Progress in
Physical Geography

Scale mismatches between predictor and response variables in species distribution modelling: a review of practices for appropriate grain selection

Abstract:	There is a lack of guidance on the choice of the spatial grain of predictor and response variables in species distribution models (SDM). This review summarises the current state of the art with regard to the following points: (i) the effects of changing the resolution of predictor and response variables on model performance; (ii) the effect of conducting multi-grain vs single-grain analysis on model performance; and (iii) the role of land cover type and spatial autocorrelation in selecting the appropriate grain size. In the reviewed literature, we found that coarsening the resolution of the response variable typically leads to declining model performance. Therefore, we recommend aiming for finer resolutions unless there is a reason to do otherwise (e.g., expert knowledge of the ecological scale). We also found that so far, the improvements in model performance reported for multi-grain models have been relatively low and that useful predictions can be generated even from single-scale models. In addition, the use of high-resolution predictors improves model performance; however, there is only limited evidence on whether this applies to models with coarser-resolution response variables (e.g. 100 km2 and coarser). Low-resolution predictors are usually sufficient for species associated with fairly common environmental conditions but not for species associated with less common ones (e.g., common vs rare land cover category). This is because coarsening the resolution reduces variability within heterogeneous predictors and leads to underrepresentation of rare environments, which can lead to a decrease in model performance. Thus, assessing the spatial autocorrelation of the predictors at multiple grains can provide insights into the impacts of coarsening their resolution on model performance. Overall, we observed a lack of studies examining the simultaneous manipulation of the resolution of predictor and response variables. We stress the need to explicitly report the resolution of all predictor and response variables.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 **Title:** Scale mismatches between predictor and response variables in species distribution modelling:

2 a review of practices for appropriate grain selection

3

4 **Abstract**

5 There is a lack of guidance on the choice of the spatial grain of predictor and response variables in species

6 distribution models (SDM). This review summarises the current state of the art with regard to the following

7 points: (i) the effects of changing the resolution of predictor and response variables on model performance; (ii)

8 the effect of conducting multi-grain vs single-grain analysis on model performance; and (iii) the role of land

9 cover type and spatial autocorrelation in selecting the appropriate grain size. In the reviewed literature, we found

10 that coarsening the resolution of the response variable typically leads to declining model performance.

11 Therefore, we recommend aiming for finer resolutions unless there is a reason to do otherwise (e.g., expert

12 knowledge of the ecological scale). We also found that so far, the improvements in model performance reported

13 for multi-grain models have been relatively low and that useful predictions can be generated even from single-

14 scale models. In addition, the use of high-resolution predictors improves model performance; however, there is

15 only limited evidence on whether this applies to models with coarser-resolution response variables (e.g. 100

16 km^2 and coarser). Low-resolution predictors are usually sufficient for species associated with fairly common

17 environmental conditions but not for species associated with less common ones (e.g., common vs rare land

18 cover category). This is because coarsening the resolution reduces variability within heterogeneous predictors

19 and leads to underrepresentation of rare environments, which can lead to a decrease in model performance.

20 Thus, assessing the spatial autocorrelation of the predictors at multiple grains can provide insights into the

21 impacts of coarsening their resolution on model performance. Overall, we observed a lack of studies examining

22 the simultaneous manipulation of the resolution of predictor and response variables. We stress the need to

23 explicitly report the resolution of all predictor and response variables.

24

25

27

1. Introduction

Species distribution models (SDMs) are widely used to assess species–environment relationships and to make predictions of species distributions in both space and time (Elith and Leathwick, 2009; Ferrier et al., 2017; Wiersma et al., 2011). To this end, SDMs relate a biodiversity-related response variable (e.g., the geographic distribution of one or more species) to explanatory variables (i.e., predictors, covariates, or features). The strength of these relationships infere species' niches, and can be used to predict a species' occurrence in unsurveyed locations. Although SDMs are a fundamental tool for answering many ecological, evolutionary, and conservation-related questions, some methodological issues remain unresolved (Araújo et al., 2019; Moudrý et al., 2017; Rocchini et al., 2011; Santini et al., 2021).

One such issue is the choice of *spatial resolution*, or *grain*, of the input data (Dungan et al., 2002). It has been hypothesized that organisms respond to their environment more strongly at some grains than at others; these grains have been referred to as 'ecological scales' (Lecours et al., 2015), 'characteristic scales' (Holland et al., 2004), 'intrinsic scales' (Wu and Li, 2006) and 'response grains' (Mertes and Jetz, 2018). This concept implies that for every species, there are one or more grains that best capture the scales at which organisms most strongly respond to specific environmental variables. For example, it is assumed that climate constrains species distributions at broader spatial scales (e.g., at the extent of a whole continent, with phenomena that can be measured at a coarse resolution like > 100 km^2). At successively finer resolutions and over smaller geographic extents, topography or biotic interactions may be the dominant variables in controlling species distribution, whereas at even finer resolutions, microclimate, vegetation structure, or the presence of individual land cover categories such as water bodies might drive local species distribution (Austin and Van Niel, 2011; Field et al., 2009; Pearson and Dawson, 2003; Wiens, 1989). However, previous studies have suggested that some of the abovementioned variables may shape species distribution across multiple grains (e.g., Alexander et al., 2015; Bütikofer et al., 2020; Wisz et al., 2013). Consequently, the choice of grain adopted in models can strongly influence our ability to detect and measure species' response to the environment (de Knegt et al., 2010; Huston, 2002; Levin, 1992; Soberón, 2007; Cord et al. 2014).

Ideally, both species occurrence data and predictor variables are available at relatively fine resolutions, allowing the researchers to coarsen the resolutions iteratively to find the best match between the predictor and response

55   variables. While the response data should preferably be available at resolutions at which the species are expected

56   to respond to the environment, predictor variables should be detailed enough to allow distinguishing important

57   features of the environment that are hypothesized to affect species distribution (e.g., a certain habitat type or

58   specific microclimatic conditions). However, this is not always the case due to limitations in data availability.

59   Usually, the original spatial resolution of different datasets that need to be integrated for modelling purposes

60   varies significantly, and thus finding an optimal match remains a significant challenge.

61   It is a common practice to modify the resolution of the input data so that it matches the resolution at which the

62   study is intended, for example, by averaging environmental variables within field plots. Both continuous (e.g.,

63   bioclimatic variables, terrain characteristics such as slope) and categorical (e.g., land cover) predictors are often

64   aggregated or resampled to match the resolution of the response variable (Grohmann, 2015; Moudrý et al.,

65   2019). While not commonly implemented, an alternative approach consists of retaining the discrepancy between

66   the grain sizes of the response and predictor variables through hierarchical modelling. This allows modelling

67   species distribution using fine-grain species data and coarse-grain environmental data (McInerny and Purves,

68   2011), coarse-grain species data using fine-grain environmental data (Keil et al., 2013, 2014), or modelling the

69   grain-dependency of the species-environment relationships. The latter can be done using an extra parameter in

70   the model to quantify the relationship across a continuum of spatial scales (Keil & Chase, 2019).

71   Any end user should know how changing the spatial resolution of predictor and response variables can affect

72   SDM performance and which data characteristics play a role in how profound the effect of changing the

73   resolution will be. Therefore, here we review methodological issues related to the choice of the spatial resolution

74   of predictor and response variables in SDM. In particular, we focus on the following issues: (i) the effects of

75   changing the resolution of predictor and response variables on model performance, (ii) the effect of conducting

76   multi-grain vs single-grain analysis on model performance, and (iii) the role of land cover type and spatial

77   autocorrelation in the selection of appropriate grain sizes. Accordingly, we aim at providing recommendations

78   for the critical assessment of the input data.

79

80

81

82

3

2. Effects of changing the resolution of predictor and response variables on model performance

Numerous studies examined the grain dependence of species-environment relationships (see the review by Moudrý and Šímová, 2012). Some authors coarsened the resolution of the response variable (section 2.1), others coarsened the resolution of predictor variables so that the resulting predictor was coarser than the response variable (section 2.2). Finally, in some studies the resolution of predictor variables was coarsened so that the resulting predictor was finer than the response variable (section 2.3). These three scenarios are shown in Figure 1. The distinction between these three approaches is often not made in the respective studies, and the effect of changing any resolution can be mistakenly understood as a single problem. We found no studies manipulating the resolution of predictors from finer to coarser resolution compared to the response variable, nor did we find studies manipulating the resolution of both the predictors and the response simultaneously (but see Tobalske, 2002).

2.1 How the resolution of the response variable affects model performance

The availability of species data at a much coarser resolution than commonly used environmental variables (e.g., species occurrence locations only available aggregated at a municipal or county level; Cheng et al., 2021; Jetz et al., 2012) can significantly limit our ability to model species-environment relationships. Studies using species data at such coarse resolutions are not uncommon, especially for less studied taxa. As examples of such data, we can name gridded atlases (Jalas and Suominen, 1988; Šťastný et al., 2021), the resolutions of which can range from hundreds of meters to tens of kilometres. However, monitoring programs collecting atlas data are organizationally and financially demanding. The choice of grid resolution then becomes a trade-off between the level of detail and the feasibility of fieldwork. It is increasingly common to supplement atlases with maps generated with SDMs (e.g., Flousek et al., 2015; Šťastný et al., 2021). As field data may nowadays be gathered with the knowledge that they will also be used for modelling, it is important to know how the resolution of the response affects model performance.

In studies specifically examining the effect of grain size of the response variable on SDM performance, response grain ranges from a few metres to hundreds of kilometres, depending on the predictors tested (Figure 1a; Table 1; see review by Miguet et al., 2015). These studies typically ask: at what scale(s) is the species distribution

most driven or constrained by specific environmental conditions? At finer resolutions, studies typically concentrate on the role of landscape structure (composition and configuration) in driving species distribution (Heikkinen et al., 2007; Holland et al., 2004; Tobalske, 2002). With coarser response grains, studies often include (bio)climatic variables (Chauvier et al., 2022; Kaliontzopoulou et al., 2008; Seo et al., 2009). Typically, such studies report declining model performance with the coarsening of the resolution of the response variable (Chauvier et al., 2022; Gábor et al., 2022a; Heikkinen et al., 2007; Kaliontzopoulou et al., 2008; Seo et al., 2009; Zarzo-Arias et al., 2022), suggesting that modelling species at coarser resolutions is not optimal. However, these studies typically focus on the general performance of the models and do not report the effect of changing the response grain on the variables' importance, which may provide valuable insights into which variables shape species distributions at individual grain sizes (but see Chauvier et al., 2022; Hanberry, 2013).

2.2 How the resolution of the predictor variable (coarser than the response variable) affects model performance?

Instead of coarsening the resolution of the response variable, some studies have coarsened the resolution of predictor variables, so that the resulting predictor is coarser than the response variable (Figure 1b; Table 2). They came to different conclusions. Ferrier and Watson (1997) concluded that coarse environmental data lead to poorer model performance. Graf et al. (2005) found that the predictive power was highest at resolutions of about 1 and 2 km^2. In contrast, Guisan et al. (2007) and Pradervand et al. (2014) concluded that coarsening the predictor variables' resolution did not substantially change model performance, meaning that refining the resolution may not be sufficient to improve the models.

2.3 How the resolution of the predictor variable (finer than the response variable) affects model performance?

Studies that manipulate the resolution of predictor variables, so that the resulting predictor was finer than the response variable (Figure 1c; Table 3), are mostly concerned with the importance of fine-scale habitat features for analyzing species-environment relationships (e.g., Gottschalk et al., 2011; Šímová et al., 2019). They

135   combine response variables at a coarse resolution with predictor variables at a fine resolution. These studies

136   typically ask: do we need fine-resolution predictors to explain species distribution at a relatively coarse

137   resolution?

138   High-resolution predictor variables suitable for modelling at multiple levels of detail may not be readily

139   available for the particular study area, their acquisition may be prohibitively expensive (especially for studies

140   conducted over large extents), and their use may require excessive data processing and significantly increase

141   computational time (Kissling et al. 2022; Moudrý et al. 2022). Hence, researchers face trade-offs between data

142   detail and availability, data processing, and analytical optimization. Several studies have examined the

143   importance of fine-grain habitat features for the analysis of species-environment relationships using a relatively

144   coarse-grained response variable (Figure 1c; Table 3). In this type of study, authors typically use predictor

145   variables of various origins, collected, for example, by remote sensing (Leitão & Santos 2019), fieldwork, or

146   crowd-sourcing (Šímová et al., 2019; Thomas et al., 2002; Venier et al., 2004). Others have coarsened the grain

147   of the original predictors to examine the grain dependency of species-environment relationships (e.g. Gottschalk

148   et al., 2011).

149   Thomas et al. (2002) found that field-collected fine-grain predictors and predictor variables derived from a 30

150   m digital elevation model lead to the same model performance at a 1 km resolution. Seoane et al. (2004) found

151   that models derived from land cover at a 250 m resolution are comparable to those based on the same variables

152   derived from satellite images at a 30 m resolution, in agreement with Venier et al. (2004). Consequently, it is

153   commonly assumed that coarse-resolution habitat predictors at continental (e.g., CORINE Land cover; Büttner

154   et al., 2004) or global (e.g., Global Consensus Land cover; Tuanmu and Jetz, 2014) geographic extents are

155   sufficient for use in combination with coarse-resolution responses.

156   However, it is essential to know if a given spatial resolution of a predictor variable captures the details that are

157   important for explaining the distribution of the species of interest. Gottschalk et al. (2011) concluded that a

158   higher spatial resolution of predictors could be essential for accurate predictions. In addition, they attributed the

159   improvement in models using detailed land cover maps to the high level of detail in the species response variable

160   (2 km diameter around survey points). This contrasts with results by Šímová et al. (2019) that demonstrated

161   improvement in model performance when using high–resolution land cover data despite the coarse resolution

162   of species data (12 x 11.2 km). They showed that the area and perimeter of water bodies derived from high–

6

163   resolution land cover datasets (raster data at 30 m resolution) explain distributions of waterbirds better than

164   predictors derived from coarser 1 km data. In line with these findings, it has been recently recommended to first

165   coarsen the resolution of the predictors to match the resolution of the assumed ecological scale before

166   calculating prediction metrics (e.g., standard deviation, Shannon-Wiener diversity index, or Rao's Q) at the

167   resolution of a response variable (Graham et al., 2019). In this context, the recent finding by Gábor et al. (2022b),

168   who showed that in the case of species inhabiting rare habitats, using simple binary predictors (i.e.

169   presence/absence of the habitat) might be sufficient, is of particular interest.

170   In conclusion, coarse–resolution land cover or terrain predictors may lack details to capture potentially suitable

171   habitats such as wetlands or cliffs. Thus, using high-resolution data could benefit models utilizing coarser-

172   resolution species data (e.g. from gridded atlases). The question of whether the need for fine-scale predictors is

173   somehow related to the resolution of the response variable or whether it can be generalized should be further

174   explored for different taxa and sets of predictors.

175

3.   Single-grain versus multi-grain analysis

177   Up to this point, we have neglected discussing the possibility of considering species-environment relationships

178   at multiple grains in a single model. Typically, experimental studies use a single grain for the response variable.

179   Therefore, they implicitly assume the existence of a common ecological scale for all predictor variables.

180   However, it has been shown that the ecological scale is variable-specific since species often respond to different

181   environmental variables at different spatial scales, and sometimes even respond differently to a single

182   environmental variable at multiple grains (Leitão et al. 2010; Lecours et al., 2020; Miguet et al., 2016; Roilo et

183   al. 2022). However, despite theoretical concepts and extensive empirical evidence that species respond to their

184   environment at different spatial grains (e.g. Bergman et al., 2012; Graf et al., 2005; Holland et al., 2004; Stuber

185   and Fontaine, 2019; Zweifel-Schielly et al., 2009), the appropriate approach to select the grain of response

186   variable remains unclear (Jackson and Fahrig, 2015; Martin and Fahrig, 2012; Stuber and Gruber, 2020). For

187   example, Mertes et al. (2020) recognized two primary spatial grains at which species typically respond to their

188   environment: they denoted the term "occupancy grain" for the grain equivalent to a species' typical home range

189   and the term "response grain" for the grain at which an individual uses an environmental resource. They also

7

190 developed an optimization procedure for their identification. However, studies usually use grains of response

191 variables coarser than the assumed occupancy and response grain, and it is unclear how to incorporate

192 occupancy and response grains in such studies (but see Graham et al., 2019).

193

194 In theory, species distributions are driven by environmental variables at a range of scales (Levin, 1992), and

195 there is no single "correct" spatial grain at which to characterize species-environment associations (Mitchell et

196 al., 2001; Wiens, 1989). Therefore, models using multiple grains should, in theory, outperform models that

197 assume a common ecological scale for all variables. However, scale-sensitive applications that aim to align the

198 grain of the response variable (or predictor variables; see Graham et al., 2019) with the ecological scale are rare

199 (McGarigal et al., 2016). In addition, studies have come up with different conclusions. Some have suggested

200 that the performance of models using multiple response variable grains is better than that of single-grain models

201 (Mertes et al., 2020), while others have not drawn similar conclusions (Martin and Fahrig, 2012). Of note is that

202 the improvements reported for multi-grain models were often relatively low, in the order of hundredths of the

203 area under the receiver operating characteristic curve (AUC) values (Boscolo and Metzger, 2009; Graf et al.,

204 2005; Kuhn et al., 2011; Mateo Sánchez et al., 2014). In other words, valuable predictions can still be generated

205 from models using a single arbitrarily selected scale. Hence, it remains unclear whether the increased complexity

206 caused by the use of multiple grains is beneficial, particularly in the case of SDMs used for the projection of

207 species distributions under future climate conditions, which are generally uncertain (e.g. Sinclair et al., 2010).

208

4. Land cover types and spatial autocorrelation

210 In an early study on the effect of spatial resolution on the performance of species-habitat relationships, Karl et

211 al. (2000) suggested that the effects of coarsening the resolution depend on the heterogeneity of the environment.

212 The difference in land cover types used in different analyses might, therefore, explain some contrasting findings.

213 For example, Seoane et al. (2004) and Venier et al. (2004) observed no improvement in models when using

214 finer-grain land cover data, while Gottschalk et al. (2011) and Šímová et al. (2019) observed a significant

215 improvement. Both Seoane et al. (2004) and Venier et al. (2004) used data on common land cover types, such

216 as the proportion of forests within mapping units. For homogeneous landscapes displaying strong spatial

217 autocorrelation (e.g., large blocks of forests), land cover information does not change much when spatially

218 aggregated to coarser resolutions. In contrast, Šímová et al. (2019) focused on water bodies, a land cover

219 category that can become virtually invisible at coarser resolutions; coarsening the resolution often leads to a

220 bias and underrepresentation of rare environments such as (especially linear) water bodies in certain landscapes.

221 Similarly, Seoane et al. (2004) observed considerable improvement in models for riparian species when finer-

222 resolutions predictors were used. This may be one of the reasons why Tuanmu and Jetz (2014) found that the

223 Global Consensus Land Cover that has a spatial resolution of 1 km² (https://www.earthenv.org/landcover; see

224 Table 2) performed worse for predicting aquatic species than species inhabiting other environments. Similarly,

225 Cord et al. (2014) showed for 30 tree species that SDM performance was significantly positively correlated with

226 the species-specific degree of association between the focal species and different land cover types.

227

228 Environmental variables are typically spatially autocorrelated (i.e., values between two locations are more

229 similar the closer the locations are in space; Legendre, 1993). This spatial autocorrelation can be quantified

230 using an empirical variogram that can be used to calculate the characteristic distance within which spatial

231 autocorrelation operates (i.e., the 'range' of an empirical variogram). Recently, Mertes and Jetz (2018)

232 highlighted the importance of considering environmental autocorrelation for the ability of SDMs to estimate

233 species-environment associations. Similar results were obtained by Kühn (2006) for species richness. More

234 recently, Smith and Santos (2020) explored the effect of the resolution of predictor variables and their

235 autocorrelation on estimates of their importance. This body of literature shows that using coarser environmental

236 data in SDMs without consideration of the autocorrelation can mischaracterize species-environment

237 relationships (see Miller, 2012, for review). This is particularly true for variables that vary rapidly over space;

238 i.e. heterogeneous landscapes characterized by spatial autocorrelation with relatively small range values (Mertes

239 and Jetz, 2018). Aggregating heterogeneous landscapes to a coarser resolution results in the loss of a portion of

240 that heterogeneity (Graham et al., 2019; Karl et al., 2000; Mertes and Jetz, 2018). Lower autocorrelation means

241 higher randomness; hence, very distinct values are aggregated together. In contrast, if there is strong

242 autocorrelation, aggregating over a larger area does not change the value much because the values were similar

243 even in the finer resolutions.

244 Importantly, the inherent spatial autocorrelation of both species occurrences and predictor variables can result

245 in models that may inadvertently capture the spatial structure rather than true functional relationships (Bahn and

246 McGill, 2007). Indeed, it has been shown that spatial autocorrelation can lead to SDMs with high discrimination

247 ability even when there is no relationship between species occurrence and environmental variables (Chapman,

248 2010; Fourcade et al., 2018) and that many SDMs, despite a good fit, are not significantly better than null models

249 (Osborne et al., 2022). Therefore, it is a question of whether the loss of explanatory power accompanying the

250 coarsening of the resolution is due to the use of an inappropriate scale (e.g. due to the lack of detail of potentially

251 suitable environmental conditions) or due to changes in the spatial structure; hence, this loss of power should

252 be further explored for different resolutions and predictors. In any case, selecting a relevant set of environmental

253 predictors based on the known ecology of the species of interest is essential to ensure fitting SDMs with an

254 appropriate ecological interpretation (Fourcade et al., 2018). In addition, it is necessary to carefully inspect

255 whether SDMs estimated from the observed data perform better than those generated from the null occurrence

256 distributions, for example by using the recently-developed "fauxcurrence" R package (Osborne et al., 2022).

257

5. The ratio between the resolution of response and predictor variables

259 A recently proposed standard protocol (Zurell et al., 2020) recommends reporting information on data,

260 modelling techniques, validation, and underlying questions (Araújo et al., 2019; Michener and Jones, 2012;

261 Rocchini and Neteler, 2012). However, many studies still lack it (see Feng et al., 2019 for a review). When

262 evaluating the effect of changing the resolution of predictor variables, it is also important to consider the

263 resolution of the response variable (i.e., species occurrences). The opposite is also true: when evaluating the

264 role of the resolution of the response variable, one should be aware of the resolution of predictor variables.

265 Although this may seem like a trivial recommendation, it remains infrequent that studies evaluating the effects

266 of changing resolutions discuss their results with respect to the ratio between the resolutions of the response

267 variable and predictor variables (but see Moudrý and Šímová, 2012). The ratio between the resolution of the

268 response and the resolution(s) of the predictor variables differs among studies and might be the reason for

269 reported contradicting results (Figure 1). For example, in studies evaluating the importance of finer–resolution

270 predictors to explain species distributions, response grains can differ considerably (Table 4). It can be expected

271 that for small ratios, coarsening of the resolution of predictor variables will have a minimal effect on model

272 performance (e.g. Seoane et al. 2004; Venier et al. 2004), while for high ratios (indicating a high difference

273 between the resolutions of the response and predictor variables), considerable effects can be expected due to the

274 aggregation of highly different values (e.g. Gottschalk et al. 2011; Šímová et al. 2019). Practices could be

275 improved by reporting the resolution of predictor variables as well as that of the response variable.

276

6. Conclusions

278 Spatial scale is one of the most critical issues in ecology and associated disciplines (Levin, 1992). Species

279 respond to their environment at different scales, and processes controlling species distribution operate at various

280 spatial scales. Unsurprisingly, the studies we reviewed found various optimal resolutions, depending on the

281 species and ecosystems analyzed. Besides, most studies analyzing multiple species usually report only a general

282 trend in models' behaviour with respect to changing resolution, and there are always some models that do not

283 conform to the general pattern (e.g., Guisan et al., 2007; Pradervand et al., 2014). Our review highlights that

284 within the typically used resolutions ($0.01 - 100$ km^2) finer-resolution models generally perform better. Besides,

285 the use of coarse-resolution response variables has implications for the predicted distribution range (Kunin,

286 1998). When the resolution of the response variable is too coarse, there is a risk of overestimating the occupied

287 area (Connor et al., 2018; Hu and Jiang, 2010; Lauzeral et al., 2013; Seo et al., 2009). Moreover, Gábor et al.

288 (2022a) recently showed that coarsening the resolution does not compensate for positional error in species

289 occurrence data. Therefore, we recommend basing the choice of the resolution of the response variable on

290 practical aspects, such as aiming for finer resolutions unless there is a reason to do otherwise (e.g., expert

291 knowledge of the ecological scale of the species under study).

292

293 Coarsening the resolution of predictor variables has been shown to negatively affect model performance as it

294 obscures fine-scale heterogeneity in environmental variables. Therefore, we recommend (1) using finer-

295 resolution environmental variables when modelling species associated with rare environmental entities (e.g., a

296 rare habitat type), even when using species occurrence data at a coarse resolution (Šímová et al., 2019). When

297 species are associated with widespread environmental conditions, using low–resolution predictors is likely

11

sufficient. However, we recommend (2) assessing spatial autocorrelation or thematic resolution of predictors at multiple grains to estimate the potential impacts of coarsening their resolution on model performance (i.e. to ensure that they preserve enough detail to distinguish environmental features that affect species distribution at a given resolution). Thirdly, (3) studies may benefit from considering multiple grains of the response variable within a single model, even though the improvements reported for multi-grain models have so far been *relatively* low, and we recognize that useful predictions can still be generated from single-scale models. Finally, (4) studies should explicitly report the resolutions of the predictor and response variables, following the standard ODMAP protocol recently proposed by Zurell et al. (2020).

7. References

Alexander JM, Diez JM and Levine JM (2015) Novel competitors shape species' responses to climate change. Nature 525(7570): 515–518. DOI: 10.1038/nature14952.

Araújo MB, Anderson RP, Márcia Barbosa A, et al. (2019) Standards for distribution models in biodiversity assessments. Science Advances 5(1): eaat4858. DOI: 10.1126/sciadv.aat4858.

Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. Journal of biogeography 38(1):1-8. https://doi.org/10.1111/j.1365-2699.2010.02416.x

Bahn V and McGill BJ (2007) Can niche-based distribution models outperform spatial interpolation? Global Ecology and Biogeography 16(6): 733–742. DOI: 10.1111/j.1466-8238.2007.00331.x.

Bergman K-O, Jansson N, Claesson K, et al. (2012) How much and at what scale? Multi-scale analyses as decision support for conservation of saproxylic oak beetles. Forest Ecology and Management 265: 133–141. DOI: 10.1016/j.foreco.2011.10.030.

Boscolo D and Metzger JP (2009) Is bird incidence in Atlantic forest fragments influenced by landscape patterns at multiple scales? Landscape Ecology 24(7): 907–918. DOI: 10.1007/s10980-009-9370-8.

Bütikofer L, Anderson K, Bebber DP, et al. (2020) The problem of scale in predicting biological responses to climate. Global Change Biology 26(12): 6657–6666. DOI: 10.1111/gcb.15358.

Büttner G, Feranec J, Jaffrain G., et al. (2004) The CORINE land cover 2000 project. EARSeL eProceedings. 3: 331–346.

325 Chauvier Y, Descombes P, Guéguen M, et al. (2022) Resolution in species distribution models shapes spatial

326 patterns of plant multifaceted diversity. Ecography. DOI: 10.1111/ecog.05973.

327 Cheng Y, Tjaden NB, Jaeschke A, et al. (2021) Using centroids of spatial units in ecological niche modelling:

328 Effects on model performance in the context of environmental data grain size. Global Ecology and

329 Biogeography Peres-Neto P (ed.) 30(3): 611–621. DOI: 10.1111/geb.13240.

330 Connor T, Hull V, Viña A, et al. (2018) Effects of grain size and niche breadth on species distribution

331 modeling. Ecography 41(8): 1270–1282. DOI: 10.1111/ecog.03416.

332 Cord AF, Klein D, Mora F, Dech S (2014) Comparing the suitability of classified land cover data and remote

333 sensing variables for modeling distribution patterns of plants. Ecological Modelling 24(272):129-40.

334 Corsi F, De Leeuw J and Skidmore AK (2000) Modelling species distribution with GIS. Research techniques

335 in animal ecology; controversies and consequences (Boitani L and Fuller TK, eds). Columbia University

336 Press, New York, USA.

337 de Knegt HJ, van Langevelde F, Coughenour MB, et al. (2010) Spatial autocorrelation and the scaling of

338 species–environment relationships. Ecology 91(8): 2455–2465. DOI: 10.1890/09-1359.1.

339 Dungan JL, Perry JN, Dale MRT, et al. (2002) A balanced view of scale in spatial statistical analysis.

340 Ecography 25(5): 626–640. DOI: 10.1034/j.1600-0587.2002.250510.x.

341 Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space

342 and time. Annual Review of Ecology, Evolution and Systematics 40(1): 677-97.

343 Feng X, Park DS, Walker C, et al. (2019) A checklist for maximizing reproducibility of ecological niche

344 models. Nature Ecology & Evolution 3(10): 1382–1395. DOI: 10.1038/s41559-019-0972-5.

345 Ferrier S and Watson G (1997) An Evaluation of the Effectiveness of Environmental Surrogates and

346 Modelling Techniques in Predicting the Distribution of Biological Diversity. Environment Australia.

347 Ferrier S, Jetz W and Scharlemann, J (2017) Biodiversity modelling as part of an observation system. In The

348 GEO handbook on biodiversity observation networks. (239-257). Springer, Cham.

349 Field R, Hawkins BA, Cornell HV, et al. (2009) Spatial species-richness gradients across scales: a meta-

350 analysis. Journal of Biogeography 36(1): 132–147. DOI: 10.1111/j.1365-2699.2008.01963.x

351 Flousek J, Gramsz B and Telenský T (2015). Ptáci Krkonoš – atlas hnízdního rozšíření 2012–2014 / Ptaki

352 Karkonoszy – atlas ptaków lęgowych2012–2014. Správa KRNAP Vrchlabí, Dyrekcja KPN Jelenia Góra: 480.

353 http://ptacikrkonos.krnap.cz/files/PTACI-KRKONOS.pdf

354 Fourcade Y, Besnard AG and Secondi J (2018) Paintings predict the distribution of species, or the challenge

355 of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography 27(2): 245–

356 256. DOI: 10.1111/geb.12684

357 Gábor L, Jetz W, Lu M, et al. (2022a) Positional errors in species distribution modelling are not overcome by

358 the coarser grains of analysis. Methods in Ecology and Evolution: 2041–210X.13956. DOI: 10.1111/2041-

359 210X.13956.

360 Gábor L, Šímová P, Keil P, et al. (2022b) Habitats as predictors in species distribution models: Shall we use

361 continuous or binary data? Ecography 2022(7). DOI: 10.1111/ecog.06022.

362 Gottschalk TK, Aue B, Hotes S, et al. (2011) Influence of grain size on species–habitat models. Ecological

363 Modelling 222(18): 3403–3412. DOI: 10.1016/j.ecolmodel.2011.07.008.

364 Graf RF, Bollmann K, Suter W, et al. (2005) The Importance of Spatial Scale in Habitat Models: Capercaillie

365 in the Swiss Alps. Landscape Ecology 20(6): 703–717. DOI: 10.1007/s10980-005-0063-7.

367 Graham LJ, Spake R, Gillings S, et al. (2019) Incorporating fine-scale environmental heterogeneity into

368 broad-extent models. Methods in Ecology and Evolution Isaac N (ed.) 10(6): 767–778. DOI: 10.1111/2041-

369 210X.13177.

370 Grohmann CH (2015) Effects of spatial resolution on slope and aspect derivation for regional-scale analysis.

371 Computers & Geosciences 77. Elsevier: 111–117.

372 Guisan A, Graham CH, Elith J, et al. (2007) Sensitivity of predictive species distribution models to change in

373 grain size. Diversity and Distributions 13(3): 332–340. DOI: 10.1111/j.1472-4642.2007.00342.x.

374 Hanberry BB (2013) Finer grain size increases effects of error and changes influence of environmental

375 predictors on species distribution models. Ecological Informatics : 15: 8-13.

376 Heikkinen RK, Luoto M, Kuussaari M, et al. (2007) Modelling the spatial distribution of a threatened

377 butterfly: Impacts of scale and statistical technique. Landscape and Urban Planning 79(3–4): 347–357. DOI:

378 10.1016/j.landurbplan.2006.04.002.

14

379 Holland JD, Bert DG and Fahrig L (2004) Determining the Spatial Scale of Species' Response to Habitat.

380 BioScience 54(3): 227. DOI: 10.1641/0006-3568(2004)054[0227:DTSSOS]2.0.CO;2.

381 Hu J and Jiang Z (2010) Predicting the potential distribution of the endangered Przewalski's gazelle. Journal

382 of Zoology 282(1): 54–63. DOI: 10.1111/j.1469-7998.2010.00715.x.

383 Huston MA (2005) Introductory essay: critical issues for improving predictions. Predicting species

384 occurrences: issues of accuracy and scale. (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G.

385 Raphael, W.A. Wall and F.B. Samson), pp. 7– 21. Island Press, Covelo, CA.

386 Jackson HB and Fahrig L (2015) Are ecologists conducting research at the optimal scale?: Is research

387 conducted at optimal scales? Global Ecology and Biogeography 24(1): 52–63. DOI: 10.1111/geb.12233.

388 Jalas J and Suominen J (Eds.) (1988) Atlas Florae Europaeae: volume 3: distribution of vascular plants in

389 Europe (Vol. 3). Cambridge University Press.

390 Jetz W, McPherson JM and Guralnick RP (2012) Integrating biodiversity distribution knowledge: toward a

391 global map of life. Trends in Ecology & Evolution 27(3): 151–159. DOI: 10.1016/j.tree.2011.09.007.

392 Kaliontzopoulou A, Brito JC, Carretero MA, et al. (2008) Modelling the partially unknown distribution of

393 wall lizards (Podarcis) in North Africa: ecological affinities, potential areas of occurrence, and methodological

394 constraints. Canadian Journal of Zoology 86(9): 992–1001. DOI: 10.1139/Z08-078.

395 Karl JW, Heglund PJ, Garton EO, et al. (2000) SENSITIVITY OF SPECIES HABITAT-RELATIONSHIP

396 MODEL PERFORMANCE TO FACTORS OF SCALE. Ecological Applications 10(6): 1690–1705. DOI:

397 10.1890/1051-0761(2000)010[1690:SOSHRM]2.0.CO;2.

398 Keil P and Chase JM (2019) Global patterns and drivers of tree diversity integrated across a continuum of

399 spatial grains. Nature Ecology & Evolution 3(3): 390–399. DOI: 10.1038/s41559-019-0799-0.

400 Keil P, Belmaker J, Wilson AM, et al. (2013) Downscaling of species distribution models: a hierarchical

401 approach. Methods in Ecology and Evolution Freckleton R (ed.) 4(1): 82–94. DOI: 10.1111/j.2041-

402 210x.2012.00264.x.

403 Keil P, Wilson AM and Jetz W (2014) Uncertainty, priors, autocorrelation and disparate data in downscaling

404 of species distributions. Diversity and Distributions Brotons L (ed.) 20(7): 797–812. DOI: 10.1111/ddi.12199.

405   Kissling WD, Shi Y, Koma Z, Meijer C, Ku O, Nattino F, Seijmonsbergen AC, Grootes MW (2022) Country-

406   wide data of ecosystem structure from the third Dutch airborne laser scanning survey. Data in Brief:108798.

407   https://doi.org/10.1016/j.dib.2022.108798

408   Kuhn A, Copeland J, Cooley J, et al. (2011) Modeling Habitat Associations for the Common Loon (Gavia

409   immer) at Multiple Scales in Northeastern North America. Avian Conservation and Ecology 6(1): art4. DOI:

410   10.5751/ACE-00451-060104.

411   Kühn I (2006) Incorporating spatial autocorrelation may invert observed patterns. Diversity & Distributions

412   0(0): 061117052025001-??? DOI: 10.1111/j.1472-4642.2006.00293.x.

413   Kunin WE (1998) Extrapolating species abundance across spatial scales. Science 281(5382). American

414   Association for the Advancement of Science: 1513–1515.

415   Lauzeral C, Grenouillet G and Brosse S (2013) Spatial range shape drives the grain size effects in species

416   distribution models. Ecography 36(7): 778–787. DOI: 10.1111/j.1600-0587.2013.07696.x.

417   Lecours V, Devillers R, Schneider DC, et al. (2015) Spatial scale and geographic context in benthic habitat

418   mapping: review and future directions. Marine Ecology Progress Series, 535: 259-284.

419   Lecours V, Gábor L, Edinger E, et al. (2020) Fine-scale habitat characterization of The Gully, the Flemish

420   Cap, and the Orphan Knoll, Northwest Atlantic, with a focus on cold-water corals. In: Seafloor

421   Geomorphology as Benthic Habitat. Elsevier, pp. 735–751.

422   Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74(6). Wiley Online Library:

423   1659–1673.

424   Leitao PJ, Moreira F, and Osborne PE (2010) Breeding habitat selection of steppe birds in Castro Verde: A

425   remote sensing and advanced statistics approach. Ardeola, 57, 93-116.

426   Leitão PJ, Santos MJ. Improving models of species ecological niches: a remote sensing overview. Frontiers in

427   Ecology and Evolution. 2019 Jan 29;7:9.

428   Levin SA (1992) The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture.

429   Ecology 73(6): 1943–1967. DOI: 10.2307/1941447.

430   Manzoor SA, Griffiths G and Lukac M (2018) Species distribution model transferability and model grain size

431   – finer may not always be better. Scientific Reports 8(1): 7168. DOI: 10.1038/s41598-018-25437-1.

432 Martin AE and Fahrig L (2012) Measuring and selecting scales of effect for landscape predictors in species–

433 habitat models. Ecological Applications 22(8): 2277–2292. DOI: 10.1890/11-2224.1.

434 Mateo Sánchez MC, Cushman SA and Saura S (2014) Scale dependence in habitat selection: the case of the

435 endangered brown bear (Ursus arctos) in the Cantabrian Range (NW Spain). International Journal of

436 Geographical Information Science 28(8): 1531–1546. DOI: 10.1080/13658816.2013.776684.

437 McGarigal K, Wan HY, Zeller KA, et al. (2016) Multi-scale habitat selection modeling: a review and outlook.

438 Landscape Ecology 31(6): 1161–1175. DOI: 10.1007/s10980-016-0374-x.

439 McInerny GJ and Purves DW (2011) Fine-scale environmental variation in species distribution modelling:

440 regression dilution, latent variables and neighbourly advice: Regression dilution in species distribution

441 models. Methods in Ecology and Evolution 2(3): 248–257. DOI: 10.1111/j.2041-210X.2010.00077.x.

442 Mertes K and Jetz W (2018) Disentangling scale dependencies in species environmental niches and

443 distributions. Ecography 41(10): 1604–1615. DOI: 10.1111/ecog.02871.

444 Mertes K, Jarzyna MA and Jetz W (2020) Hierarchical multi-grain models improve descriptions of species'

445 environmental associations, distribution, and abundance. Ecological Applications 30(6). DOI:

446 10.1002/eap.2117.

447 Michener WK and Jones MB (2012) Ecoinformatics: supporting ecology as a data-intensive science. Trends

448 in Ecology & Evolution 27(2): 85–93. DOI: 10.1016/j.tree.2011.11.016.

449 Miguet P, Jackson HB, Jackson ND, et al. (2016) What determines the spatial extent of landscape effects on

450 species? Landscape Ecology 31(6): 1177–1194. DOI: 10.1007/s10980-015-0314-1.

451 Miller JA (2012) Species distribution models: Spatial autocorrelation and non-stationarity. Progress in

452 Physical Geography: Earth and Environment 36(5): 681–692. DOI: 10.1177/0309133312442522.

453 Mitchell MS, Lancia RA, Gerwin JA (2001) Using landscape-level data to predict the distribution of birds on

454 a managed forest: effects of scale. Ecological Applications: 11(6):1692-708.

455 Moudrý V and Šímová P (2012) Influence of positional accuracy, sample size and scale on modelling species

456 distributions: a review. International Journal of Geographical Information Science 26(11): 2083–2095. DOI:

457 10.1080/13658816.2012.721553.

458 Moudrý V, Komárek J and Šímová P (2017) Which breeding bird categories should we use in models of

459 species distribution? Ecological Indicators 74: 526–529. DOI: 10.1016/j.ecolind.2016.11.006.

17

460 Moudrý V, Lecours V, Malavasi M, et al. (2019) Potential pitfalls in rescaling digital terrain model-derived

461 attributes for ecological studies. Ecological Informatics 54: 100987. DOI: 10.1016/j.ecoinf.2019.100987.

462 Moudrý V, Cord AF, Gábor L, Laurin GV, Barták V, Gdulová K, Malavasi M, Rocchini D, Stereńczak K,

463 Prošek J, Klápště P (2022) Vegetation structure derived from airborne laser scanning to assess species

464 distribution and habitat suitability: The way forward. Diversity and Distributions.

465 https://doi.org/10.1111/ddi.13644

466 Osborne OG, Fell HG, Atkins H, et al. (2022) Fauxcurrence: simulating multi-species occurrences for null

467 models in species distribution modelling and biogeography. Ecography 2022(7). DOI: 10.1111/ecog.05880.

468 Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species:

469 are bioclimate envelope models useful?: Evaluating bioclimate envelope models. Global Ecology and

470 Biogeography 12(5): 361–371. DOI: 10.1046/j.1466-822X.2003.00042.x.

471 Pradervand J-N, Dubuis A, Pellissier L, et al. (2014) Very high resolution environmental predictors in species

472 distribution models: Moving beyond topography? Progress in Physical Geography: Earth and Environment

473 38(1): 79–96. DOI: 10.1177/0309133313512667.

474 Rengstorf AM, Grehan A, Yesson C, et al. (2012) Towards High-Resolution Habitat Suitability Modeling of

475 Vulnerable Marine Ecosystems in the Deep-Sea: Resolving Terrain Attribute Dependencies. Marine Geodesy

476 35(4): 343–361. DOI: 10.1080/01490419.2012.699020.

477 Rocchini D and Neteler MG (2012) Let the four freedoms paradigm apply to ecology. GB.

478 Rocchini D, Hortal J, Lengyel S, et al. (2011) Accounting for uncertainty when mapping species distributions:

479 The need for maps of ignorance. Progress in Physical Geography: Earth and Environment 35(2): 211–226.

480 DOI: 10.1177/0309133311399491.

481 Roilo S, Engler JO, Václavík T, Cord AF (2022) Landscape-level heterogeneity of agri-environment measures

482 improves habitat suitability for farmland birds. Ecological Applications. Aug 6:e2720.

483 Santini L, Benítez-López A, Maiorano L, et al. (2021) Assessing the reliability of species distribution

484 projections in climate change research. Diversity and Distributions Fourcade Y (ed.) 27(6): 1035–1050. DOI:

485 10.1111/ddi.13252.

486 Seo C, Thorne JH, Hannah L, et al. (2009) Scale effects in species distribution models: implications for

487 conservation planning under climate change. Biology letters 5(1). The Royal Society London: 39–43.

488 Seoane J, Bustamante J and Díaz-Delgado R (2004) Are existing vegetation maps adequate to predict bird

489 distributions? Ecological Modelling 175(2): 137–149. DOI: 10.1016/j.ecolmodel.2003.10.011.

490 Sillero N and Barbosa AM (2021) Common mistakes in ecological niche models. International Journal of

491 Geographical Information Science 35(2): 213–226. DOI: 10.1080/13658816.2020.1798968.

492 Šímová P, Moudrý V, Komárek J, et al. (2019) Fine scale waterbody data improve prediction of waterbird

493 occurrence despite coarse species data. Ecography 42(3): 511–520. DOI: 10.1111/ecog.03724.

494 Sinclair SJ, White MD and Newell GR (2010) How useful are species distribution models for managing

495 biodiversity under future climates? Ecology and Society 15(1). JSTOR.

496 Smith AB and Santos MJ (2020) Testing the ability of species distribution models to infer variable

497 importance. Ecography 43(12): 1801–1813. DOI: 10.1111/ecog.05317.

498 Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. Ecology letters,

499 10(12), 1115-1123. https://doi.org/10.1111/j.1461-0248.2007.01107.x

500 Šťastný K, Bejček V, Mikuláš I. and Telenský T (2021) Atlas Hnízdního Rozšíření Ptáků v České Republice

501 2014-2017 (Aventinum, 2021)

502 Stockwell DRB and Peterson AT (2002) Controlling bias in biodiversity data. Predicting species occurrences:

503 issues of accuracy and scale. Edited by JM Scott, PJ Heglund, F. Samson, J. Haufler, M. Morrison, M.

504 Raphael, and B. Wall. Island Press, Washington, DC, 537-546.

505 Stuber EF and Fontaine JJ (2019) How characteristic is the species characteristic selection scale? Global

506 Ecology and Biogeography Issac N (ed.) 28(12): 1839–1854. DOI: 10.1111/geb.12998.

507 Stuber EF and Gruber LF (2020) Recent Methodological Solutions to Identifying Scales of Effect in Multi-

508 scale Modeling. Current Landscape Ecology Reports 5(4): 127–139. DOI: 10.1007/s40823-020-00055-8.

509 Thomas K, Keeler-Wolf T and Franklin J (2002) A comparison of fine-and coarse-resolution environmental

510 variables toward predicting vegetation distribution in the Mojave desert. Predicting species occurrences:

511 issues of accuracy and scale. Edited by JM Scott, PJ Heglund, F Samson, J Haufler, M Morrison, M Raphael,

512 and B Wall. Island Press, Washington, DC, 133-139.

513 Tobalske C (2002) Effects of spatial scale on the predictive ability of habitat models for the green woodpecker

514 in Switzerland. Predicting species occurrences: issues of accuracy and scale. Edited by JM Scott, PJ Heglund,

515 F Samson, J Haufler, M Morrison, M Raphael, and B Wall. Island Press, Washington, DC, 197-204.

19

516 Tuanmu M-N and Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem

517 modelling: Consensus land cover. Global Ecology and Biogeography 23(9): 1031–1045. DOI:

518 10.1111/geb.12182.

519 Vale CG, Tarroso P and Brito JC (2014) Predicting species distribution at range margins: testing the effects of

520 study area extent, resolution and threshold selection in the Sahara-Sahel transition zone. Diversity and

521 Distributions Robertson M (ed.) 20(1): 20–33. DOI: 10.1111/ddi.12115.

522 Venier LA, Pearce J, McKee JE, et al. (2004) Climate and satellite-derived land cover for predicting breeding

523 bird distribution in the Great Lakes Basin: Climate and land cover for predicting bird distribution. Journal of

524 Biogeography 31(2): 315–331. DOI: 10.1046/j.0305-0270.2003.01014.x.

525 Wiens JA (1989) Spatial scaling in ecology. Functional ecology. 3: 385-397.

526 Wiersma YF, Huettmann F and Drew CA (2011) Introduction, landscape modeling of species and their

527 habitats: history, uncertainty, and complexity. In: Drew CA, Wiersma YF, Huettmann F (Eds.), Predictive

528 Species and Habitat Modeling in Landscape Ecology. Springer, New York, pp. 1–6.

529 Wisz MS, Pottier J, Kissling WD, et al. (2013) The role of biotic interactions in shaping distributions and

530 realised assemblages of species: implications for species distribution modelling. Biological Reviews 88(1):

531 15–30. DOI: 10.1111/j.1469-185X.2012.00235.x.

532 Wu J and Li H (2006) Concepts of scale and scaling. In: Jianguo Wu, Jones KB, Li H, Louck OL (eds)

533 Scaling and uncertainty analysis in ecology: methods and applications. Springer, Dordrecht, Netherlands, pp

534 3–15

535 Zarzo-Arias A, Penteriani V, Gábor L, Šímová P, Grattarola F, Moudrý V (2022) Importance of data selection

536 and filtering in species distribution models: a case study on the Cantabrian Brown bear. Ecosphere.

537 Zurell D, Franklin J, König C, et al. (2020) A standard protocol for reporting species distribution models.

538 Ecography 43(9): 1261–1277. DOI: 10.1111/ecog.04960.

539 Zweifel-Schielly B, Kreuzer M, Ewald KC, et al. (2009) Habitat selection by an Alpine ungulate: the

540 significance of forage characteristics varies with scale and season. Ecography 32(1): 103–113. DOI:

541 10.1111/j.1600-0587.2008.05178.x.
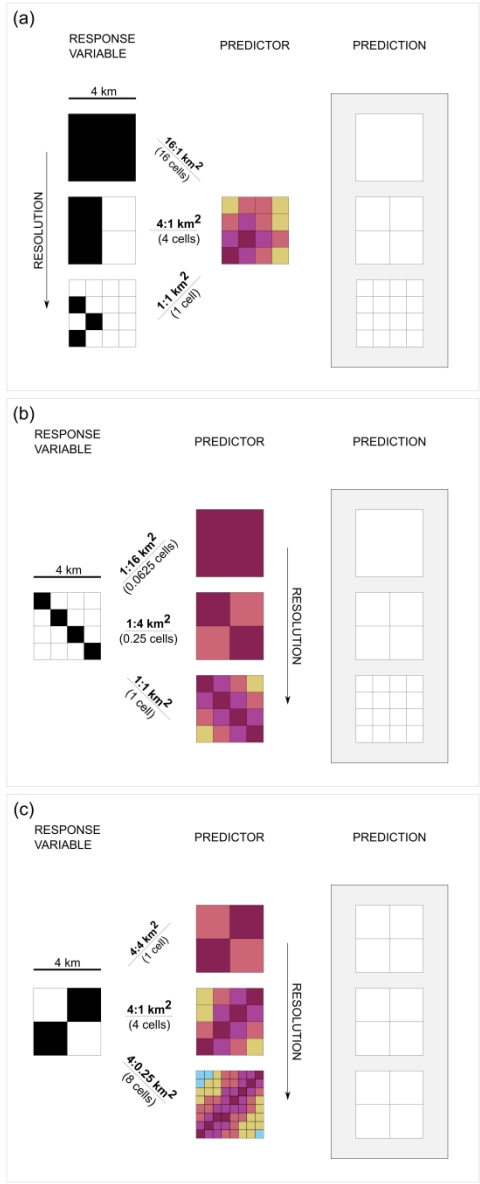
1016x2477mm (100 x 100 DPI)

Figure 1. Three approaches typically used for testing the grain dependence of species–environment relationships. (a) Manipulation of the resolution of the response variable with a fixed resolution of the predictors (b) Manipulation of the resolution of predictors, so that the resulting predictor was coarser than the response variable. (c) Manipulation of the resolution of predictors, so that the resulting predictor was finer than the response variable. The ratios shown in the figure (i.e. 16:1, 1:16, 4:4) express the resolution ratio that quantifies the magnitude of the difference between the resolution of the species data and that of environmental variables (this ratio quantifies how many cells of a particular predictor lie within a single cell of the response variable – the number shown in brackets). Note that if the fitted relationship is to be used for prediction, it is always limited by the coarsest grain used (either of predictor or response).

Table 1. Examples of studies focused on resolution of the response variable.

Study	Species	Resolution of the response	Predictors	Resolution of predictors
Tobalske (2002)	1 bird species	1 and 4 km^2	Landcover, elevation, edge density	25 m (50 m for elevation)
Heikkinen et al. (2007)	1 butterfly species	0.01, 0.25, and 1 km^2	Landcover, habitat connectivity	25 m
Kaliontzopoulou et al. (2008)	1 lizard genus	1 and 100 km^2	Climate, topography, land cover	from 90 m to 1 km depending on the variable
Seo et al. (2009)	9 plant species	from 1 km^2 up to 4096 km^2	Climate	1 km
Rengstorf et al. (2012)	Cold water corals	From 2500 m^2 up to 1 km^2	Bathymetry	50 m
Gábor et al. (2022a)	1 virtual species	25 m^2 to 0.25 km^2	Topography, canopy height	5 m
Zarzo-Arias et al. (2022)	1 mammal species	0.25, 1 and 25 km^2	Landcover, topography, human activity	100 m (25 m for topography)
Chauvier et al. (2022)	1180 plant species	From 0.01 km^2 up to 1600 km^2	Climate, soil	100 m

Table 2. Examples of studies that coarsened the resolution of predictors beyond the resolution of the response variable.

Study	Species	Resolution of the response	Predictors	Resolution of the predictors
Ferrier and Watson (1997)	56 species	20 x 50 m	climatic, topographic, soil, vegetation	200 m and 5 km
Graf et al. (2005)	1 bird species	0.01 km^2	topographic, climatic, land cover, human disturbance	1 ha up to ~11 km^2
Guisan et al. (2007)	50 species	0.01 - 1 km^2	climatic, topographic, soil	0.01 - 1 km^2 and 1 - 100 km^2
Pradervand et al. (2014)	239 Plant species	4 m^2	topographic, climatic	4 m^2 up to 0.01 km^2
Vale et al. (2014)	3 vertebrate species	Point occurrences	climatic, topographic, habitat	1 km^2 and 100 km^2
Manzoor et al. (2018)	1 plant species	Point occurrences	climatic, topographic, land cover	50 m, 300 m, and 1 km

Table 3. Examples of studies focused on the role of resolution of environmental predictors.

Study	Species	Resolution of response	Predictors	Resolution of predictors	Source of data
Thomas et al. (2002)	Vegeta-tion types	1 km²	Terrain, landform, rock/sediment composition	Field measured, 30 m	Various
Tobalske (2002)	1 bird species	1 km²	Landcover, elevation, edge density	25 m, 100m	Coarsening resolution of original data
Seoane et al. (2004)	54 bird species	350 m diameter around survey points	Area of shrubs or forests	30, 50, and 250 m	Various
Venier et al. (2004)	10 forest birds	5 km²	Climate, habitat	200, and 1000 m	Various
Gottschalk et al. (2009)	13 bird species	2 km diameter around survey points	Terrain, land-use	from 1 m up to 1 km	Coarsening resolution of original data
Šímová et al. (2019)	7 Water birds	12 x 11.2 km mapping quadrats	Area and perimeter of water bodies	from 30 m up to 1 km	Various
Connor et al. (2019)	1 mammal species	2 km² grid cells	Terrain, land cover, climate, phenology	From 30 m up to 2 km	Coarsening resolution of original data

Table 4. Resolutions adopted in studies evaluating the importance of fine–resolution predictors. The Resolution ratio quantifies how many cells of the predictor lie within a single cell of the response variable.

Study	Resolution of response	Resolution of predictors	Resolution Ratios
Seoane et al. (2004)	96,000 m^2 (350 m diameter around survey points)	900 m^2; 2500 m^2, and 62,500 m^2	96,000 : 900 m^2 (~ 107 cells) 96,000 : 2500 m^2 (~ 38 cells) 96,000 : 62,500 m^2 (~ 1.5 cells)
Venier et al. (2004)	5 km^2	0.04 km^2, and 1 km^2	5 : 0.04 km^2 (125 cells) 5 : 1 km^2 (5 cells)
Gottschalk et al. (2009)	3.14 km^2 (2 km diameter around survey points)	from 1 m^2 up to 1 km^2	3.14 : 1x10^{-6} km^2 (3,140,000 cells) 3.14 : 1x10^{-5} km^2 (314,000 cells) 3.14 : 1x10^{-4} km^2 (31,400 cells) 3.14 : 1x10^{-3} km^2 (3,140 cells) 3.14 : 0.01 km^2 (314 cells) 3.14 : 0.1 km^2 (31.4 cells) 3.14 : 1 km^2 (3.14 cells)
Šímová et al. (2019)	134 km^2 (12 x 11.2 km mapping quadrats)	from 900 m^2 up to 1 km^2	134 : 0.0009 km^2 (~ 150 000 cells) 134 : 0.01 km^2 (13,400 cells) 134 : 1 km^2 (134 cells)