and plate tectonics. *Glob. Ecol. Biogeogr.* 26, 1211–1221

464   3   Eisenhauer, N. *et al.* (2019) Recognizing the quiet extinction of invertebrates. *Nat. Commun.*

465   10, 1–3

466   4   Davies, T. *et al.* (2018) Popular interest in vertebrates does not reflect extinction risk and is

467   associated with bias in conservation investment. *PLOS ONE* 13, e0203694

468   5   Asmyhr, M.G. *et al.* (2014) Systematic Conservation Planning for Groundwater Ecosystems

469   Using Phylogenetic Diversity. *PLoS ONE* 9, e115132

470   6   Rosauer, D.F. *et al.* (2018) Real-world conservation planning for evolutionary diversity in

471   the Kimberley, Australia, sidesteps uncertain taxonomy. *Conserv. Lett.* 11, e12438

472   7   Delgado-Baquerizo, M. *et al.* (2016) Microbial diversity drives multifunctionality in

473   terrestrial ecosystems. *Nat. Commun.* 7, 10541

474   8   Violle, C. *et al.* (2017) Functional Rarity: The Ecology of Outliers. *Trends Ecol. Evol.* 32,

475   356–367

476   9   Harvey, E. *et al.* (2017) Bridging ecology and conservation: from ecological networks to

477   ecosystem function. *J. Appl. Ecol.* 54, 371–379

478   10 Thompson, R.M. *et al.* (2012) Food webs: reconciling the structure and function of

479   biodiversity. *Trends Ecol. Evol.* 27, 689–697

480   11 Marshall, E. *et al.* (2020) What are we measuring? A review of metrics used to describe

481   biodiversity in offsets exchanges. *Biol. Conserv.* 241, 108250

482   12 Hobbs, R.J. (2016) Degraded or just different? Perceptions and value judgements in

483   restoration decisions. *Restor. Ecol.* 24, 153–158

484   13 Beier, P. *et al.* (2015) A review of selection-based tests of abiotic surrogates for species

485   representation. *Conserv. Biol.* 29, 668–679

486   14 González-del-Pliego, P. *et al.* (2019) Phylogenetic and Trait-Based Prediction of Extinction

487   Risk for Data-Deficient Amphibians. *Curr. Biol.* 29, 1557-1563.e3

488   15 Rapacciuolo, G. *et al.* (2019) Species diversity as a surrogate for conservation of

489   phylogenetic and functional diversity in terrestrial vertebrates across the Americas. *Nat.*

490   *Ecol. Evol.* 3, 53–61

491   16 Pollock, L.J. *et al.* (2017) Large conservation gains possible for global biodiversity facets.

492   *Nature* 546, 141–144

493   17 Rosauer, D.F. *et al.* (2017) Phylogenetically informed spatial planning is required to

494   conserve the mammalian tree of life. *Proc R Soc B* 284, 20170627

495   18 Pearson, R.G. (2016) Reasons to Conserve Nature. *Trends Ecol. Evol.* 31, 366–371

496   19 Pouzols, F.M. *et al.* (2014) Global protected area expansion is compromised by projected

497   land-use and parochialism. *Nature* 516, 383–386

498   20 Di Marco, M. *et al.* (2018) Changes in human footprint drive changes in species extinction

499   risk. *Nat. Commun.* 9, 1–9

500   21 Sax, D.F. and Gaines, S.D. (2003) Species diversity: from global decreases to local

501   increases. *Trends Ecol. Evol.* 18, 561–566

502   22 Vellend, M. *et al.* (2013) Global meta-analysis reveals no net change in local-scale plant

503   biodiversity over time. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19456–19459

504   23 Cardinale, B.J. *et al.* (2018) Is local biodiversity declining or not? A summary of the debate

505   over analysis of species richness time trends. *Biol. Conserv.* 219, 175–183

506   24 Wintle, B. *et al.* (2018) Global synthesis of conservation studies reveals the importance of

507   small habitat patches for biodiversity. *Proc. Natl. Acad. Sci.* 116, 201813051

508   25 Kobori, H. *et al.* (2016) Citizen science: a new approach to advance ecology, education, and
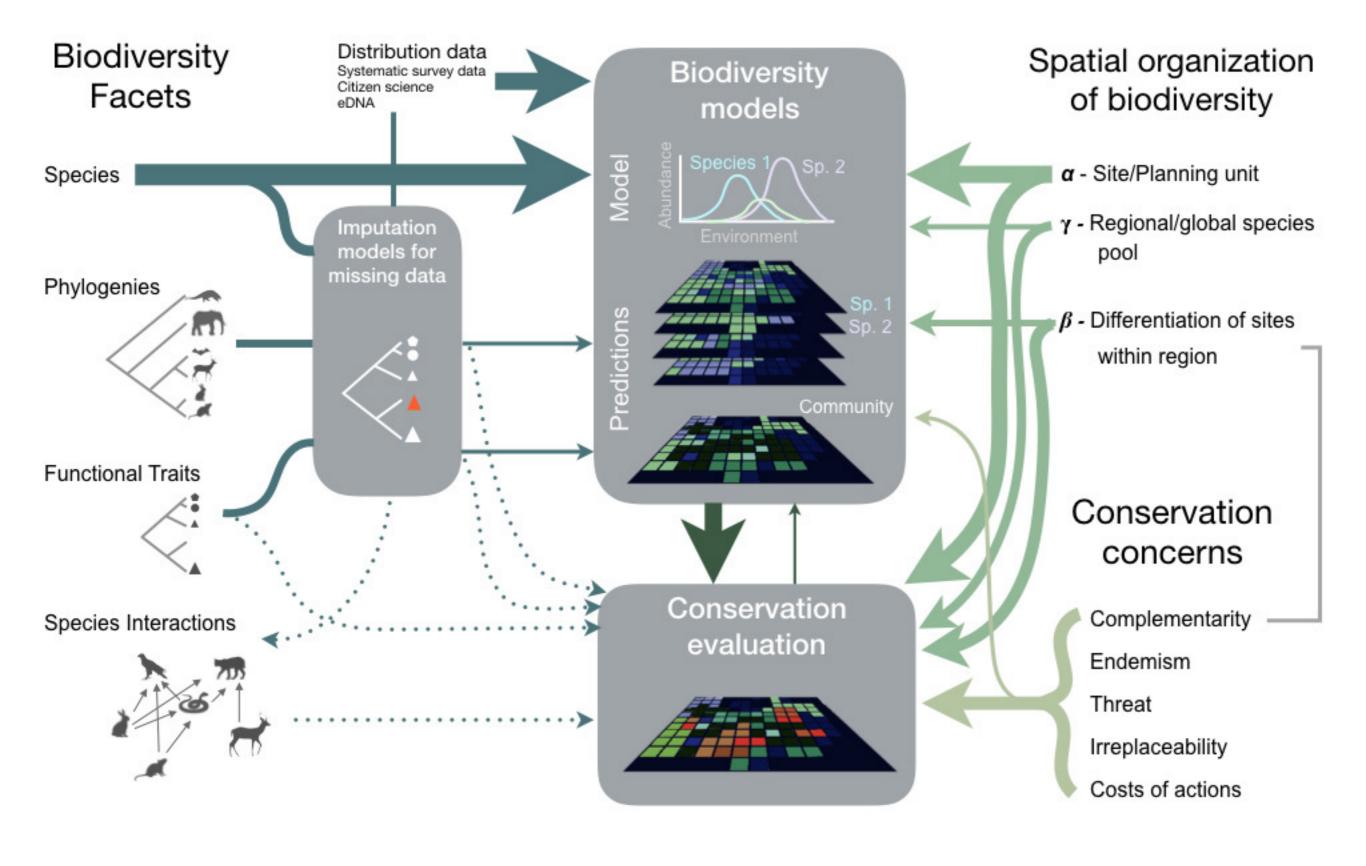
509   conservation. *Ecol. Res.* 31, 1–19

510   26 Hortal, J. *et al.* (2015) Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity.

511   *Annu. Rev. Ecol. Evol. Syst.* 46, 523–549

512   27 Wilman, H. *et al.* (2014) EltonTraits 1.0: Species-level foraging attributes of the world's

513   birds and mammals. *Ecology* 95, 2027–2027

514   28 Swenson, N.G. (2014) Phylogenetic imputation of plant functional trait databases.

515   *Ecography* 37, 105–110

516   29 Penone, C. *et al.* (2014) Imputation of missing data in life-history trait datasets: which

517   approach performs the best? *Methods Ecol. Evol.* 5, 961–970

518   30 Bhattacharjee, A. and Bayzid, Md.S. (2020) Machine learning based imputation techniques

519   for estimating phylogenetic trees from incomplete distance matrices. *BMC Genomics* 21, 497

520   31 von Haeseler, A. (2012) Do we still need supertrees? *BMC Biol.* 10, 13

521   32 Gravel, D. *et al.* (2013) Inferring food web structure from predator–prey body size

522   relationships. *Methods Ecol. Evol.* 4, 1083–1090

523   33 Morales-Castilla, I. *et al.* (2015) Inferring biotic interactions from proxies. *Trends Ecol.*

524   *Evol.* 30, 347–356

525   34 Braga, J. *et al.* (2019) Spatial analyses of multi-trophic terrestrial vertebrate assemblages in

526   Europe. *Glob. Ecol. Biogeogr.* 28, 1636–1648

527   35 Guisan, A. *et al.* (2013) Predicting species distributions for conservation decisions. *Ecol.*

528   *Lett.* 16, 1424–1435

529   36 Warton, D.I. *et al.* (2015) So Many Variables: Joint Modeling in Community Ecology.

530   *Trends Ecol. Evol.* 30, 766–779

531   37 Pollock, L.J. *et al.* (2012) The role of functional traits in species distributions revealed

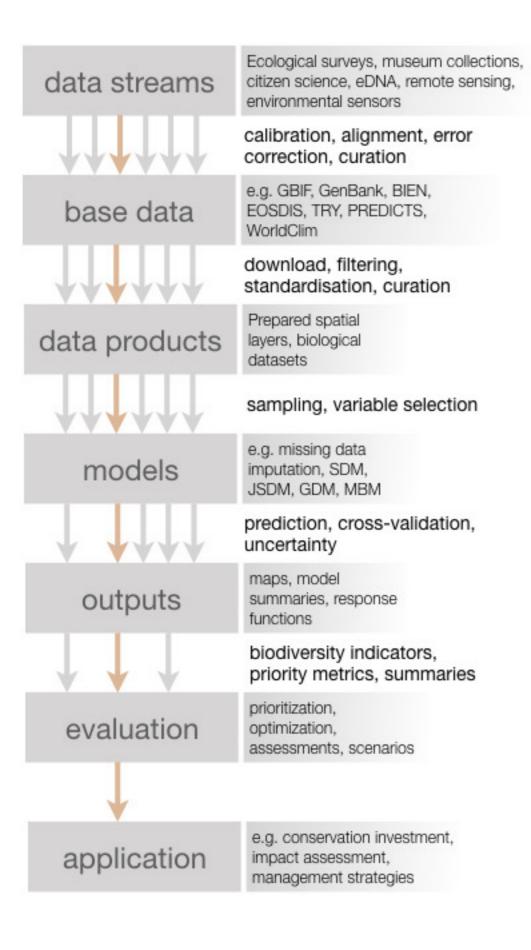532   through a hierarchical model. *Ecography* 35, 716–725

533   38 Ovaskainen, O. and Soininen, J. (2011) Making more out of sparse data: hierarchical

534   modeling of species communities. *Ecology* 92, 289–295

535   39 Clark, J.S. *et al.* (2014) More than the sum of the parts: forest climate response from joint

536   species distribution models. *Ecol. Appl.* 24, 990–999

537   40 Pollock, L.J. *et al.* (2014) Understanding co-occurrence by modelling species simultaneously

538   with a Joint Species Distribution Model (JSDM). *Methods Ecol. Evol.* 5, 397–406

539   41 Ovaskainen, O. *et al.* (2010) Modeling species co-occurrence by multivariate logistic

540   regression generates new hypotheses on fungal interactions. *Ecology* 91, 2514–2521

541   42 Ives, A.R. and Helmus, M.R. (2011) Generalized linear mixed models for phylogenetic

542   analyses of community structure. *Ecol. Monogr.* 81, 511–525

543   43 Ovaskainen, O. *et al.* (2016) Uncovering hidden spatial structure in species communities

544   with spatially explicit joint species distribution models. *Methods Ecol. Evol.* 7, 428–436

545   44 Schliep, E.M. *et al.* (2018) Joint species distribution modelling for spatio-temporal

546   occurrence and ordinal abundance data. *Glob. Ecol. Biogeogr.* 27, 142–155

547   45 Tobler, M.W. *et al.* (2019) Joint species distribution models with species correlations and

548   imperfect detection. *Ecology* 100, e02754

549   46 Jarzyna, M.A. and Jetz, W. (2016) Detecting the Multiple Facets of Biodiversity. *Trends*

550   *Ecol. Evol.* 31, 527–538

551   47 Staniczenko, P.P.A. *et al.* (2017) Linking macroecology and community ecology: refining

552   predictions of species distributions using biotic interaction networks. *Ecol. Lett.* 20, 693–707

553   48 Taylor-Rodríguez, D. *et al.* (2017) Joint Species Distribution Modeling: Dimension

554   Reduction Using Dirichlet Processes. *Bayesian Anal.* 12, 939–967
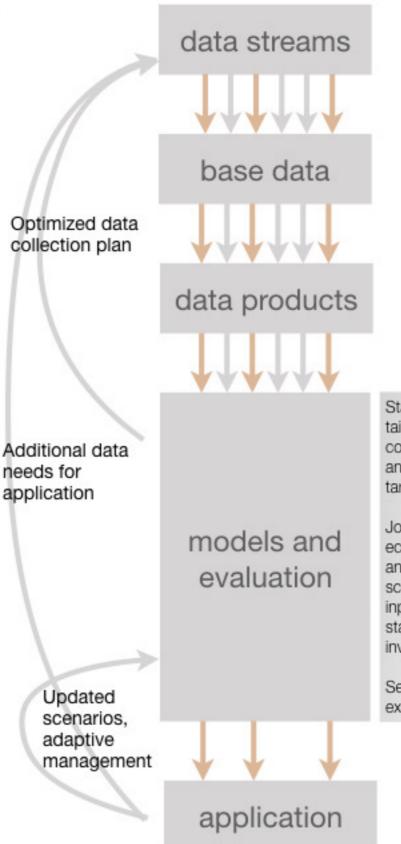
555   49 Talluto, M.V. *et al.* (2018) Multifaceted biodiversity modelling at macroecological scales

556   using Gaussian processes. *Divers. Distrib.* 24, 1492–1502

557   50 Margules, C.R. and Pressey, R.L. (2000) Systematic conservation planning. *Nature* 405,

558   243–252

559   51 Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61, 1–

560   10

561   52 Pollock, L.J. *et al.* (2015) Phylogenetic diversity meets conservation policy: small areas are

562   key to preserving eucalypt lineages. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370, 20140007

563   53 Strecker, A.L. *et al.* (2011) Defining conservation priorities for freshwater fishes according

564   to taxonomic, functional, and phylogenetic diversity. *Ecol. Appl.* 21, 3002–3013

565   54 Strimas-Mackey, M. *prioritizr: Systematic Conservation Prioritization in R*,

566   55 McDonald-Madden, E. *et al.* (2016) Using food-web theory to conserve ecosystems. *Nat.*

567   *Commun.* 7, 1–8

568   56 Decker, E. *et al.* (2017) Incorporating ecological functions in conservation decision making.

569   *Ecol. Evol.* 7, 8273–8281

570   57 Punt, A.E. *et al.* (2016) Exploring the implications of the harvest control rule for Pacific

571   sardine, accounting for predator dynamics: A MICE model. *Ecol. Model.* 337, 79–95

572   58 Nicholson, E. *et al.* (2019) Scenarios and Models to Support Global Conservation Targets.

573   *Trends Ecol. Evol.* 34, 57–68

574   59 Hays, G.C. *et al.* (2019) Translating Marine Animal Tracking Data into Conservation Policy

575   and Management. *Trends Ecol. Evol.* 34, 459–473

576   60 Rapacciuolo, G. (2019) Strengthening the contribution of macroecological models to

577   conservation practice. *Glob. Ecol. Biogeogr.* 28, 54–60

578   61   Golding, N. *et al.* (2018) The zoon r package for reproducible and shareable species

579   distribution modelling. *Methods Ecol. Evol.* 9, 260–268

580   62   White, E.P. *et al.* (2019) Developing an automated iterative near-term forecasting system for

581   an ecological study. *Methods Ecol. Evol.* 10, 332–344

582   63   Lowndes, J.S.S. *et al.* (2017) Our path to better science in less time using open data science

583   tools. *Nat. Ecol. Evol.* 1, 0160

584   64   Plagányi, É.E. *et al.* (2014) Multispecies fisheries management and conservation: tactical

585   applications using models of intermediate complexity. *Fish Fish.* 15, 1–22

586   65   Scroggie, M.P. *et al.* (2019) Optimizing habitat management for amphibians: From simple

587   models to complex decisions. *Biol. Conserv.* 236, 60–69

588   66   Thorson, J.T. *et al.* (2019) Spatio-temporal models of intermediate complexity for ecosystem

589   assessments: A new tool for spatial fisheries management. *Fish Fish.* 20, 1083–1099

590   67   Morris, W.K. *et al.* (2015) The neglected tool in the Bayesian ecologist's shed: a case study

591   testing informative priors' effect on model accuracy. *Ecol. Evol.* 5, 102–108

592   68   Guillera-Arroita, G. *et al.* (2015) Is my species distribution model fit for purpose? Matching

593   data and models to applications. *Glob. Ecol. Biogeogr.* 24, 276–292

594   69   Bush, A. *et al.* (2016) Current Uses of Beta-Diversity in Biodiversity Conservation: A

595   response to Socolar et al. *Trends Ecol. Evol.* 31, 337–338

596   70   Ohlmann, M. *et al.* (2019) Diversity indices for ecological networks: a unifying framework

597   using Hill numbers. *Ecol. Lett.* 22, 737–747

598   71   Zhang, C. *et al.* (2020) Improving prediction of rare species' distribution from community

599   data. *Sci. Rep.* 10, 12230

600 72 Flores-Tolentino, M. *et al.* (2020) Distribution and conservation of species is misestimated if

601 biotic interactions are ignored: the case of the orchid Laelia speciosa. *Sci. Rep.* 10, 9542

602 73 Ferrier, S. *et al.* (2007) Using generalized dissimilarity modelling to analyse and predict

603 patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* 13, 252–264

604 74 Rosauer, D.F. *et al.* (2014) Phylogenetic generalised dissimilarity modelling: a new approach

605 to analysing and predicting spatial turnover in the phylogenetic composition of communities.

606 *Ecography* 37, 21–32

607 75 Mokany, K. *et al.* (2011) Combining α - and β -diversity models to fill gaps in our

608 knowledge of biodiversity: Filling gaps in biodiversity knowledge. *Ecol. Lett.* 14, 1043–1051

609 76 Guisan, A. and Rahbek, C. (2011) SESAM – a new framework integrating macroecological

610 and species distribution models for predicting spatio-temporal patterns of species

611 assemblages. *J. Biogeogr.* 38, 1433–1444

612 77 Hoskins, A.J. *et al.* (2019) Supporting global biodiversity assessment through high-

613 resolution macroecological modelling: Methodological underpinnings of the BILBI

614 framework. *Environ. Model. Softw.* 132, 104806

615 78 Allnutt, T.F. *et al.* (2008) A method for quantifying biodiversity loss and its application to a

616 50-year record of deforestation across Madagascar. *Conserv. Lett.* 1, 173–181

617 79 Prober, S.M. *et al.* (2012) Combining community-level spatial modelling and expert

618 knowledge to inform climate adaptation in temperate grassy eucalypt woodlands and related

619 grasslands. *Biodivers. Conserv.* 21, 1627–1650

620

Left diagram (top to bottom):

data streams — Ecological surveys, museum collections, citizen science, eDNA, remote sensing, environmental sensors

calibration, alignment, error correction, curation

base data — e.g. GBIF, GenBank, BIEN, EOSDIS, TRY, PREDICTS, WorldClim

download, filtering, standardisation, curation

data products — Prepared spatial layers, biological datasets

sampling, variable selection

models — e.g. missing data imputation, SDM, JSDM, GDM, MBM

prediction, cross-validation, uncertainty

outputs — maps, model summaries, response functions

biodiversity indicators, priority metrics, summaries

evaluation — prioritization, optimization, assessments, scenarios

application — e.g. conservation investment, impact assessment, management strategies

Right diagram (top to bottom):

data streams

base data

data products

models and evaluation — Statistical models are tailored to conservation objective and biodiversity targets

Joint optimization of ecological models and conservation scenarios based on input data and stakeholder involvement

See Box 2 for examples

application

Connecting arrows (between diagrams):

Optimized data collection plan

Additional data needs for application

Updated scenarios, adaptive management