How to predict biodiversity in space? An evaluation of modelling approaches in marine ecosystems

Chongliang Zhang1, Yong Chen2, Binduo Xu1, Ying Xue1, Yiping Ren1,3

1 College of Fisheries, Ocean University of China, Qingdao, China
2 School of Marine Sciences, University of Maine, Orono, ME, USA
3 Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

Abstract

Aim: Biodiversity prediction becomes increasingly important in the face of global diversity loss, whereas substantial challenges still exist in both conceptual and technical aspects. There exist many predictive models, and an integrative evaluation can help understand their performance in handling the multifacets of biodiversity. This study aims to evaluate the performance of these modelling approaches to predict both $\alpha$- and $\beta$-diversity in diverse ecological contexts.

Location: North Yellow Sea, China.

Methods: The biodiversity models follow three strategies, "assemble first, predict later", "predict first, assemble later" and "assemble and predict together". Hill diversity profile, Fisher’s log-series parameter and the distance decay of similarity are used to measure $\alpha$- and $\beta$-diversity. The evaluation study is conducted based on seasonal bottom trawl surveys from October 2016 to August 2017 in North Yellow Sea, China, allocated to coastal and offshore areas. We evaluate the predictive power of the models using cross-validation.

Results: Following the "assemble first, predict later" approach, macroecological model (MEM) provided the most accurate prediction overall, whereas stacked species distribution model (SSDM) and joint species distribution model (JSDM), following the second and third modelling approaches, tended to overestimate $\alpha$-diversity and underestimate $\beta$-diversity. The performances of SSDM and JSDM could be improved by moderately down-weighting rare species. The relative performances of the three modelling approaches were consistent among seasons and spatial regions.

Main conclusions: The superior performances of MEM in a range of temporal and spatial contexts favour the "assemble first, predict later" approach and imply a tight community assembly in the studied area. The overall predictive powers of varying models suggest that the spatial pattern of marine biodiversity could be fairly well predicted with commonly accessible hydrologic data in a mesoscales. The approach of multi-model evaluations is applicable to a variety of ecosystems for biodiversity prediction.

Keywords

GDM, Hill diversity profile, HMSC, Random Forest, SSDM, $\beta$-diversity
Biodiversity is a central topic of ecology, and there is a large body of literature demonstrating its critical role in ecosystem functions and services (Ibsell et al., 2018; Naeem, Chazdon, Duffy, Prager, & Worm, 2016). From a pragmatic viewpoint, predicting biodiversity is critical to identify management priorities and biological conservation (Certain & Planque, 2015; Coll et al., 2016; Keith, Martin, McDonald-Madden, & Walters, 2011; Lindenmayer, Margules, & Botkin, 2000), which is drawing increasing attention as a result of prevalent ecological threats from climate changes, marine pollution, overexploitation and habitat degradation worldwide (Agardy, 2000; Cheung et al., 2009; Defeo et al., 2017; Magurran, Dornelas, Moyes, Gotelli, & McGill, 2015; Urban et al., 2016). Although the topic is appealing, the challenge of biodiversity prediction is daunting. The concept of biodiversity is complicated, involving richness, abundance, evenness and variations, as well as components in the genetic, taxonomic, and functional levels (Jarzyna & Jetz, 2016; Ricotta, 2005). Although there are numerous measurements of diversity (Magurran & McGill, 2011), no single metric can depict the full scope of biodiversity (Santini et al., 2017). Moreover, influencing factors and driving forces for biodiversity are not always clear and observable when measurement errors and ecological stochasticity largely blur the relationships between biodiversity and predictors, leading to limited predictive powers in biodiversity models (Iknayan, Tingley, Furnas, & Beissinger, 2014; Jarzyna & Jetz, 2016). Therefore, increasing research efforts are needed for the methodological development as well as comparing, evaluating and applying biodiversity models.

Many models have been developed to address the challenges of biodiversity prediction in space. As reviewed by Ferrier and Guisan (2006) and D’Amen, Rahbek, Zimmermann, and Guisan (2017), existing spatial modelling methods can be classified into three strategies according to their conceptual and statistical bases. The first strategy is to “assemble first, predict later”, in which species data are first assembled into community-level attributes, such as species richness, then modelled with environmental predictors. The second strategy is to “predict first, assemble later”, in which each species is modelled individually then species-level predictions are stacked to community-level attributes. The third strategy is to “assemble and predict together”, in which all species are modelled simultaneously in an integrated model, and biodiversity is derived from species-level predictions accordingly. The performance of the three modelling strategies depends on specific ecological contexts, as ecological drivers of community structure change substantially in different spatial scales (Rahbek & Graves, 2001). For example, environmental filters often take effects in a large spatial scale by limiting the range of species distributions, whereas biotic interactions are more influential in a fine scale via predation and competition (Araújo & Rozenfeld, 2014; Wisz et al., 2013). However, there are also contrary evidences that microclimates may affect local communities and biotic interactions influence species distribution in large ranges (Araújo & Luoto, 2007; D’Amen et al., 2017). In addition, the effects of biotic interaction may depend on abiotic gradients, as environmental effects often dominate communities in stressful areas and species interactions are more influential in benign areas (i.e. the “species interactions–abiotic stress” hypothesis, Louthan, Doak, & Angert, 2015). As such, the driving forces of community assembly depend on the spatial scale of observations, suitability of environmental conditions, population density and species’ dispersal/migration ability. Therefore, it is unlikely that any one approach to biodiversity modelling and prediction will outperform other approaches in all situations (Ferrier & Guisan, 2006).

With respect to the challenges identified above, it is beneficial to incorporate the strengths of multiple modelling strategies to handle the multifacets of biodiversity. Although modelling methods of the three strategies have been well established, they are rarely compared and evaluated on the same stage. In addition, most studies have been focused on occurrence data and species richness, and biodiversity models are rarely examined based on abundance data (Algar, Kharouba, Young, & Kerr, 2009; D’Amen et al., 2017; Santini et al., 2017). The present study evaluates the performances of various existing modelling approaches to predict the spatial patterns of α- and β-diversity based on abundance data. With respect to the multifacets of biodiversity, multiple metrics are adopted to reflect both α- and β-diversity (Jarzyna & Jetz, 2016), and their spatial distributions serve as the targets of biodiversity prediction. We evaluate the predictive power of different modelling approaches using cross-validation and compare their performances in different ecological contexts. The study is implemented in a marine ecosystem in Yellow Sea, China based on seasonal surveys, whereas the evaluation approaches are applicable to a variety of ecosystems. The objective of this study is to improve our understanding of community assemblies and provide guidance for the improvement of biodiversity predictions.

2 | METHODS

2.1 | Study area and data

The study was based on a marine survey in North Yellow Sea (NYS), China (Figure S1). The surveys were conducted seasonally from October 2016 to August 2017. A total of 118 sampling sites were investigated in each of the four cruises using otter trawls. The survey area was stratified into three sub-regions, south coastal area A, north coastal area B and offshore area, in which 40, 43, and 35 stations were sampled, respectively (Figure 1). The trawl had a net width 15 m and cod-end mesh 20 mm, hauling for 1 hr at a speed 3 knots in each sampling site. Environmental variables including temperature, salinity and depth were measured using CTD (XR-420) at each sampling station after hauling.

We included 105 nekton species in our analyses after excluding occasional catches (rare fish or benthos caught with a few individuals), which might cause the instability of model fitting. It should be noted that removing those rare species could lead to underestimation of species richness; however, the influence was limited for
abundance-based diversity metrics when evaluating models’ relative performances. The abundance of each species that was standardized to the same sampling efforts were used as response variables directly or aggregated to diversity metrics before modelling. The water depth, temperature and salinity measured at the sea bottom at each sampling site were used as the predictive variables for spatial biodiversity modelling (environmental data were summarized in Table S1 and Figure S2).

2.2 Biodiversity metrics

Instead of the contemporary focuses on occurrence and species richness, this study used abundance (number of individuals) as the currency to measure biodiversity (Algar et al., 2009; D’Amen et al., 2017; Santini et al., 2017). A range of diversity metrics was adopted in this study to reflect both α- and β-diversity. Hill’s number was used to depict α-diversity synthetically (Hill, 1973),

\[
H_r = \left( \sum_i p_i^r \right)^{1/(1-r)}
\]

where \(p_i\) is the proportion of species \(i\) in the sample, and \(r\) denotes the order of diversity. The value of \(r\) ranges from zero to infinity, and an increasing \(r\) reduces the weight of rare species (Figure S4). By changing \(r\), this equation yields most well-known diversity indices, for example, \(H_0\) is simply species richness (when \(r = 0\), \(H_0\) equals the exponential of Shannon diversity index, and \(H_2\) is the inverse Simpson index. Exhibiting different \(r\) with the corresponding \(H_r\) in a plot leads to Hill diversity profile, which provides more information than any individual diversity index. In this study, the order \(r\) (referred to as level in this study) ranges from 0 to 4 in the diversity profile (Certain & Planque, 2015).

Species abundance distribution (SAD) is an alternative approach to reflect α-diversity with the preservation of more information compared to ecological indices (McGill et al., 2007). The SADs describes the relative abundance (commonness and rarity) of different species, often exhibiting a hyperbolic curve (i.e. few abundant species and many rare species, Figure S5). A range of statistical models have been used to depict the shape of SADs, including log-series, Gamma, lognormal, Weibull, Broken stick and zero-sum multinomial distribution (Magurran & McGill, 2011). We adopted Fisher’s log-series model (Fisher, Corbet, & Williams, 1943) to fit abundance data and used the log-series parameter as a diversity metric (referred to as \(f\) below).

β-diversity is often defined as the variation of species composition in a community over space or time (Whittaker, 1960). Among a wide range of optional metrics (Tuomisto, 2010), this study measures β-diversity as the dissimilarity of species composition among sampling units in different geographical distances, that is, distance decay of similarity (Barwell, Isaac, & Kunin, 2015). Two popular metrics, Bray–Curtis index (BC) and Cao’s similarity index (CY) were used to indicate the dissimilarities in species abundance (Figure S6). The former is a traditional measure of similarity in ecology (Bray & Curtis, 1957),

\[
BC_f = \frac{\sum_i |x_i - x_k|}{\sum_i (x_i + x_k)}
\]

The latter index, developed by Cao, Williams, and Bark (1997), shows satisfactory performances in measuring β-diversity (Cao & Epifanio, 2010; Barwell et al., 2015).

\[
CY_f = \frac{1}{5} \sum_{i=1}^{n} \left[ \log \left( \frac{x_i + x_k}{2} \right) - \frac{x_i \log x_i + x_k \log x_k}{x_i + x_k} \right]
\]
In both equations, $x_{ik}$ is the abundance of species $k$ in sample $i$, and $S$ is the total number of species. The Bray–Curtis and Cao’s index were used to illustrate the distance decay of similarity. The Hill diversity profile, Fisher’s log-series model and similarity metrics were calculated using the "vegan" R package.

2.3 Modelling approaches

We followed the paradigm of Ferrier and Guisan (2006) to adopt three conceptually different modelling approaches. In the first “assemble first, predict later” approach, biodiversity metrics were modelled directly with environmental variables. This approach was also referred to as MEM (Algar et al., 2009; Guisan & Rahbek, 2011; Certain, Dormann, & Planque, 2014). We adopted two flexible statistical techniques to tackle possible nonlinear relationships between biodiversity and environmental predictors.

The first MEM considered is Random Forest (RF) model for predicting $\alpha$-diversity. The RF is a machine-learning technique based on the ensemble of a large number of bootstrapped classification and regression trees (Breiman, 2001), in which each tree is structured by splitting nodes with optimal variables from a random subset of predictors. The prediction of RF is derived as an average over all trees in the forest. The number of variables selected at each node (mtry) and the number of trees in a forest (ntree) needs to be specified in the algorithm (Degenhardt, Seifert, & Szymczak, 2017). Random Forest model is evident to be insensitive to underlying error distributions, collinearity and interactions among predictors (Di Franco et al., 2016; Duffy, Lefcheck, Stuart-Smith, Navarrete, & Edgar, 2016). We used spatial coordinates in addition to environmental variables as predictors to facilitate spatial predictions. The second MEM considered is Generalized Dissimilarity Model (GDM) for predicting $\beta$-diversity. The GDM uses the distance in environmental space to explain biological variations (Ferrier, Manion, Elith, & Richardson, 2007). That is, the response variable in GDM is a matrix of compositional dissimilarities (e.g. Bray–Curtis similarity index), and the explanatory variables are environmental predictors transformed to distance matrices among every pairs of sampling sites (Ferrier et al., 2007; Lasram, Hattab, Halouani, Romdhane, & Le Lôc’h, 2015). This algorithm is technically an extension of permutation matrix regression that uses splines to accommodate nonlinearity relationships. The two algorithms were implemented using the "randomForest" (version 4.6 on CRAN) and "GDM" package (version 1.3.7 on GitHub, https://github.com/fitzLab-AL/GDM), respectively.

In the second approach, $\alpha$- and $\beta$-diversity are derived from species-level predictions, for which the abundance of individual species is modelled with environmental predictors using species distribution models, such as GLM, GAM and Maxent (Elith & Leathwick, 2009). This approach is referred to as stacked species distribution models (SSDM). Generalized additive model (GAM) was used to build SSDM in this study regarding its flexibility in tackling irregular data and nonlinear ecological relationships. Generalized additive model is one of the most widely used methods in species distribution modelling (Hastie & Tibshirani, 1990; Guisan, Edwards, & Hastie, 2002), using regression techniques, semi-parametric smoothing functions and various types of error distributions to handle complex species-environment relationships. In the present study, a two-step GAM was adopted to account for the many zeros in multispecies surveys which caused failures in many modelling techniques, by modelling absence-presence data and positive catches separately (O’Neill & Faddy, 2003; Zhang & Chen, 2015). A alternative Tweedie-GAM was used instead when the two steps failed for rare (unfeasible in the second step) or prevalent species (unfeasible in the first step) (Candy, 2004). Spatial coordinates (latitude and longitude) were also included in modelling to facilitate spatial predictions according to preliminary analyses of model predictability (Algar et al., 2009). The algorithm was implemented using the "mgcv" package in R.

In the third modelling approach, we adopt a newly developed "joint species distribution model" (JSDM) to predict the distribution of all species and their associations altogether (Warton et al., 2015). The JSDMs commonly relate multivariate response variables to abiotic predictors using regression methods (but see neural network in Harris, 2015) while accounting for the residual biotic interactions using random effects (Pollock et al., 2014; Hui, 2016; Clark, Nemergut, Seyednasrollah, Turner, & Zhang, 2017; Ovaskainen et al., 2017; Thorson & Barnett, 2017; Schliep et al., 2018). The random effects are specified by either unstructured variance-covariance matrices or unobserved "latent variables" (Skonradal & Rabe-Hesketh, 2004). This study followed the latent-variable approach and adopted a JSDM framework "Hierarchical Modelling of Species Communities" (HMSC) (Ovaskainen et al., 2017). Hierarchical Modelling of Species Communities uses a Bayesian hierarchical method to incorporate environmental variables, species traits, phylogeny, spatial and temporal structures in community modelling (Ovaskainen & Solininen, 2011). The model uses sparse infinite factor priors for the loadings of latent variables to account for uncertainty in the number of latent factors and sparsity structure (Bhattacharya & Dunson, 2011). We assumed spatial autocorrelation in the structure of latent variables to facilitate spatial prediction (Ovaskainen et al., 2017). The algorithm was implemented using the "HMSC" package (version 2.1.1 on GitHub, https://github.com/guiblanche/HMSC).

In summary, the abundance of multiple species were used as the response variables in SSDM and JSDM, and the derived diversity metrics from abundance data were used as the response variables in MEM. The same predictive variables were used for different modelling approaches. In addition to hydrologic variables, the latitude and longitude of each sampling site were used to be predictive variables in MEM and SSDM and to formulate spatial autocorrelation in JSDM.

2.4 Performance evaluation

We used threefold cross-validations to examine the predictability of the three modelling approaches. Two thirds of the sampling sites were randomly drawn from the survey data as the training set for
model fitting, and the remaining 1/3 of the data were used as the testing set to evaluate model predictions. Two methods were used to evaluate the accuracy of the predictions on α- and β-diversity. A linear regression analysis was conducted between observations and model outputs, and the regression slope and the determination coefficient were used to describe model accuracy. Normalized root mean square error (NRMSE), calculated as below, was also used to measure the divergence of prediction from observation.

\[
\text{NRMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2 / (x_{\text{max}} - x_{\text{min}})}
\]

where \(\hat{x}_i\) is the predicted value in the \(i\)th sampling site, and \(x_i\) is the observed value in the same site. \(x_{\text{max}}\) and \(x_{\text{min}}\) denote the maximum and minimum values in the observation data. Normalized root mean square error ranges from 0 to 1 with a lower value indicating better predictions.

The processes of data splitting, model fitting and accuracy examination were repeated 1000 times in cross-validation. The three modelling approaches were implemented and examined independently, from which we could compare and identify the optimal strategy for biodiversity prediction. Regarding the context-dependent performances of different models (Ferrier & Guisan, 2006), the cross-validation was conducted for varying survey seasons and regions, respectively. The intensive evaluations aimed to tackle the seasonality of NYS ecosystem driven by monsoon climate. Besides, fishing activities in this region, regulated by an annual moratorium policy in summer and adverse climate in winter, might also contribute to the substantial changes in community structure among seasons (Jiang, Cheng, & Li, 2009). In addition, the biotic community showed significant differences between coastal and offshore waters (unpublished data from our survey), which should be accounted for in model evaluation. Therefore, we treated the four seasons and four regions (whole area, coastal area A, B and offshore, Figure 1) as different ecological contexts for model evaluation, and the model fitting and evaluation analyses were developed for a combination of seasons and regions separately, resulting in a total of 16,000 sets of evaluation results (4 seasons × 4 regions × 1000 cross-validation iterations). The overall design of model fitting and evaluation was illustrated in a flow chart (Figure S3).

### 3 | RESULTS

#### 3.1 | α-diversity

The three modelling approaches (MEM, SSDM, JSDM) showed different accuracies in the predictions of Fisher’s log-series parameter \(f\). According to the linear regression between prediction and observation over the study area (examples shown in Figure S7), MEM showed the highest predictive accuracy, with the regression slope close to 1.0 and the coefficient of determination (\(R^2\)) over 0.9, followed by SSDM (Figure S10, using data in autumn as an example). Fisher’s \(f\) tended to be underestimated by MEM and largely overestimated by SSDM and JSDM. The NRMSE also showed accurate predictions by MEM, followed by SSDM. The same analyses were conducted across the four seasons and the results were generally consistent, that is, MEM showed superior performances over SSDM and JSDM with substantially high precision. For the three modelling approaches, NRMSE ranged from 0.2 to 0.6 (Figure 2, based on data of the whole area) and \(R^2\) ranged from 0.8 to 0.95 (Figures S11, S12), indicating an overall satisfactory predictability, although there were remarkable variations of prediction among models and seasons. Generally, Fisher’s \(f\) was better predicted in spring and fall compared to that in summer and winter, especially for SSDM.

Hill diversity profile provided additional information on the variation of biodiversity predictions. Macroecological model showed

![FIGURE 2](https://example.com/f2.png) The predictive accuracy of Fisher’s \(f\) by three modelling approaches in the four seasons. The MEM, SSDM and JSDM denote macroecological model, stacked species distribution model, and joint species distribution model, respectively. The NRMSE denotes the normalized root mean square error among observation and prediction values.
constantly better performances among the three modelling approaches with respect to regression Slope, $R^2$ and NRMSE (Figures S8, S13 and S14), but underestimated Hill’s number in all levels. Stacked species distribution model and JSDM tended to overestimate Hill’s number in low levels whereas the bias decreased with the increasing levels, minimized at $H_1$ (Figure 3). All modelling approaches tended to underestimate Hill’s number when the level $r$ was large. Different from the results of Fisher’s $f$, JSDM tended to outperform SSDM in some situations (e.g. summer in Figure 3).

### 3.2 | $\beta$-diversity

Both the observations and predictions of $\beta$-diversity showed remarkable distance decay of similarity in community compositions. There was a constant trend in Bray–Curtis indices that increased with the geographical distances among sampling sites, ranging from 0.60 to 0.95, although the similarity varied substantially in similar distances (Figure S9). The prediction of MEM closely matched the observed patterns of distance decay. Stacked species distribution model and JSDM predicted less dissimilarity in small distances but proper estimation in large distances (Figure 4). The relative performances of the three modelling approaches were consistent among the four seasons. The similarity decay was nonlinear across the geographical range of this study, and the rate of decay generally changed at a geographical distance of 1.5–2.0 units of the scaled spatial coordinates (divided by standard deviations). Cao’s similarity index showed a similar pattern to BC in distance decay but the divergences were more remarkable among the three modelling approaches (Figure S15).

**FIGURE 3** The predictive accuracy of Hill diversity profile by three modelling approaches in four seasons. The levels in X-axis indicates the order of Hill’s number. MEM, SSDM and JSDM denote macroecological model, stacked species distribution model and joint species distribution model, respectively. Slope is the regression coefficient between the observation and prediction of Hill’s number.
3.3 | Regional evaluation

Model performances in predicting $\alpha$- and $\beta$-diversity were evaluated in sub-regions of the study area, including two coastal areas (A and B) and the offshore area. The superior performance of MEM was consistent with the results in the whole area, whereas the three modelling approaches showed larger divergences in their performances (Figure 5). In the south coastal area A, MEM showed increasing NRMSE with the increasing levels of Hill’s number, whereas NRMSE firstly decreased then stabilized in SSDM and JSDM. In the north coastal area B, NRMSE was relatively stable in MEM and tended to decease in the other models except the hump of JSDM at the level of 0.4. The pattern of Hill profile in offshore area was similar to area A, and SSDM and JSDM yielded close predictions.

The MEM also provided better predictions of $\beta$-diversity in the subareas, whereas SSDM and JSDM substantially underestimated $\beta$-diversity in short distances (Figure 5). The community dissimilarity ranged from 0.5 to 0.8 in coastal areas, substantially less than that of the offshore area. There was a large change of similarity decay rates at the geographical distance of 1.6 units (spatial coordinates scaled by standard deviations) in the offshore area, captured by all three models. Such a pattern was obscure in the coastal areas.

4 | DISCUSSION

Predicting the changes of biodiversity has received growing attention given increasing concerns about habitat and climate changes. As a step towards an integrative research framework, this study evaluated a range of modelling approaches to predict the spatial pattern of $\alpha$- and $\beta$-diversity based on abundance data. Particularly, abundance-based diversity and $\beta$-diversity should be highlighted regarding the requirement of multi-functionality and ongoing biotic homogenization of marine ecosystems (Olden, 2006; Magurran et al., 2015) and the prevalent research focuses on species richness (Mori, Isbell, & Seidl, 2018). In general, our study showed that the three modelling approaches evaluated in this study had fair predictive power on biodiversity, implying that marine biodiversity could be reasonably predicted with commonly accessible hydrologic data (i.e. temperature, salinity and depth), although the driving forces behind biodiversity are complex. Additionally, MEM provided most accurate predictions in a range of temporal and spatial contexts, whereas SSDM and JSDM tended to overestimate $\alpha$-diversity and underestimate $\beta$-diversity in certain circumstances. Our results favoured the “assemble first, predict later” modelling approach. The study was conducted in the specific areas of NYS, whereas the consistent results among seasons and regions suggested the modelling approaches might be applicable to diverse ecological contexts for spatial prediction of biodiversity.

Although the results were encouraging, it should be noted that the cross-validation of this study was optimistic, as the predictions were interpolated into the same area of model fitting where underlying ecological relationships tended to be consistent. In other words, the modelling methods were based on correlational relationships which might be less reliable when model fitting and prediction were conducted in different areas and/or temporal scales (Petchey et al.,...
such as the case of biodiversity in climate changes. In fact, biodiversity is an emergent result of community assembly, driven by diverse processes including biogeographic history, evolutionary adaptation, environmental filtering, biotic interaction and stochastic dispersal (Thuiller et al., 2013; Ovaskainen, Abrego, Halme, & Dunson, 2016a; D’Amen et al., 2017). Therefore, the transferability of statistical models is worthy of future investigation. We suggest that correlative models may be preferable for spatial prediction of biodiversity in the scope of marine reserves (Parnell, Dayton, Lennert-cody, Rasmussen, & Leichter, 2013), and mechanistic models are necessary to tackle ecosystem shifts under environmental changes (Singer et al., 2016; Urban et al., 2016; Cabral, Valente, & Hartig, 2017).

Comparing performances of multiple modelling approaches contributed to our understanding of their strengths and weaknesses, which could serve as a benchmark for model improvement (Ferrier & Guisan, 2006; Certain et al., 2014). For example, the superior performance of MEM might be attributed to the aggregation of species-level information, that is, modelling collective properties at the community-level avoided the challenges of handling individual species especially rare species, in addition to the uncertainty in deriving biodiversity metrics from species-level predictions. Besides, this modelling approach might also benefit from the RF and GDM algorithms, which were evidenced to be powerful for prediction among many statistical techniques (Cutler et al., 2007; Boulesteix, Janitza, Kruppa, & König, 2012). On the other hand, MEM relied on elaborate community-level inventories along environmental gradients, which limited it to predict beyond known communities (Ferrier & Guisan, 2006). As such, SSDM and JSDM were more flexible for biodiversity extrapolation, when ecological niches of individual species were reliable in predictions (Schmitt, Pouteau, Justeau, Boissieu, & Birnbaum, 2017). It is worthy to note that their predictions tended to be biased without considering macroecological constraints (D’Amen et al., 2018), for example, overestimating α-diversity (Guisan & Rahbek, 2011; Schmitt et al., 2017). Our results suggested that the performances of SSDM and JSDM were likely undermined by rare species, as both models showed improved performances with increasing levels in Hill diversity profile which reduced weights on rare species. Besides, rare species also prevented the use of powerful algorithms such as RF and ANN in SSDM, for which only the GAM was flexible to avoid fallacious predictions. An extensive survey that
moderately covered rare species might benefit the predictions of SSDM and JSDM.

Hierarchical Modelling of Species Communities, the algorithm used in the third strategy, is versatile in handling environmental factors and species interactions in species distribution modelling (Ovaskainen et al., 2017). In particular, the inclusion of spatial structure of latent variables might greatly improve models' predictive performances (Ovaskainen, Roy, Fox, & Anderson, 2016b; Ovaskainen et al., 2017). However, the spatial-structured HMSC in this study showed no better performances than MEM and SSDM in most situations. The results suggested that environmental gradients dominated the spatial pattern of biodiversity in the study area while biotic interactions played a minor role (also evidenced in variation partition of HMSC). In addition, from a technical viewpoint, HMSC is essentially a multispecies generalized linear mixed model. The algorithm makes it convenient to model all species simultaneously but constrains the capability to deal with nonlinearity, massive zeros and interactions (Zhang, Chen, Xu, Xue, & Ren, 2018), which can be better handled in GAM and RF. Nevertheless, the model's less performance can still help understand the patterns of biodiversity, as the model provides evidences of limited biotic interactions which could not be assured a priori. In addition, the flexible Bayesian hierarchical structure (Ovaskainen & Soininen, 2011) may facilitate the improvement of HMSC in diverse ecological contexts (c.f. Harris, 2015; Thorson & Barnett, 2017).

The choice of diversity metrics and available data are critical for the predictive power of biodiversity models (Certain et al., 2014; Riera et al., 2017). For example, Certain et al. (2014) suggested that high predictive performances could be obtained by modelling with species biomass and using diversity measures that give high weight to rare species. The conclusion was consistent with the results of MEM in our study but opposite to that of SSDM and JSDM, which were improved with increasing r in the Hill's diversity profile (down-weighting rare species) to certain extents. Our evaluations are conducted in different marine areas (coastal and offshore), and the consistent results among regions suggest that the superior performance of MEM may be applicable to diverse ecological contexts. Nevertheless, we evaluated performances of the modelling approaches based on collective biodiversity properties in this study, and the predictive ability could be quite different in terms of individual species distributions. In addition, the type of available data needs to be considered for model development regarding different spatial scales. For example, the "assemble first, predict later" approach relies on local inventories of biotic communities that are often obtained at a fine spatial resolution, whereas the "predict first, assemble later" approach may be based on species-occurrence records which are collected individually at coarse spatial resolutions. Moreover, the relative performances of different models may rely on spatial scales. Given the impact of training data on prediction success, future evaluation studies of biodiversity models should account for the spatial characteristics of survey data, including spatial resolution and extent, survey design, spatial autocorrelation, species prevalence, and sample size as well as their interactive effects (Stockwell & Peterson, 2002; Edwards, Cutler, Zimmermann, Geiser, & Moisen, 2006; Tessarolo, Rangel, Araújo, & Hortal, 2014).

In future studies, the modelling approaches evaluated in this study can be improved by tuning available model algorithms, while some metrics may be considered with priority in conservation with respect to their predictability (e.g., H), as policymakers often want a single number. Meanwhile, additional studies are needed to improve the prediction of biodiversity regarding diverse research objectives, ecological contexts and quality of data, covering the facets of taxonomic, functional and phylogenetic components of α- and β-diversity (Mokany, Harwood, Overton, Barker, & Ferrier, 2011; Riera et al., 2018). We suggest that the comparison of multiple predictive models and multiple diversity metrics can be useful to provide additional information of biodiversity. However, facing the challenges of changes in biodiversity over broad environmental gradients and climate changes, structural integrations of different modelling modules are desired to combine both mechanistic and correlative approaches (Cabral et al., 2017; D’Amen et al., 2017), for example SESAM (Guisan & Rahbek, 2011), Dynamic FOAM (Mokany et al., 2011) and DBEM (Fernandes et al., 2013). Meanwhile, we emphasize that fully exploring the existing models may be as important as, if no more than, developing new ones. In any case, biodiversity is more than just species number or abundance but should be embedded in variations, interactions and dynamics, which calls for the development of comprehensive research frameworks.

ACKNOWLEDGEMENTS

Funding for this study was provided by the National Natural Science Foundation of China (31802301, 31772852).

DATA AVAILABILITY STATEMENT

R code and further data supporting the results are provided on Dryad: https://doi.org/10.5061/dryad.2f7mt8

ORCID

Chongliang Zhang https://orcid.org/0000-0002-4753-9973

REFERENCES

Agardy, T. (2000). Effects of fisheries on marine ecosystems: A conservationist's perspective. ICES Journal of Marine Science, 57, 761-765. https://doi.org/10.1006/jmsc.2000.0721

Algar, A. C., Kharouba, H. M., Young, E. R., & Kerr, J. T. (2009). Predicting the future of species diversity: Macroecological theory, climate change, and direct tests of alternative forecasting methods. Ecography, 32, 22-33. https://doi.org/10.1111/j.1600-0587.2009.05832.x

Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography, 16, 743-753. https://doi.org/10.1111/j.1466-8238.2007.00359.x

Agardy, T. (2000). Effects of fisheries on marine ecosystems: A conservationist's perspective. ICES Journal of Marine Science, 57, 761-765. https://doi.org/10.1006/jmsc.2000.0721

Algar, A. C., Kharouba, H. M., Young, E. R., & Kerr, J. T. (2009). Predicting the future of species diversity: Macroecological theory, climate change, and direct tests of alternative forecasting methods. Ecography, 32, 22-33. https://doi.org/10.1111/j.1600-0587.2009.05832.x

Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography, 16, 743-753. https://doi.org/10.1111/j.1466-8238.2007.00359.x
Tuomisto, H. (2010). A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecology, 33, 23–45. https://doi.org/10.1111/j.1600-0587.2009.06148.x

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Peer, G., Singer, A., ... Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. Science, 353, aad8466. https://doi.org/10.1126/science.aad8466

Warton, D. I., Blanchet, F. G., O’Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. C. (2015). So many variables: Joint modeling in community ecology. Trends in Ecology and Evolution, 30, 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs, 30, 280–338. https://doi.org/10.2307/1943563

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., ... Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. Biological Reviews, 88, 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x

Zhang, C., & Chen, Y. (2015). Development of abundance indices for Atlantic Cod and Cusk in the Coastal Gulf of Maine from their Bycatch in the Lobster Fishery. North American Journal of Fisheries Management, 35, 708–719. https://doi.org/10.1080/02755947.2015.1043413

Zhang, C., Chen, Y., Xu, B., Xue, Y., & Ren, Y. (2018). Comparing the prediction of joint species distribution models with respect to characteristics of sampling data. Ecology, 41, 1876–1887. https://doi.org/10.1111/ecog.03571

BIOSKETCH

Chongliang Zhang is a marine fisheries scientist. His research interests are ecosystem modelling to support marine and fisheries management.

Author contributions: C.Z. conceived the ideas; C.Z. and Y.C. led the analysis and writing, assisted by B.X., Y.X. and Y.R. All authors contributed to the interpretation of analyses.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zhang C, Chen Y, Xu B, Xue Y, Ren Y. How to predict biodiversity in space? An evaluation of modelling approaches in marine ecosystems. Divers Distrib. 2019;25:1697–1708. https://doi.org/10.1111/ddi.12970