frontiers in Ecology and Evolution

TYPE Original Research
PUBLISHED 09 March 2026
DOI 10.3389/fevo.2026.1684304

OPEN ACCESS

EDITED BY
Nicholas A. Som,
California Polytechnic University,
United States

REVIEWED BY
Vinicius Albani,
Federal University of Rio de Janeiro, Brazil
Nadav Shashar,
Ben-Gurion University of the Negev, Israel

*CORRESPONDENCE
Yanli Tang
tangyanli@ouc.edu.cn
Tao Zhang
tzhang@qdio.ac.cn

†PRESENT ADDRESS
Haolin Yu,
MAST-FOCUS, Astrophysics, Geophysics and Oceanography Department,
University of Liege, Liege, Belgium

RECEIVED 12 August 2025
REVISED 02 February 2026
ACCEPTED 16 February 2026
PUBLISHED 09 March 2026

CITATION
Yu H, Rose KA, Fang G, Tang Y, Becker A, Feng J and Zhang T (2026) Simulation analysis of the ecological performance of artificial reefs using species distribution models. *Front. Ecol. Evol.* 14:1684304. doi: 10.3389/fevo.2026.1684304

Simulation analysis of the ecological performance of artificial reefs using species distribution models

Haolin Yu[1,2†], Kenneth A. Rose[3], Guangjie Fang[4], Yanli Tang[5*], Alistair Becker[6], Jie Feng[7] and Tao Zhang[1,2*]

[1]Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2]Laboratory for Marine Ecology and Environmental Science, Qingdao Marine Science and Technology Center, Qingdao, China, [3]Horn Point Laboratory, University of Maryland Center for Environmental Science, Cambridge, MD, United States, [4]Fishery Resources and Ecology Research Department, Zhejiang Marine Fisheries Research Institute, Zhoushan, China, [5]College of Fisheries, Ocean University of China, Qingdao, China, [6]New South Wales Department of Primary Industries and Regional Development, Port Stephens Fisheries Institute, Taylors Beach, NSW, Australia, [7]North China Sea Marine Forecasting and Hazard Mitigation Center, Ministry of Natural Resources, Qingdao, China

Introduction: Artificial reefs (ARs) have been widely used to provide ecological and fisheries benefits. Quantitative evaluation of how the numbers and spatial layout of ARs affect fish responses can provide critical insights for management.
Methods: We used survey data from an area in the Bohai Strait (China) that has ARs deployed and developed species distribution models (SDMs) predicting the occurrence probabilities of three commercially important species (*Charybdis japonica*, *Sebastes schlegelii*, and *Hexagrammos otakii*) and one undesirable focal species (*Asterias amurensis* and *Asterina pectinifera* combined). Three versions of SDMs were developed: a Best-fit model based on survey data, and two modified versions (Intermediate and Extreme), that incorporated increasing levels of AR-to-AR interactions, reflecting competitive and connectivity effects. A 24×24 grid with 50-m cells was populated with bottom habitat type (mud, gravel, rubble, boulder, or AR) and key environmental variables affecting occurrence probabilities. Using simulations of 10,000 randomly-generated spatial layouts of added ARs (bottom type switched to AR), we compared predicted occurrence probabilities from the Best-fit, Intermediate, and Extreme SDMs when 5, 10, 25, and 50 ARs were added to the existing ARs. Performance was evaluated using the predicted species-specific occurrence probabilities, with occurrence of the undesirable species treated inversely (i.e., lower occurrence probability indicated higher performance).
Results: Increasing AR numbers increased the percentage of grid cells supporting good habitat, but saturation and interference effects caused similar values for 25-50 ARs for several species, while reducing variation across spatial layouts. The three desirable species showed similar patterns with the representative layouts categorized into good and bad performing: increasing spread of good ARs on the grid with number of ARs deployed, shift of good ARs from upper to lower triangle, and an increasingly bad-performing central area with 50 ARs. The spatial overlap

between high-performing cells for desirable species and elevated occurrence of undesirable Sea star illustrated an inherent trade-off that should be considered in management objectives.

Discussion: We conclude with a discussion of the need for AR layout-specific evaluation, consideration of AR interaction strength, broader applicability to other systems, potential pathways to expand the analysis (e.g., add hydrodynamic models), and caveats and recommendations for further development.

KEYWORDS

artificial reefs, layout sensitivity, reef interaction, simulation analysis, species distribution model

1 Introduction

Artificial reefs (ARs) have a long history of enhancing commercial and recreational fisheries through their provision of complex habitats in regions predominantly consisting of bare bottom substrate (Baynes and Szmant, 1989; Charbonnel et al., 2002; Lima et al., 2019; Strelcheck et al., 2005). The ecological functioning of ARs is complex; for example, in addition to providing fish shelter from their predators, AR modules also potentially modify nutrient transport by altering current flow and entraining nutrients and thereby affecting local biological production available to fish through trophic links (Castège et al., 2016). Quantifying the benefits that ARs provide to local fisheries depends on module design of each AR and the appropriate spatial arrangement of the modules (Becker et al., 2019; Lan and Hsui, 2006).

Several studies have examined the influence of AR placement on the effectiveness of a reef to enhance finfish populations (Grove et al., 1991; Nakamura, 1985) and epibenthic organisms (Ushiama et al., 2016). However, they did not quantify these effects and did not directly compare ecological performance offered by systematic variation of the numbers of ARs and their spatial arrangements. The ability of ARs to enhance local productivity depends on their connectivity with neighboring ARs as well their proximity to nearby natural reefs (Campbell et al., 2011; Logan and Lowe, 2018). Therefore, the positioning and layout of the ARs is critical in determining their effectiveness (Becker et al., 2019). Many researchers that examined ARs showed that benthic-oriented fish with affinities for complex habitat typically show residency on ARs but also travel to and from the surrounding areas, including other ARs and natural reefs (Keller et al., 2017; Logan and Lowe, 2018; Puckeridge et al., 2021). Natural reefs often consist of boulder habitat and is a common coastal hardbottom habitat substrate (Kuklinski et al., 2006) that often supports high fish biomass (Kilfoyle et al., 2013). Many species also show a range of affinities for various bottom habitat types such as cobble and gravel. For example, using acoustic telemetry, Zhang et al. (2015) showed that the rockfish (*Sebastes schlegelii*) regularly utilized these habitats as well as ARs, highlighting the potentially complex connectivity that exists among habitat types and how spatial placement can influence their effectiveness.

We suggest further analyses are needed to better understand the influence of AR placement on the response of fisheries-related species and their associated productivity. Such information complements the ecological studies that compare more aggregate measures such as community structure and biodiversity between artificial and natural reefs (Kilfoyle et al., 2013; Lowry et al., 2014; Vicente et al., 2008). Determining the effects of AR layout on fish distributions and productivity in order to optimize the spatial deployment of multiple ARs is challenging. Different hydrodynamic approaches have been used to assess the impacts of AR layouts on surrounding flow fields to identify the suitable intervals and arrangements of AR monomers (Liu and Su, 2013; Tang et al., 2019; Vijay et al., 2021). Many studies have assessed the relationship between environmental factors and species present/absence, density, or community composition in areas with ARs (Chittaro et al., 2005). However, our ability to quantitatively link ecological responses to specific positioning of the ARs, often in places that never previously had ARs, is still uncertain. Quantification relies on being able to predict marine fish and crustacean species response on small spatial scales, enabling the capture of the local responses to habitat changes such as the introduction of ARs (Liu and Su, 2013; Parsons et al., 2004; Wilborn et al., 2018). Modeling tools exploring altered distributions of focal species in response to AR placement can provide important information to guide management. Pilot studies with a few ARs deployed provide an approach for model specification and estimation that then is usable in the design of larger scale AR deployments; such analyses can help ensure that deployments achieve target responses and ecological goals related to fish productivity (Becker et al., 2017).

One approach to model development is to select a set of representative focal species, and then using species distribution models (SDMs) estimated from local data and computer simulation methods, compare predicted habitat and species occurrence probabilities among various AR spatial layouts. SDMs construct relationships between species' current distributions and associated environmental factors, enabling prediction of species distributions under new environmental conditions (Velásquez-Tibatá and Graham, 2013). These environmental conditions include the introduction of ARs in different spatial layouts. There is some criticism that SDMs are overly simplistic (Araújo and Luoto, 2010) and that they do not consider long-term population viability (Keith et al., 2008). SDMs, when estimated and used appropriately, are well-suited for projecting certain types of responses, such as potential range shifts of species (10–100 km) and potential

utilization of natural and introduced habitat (0.01–1 km) under changing environmental conditions (Cannizzo et al., 2020). SDMs also have the advantage of well-developed methods for statistical estimation using a variety of types of monitoring and survey data (Renwick and Johnston, 2012; Velásquez-Tibatá and Graham, 2013).

The aim of this study is to use fish survey data and simulation methods to explore how focal species respond to the number and spatial arrangement of ARs. We previously performed a survey in an area in the Bohai Strait (China) that had an initial cluster of ARs deployed and used statistical analysis to identify species archetypes (Yu et al., 2022); we use the first two-years of the survey data here and fit SDMs for four focal species. The re-analyzed field survey data are then used to define SDMs that reflect the system as it approaches steady-state conditions after the introduction of the ARs. In addition to the environmental variables, the SDMs used here to evaluate spatial layouts also needed to consider the influence on ecological performance of the distances to the nearby ARs. We therefore analyzed the survey data (focused on the extreme observations) to define an alternative set of SDMs that explicitly included the influence (positive from connectivity and negative from competition) of ARs on neighboring ARs. The realistic SDMs (labelled as Best-fit) were based on all of the survey data and are subject to data limitations, while two sets that increased the nearest neighbor effects (labelled Intermediate and Extreme) was aimed to verify the general applicability and replicability of this simulation approach for AR layout optimization if there were ecologically strong interaction effects among ARs.

Simulations of randomly-generated spatial layouts for a given number of total ARs were separately analyzed using the Best-fit, Intermediate, and Extreme SDMs. We use the simulation results to identify how the number of ARs and their spatial layout affect the occurrences of three representative (focal) commercially-important species and one undesirable species, and how those responses might be affected under conditions of stronger interactions from neighboring ARs. We conclude with a discussion of the implications of our results for future deployments of ARs in the Bohai Strait, how our approach is applicable to other systems, and some caveats and ideas for further development of the approach.

2 Material and methods

2.1 Overview

Our analysis approach consisted of ten steps (Figure 1). These steps were: (1) Assemble spatially-resolved biological, environmental, and substrate type data for all survey sampling sites and identify economically and ecologically important and undesirable focal species; (2) Use redundancy canonical analysis (RDA) and variance inflation factor (VIF) to screen observed environmental variables and ecological constraint variables (i.e., distance-to-substrate variables, Dis.1 to Dis.5, Table 1), and retain important and weakly collinear predictors. Then, use those filtered environmental variables (excluding distance-to-substrate variables that were automatically included in models) with associated species presence/absence data to fit random forest models; this resulted in further selection of important species-specific environmental variables; (3) For each focal species, use the final set of species-specific environmental variables together with ecological constraint variables (Dis.2 and Dis.5) to fit four SDM algorithms and one ensemble model. All survey observations were treated as independent presence/absence samples, and the selected predictors were used to generate species-specific occurrence probability estimates at all sampling sites; (4) Evaluate model performance using cross-validation, and the best-performing algorithm among the five candidate models was selected as the final Best-fit SDM for each focal species; (5) Build the Intermediate and Extreme versions of the dataset and re-estimate SDMs based on the selected best model from Step (4) that now represent ecologically moderate (Intermediate SDM) and strong (Extreme SDM) interaction effects of neighboring ARs; (6) Grid and estimate environmental variables for all cells on the grid needed by the selected (best) SDM to generate grid-wide estimates of occurrence probabilities; (7) For 10,000 realizations with each AR-number group (e.g., 50 ARs deployed), randomly select cells for AR placement and change the value of the bottom substrate to denote the presence of the AR; (8) Calculate the new occurrence probabilities of each AR layout using updated bottom substrate values and average using only cells influenced by an AR (local response) and report the percent of cells influenced; (9) Identify good (high performing) and bad (low performing) cells for ARs based on how frequently each cell contributed to high- and low-performing layouts; present results as cell sensitivities and the three best and three worst performing representative layouts; and (10) Examine layouts for connectivity to other ARs and report the usage of habitat types among high- and low-performing layouts.

All simulation procedures and fitted SDMs were performed using *ROCR*, *gbm*, *randomForest*, *e1071*, and *mda* packages in the R programming language (version 4.3.2). Generating spatial grid and simulated layouts were performed using *sp*, *raster*, *gstat*, and *ggplot2* packages in R.

2.2 Study area

The study area was located in the south-central Bohai Strait, China (Figure 2A), a region of rocky and muddy substrates with water depths ranging from 9 to 32 m (Supplementary Figure 1). The area was approximately 146 hm^2, with natural reefs covering nearly one-half of the area (Figure 2B). Due to good accessibility and well-known fishing activities in the area, twenty-four AR monomers, each with a single 7.1 m × 7.1 m × 3.4 m (length × width × height) concrete structure, were deployed in June 2017 to enhance local fisheries for reef-associated species such as black rockfish *Sebastes schlegelii* (Supplementary Figures 2, S3). These 24 ARs were deployed on the bottom about 500 m northwest of Xiaozhushan Island (38°1'45" N, 120°51'56" E). Previous studies of ARs have shown the benefits of clustering individual modules (Thanner et al., 2006). Based on this, the AR deployment was conducted within a

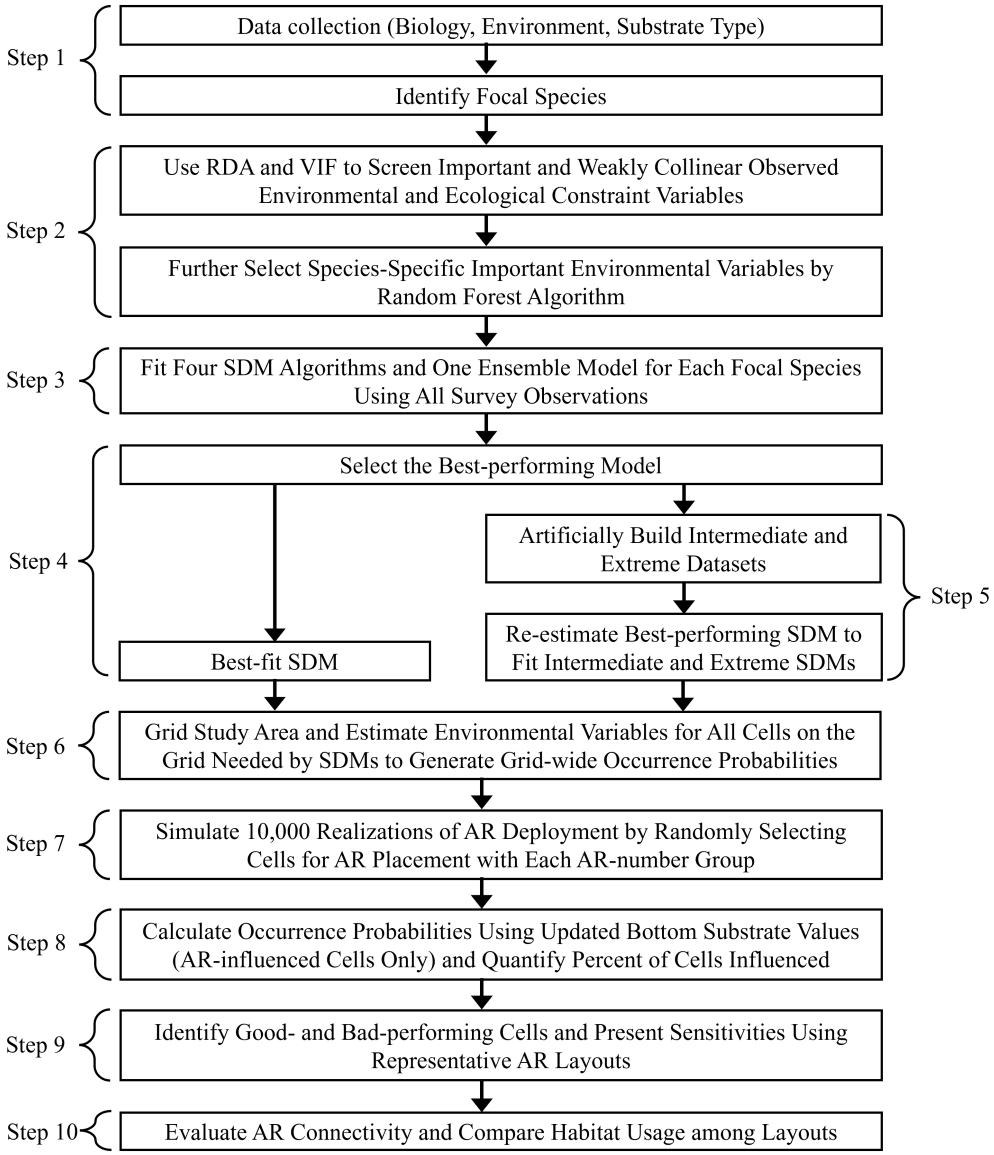

FIGURE 1
Schematic overview of the modeling workflow.

limited spatial test case as a clustering of AR monomers, with monomers released in close proximity to each other during two deployment events. During each deployment, 12 AR monomers were transported by a barge and sequentially deployed using auxiliary crane vessels (Supplementary Figure 3). Only the geographic positions of the barges were recorded at the time of deployment, and the precise seafloor positions of individual monomers were not directly controlled. Post-deployment underwater video surveys indicated that the AR monomers were separated by distances of approximately 10 m, with some monomers directly adjacent or showing partial lateral contact between structural elements. The result of the deployment was a clustered reef complex rather than a set of evenly spaced and discrete monomers. In addition, three AR monomers were deployed on top of existing natural reef structures.

2.3 Surveys

Field surveys were conducted in June, August, and December of 2017, and in May, June, September, and November of 2018 (Yu et al., 2022), covering all four major seasons across the two-year period. Nektonic and benthic sampling was done using accordion-shaped traps and underwater visual census, resulting in the collection of a variety of taxa, including fish, crustaceans, cephalopods, and echinoderms. A total of 42 sites were sampled with traps and 40 sites were sampled by visual census (total of 82 sites) across the two-year study period (Supplementary Table 1; Supplementary Figure 4). A stratified random sampling design was used (to the extent weather allowed) to select the specific sites for each survey, with strata being concentric circles (200, 400, and 600 m radii) centered on the AR cluster (Figure 2A). We used Pinkas'

TABLE 1 Abbreviations for the 15 environmental variables and ecological constraint variables determined to be included by the RDA analysis.

Category	Abbreviation	Full name	Distribution type
Environmental Variables	S. Chla	Chlorophyll *a* Concentration at Surface	Continuous
	S. sus	Suspended Solid Concentration at Surface	Continuous
	S. WQI	WQI of Surface	Continuous
	B. Chla	Chlorophyll *a* Concentration at Bottom	Continuous
	B. sus	Suspended Solid Concentration at Bottom	Continuous
	B. Temp	Temperature of Bottom	Continuous
	B. Sal	Salinity of Bottom	Continuous
	B. pH	pH value of Bottom	Continuous
	B. WQI	WQI of Bottom	Continuous
	Depth	Water Depth	Continuous
	Trans	Transparency	Continuous
	Method	trapping or Visual Survey	Classified
	BotType	Bottom Type Score	Continuous
Ecological Constraints	Dis.1	The distance of cell with the nearest mud	Continuous
	Dis.2	The distance of cell with the nearest gravel	Continuous
	Dis.3	The distance of cell with the nearest rubble	Continuous
	Dis.4	The distance of cell with the nearest boulder	Continuous
	Dis.5	The distance of cell with the nearest AR	Continuous
Focal Species	*C. japonica*	*Charybdis japonica*	Bernoulli
	S. schlegelii	*Sebastes schlegelii*	Bernoulli
	H. otakii	*Hexagrammos otakii*	Bernoulli
	Sea star	*Asterina pectinifera*	Bernoulli
		Asterias amurensis	

index of relative importance (*IRI*) to select focal species (detailed in Supplementary Methods section 1.1).

All analyses of the field data used in this study were for the first two years (2017-2018) of a four-year field survey. We limited the SDMs used here to the first two years because those years had a complete set of measurements of environmental variables collected using a consistent approach. Due to changes in project objectives during years 3 and 4, certain variables (e.g., suspended solids concentration or *WQI*) and sampling depths (e.g., surface level) were not routinely measured in surveys.

In an earlier analysis that informed the analyses in this paper, Yu et al. (2022) used the full four-year dataset (2017-2020) to fit species archetype models, which revealed consistent spatial distribution patterns for reef-dependent species classified within the same species archetype (SA4). In particular, strong and consistent relationships were identified between occurrence of species and the bottom type score and the distance to the nearest AR. Those analyses provided first-level empirical evidence for strong associations between reef-dependent species and reef-related habitat characteristics. The predicted occurrence probability of the reef-dependent archetype was highest in the central AR area (approximately 0.8–0.9), decreased moderately in nearby boulder habitats (approximately 0.7–0.8), and decreased sharply in less structurally complex substrates (approximately 0.2–0.5).

2.4 Spatial grid

A spatial grid of the study area was configured for the SDMs that consisted of 521 square (50 m x 50 m) cells (Figure 2B). Becker et al. (2019, 2023) showed that the ecological effectiveness of the ARs was about 10 to 50 m for both pelagic and benthic species; when within these distances, ARs created a contiguous fish assemblage across the local area. dos Santos et al. (2010) and Puckeridge et al. (2021) both observed that ARs significantly affected fish assemblages when sampled within 0 to 50m of an AR. We therefore assumed a 50-m interval for the spatial grid, as most deployments in practice avoid positioning AR monumers closer than 50 m (i.e., adjacent cells on our grids).

We used the environmental data of water transparency, chlorophyll *a* concentration, suspended solids concentration, temperature, dissolved oxygen, salinity, and pH measured at the sampling (surface and bottom) sites to populate the grid cells (detailed methods see Yu et al., 2020). Concentration of six water quality variables (dissolved inorganic nitrogen, silicate, ammonia, nitrite, nitrate, phosphate) were highly correlated. Therefore, a

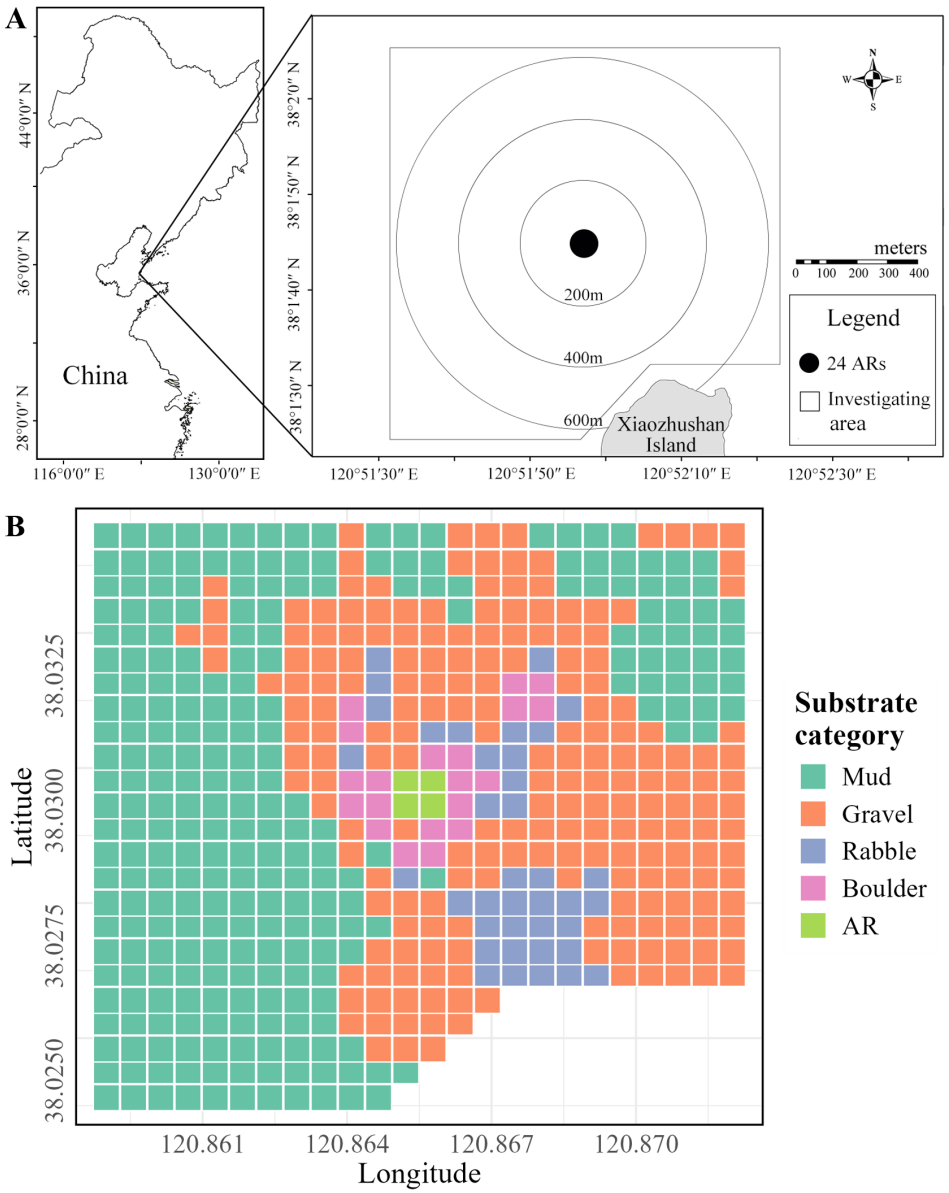

FIGURE 2
Map of the study area. **(A)** Study area along the China coast and the center of the cluster of ARs (black circle) within the area. The three concentric circles (radii of 200 m, 400 m, and 600 m) show the strata used in the sampling survey. **(B)** The spatial grid of the study area showing the bottom types for each cell.

water quality index that combined the six variables (*WQI*, Fabricius et al., 2005) was calculated for each sample. Each of the six variables was standardized to a mean of zero and standard deviation of one (z-scores) and then standardized values were summed to form the *WQI*. In addition, depth of each cell was interpolated from available data (Supplementary Figure 1).

A bottom type score was generated for each cell. We first categorized the substrate type as mud, gravel, rubble, and boulder (Figure 2B), which were separately assigned scores of 0, 5, 20, and 50 (Supplementary Figure 5, Mellin et al., 2007). Interpolation (Inverse distance weighting, IDW method) was then used to obtain bottom type scores for all cells in the grid as a continuous variable. AR layouts were simulated on the grid by placing a single AR in a cell and changing the bottom type value. The receiving cell was assigned a new score of 100, no matter what the current substrate type was. For presentation, we classified the continuous bottom type back to categories using 0-2.5 for mud (257 cells), 2.5-12.5 for gravel (202 cells), 12.5–35 for rubble (38 cells), 35-99.9 for boulder (20 cells), and 100 for AR (4 cells, Figure 2B).

We calculated the distance between each AR and the nearest cells of mud, gravel, rubble, boulder, and other ARs (termed ecological constraint variables: Dis.1-5, Table 1). We focused on Dis.5, the distance to the nearest AR, because that represented interactions between neighboring ARs that likely have a strong effect on the performance of a layout. We considered the interactions between ARs to be a mix of the degree of connectivity (positive) and possible interference by competition (negative). A cluster of ARs suitably spaced, can provide sufficient shelter, prey, and suitable flow field conditions for rocky reef associated species (Caddy, 2011; Liu and Su, 2013; Nakamura,

1985). Thus, ARs can be too far apart with a loss of some benefits of connectivity. However, if ARs are too close to each other, they can interfere (e.g., competition for prey and space) and diminish the ecological benefits (Jaxion-Harm and Szedlmayer, 2015; Leitão et al., 2008; Strelcheck et al., 2005). In both cases (too far and too close), ecological performance of the layout is decreased from the level performance expected under no interactions. Thus, maximum performance occurred at an intermediate distance, and these interaction effects between AR monomers essentially prevent the model from just adding huge numbers of AR monomers.

2.5 Selecting optimal species distribution models

We first used redundancy canonical analysis (RDA, Gaudreau et al., 2018) to explore the relationship between focal species, environmental variables, and ecological constraints (i.e., variables Dis.1-Dis.5) across sampling sites. Focal species were each analyzed using their catch per unit effort (CPUE, as number of individuals caught in a unit of sampling adjusted to the number expected in a 24-hour day) that was transformed by Hellinger distance (Parris, 2004) for use in the RDA. Collinear variables were removed based on the variance inflation factor (VIF, Lawesson, 2010). A progressive bidirectional stepwise selection was performed to retain significant variables based on the Akaike information criterion (AIC) and statistical significance (p-values < 0.05) (Dixon, 2003). The final RDA models that considered all four focal species included only significant and noncollinear (VIF<10, Belsley et al., 1980) environmental and ecological constraint variables (Table 1).

Four machine learning models and one ensemble model that combined the predictions of the four machine learning models were used as the five possible SDMs. All of the SDMs predicted the occurrence probability of each focal species as a function of the subset of environmental variables identified with the RDA plus the distance to the nearest gravel cell (Dist.2) or distance to the nearest AR (Dist.5). The four machine learning models were: random forest (RF, Mi et al., 2017); boosting regression tree (BRT, Elith et al., 2008); support vector machine (SVM, Vapnik, 2000); and multivariate adaptive regression splines (MARS, Elith and Leathwick, 2007). We developed the ensemble model prediction by combining (weighted average) of the predicted occurrence probabilities of the four models (Thuiller et al., 2009).

Before fitting the SDMs for use with the simulation analysis of alternative layouts, we further reduced the environmental variables identified by the RDA for each focal species using the random forest SDM. The random forest model was used because, unlike the other SDMs, it provides a relatively accurate measure of the importance of each environmental variable, calculated as the mean decrease in accuracy (MDA). Environmental variables from the initial RDA that had MDA values > 2 (Wang et al., 2016) from the random forest model were retained on a species-specific basis. We did not include the distance to the nearest gravel cell (Dist.2) or distance to the nearest AR (Dist.5) in the filtering using the random forest model because we wanted to force their inclusion into the final selected SDMs to ensure we captured the influence of these two nearest distance effects in simulations.

We then fitted the five SDMs for each focal species (without Dis.2 and Dis.5) and identified the best fitted model for each of the four focal species. All species abundance data were transformed into presence-absence type (presence = 1, absence = 0) for use in fitting the SDMs. Models were estimated using the data from all 82 sampling sites and 5-fold cross validation (CV), which was repeated 1000 times (Perlich and Swirszcz, 2011). We computed the area under the receiver operating characteristic (ROC) curve (AUC, Pearson et al., 2010) and root mean squared error (RMSE) to assess the goodness-of-fit of each fitted model. ROC curve is a measurement of two-dimensional classification performance based on a plot of the correct positive samples against the proportion of incorrectly classified true negative examples. False Positive and False Negative rates of predictions are obtained by applying different thresholds to model outputs to configure the ROC curve (Fielding and Bell, 1997). For the ensemble model, we averaged the occurrence probabilities of the four SDMs weighted by their AUC values. Based on the similarities among the best models of each of the four SDMs and the ensemble model, we selected one of the SDM approaches to use for all species. We then added the Dis.2 and Dis.5 and refitted the final models for each species.

2.6 Intermediate and extreme versions of the dataset

According to the responding curves of focal species to each filtered environmental variable in the optimal (Best-fit) versions of the SDMs based on the survey data (Supplementary Figures 6-S9), Dis.5 (distance to nearest AR) showed low model explanation and relatively low influence. This indicated that, within the spatial configuration observed during the field surveys, strong interference or connectivity effects among neighboring ARs were not present. Nevertheless, ecological theory and empirical studies suggest that AR monomers placed too closely together may experience intensified competition or attraction overlap, whereas ARs spaced too far apart may suffer from reduced functional connectivity (Jaxion-Harm and Szedlmayer, 2015; Leitão et al., 2008; Strelcheck et al., 2005). Reef-dependent species are therefore often expected to occur at higher densities when ARs are spaced near or at an optimal distance. To account for the possibility that interactions between AR monomers may become more significant under alternative AR layouts that deviate from the limited layout represented in the field survey, we developed the Intermediate and Extreme SDMs. These models were not intended to represent observed conditions, but rather to define plausible bounds on interaction strength between AR monomers by progressively increasing the influence of Dis.5. In this way, the Intermediate and Extreme SDMs provide a structured sensitivity analysis to evaluate how layout performance and recommendations may change under increasing levels of interaction effects among ARs.

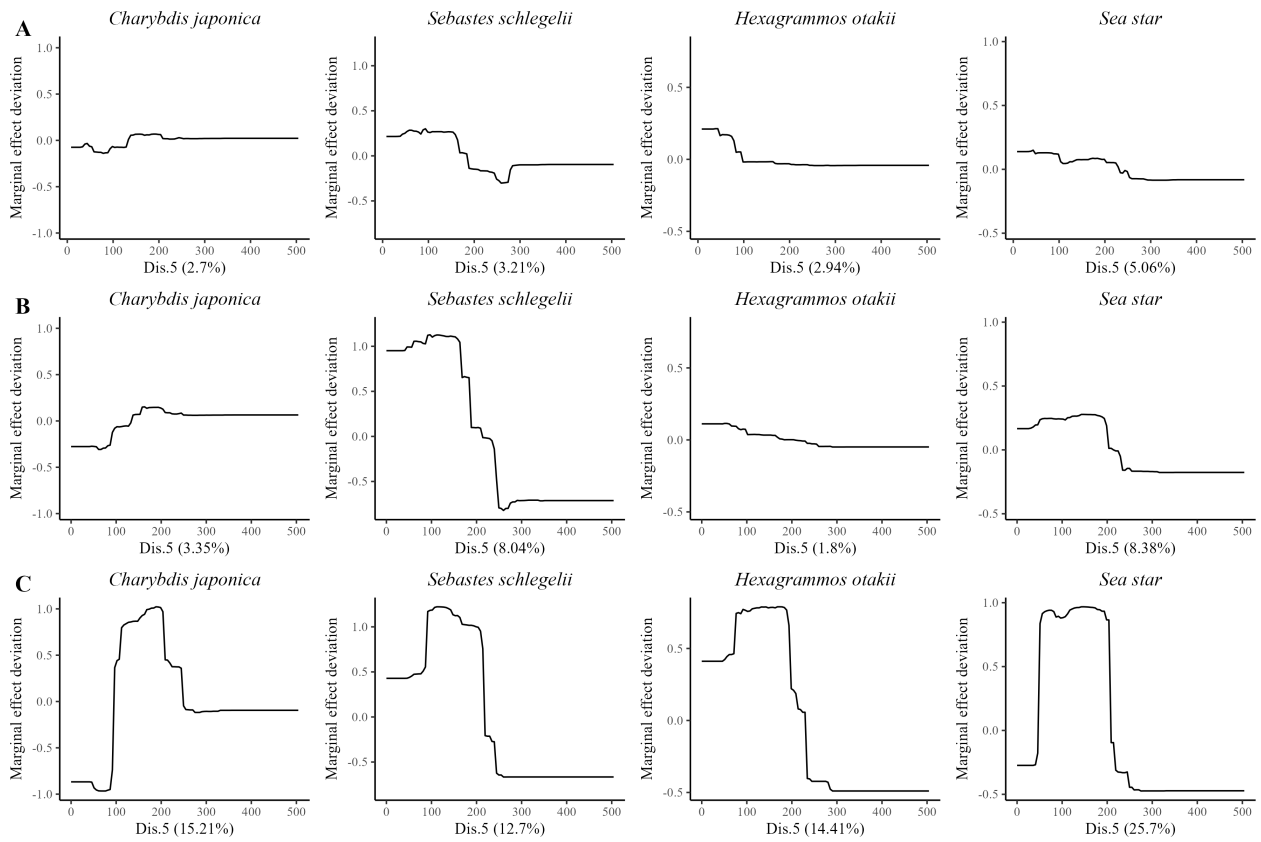

FIGURE 3
Responding curves of Dis.5 (distance to nearest AR) from four focal species and percent explained from the fitted model. **(A)** Best-fit SDMs; **(B)** Intermediate SDMs; **(C)** Extreme SDMs.

The Intermediate and Extreme versions of the dataset were derived by modifying the same observational dataset used to estimate the Best-fit SDMs, with changes applied exclusively to the variable Dis.5 (distance to nearest AR). To induce a bell-shaped response curve to Dist.5 within the SDMs, we applied mathematical transformations to the Dis.5 values in the dataset. This involved reordering the observational dataset and reassigning (overwriting original Dis.5 values with new values). The extent of AR interaction was controlled by adjusting the proportion of observation sites for which Dis.5 was changed. A detailed description of this process is provided in the Supplementary Materials (Supplementary Methods 1.2; Supplementary Figure SA1; Supplementary Table 2). To avoid ecologically implausible situations where the interference interaction dominated the performance of the ARs, even the Extreme version was designed to retain unobserved but plausible spatial patterns of occurrence probabilities in model simulations. For example, when interference was too high, highly defined doughnut shapes or low values uniformly throughout the grid were obtained. Ultimately, two levels of modification were implemented: 20% of observation sites with updated Dis.5 values, referred to as the Intermediate version, and 50% labelled the Extreme version. The Extreme version maintained reasonable patterns but was near the limit of realistic interactions. The resulting SDM response curves for Dis.5 in these modified datasets, compared to the Best-fit version, are illustrated in Figure 3.

2.7 Design of simulations

2.7.1 Use of the selected SDMs in the evaluation of AR layouts

We applied the selected SDMs to the complete spatial map of depth and substrate type for each of the four focal species. For the environmental variables included in the SDM besides the Dis.1–5 variables, depth, and substrate type (Table 1), we normalized their influence by setting their values of environmental variables to where the y-axis of plots equaled zero (Supplementary Figures 6-S9). This was necessary because many of the environmental variables showed high variation in time and space thereby preventing realistic interpolation to the cells of the grid or were relatively unimportant compared to substrate type. An exception to the normalization of using the value of zero on the y-axis was for the variable ‘method’. Method was either a 0 or 1 (trapping versus visual census) and its inclusion in the fitting was used to assess gear effects (Yu et al., 2020). For the evaluation of simulated deployments, the variable ‘method’ was set to 0 for all cells.

2.7.2 Simulation experiments

We simulated the deployment of an increasing number of ARs (5, 10, 25, and 50) based on adding these ARs to the existing 24 ARs

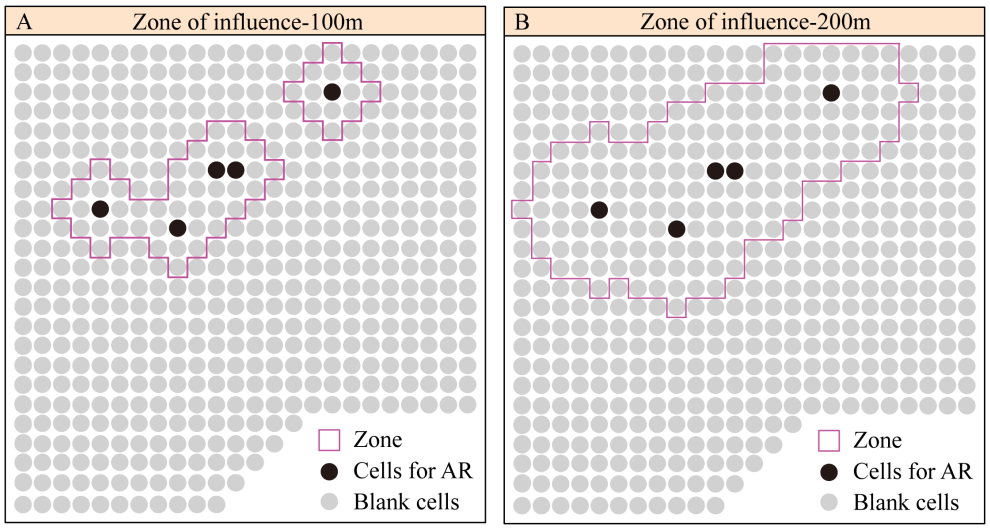

FIGURE 4
Sketch map illustrating examples of how the zone of influence determines the cells affected by the placement of an AR. The cells within the zone were all included to compute the arithmetic occurrence probability of each focal species. Zone of influence using distances of **(A)** 100 m and **(B)** 200 m.

and assessing changes in the occurrence probabilities using the Best-fit, Intermediate, and Extreme SDMs. This represents the often encountered management situation of adding ARs to the current state of the system that already has existing ARs. For each deployment, cells among the 521 available in the grid were randomly selected without replacement (one AR allowed per cell), and their bottom habitat type were changed to AR (i.e., value of 100). If cells with existing ARs were selected, we deployed the AR to that cell and the substrate type score of the cell was retained as indicating an AR was present. The selection of cells was random, so that each time the number of ARs were independently assigned to cells. However, for the same number of ARs, the same generated AR layouts were used for the Best-fit, Intermediate and Extreme SDMs to ensure a direct comparison that only differed by the strength of the interactions between ARs. For each focal species, 10,000 simulations were performed for each fixed number of added ARs.

We first established baseline occurrence maps for each focal species for the Best-fit, Intermediate, and Extreme SDMs, and the baseline values were calculated by taking the arithmetic mean occurrence probability computed across all cells on the grid. We used the original bottom type scores from no added ARs (but including the original 24 ARs) to display the spatial distribution of the occurrence probabilities relative to bottom type.

In each simulation of layout, cells were selected for the desired number of total ARs, their bottom habitat type was switched to 100, and the values of Dis.1–5 for all cells on the grid were re-estimated. As part of adding ARs, we included the effect of ARs on the occurrences of the focal species in nearby cells (Becker et al., 2019; Reeds et al., 2018). Distances from an AR that determined they affected nearby cells were estimated from the SDMs. The nearest distances included both the most positive and negative responses in Dis. 5 was selected as the zone of influence (Figure 4). We used 200 m, 270 m, 100 m, and 300 m for *C. japonica*, *S. schlegelii*, *H. otakii*, and Sea star in the Best-fit, and used 280 m for Intermediate and Extreme SDMs for all species (Figure 3). This allows for spillover effects of ARs (beyond the immediate cells) to be included in performance when optimizing layouts (Langhamer et al., 2016; Punt et al., 2009). The nearest AR was used for each AR placed in a cell and so if multiple ARs were within the zone of influence, the effect of only the nearest was included. For each species and layout, a new set of predicted occurrences for potentially all cells on the grid were obtained. Finally, we reported the arithmetic average occurrence for each focal species for all cells affected by a deployment (i.e., all cells within a zone of influence). We then used the arithmetic mean of occurrence probability in the 50 AR case as a baseline threshold for each species and each SDM (Best-fit, Intermediate, Extreme), and then calculated the percentage of grid cells with occurrence probabilities exceeding the threshold across different numbers of ARs. We used a model-generated threshold (mean for 50 ARs) so that we could directly compare how the percentage of cells changed across different numbers of ARs deployed. Because all thresholds were between 0.44 and 0.76, the percentage of cells affected reflected how ARs would support good habitat conditions for that species. Moreover, to summarize across focal species and obtain a single measure of performance across the four species for a layout, we computed the geometric mean of the 4 average occurrences.

We also reported the average percentage of grid cells affected by the spillover effect for each layout for each focal species. We calculated the average (over layouts) of the percent of the 521 cells affected by an AR being added. We used only the cells affected by the deployment to account for the different number of cells involved with adding different numbers of ARs (e.g., 5 versus 50). Area affected by a layout is also of interest to assessing the performance of a layout and to management.

We used a personal computer (12th-generation Intel Core i7 CPU, 32 GB RAM) for all analyses and simulating 10,000 realizations of AR layouts for each AR number group for each species required 30 minutes.

2.8 Representative layouts and cell sensitivity

Determining which cells tended to occur in high-performing versus low-performing layouts is useful for designing layouts. We used two measures of layout performance: representative layouts and cell sensitivities. For both measures, we grouped the 10,000 layout realizations evenly into five ordered categories (Good, Moderate-good, Medium, Moderate-bad, and Bad) using the species-specific arithmetic mean of occurrence probabilities (based only cells affected by added ARs) for each layout realization; each category contained 2,000 layout realizations. To select representative layouts, we used the top three ranked in the good category and the bottom three ranked in the bad category (i.e., the three best performing and the three worst performing layouts).

Calculation of the sensitivity of each cell on the grid was based on randomly selecting 100 of 2,000 layouts from each category. The index was the fraction of the 100 layouts for each cell that were selected to host an AR. A high sensitivity index in certain cells for the good performing AR category suggests these cells should be examined for possible AR placement because they often occurred in high-performing layouts. Similarly, high sensitivity cells in the bad performing category indicated specific cells to avoid. For the undesirable Sea star, we reversed the occurrence probability (1 – calculated probability) so that the good category for Sea star corresponded to low occurrence probabilities.

In addition to layouts categorized independently based on the average occurrence probability for each of the four species, we also present representative layouts and sensitivity index results based on the geometric mean of occurrence probabilities that combined the four species. Comparison of the spatial patterns of ARs in good and bad layouts between rankings based on a single species versus the geometric mean identifies to what degree there were tradeoffs among the species. If the patterns were similar for individual species and all species combined, then species were responding similarly to the spatial arrangements of the ARs.

In order to indicate the connectivity of the deployed ARs to the existing and new ARs and to natural reefs (boulder), we reported the averaged distance from each AR to its nearest AR neighbor (D_k) and the average distanced from each AR to its nearest boulder (B_k) for each layout. We wanted to see if placing ARs near boulders affected performance, as boulders are a known good habitat and often easily identifiable. We confirmed that the two output variables (D_k and B_k) showed reasonable dependencies across the five good-to-bad categories. Therefore, for each focal-species categorization, the values of D_k and B_k from the layouts were only plotted on the two extremes of good and bad categories by box plots. We also examined the percentage usage of original (before any switching to 100 by a deployed AR) habitats (mud, gravel, rubble, and boulder) of cells housing an AR using good versus bad categories. The good and bad categories scheme adopted the same sampling method (100/2000 layout realizations) as the good and bad in the five ordered categories scheme. Thus, good refers to the top 20% and bad to the bottom 20% in both schemes.

3 Results

3.1 Focal species and environmental variables

A total of 59 species were identified in the survey using the two sampling methods, and there was a total of 23 environmental and ecological constraint variables included in the RDA. As described above we identified the four focal species: *Charybdis japonica* (Crustacea), *Sebastes schlegelii*, *Hexagrammos otakii* (rocky fish), and the undesirable sea stars (*Asterias amurensis* plus *Asterina pectinifera*). A total of 13 environmental variables and 2 ecological constraint variables filtered by RDA were included (Supplementary Table 3). The 13 environmental variables were further reduced for each focal species using the Random Forest algorithm (Supplementary Table 4).

3.2 Selection of the optimal SDMs

We selected the BRT model approach for all four focal species. The fits of the four-algorithmic SDMs and the ensemble version (assessed by RMSE and AUC) fitted using the 13 environmental variables are shown in Supplementary Table 5 and the influence of each environmental variable is shown in Supplementary Figure 10. We selected the BRT model because results for habitat substrate type were generally similar across the models and BRT showed the highest sensitivity to substrate type. Weight values of substrate type for the focal species in the BRT models were 1.73, 1.71, 2.78 and 2.39 and their RMSE values were relatively low and their AUC values were high (Supplementary Table 5).

3.3 Response curves for best-fit, intermediate, and extreme SDMs

Response curves from the Best-fit SDMs revealed differences among focal species in their responses to their four most important environmental variables (Figure 5). For all species, bottom type score consistently showed positive effects and ranked among the first two most important predictors, whereas the remaining variables generally exhibited threshold-type or weaker responses. In addition, species differed in which environmental variables, besides bottom type, were important. For example, temperature and depth were important (ranked 3 and 4) for *C. japonica*, bottom chlorophyll-a concentration and survey method contributed strongly (ranked 1 and 3) to *S. schlegelii*, while depth and the water quality index (*WQI*) played a more prominent role (ranked 2 and 3) for *H. otakii*. For sea stars, responses were sensitive to bottom type (like the other species) but showed limited sensitivity to other environmental variables.

A clear, gradually expanding interaction effect of distance to nearest AR (effect of Dist.5 variable) was evident in the increased range of marginal effect deviation across four species from the Best-fit to Intermediate to Extreme versions (Figure 3). The Intermediate version relationships retained shape characteristics similar to those

10.3389/fevo.2026.1684304

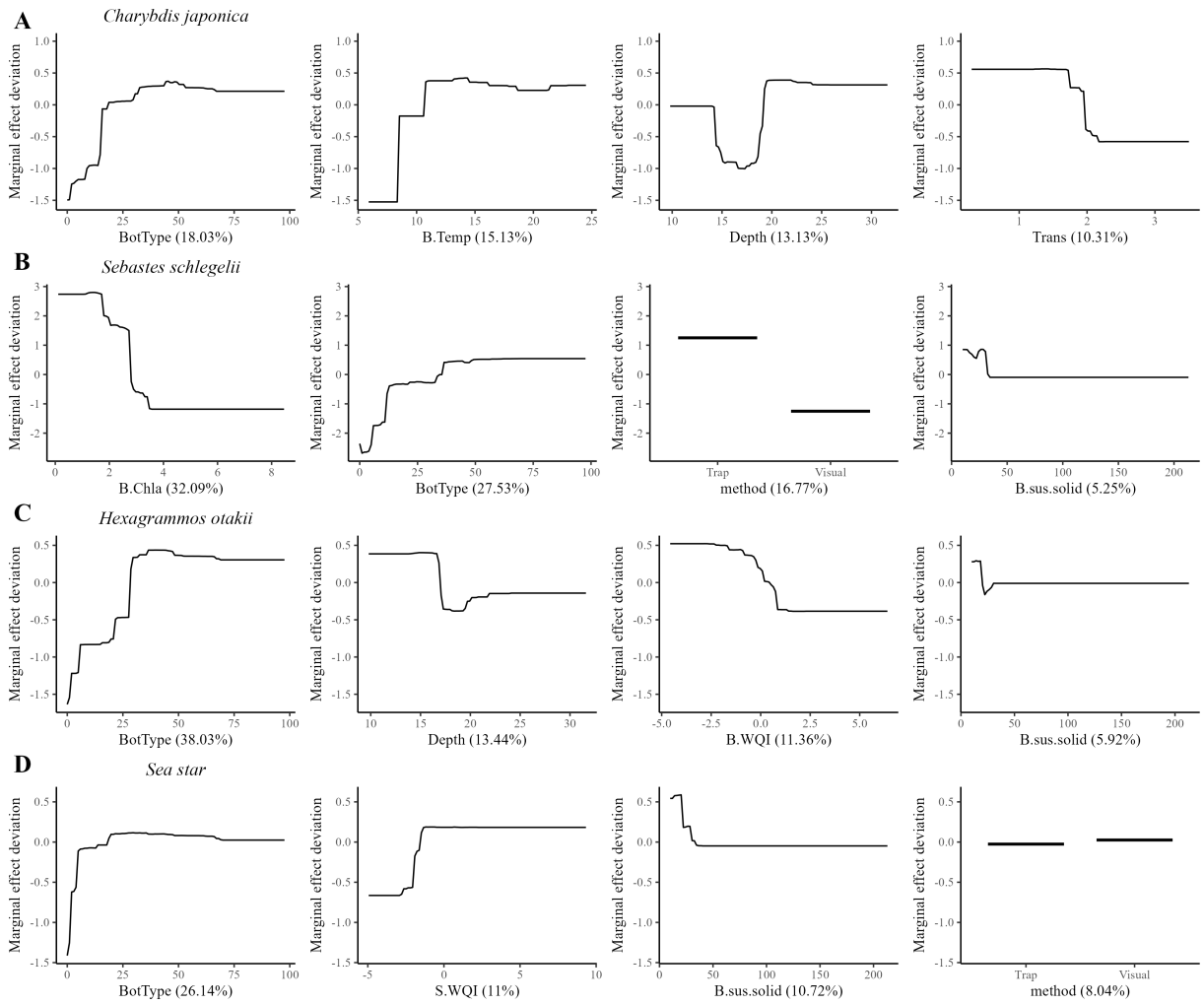

FIGURE 5
Responding curves of four important environmental variables (abbreviations defined in Table 1) for the four focal species, and the percent explained, from the Best-fit models. **(A)** *C. japonica*; **(B)** *S. schlegelii*; **(C)** *H. otakii*; **(D)** Sea star.

in the Best-fit model (Figure 3B). The Extreme version further exaggerated the deviations from the Best-fit and resulted in shapes often unlike those estimated for Best-fit, particularly for *C. japonica* and the sea star (Figure 3C).

3.4 Simulation experiments and evaluation of layouts

3.4.1 Baseline occurrence probability of focal species

For the Best-fit models, baseline occurrence probabilities were high in the central region for all four species (row A of Figure 6). Sea star showed the broadest spread from the center, while *H. otakii* only displayed high occurrence probabilities (> 0.6) in the central area of the grid. High probabilities of occurrence spread into the upper triangle for *C. japonica* and spread from the central region outward for *S. schlegelii*.

With the Intermediate and Extreme models, the increased effect of distance between ARs sharpened the areas of low and high probabilities and somewhat reduced the spatial extent of high probability areas (Figure 6). The spatial map of occurrence probabilities was least affected for *C. japonica*, (left column of Figure 6) while for the other three species, the spatial patterns were clearly sharpened and the high probability areas were somewhat reduced in extent (rows B and C of Figure 6). *S. schlegelii* exhibited a very clear sharpening (second from left column), while *H. atakii* illustrated the contraction of the high probability central area (third from left column). For the Sea Star, a fuzzy halo effect (doughnut shape) also emerged, with high occurrence probabilities concentrated within the band 100m to 300m from the originally located central AR cells. The halo was attributed to the existing ARs causing negative effects on ARs cells very near the center but positive effects for ARs located optimal distances away in the central area encompassing the existing ARs.

The general sharpening effect between high and low probability areas and reduced extent of high probability areas resulted in minimal or small changes in the overall baseline average occurrence probability relative to the Best-fit models. Average probabilities of occurrence were very similar between Best-fit and Intermediate, while were slightly lower for Extreme (e.g., 0.52 versus 0.53 versus 0.48 for *C. japonica*). As a rough check on whether the probabilities were related to fish densities, baseline occurrence

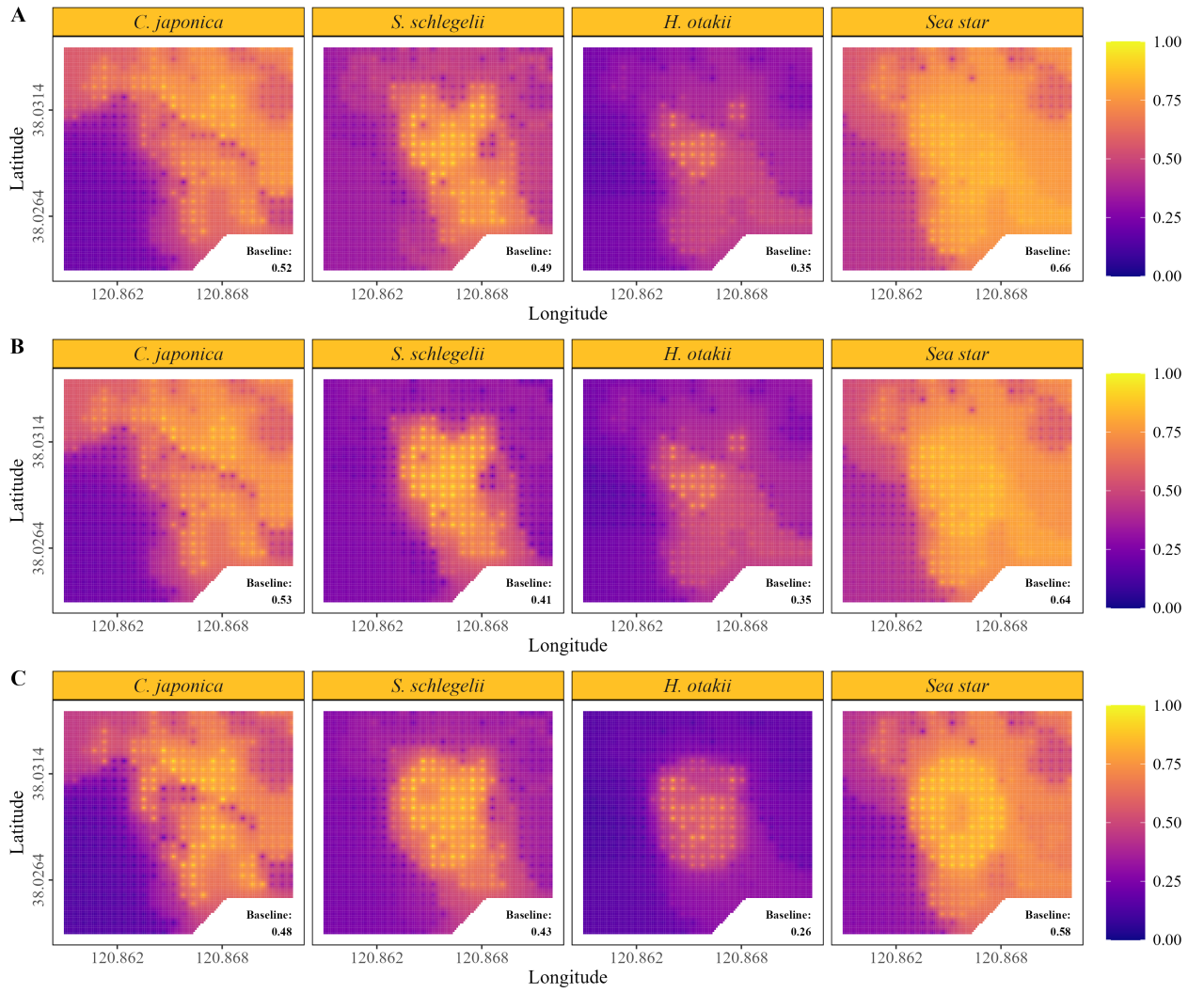

FIGURE 6
Spatial distribution of occurrence probabilities for the four focal species using the Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs. The value shown on the bottom right of each plot is the averaged occurrence probabilities computed across all cells.

probabilities of the sampling sites showed some agreement with corresponding survey CPUE for the four focal species (Supplementary Figure 11). The general pattern was low occurrences matched low CPUE values and high occurrences were associated with somewhat higher but very variable CPUE values.

3.4.2 Occurrence probabilities of focal species with increasing ARs

Overall averaged occurrence probabilities (computed across the average value for each layout) of focal species in affected cells with the Best-fit models were unaffected by the number of ARs deployed (horizontal lines in row A of Figure 7). The increase in overall averaged occurrence probabilities from 5 to 50 ARs, while consistent, were not ecologically meaningful, with the increase (absolute change from 50 to 5 ARs) in average occurrence probability being 0.0071, 0.0398, 0.0184, and 0.0191 for *C. japonica*, *S. schlegelii*, *H. otakii*, and Sea star.

While overall averaged occurrence probabilities were unaffected, variation in occurrence probabilities across layouts (i.e., the average values for each layout) with the Best-fit models decreased with increasing number of ARs for all focal species (boxes and vertical lines in row A of Figure 7). The range of the averaged occurrence probabilities was the highest for 5 ARs and gradually reduced with increasing number of ARs to minimal variation at 50 ARs. Minimum to maximum values across realizations for *C. japonica* were 0.21 to 0.73 for 5 ARs versus 0.51 to 0.56 for 50 ARs, 0.37 to 0.71 versus 0.55 to 0.58 for *S. schlegelii*, 0.25 to 0.59 versus 0.39 to 0.47 for *H. otakii*, and 0.52 to 0.84 versus 0.71 to 0.73 for Sea star. Thus, occurrence probabilities viewed on a local response scale (within zones of influence) were sensitive to the differences among layouts and this sensitivity decreased as more ARs in layouts averaged out these differences. When relatively few ARs were deployed (i.e., 5 ARs), some had multiple ARs located in high habitat quality cells, while other deployments involved mostly all poor habitat cells. Once 50 ARs were deployed, all realizations contained a relatively consistent mix of good and bad habitat cells

10.3389/fevo.2026.1684304

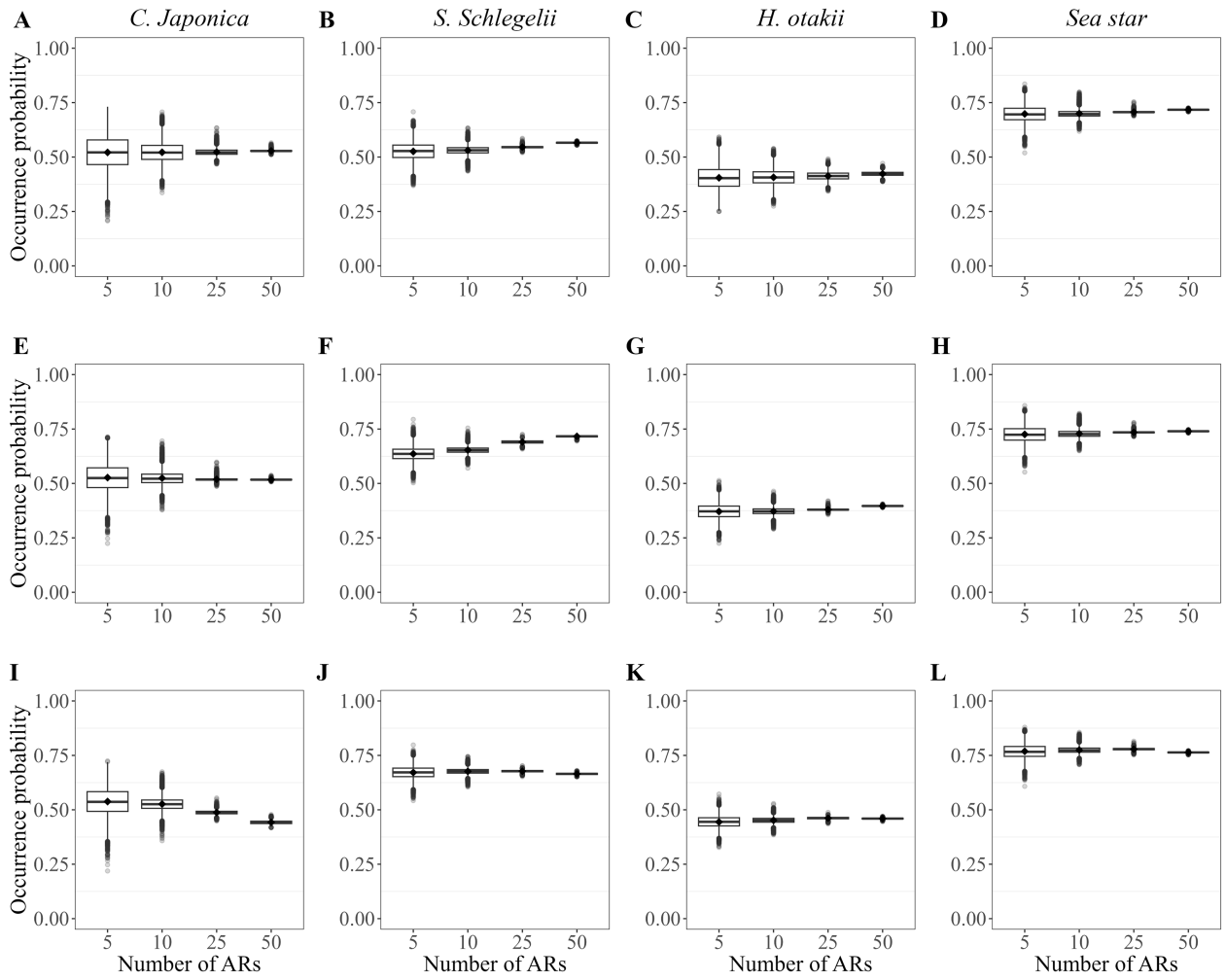

FIGURE 7
Occurrence probabilities of four focal species with the addition of 5, 10, 25, and 50 ARs in Best-fit **(A–D)**, Intermediate **(E–H)**, and Extreme **(I–L)** SDMs. Panels **(A–D)** each refer to *C. japonica*, *S. schlegelii*, *H. otakii*, and Sea star, which had zone of influences defined by 200m, 270m, 100m, and 300m, respectively; Panels **(E–L)** for the Intermediate and Extreme SDMs used a uniform zone of influences defined by 280m for all four species. Boxplots display the overall mean as a diamond-shaped point. The box represents the interquartile range (IQR), with the lower edge corresponding to the 25th percentile and the upper edge to the 75th percentile. The whiskers extend from the box to show the range of data within a certain distance from the quartiles (1.5 times the IQR), and any data points beyond this range are plotted individually as dots.

that generated similar average occurrence probabilities among layouts (i.e., converged to the overall average values).

The Intermediate and Extreme models produced similar patterns in overall average and variability of layout-specific average occurrence probabilities (rows B and C in Figure 7) as predicted with the Best-fit models (row A of Figure 7). With the Extreme models, occurrence probabilities showed a slightly stronger decrease in overall average occurrence probabilities for *C. japonica* going from 5 to 50 ARs, but this was still small in magnitude. This weak downward pattern in overall average occurrence probabilities compared to the Best-fit model provides some evidence of the increased in interference interactions by neighboring ARs. Statistical evidence for the increased importance of this effect of neighboring AR interactions (Dis.5 variable) under Extreme was that the model explanation for Dis.5 was 15.21%, 12.7%, 14.41%, and 25.7% for *C. japonica*, *S. schlegelii*, *H. otakii*, and Sea star compared to little importance for Best-fit (Figures 3A, C). Lowered overall averaged occurrence probabilities with *C. japonica* were caused by the low values of the function before the peak (competition) for less than 25 ARs and declining post-peak values (reduced connectivity) at 50 ARs.

For all Best-fit, Intermediate and Extreme models, the percent of cells affected by deployments increased with increasing numbers of ARs (Table 2). The degree of the grid affected depended on the distance used to define the zone of influence, which was similar among the three models for *S. schlegelii* (270 versus 280 m), but larger for the Intermediate and Extreme models, especially for *H.otakii* (100 to 280 m), for the other focal species. Except for *H. otakii* with Best-fit (70.5%), all layouts resulted in nearly all of the cells on the grid being affected by the addition of 50 ARs. In addition, for the Extreme, the similarity in the distances of the zones of influence resulted in the general pattern of about 55% of the grid affected with the addition of 5 ARs, 78% for 10 ARs, and 97.2 to 99.8% for 25 and 50 ARs. Once most all cells were affected by being

10.3389/fevo.2026.1684304

TABLE 2 Average percentage of grid cells affected by spillover effect distances for different focal species.

Models	Number of AR	Focal species			
		C. japonica	*S. schlegelii*	*H. otakii*	Sea star
Best-fit	5	34.3% ± 4.3%	59.8% ± 8.3%	11.1% ± 0.9%	59.8% ± 8.3%
	10	56.6% ± 5.5%	82.9% ± 7.1%	21.1% ± 1.5%	82.9% ± 7.1%
	25	87.1% ± 4.4%	98.2% ± 2.2%	45.1% ± 2.5%	98.2% ± 2.2%
	50	98.0% ± 1.6%	99.9% ± 0.3%	70.5% ± 2.8%	99.9% ± 0.3%
Intermediate	5	54.8% ± 7.6%	Same as *C. japonica*	Same as *C. japonica*	Same as *C. japonica*
	10	78.8% ± 7.1%			
	25	97.2% ± 2.7%			
	50	99.8% ± 0.5%			
Extreme	5	54.8% ± 7.6%	Same as *C. japonica*	Same as *C. japonica*	Same as *C. japonica*
	10	78.8% ± 7.1%			
	25	97.2% ± 2.7%			
	50	99.8% ± 0.5%			

In the Best-fit model, the spillover effect distance is species-specific: *C. japonica* (200m), *S. schlagelii* (270m), *H. otakii* (100m), and Sea star (300m). In the Intermediate and Extreme models, a uniform spillover effect distance of 280m is applied to all four focal species.

within a zone of influence of deployed ARs, the layout-specific average occurrence probabilities converged to very similar values (i.e., to the overall average value) (Figure 7).

The percentage of grid cells with occurrence probabilities exceeding species- and model-specific threshold values increased with the number of ARs and leveled off when interference among AR monomers became important (Figure 8). Under the Best-fit SDM (first row in Figure 8), this percentage increased monotonically for all species, indicating a strong positive response to AR deployment. For the Best-fit SDM, *C. japonica* exhibited the largest increase in area supporting good habitat conditions (from a mean of 20.2% at 5 ARs to 58.9% at 50 ARs), whereas *S. schlegelii* showed the smallest increase (from 17.8% to 35.2%). Under the Intermediate SDM (second row in Figure 8), the increasing trend with AR number remained evident across species, but the rate of increase was reduced compared to the Best-Fit results, suggesting the benefit of adding more ARs diminishes as interference related to AR proximity increases (e.g., Figures 8A, E). In the Extreme SDM (third row in Figure 8), the interference effect became more apparent; the 50 AR case resulted in lower percentages of high-occurrence (good habitat) cells than the 25 AR case for all four species. Nevertheless, both the 25 AR and 50 AR cases generated percentage areas greater than from their 5AR case. High interference effect did not substantially influence layout variability at low AR numbers, as high variability among layouts persisted under the 5 AR case across all species and SDMs.

3.4.3 Representative layouts and cell sensitivity

Results for cell sensitivity grid maps resembled the results illustrated below using representative layouts, and are presented in the Supplementary Figures 12-S23. We focused on presenting the representative layouts results because they provided specific layouts that were easier to visualize rather than sensitivities that were computed over layouts.

In general, the representative layouts illustrated that the locations of cells that were consistently in high-performing (or analogously in low-performing) layouts were similar for the Best-fit, Intermediate, and Extreme models for 5 and 10 ARs, while sometimes showed increased uncertainty or converse patterns with 25 and 50 ARs. For *C. japonica*, good cells were concentrated in the upper triangle of the grid and extended toward the left and bottom edges going from 5 to 50 ARs in the Best-fit model (left to right columns of the first row in Figure 9A Good.5 to Good.50). Poor performing ARs were, as expected, found in cells outside of the center and concentrated in the lower triangle (second row of Figure 9A Bad.5 to Bad.50). The Intermediate models generated similar patterns (Figure 9B). However, this pattern was reversed with 50 ARs in the Extreme model (Figure 9C). In this case, better performing ARs were found towards the lower left corner (lower triangle) and bad performers were almost uniformly distributed throughout the grid. This reversal occurred due to a trade-off between ARs being placed in areas with higher occurrence probabilities (when viewed independently) and the increasing interference effects of additional neighboring ARs.

S. schlegelii and *H. otakii* showed similar patterns of good and bad performing ARs as *C. japonica*. For both species and all three SDMs, the 5 and 10 AR layouts showed that good performing ARs were concentrated in the center towards the upper right corner (upper triangle) and bad performed ARs towards the left-side or lower triangle (Figures 10, 11). Under 25 and especially 50 ARs, and for all three models (third and fourth columns in Figures 10, 11), the spatial patterns for good and bad performing ARs appeared more uncertain (i.e., ARs more evenly scattered across the grid) and, similar to *C. japonica*, some reversal with 50 ARs compared to 5 and 10 ARs. Specifically, for *S. schlegelii*, increasing ARs shifted the central area from good performing ARs to bad performing ARs (e.g., top row of Figure 10A), while for *H. otakii* this reversal occurred at 50 ARs and was clear with the Intermediate and

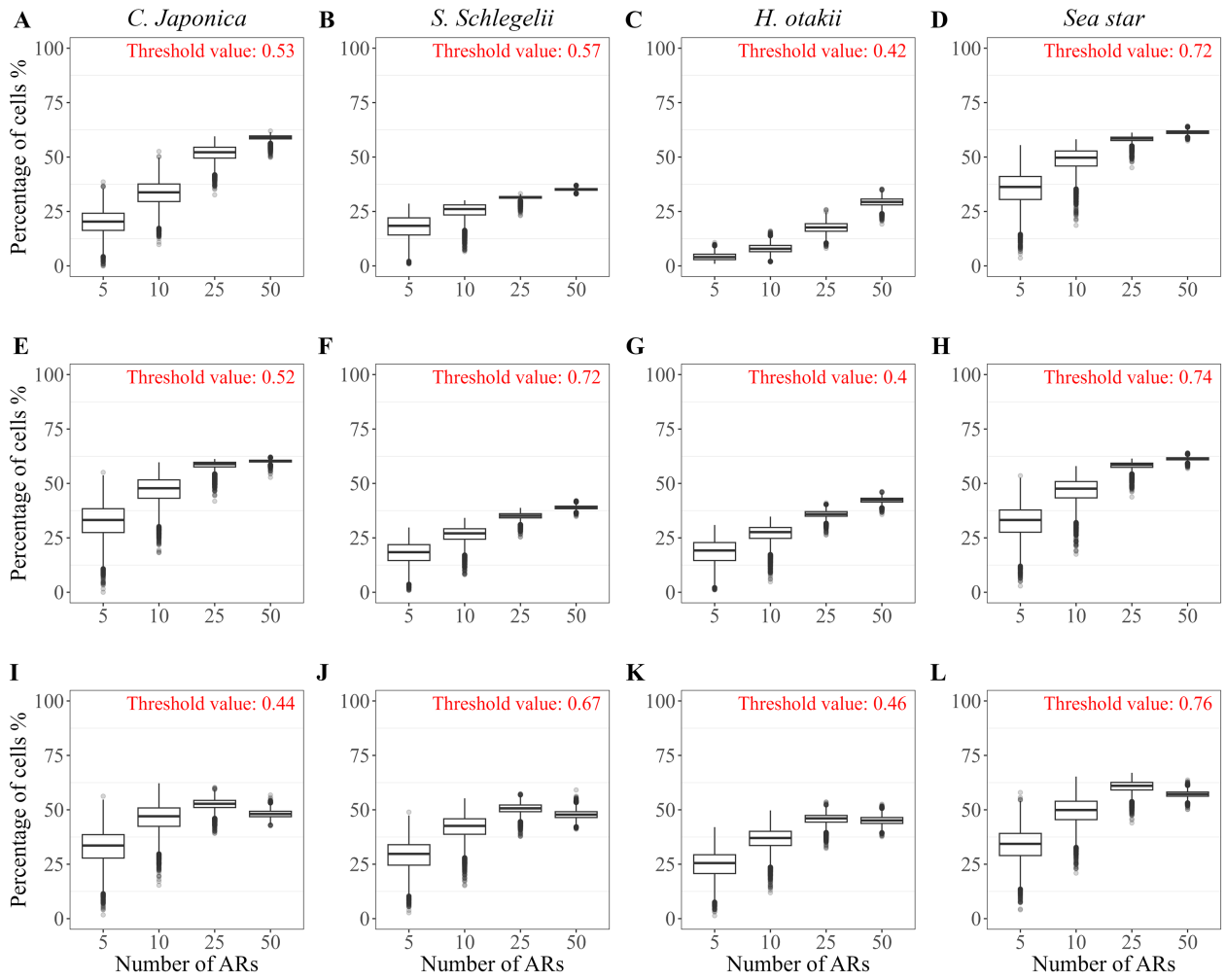

FIGURE 8
Percentage of grid cells with occurrence probabilities exceeding species-specific and model-specific (Best-fit, Intermediate, Extreme) threshold values for increasing numbers of ARs in deployments. Columns represent individual species, while rows correspond to the Best-fit SDM **(A–D)**, Intermediate SDM **(E–H)**, and Extreme SDM **(I–L)**. Threshold values (shown in red on the top right) were defined as the mean occurrence probability under the 50-AR case for each species and SDM.

Extreme (but not Best-fit) models (rightmost column in Figure 11). One difference between *S. schlegelii* and *H. otakii* compared to *C. japonica* was the elbow shaped pattern for bad performing ARs with 25 ARs that created overlap between regions of cells associated with good and bad performing ARs (second from right columns in Figures 10, 11). Thus, the effects of increasing ARs and interaction effects with *S. schlegelii* and *H. otakii* showed similar spatial patterns that resembled those for *C. japonica*: increasing spread of good ARs on the grid with number of ARs deployed, shift of good ARs from upper to lower triangle, and a strong suggestion of an increasingly bad performing central area with 50 ARs (Best-fit for *S. schlegelii*, Extreme for *H. otakii*).

Sea star, as an undesirable species with reversed scoring (good meant low occurrence probability), exhibited a generally contrasting spatial pattern relative to the other species. In the Best-fit model and 5 and 10 ARs (Figure 12A), good cells were absent from the center of the grid and instead clustered along the left side of the grid, indicating a good location for deployment because of lower probabilities of occurrence. Good performing ARs spread to the right with 25 ARs (second from right column in Figure 12), and with 50 ARs suggested a reversal and became less present along the left side (leftmost versus rightmost columns in Figure 12). Thus, the spatial distributions of good and bad performing ARs for Sea stars showed analogous reversals as the three desirable species. Cells supporting high occurrence probabilities for the three desired focal species (good performing ARs) therefore overlapped with cells providing high probabilities for Sea stars (bad performing ARs).

The geometric mean, used as an overall indicator across species, reflected the patterns of good and bad performing locations for *C. japonica*, *S. schlegelii* and *H. otakii*, mostly ignoring the opposing influences of Sea star alone. A general pattern of central to upper triangle concentration of good cells and peripheral or lower triangle distribution of bad cells were apparent with the Best-fit model (Figure 13A). Consistent with the individual desirable species results (and in contrast to Sea Star), this pattern generally became

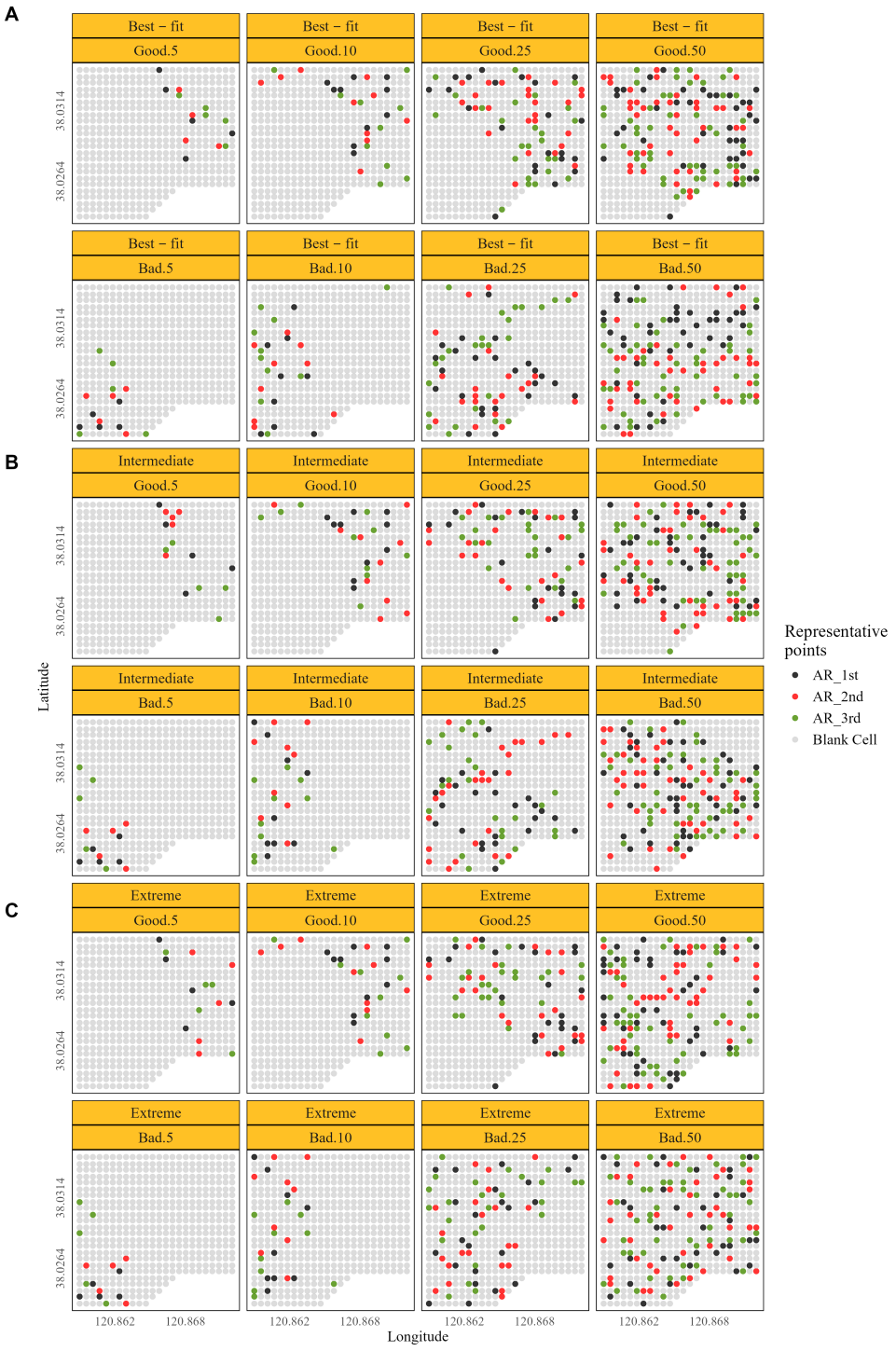

FIGURE 9
Representative AR layouts selected based on being the highest, second highest, and third highest ranked layouts in the good performing (top 20%) and the lowest, second lowest, and third lowest ranked layouts in the bad performing (worst 20%) categories for *C. japonica*. Results are shown for the addition of 5, 10, 25, and 50 ARs under Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs.

more dispersed and reversed with 50 ARs in the Intermediate and Extreme models (Figures 13B, C). The geometric mean treated all four species equally and so was weighted towards mimicking the results of mostly similar-behaving three desirable species. The collapsing over all four focal species therefore hid the effects of increased Sea star occurrence probabilities that would accompany only examining the geometric mean.

3.5 Dependencies of D_k, B_k, and habitat usage on performance

Distance to ARs (D_k) decreased with the increasing number of ARs for Best-fit, Intermediate, and Extreme models for all four species, while distance to boulder (B_k) was relatively constant across the three models (left pair of bars and right pair of bars in

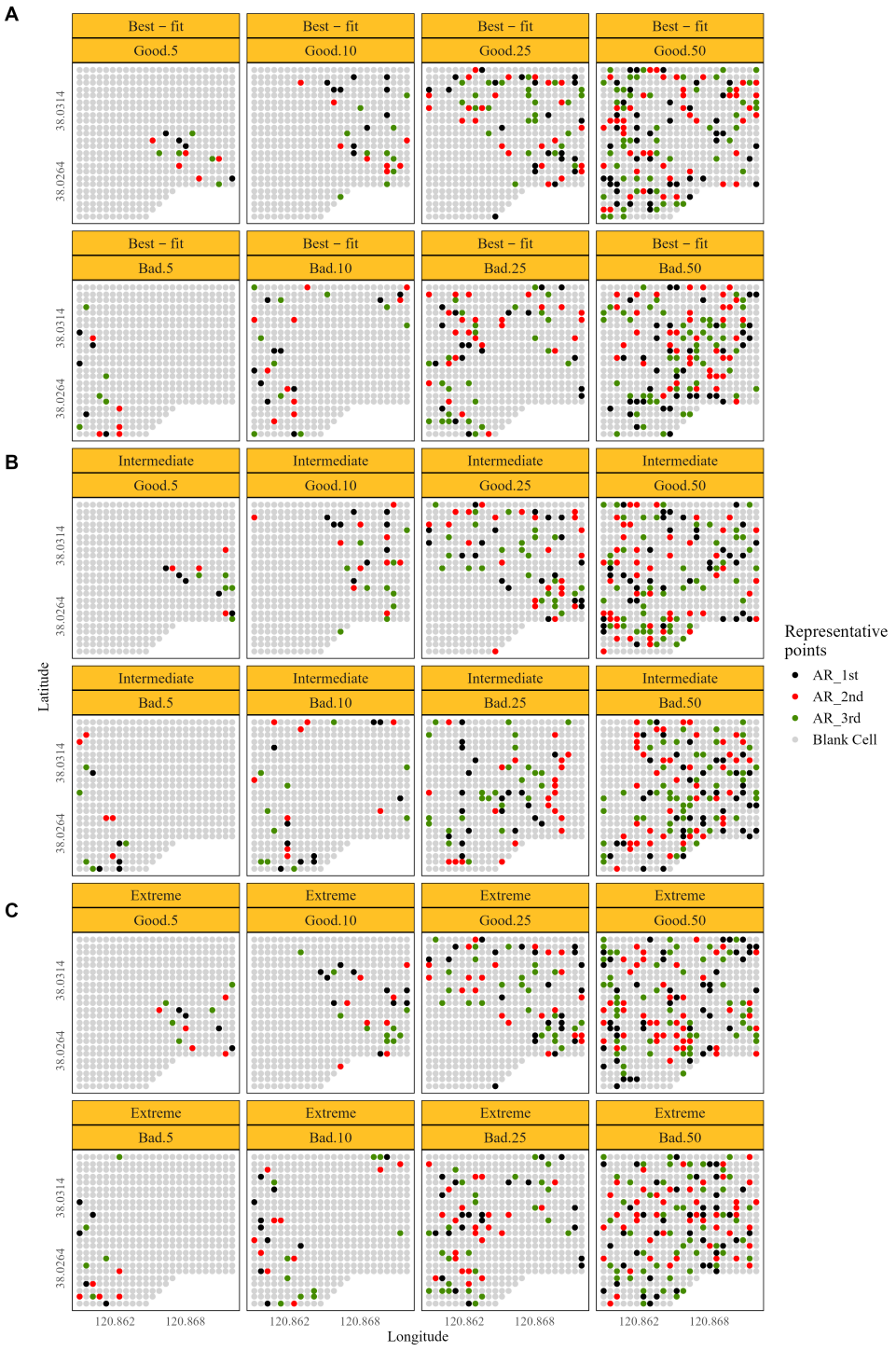

FIGURE 10
Representative AR layouts selected based on being the highest, second highest, and third highest ranked layouts in the good performing (top 20%) and the lowest, second lowest, and third lowest ranked layouts in the bad performing (worst 20%) categories for *S. schlegelii*. Results are shown for the addition of 5, 10, 25, and 50 ARs under Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs.

each row of Figure 14 for Best-fit, Supplementary Figure S23 for Intermediate, and Supplementary Figure S24 for Extreme models). This was expected, as cells were randomly selected so that more ARs increased the chance of ARs being closer to other ARs but there were too few boulders on the grid for their rare proximity to an AR to cause major differences among layouts.

For the Best-fit models of *C. japonica*, *S. schlegelii*, and *H. otakii*, the values of D_k and B_k for good performing ARs were similar or lower than those for bad performing ARs for 5 and 10 ARs (Figures 14A–L). However, deployments of 25 and 50 ARs did not exhibit clear differences due to saturation effects. Sea Star showed the reverse pattern of generally higher values for D_k and B_k with good performing ARs (5 and 10 ARs in Figures 14M–P). Good performance for Sea stars meant low occurrence probabilities and so farther away from ARs and boulders would contribute to good performing layouts.

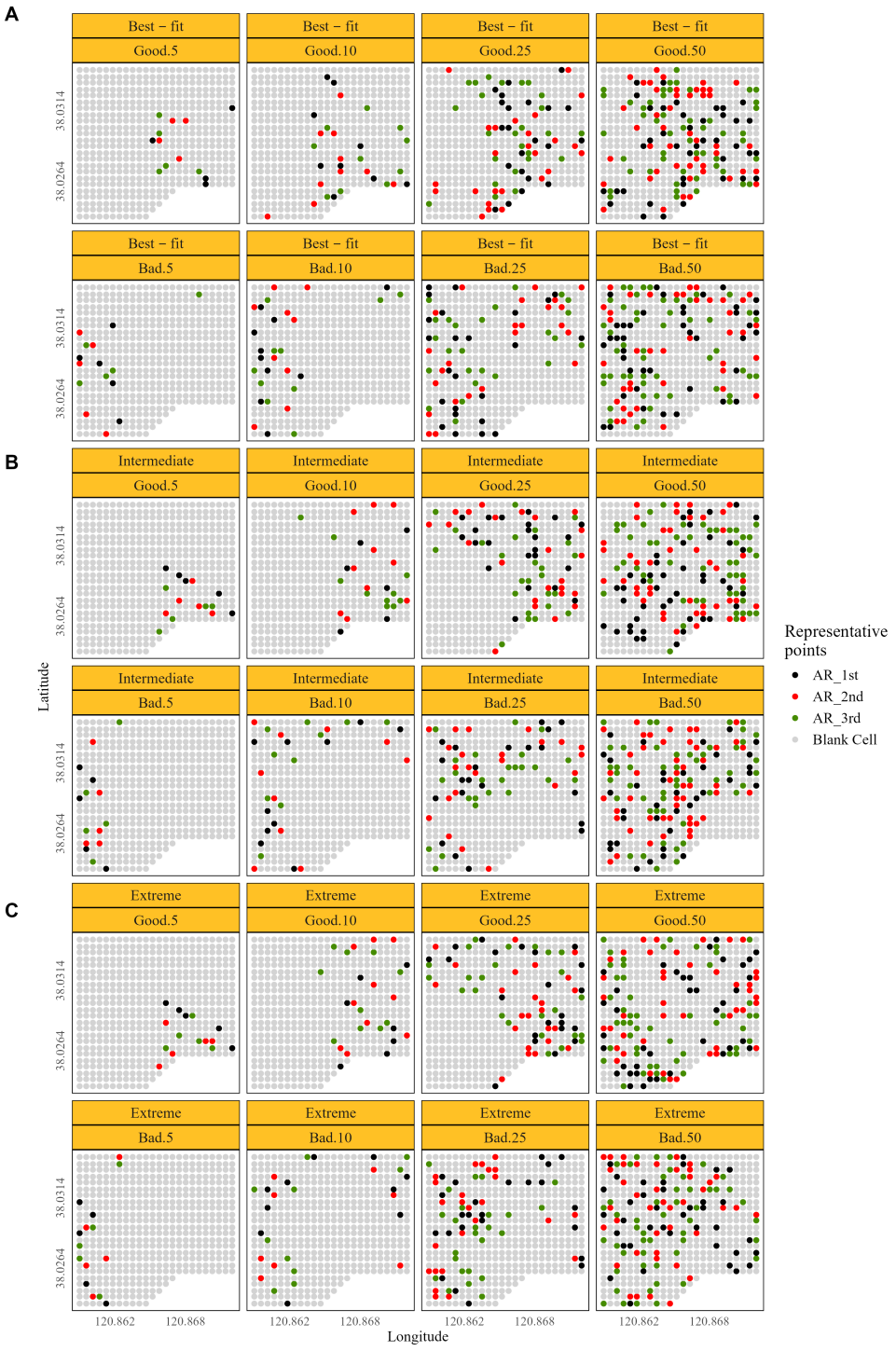

FIGURE 11
Representative AR layouts selected based on being the highest, second highest, and third highest ranked layouts in the good performing (top 20%) and the lowest, second lowest, and third lowest ranked layouts in the bad performing (worst 20%) categories for *H. otakii*. Results are shown for the addition of 5, 10, 25, and 50 ARs under Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs.

Typical D_k values for the desired focal species with the Best-fit models ranged from 200–300 m with 5 ARs and decreased to slightly below 100 m with 50 ARs. Increasing the number of ARs reduced the variation in D_k and B_k across different spatial layouts, also suggesting a compression effect driven by denser AR deployment saturating the AR effect on the grid. Lower B_k values for good performing ARs with 5 and 10 ARs indicated that proximity to local boulder areas enhanced the occurrence probabilities of the three (not Sea star) species, while reducing the occurrence probability of Sea star. The average distances were same (about 250 m) for increasing ARs because there were so few boulders on the grid. As expected, some species exhibited converse patterns with 25 and 50 ARs, reflecting the uncertainty and reversal observed in the representative layout maps. For

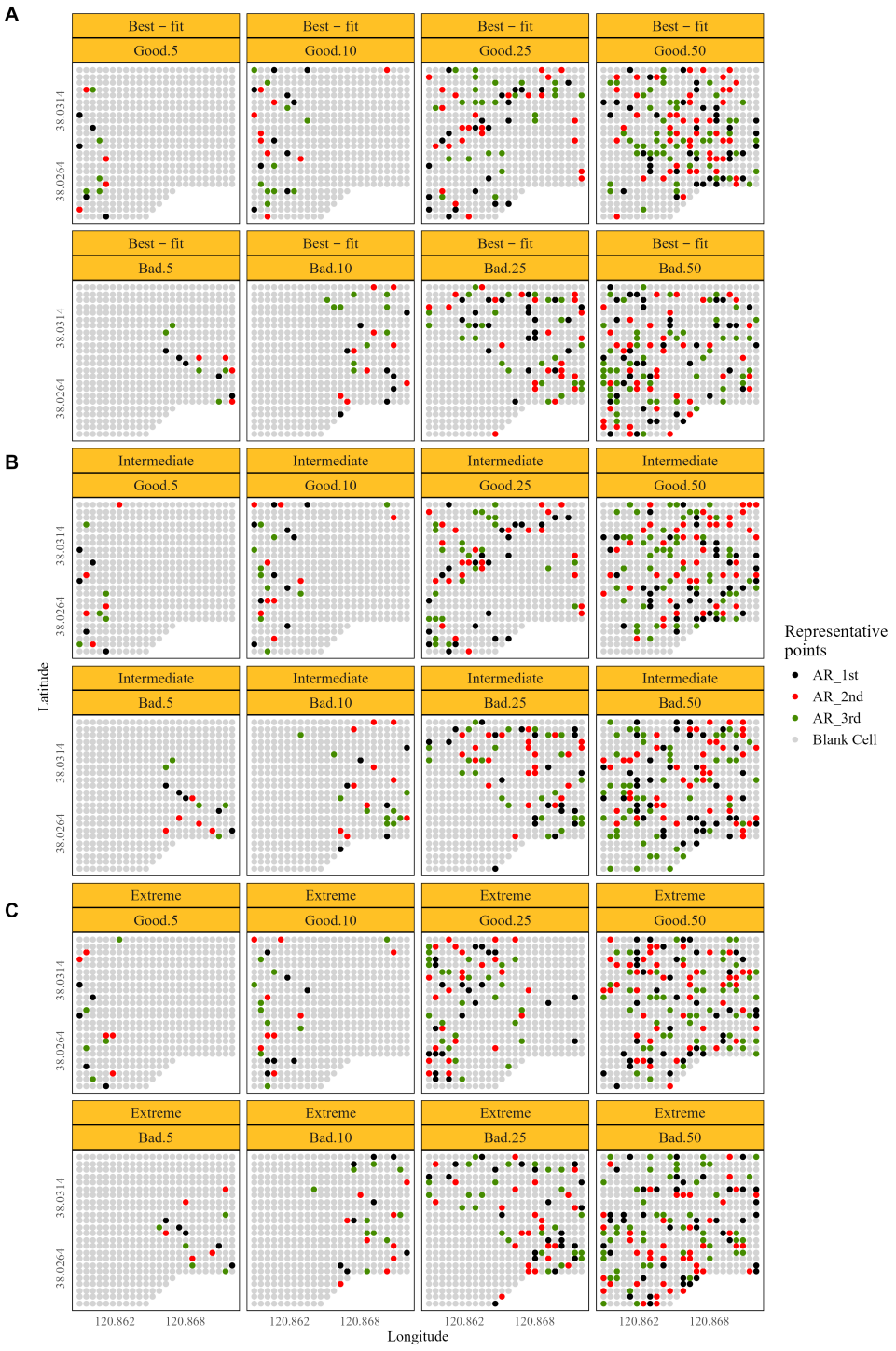

FIGURE 12
Representative AR layouts selected based on being the highest, second highest, and third highest ranked layouts in the good performing (top 20%) and the lowest, second lowest, and third lowest ranked layouts in the bad performing (worst 20%) categories for *Sea star*. Results are shown for the addition of 5, 10, 25, and 50 ARs under Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs.

example, *S. schlegelii* showed a converse pattern with the Best-fit model when comparing B_k values between 5 and 50 ARs (Figures 14H, E), mirroring trends observed in the representative layout maps.

In terms of bottom habitat usage, Best-fit models revealed similar distinctions between good and bad performing ARs across the three desirable species. Based on the representative layouts and the spatial distribution of substrate types (Figure 2B), habitat usage showed a consistent pattern of higher performance in preferred bottom habitats of gravel, rubble, and boulder and bad performance in mud (Figures 15A–L). However, Sea star exhibited an opposite pattern of habitat usage, with a preference for mud type (Figures 15M–P). This contrast was based on the different evaluation metric applied to sea star, where layout performance

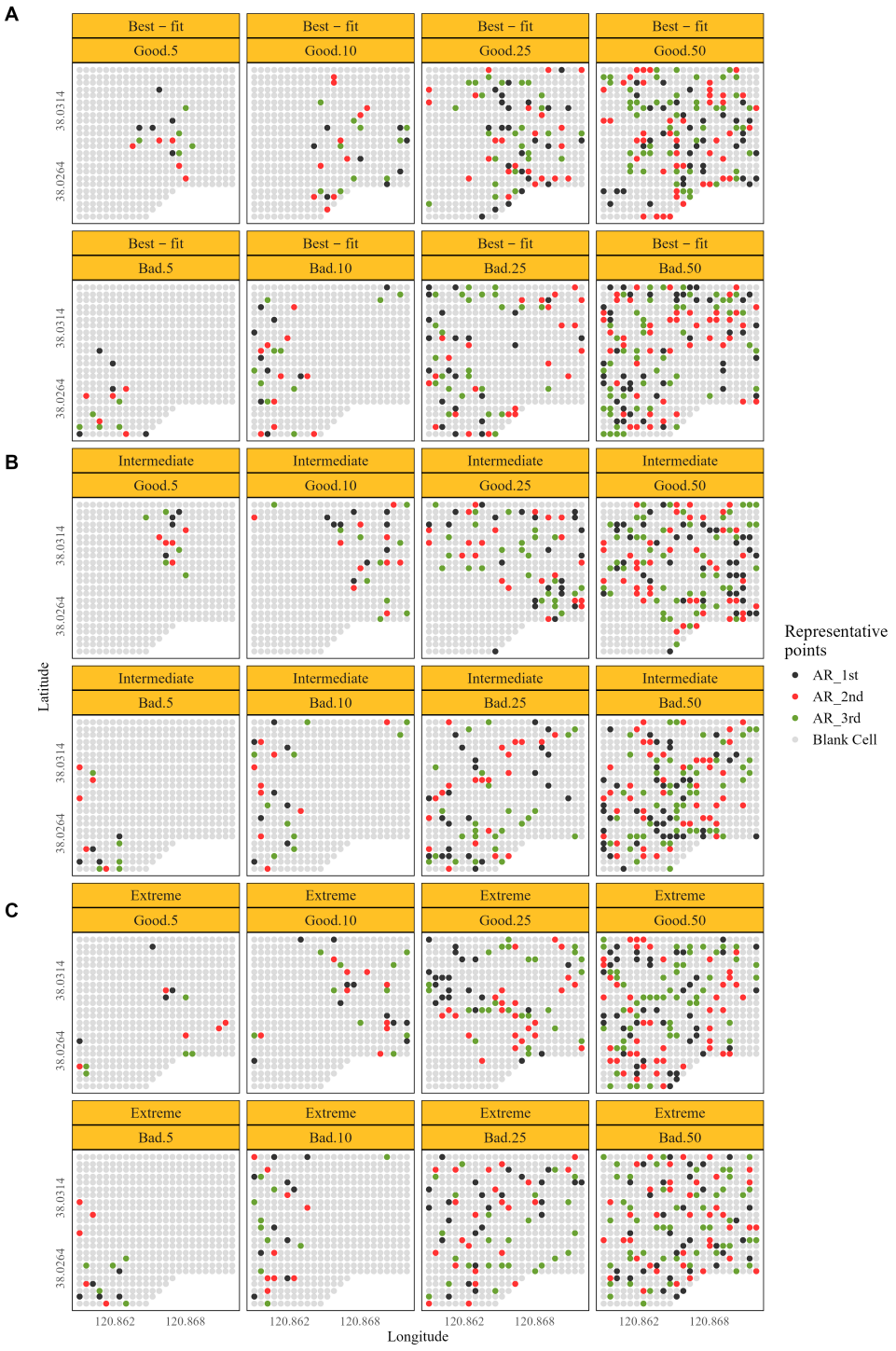

FIGURE 13
Representative AR layouts selected based on being the highest, second highest, and third highest ranked layouts in the good performing (top 20%) and the lowest, second lowest, and third lowest ranked layouts in the bad performing (worst 20%) categories for the geometric mean of occurrence probability across four focal species. Results are shown for the addition of 5, 10, 25, and 50 ARs under Best-fit **(A)**, Intermediate **(B)**, and Extreme **(C)** SDMs.

was ranked using 1 – occurrence probability, rather than occurrence probability as used for the desirable species. The difference between good and bad performing ARs was strongest with 5 ARs (>50% of changes) and progressively became flatter between good and bad with increasing ARs (<10% of changes). Thus, there was reduced ability to discern clear habitat selection at higher AR densities. The Intermediate and Extreme models for D_k, B_k, and habitat usage (Supplementary Figures 24-S27) demonstrated patterns consistent with those observed in the Best-fit model.

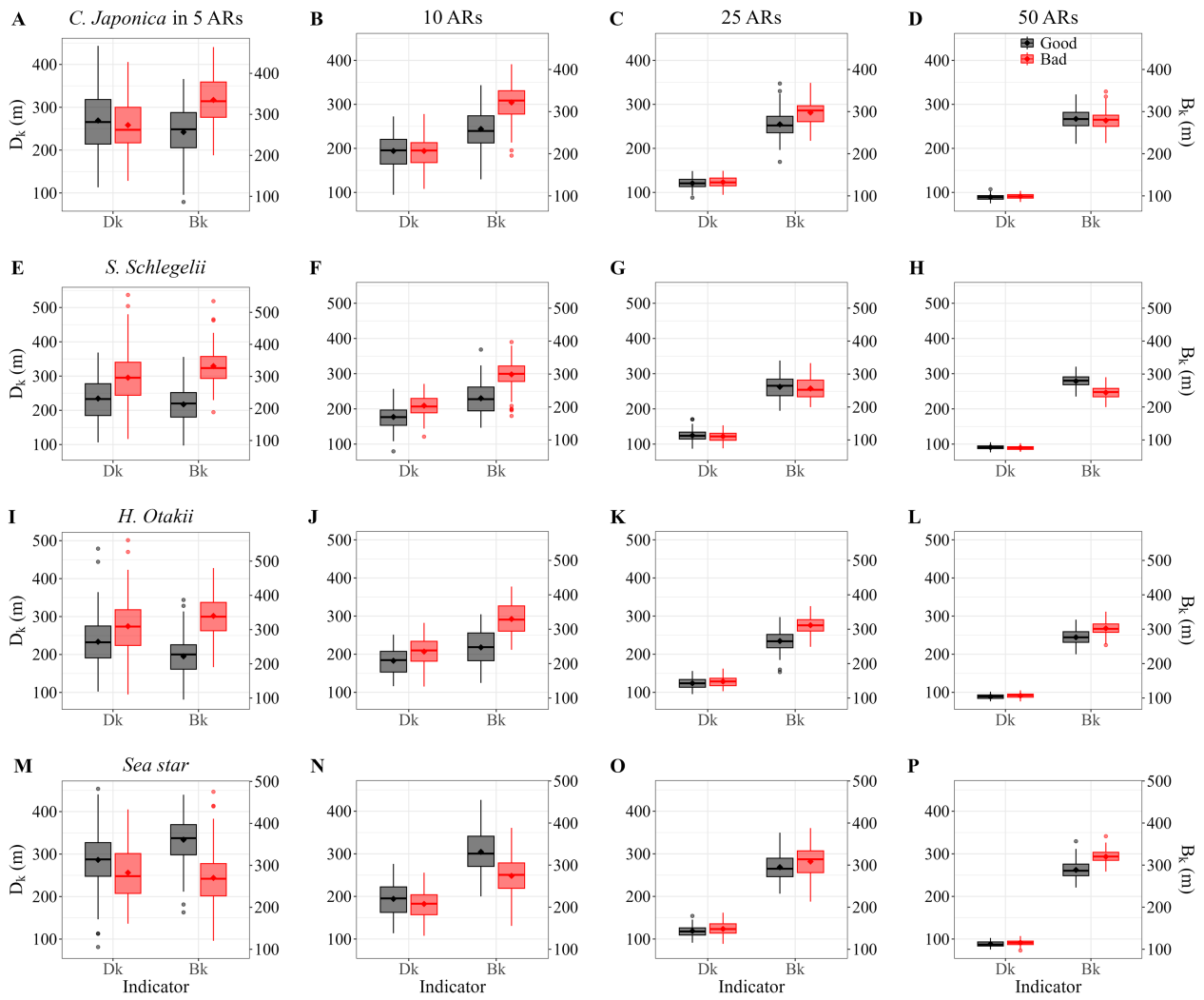

FIGURE 14
Best-fit SDMs: Distribution of D_k (left y-axis) and B_k (right y-axis) for good (top 20% of the layouts, in black) and bad (bottom 20% of the layouts, in red) categories based on the occurrence probability of *C. japonica* **(A–D)**, *S. schlegelii* **(E–H)**, *H. otakii* **(I–L)**, and Sea star **(M–P)**. Rows represent different species, and columns correspond to the addition of 5, 10, 25, and 50 ARs. Boxplots display the overall mean as a diamond-shaped point. The box represents the interquartile range (IQR), with the lower edge corresponding to the 25th percentile and the upper edge to the 75th percentile. The whiskers extend from the box to show the range of data within a certain distance from the quartiles (1.5 times the IQR), and any data points beyond this range are plotted individually as dots.

4 Discussion

One of the objectives of many of China's artificial reefs are to enhance ecological conditions and fisheries (Wang et al., 2009; Zhang et al., 2010). Our analysis explored how the number and spatial layout of ARs can affect the distribution of focal species and our results can be used to help improve the efficiency, productivity, and expectations of AR deployments (Coll et al., 1998). Advancing a method that allows for quantitatively evaluating the responses of multiple species to AR deployments has both economic and ecological importance, especially given the increasing use of ARs worldwide (Lima et al., 2019). This study started with a field survey, then established versions of the SDMs most reflective of the data (Best-fit) and that included different magnitude of interaction effects between ARs of connectivity and competition (Intermediate and Extreme), and finally used these SDMs with simulation analysis to evaluate the performance of different AR layouts on fisheries and undesirable species. We identified high-performing layouts based on the averaged probability of occurrences of each focal species and the geometric mean of their averaged probabilities. The general approach used here with the Intermediate and Extreme versions of constructing artificial datasets to reflect and amplify characteristics under-represented in the observed data is widely used in data mining (Macià et al., 2013), plant distribution probability prediction (Austin et al., 2006), and the supplication of fishery evaluation models (Mesnil et al., 2009). The overall general approach used here of simulation of alternative versions of the habitat to explore performance could also apply to other situations that involve the design of habitats such as in ecosystem restoration or the establishment of marine protected areas.

4.1 Selecting focal species

The bottom habitat type score significantly affected all focal species, making them species to focus on in this analysis. In

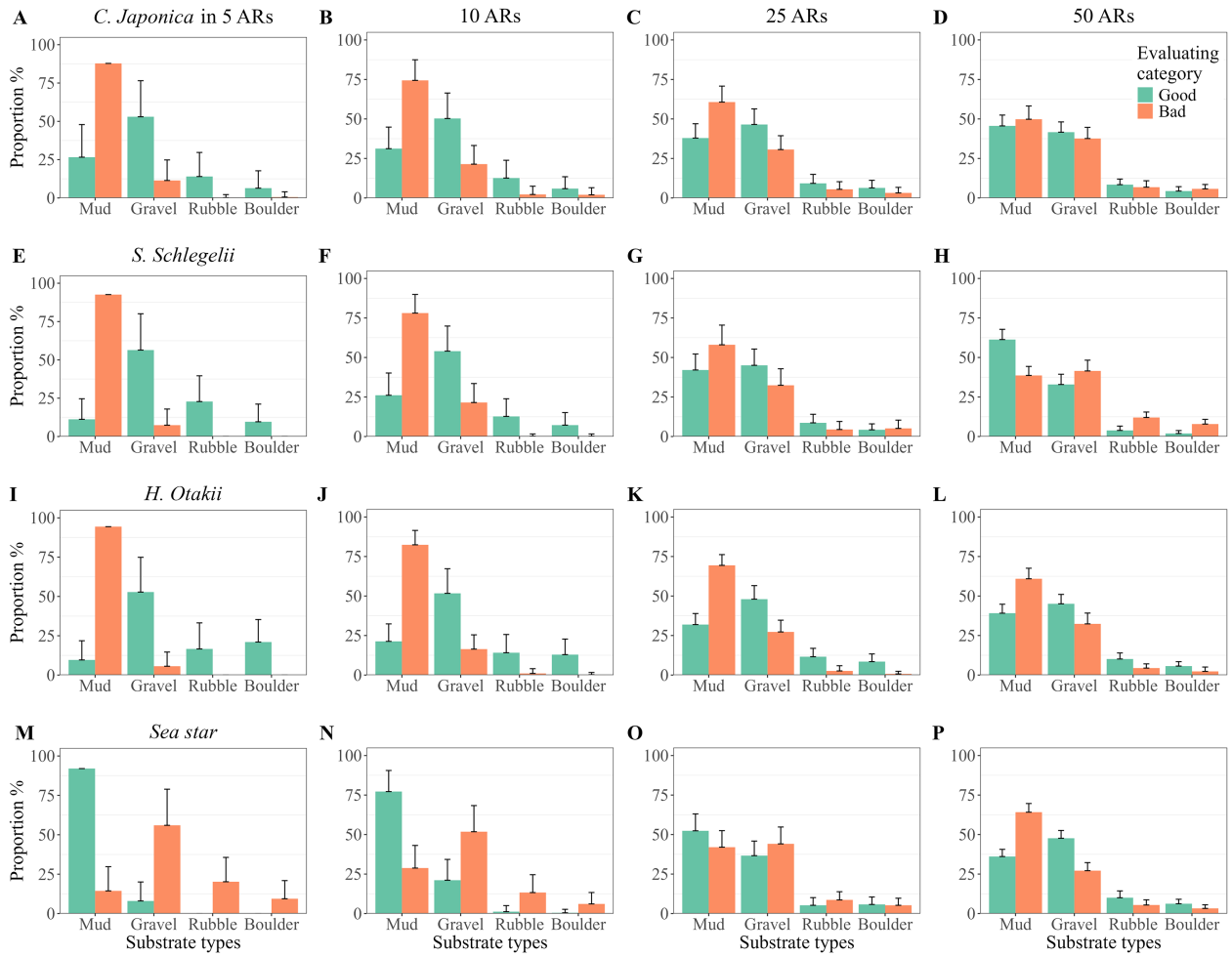

FIGURE 15
Best-fit SDMs: Usage percentages of original substrates (mud, gravel, rubble, boulder) for layouts in the good and bad performing categories for *C. japonica* **(A–D)**, *S. schlegelii* **(E–H)**, *H. otakii* **(I–L)**, and Sea star **(M–P)**. Rows represent the different species, and columns correspond to the addition of 5 ARs, 10 ARs, 25 ARs, and 50 ARs. Bar heights represent the mean values, with error bars representing the mean plus one standard deviation.

addition, *C. japonica*, *S. schlegelii*, and *H. otakii* are the main economic species in the study area and throughout northeastern China (Yu et al., 2005; Huo et al., 2002; Sungju et al., 2014; Yu et al., 2020) and their biomass is often evaluated as part of the site selection process before deploying ARs. All four of our focal species associate with rocky habitat and show high reef dependence, increased tolerance of environmental conditions, a preference for shade areas, and diverse diets (Yu et al., 2020). These species readily occupy a new deployment of ARs.

Regardless of the specific goals of the AR deployment, it is also important to consider the potential impact of AR deployments on undesirable species and how they may respond to changes to the local environment (Weijerman et al., 2018). Based on our analysis, undesirable species (represented by Sea star) had occurrence probabilities comparable to those of the desired focal species (Figures 6, 7). Therefore, future planned AR deployments should take into consideration the effect of the reef installation on the abundance of both commercially important species and undesirable species (Byrne et al., 1997; Millers, 2015).

In addition, care is needed to minimize other unwanted impacts that can be caused by deploying large numbers of AR monomers. For example, ARs can disturb invertebrate communities occupying mud or gravel substrate (Wilson, 1990) and can impact lower trophic level species that provide the food base for commercially important species (Stallings, 2009). Such negative impacts can be considered with candidate layouts to ensure these unintended effects are minimized but may need to be traded-off with any concomitant reductions in the desirable species. Additional site-specific data may be needed to ensure there is adequate information to document and sufficiently quantify potential indirect and untended impacts.

4.2 Accuracy of SDMs

The available survey data for estimating the SDMs had some limitations that could affect the accuracy of our SDMs. Classifying the bottom type was restricted to spot observations and therefore required extensive interpolation. Even though we collected data for this purpose, the resulting fitted SDMs can only reflect a snapshot of the various relationships (Guisan and Thuiller, 2010; Guisan and Zimmermann, 2000). We present in Section 4.3 the evidence supporting our use of the first 2 years of the four-year field study

for estimating the SDMs. The first two years includes the transition period of organisms adjusting to the new habitat availability. Because of the field conditions and clustered location of existing ARs (all roughly equidistant), the data did not provide much information on interspecific interactions or neighborhood effects of ARs. All simulations assumed the presence of the clustered deployment of the existing 24 ARs, which could potentially introduce bias in how species used the placed AR monomers nearby.

In particular, using the SDMs to predict probability of occurrence when a single AR monomer is assumed placed in a cell and no other ARs are allowed to be in the same cell is a strong assumption. The implication is that ARs must be at least 50 m apart. Yet, the placement of the original 24 ARs included shorter distances than 50 m (i.e., some were as close as 10 m). Granneman and Steele (2015) have previously investigated species assemblage composition over time, demonstrating similarity in usage across different distances between ARs (14 to 168 m) among single AR monomers and cluster AR deployment. There are multiple examples of other AR layouts that used inter-AR distances of 50 m or longer. Typical studies used either 50 m between ARs or clusters (e.g., Lima et al., 2020; Gatts et al., 2015) or 100 m (e.g., dos Santos et al., 2010). Tessier et al. (2014) focused on 3 types of ARs located along the Gulf of Lion coastline in the French Mediterranean Sea, and used 50 m or 200 m distances, depending on the type of AR. Our use of limiting a spatial cell to housing one AR is reasonable and needs to be evaluated for other applications.

We improved the SDM performance by filtering collinear and non-significant environmental variables using a combination of RDA and Random Forest methods. However, two goodness-of-fit measures (RMSE, AUC) of the SDMs showed only moderate performance. The AUC values for Sea star are considered relatively low (≈0.6 in Supplementary Table 5) compared to other studies (Velásquez-Tibatá and Graham, 2013; Gaudreau et al., 2018), and other candidate SDMs we investigated also showed moderate (not exceptional) performance (AUC<0.7). We attribute the relatively low performance of the Sea star models to the large number of presence observations throughout the whole area. Once transformed using a Bernoulli distribution, presence data generates many values of one, leading to a tendency to form a presence-only dataset that introduces bias and uncertainties (Stephenson et al., 2021). Although the three desirable focal species showed relatively higher values of AUC, the values for RMSE (0.39-0.48) were also relatively high. Such model fits should be viewed with some caution and require an independent assessment of their accuracy before using them to inform management decisions (Costa et al., 2014).

A likely contributor to the moderate performance of the SDMs with the rockfishes is their behavioral movement that depends greatly on the complex habitat and that created detection problems with our sampling methods. The fitted SDMs recognized the high affinity of the rockfishes to the boulder or ARs (Supplementary Figures 6-S9), while the sampling methods used in the field survey (trap and visual census) had reduced catchability for these species. Rockfishes are renowned for their habit of hiding in crevices or shaded areas within the reef (Garner et al., 2019). Our visual survey method was better positioned to detect the rockfishes but required divers to operate underwater. Due to the turbidity of water in northern China, clear visibility necessitated the use of powerful flashlight illumination, resulting in reduced detectability of the fish. Additionally, differences in the photographic proficiency of each diver likely influenced the consistency of detection. Therefore, although the SDMs for rockfishes had relatively higher AUC values than for the other species, detectability issues resulted in relatively lower RMSE values.

In general, our analysis using SDMs, like many other analyses, is also subject to several well-known uncertainties, such as point-locality data, present climate or environmental interpolations, future environmental projections, and adherence to the assumptions of the statistical modeling algorithms (Velásquez-Tibatá and Graham, 2013). Rather than attempt to use many environmental variables in the SDMs, we fit the variables, and then normalized all but depth and substrate type when evaluating the performance of ARs. While not a fully satisfying solution, we opted for using the reduced set of environmental variables for ARs because they always had a very strong influence and they were well-estimated for the cells of the grid.

4.3 Use of the field survey data

The use of only the first two years of field survey data (Yu et al., 2022) can create bias because the organisms are still adjusting to the presence of the new habitat. The sampling surveys we used included observations immediately following AR deployment and continued until two years post-initial AR installation. Other studies have investigated recruitment to ARs and showed that the highest variance was observed in the first 6 months and then densities or abundances showed only moderate increases in the subsequent years (Alevizon and Gorham, 1989; Bohnsack and Talbot, 1980; Duffy-Anderson et al., 2003; Lowry et al., 2014). Lowry et al. (2014) documented that species diversity tended to stabilize once the reef reached 21–24 months since deployment. Yu et al. (2022), using all four years of survey data, indicated the reef age has a lesser influence on fish utilization of the artificial habitat compared to the larger effects of bottom type score and distance to the nearest AR. In addition, the presence of the ARs near boulders, which were present for a long time in the system, accelerates the trajectory of fish usage of the ARs towards their longer-term usage patterns. Boulders act as buffer zones for reef-dependent species, and provide temporary habitat for species being attracted to or deterred from using the newly deployed ARs (Alevizon and Gorham, 1989; Fowler and Booth, 2012; Rosemond et al., 2018). While the first few months after deployment are a transition period, the species of this study likely approach typical usage of ARs soon after. Thus, the use of the first two years of survey data, which its high level of consistency and similar SDM variables to models fit to all four years, is a reasonable basis for estimating the SDMs used in this analysis.

During the construction of the Best-fit, Intermediate, and Extreme models, we focused on simulating AR layouts. We emphasized geographical variables such as bottom type score, Dis.5 (distance to nearest AR), and depth by altering the bottom type score to represent AR deployment, with subsequent adjustments to Dis.5. Other environmental variables like temperature, dissolved oxygen (DO), and chlorophyll a were held

constant and not adjusted for the presence of a placed AR. This approach, which solely relies on altering the bottom type score to simulate AR deployment, was also used by others in simulating AR effects on the local environment (Schuh et al., 2023; Smith et al., 2015). Effects of AR placement on these other environmental variables would not affect the results of our analysis because these variables were generally less influential than the bottom type score and Dis.5.

4.4 Considerations for AR deployments

Our results provide preliminary recommendations and considerations for deployment of ARs in the study area and, when viewed generally, topics for consideration for the deployment of ARs in other regions. The implications were similar across Best-fit, Intermediate, and Extreme versions of the SDMs for some aspects of deployment but differed for other aspects. The three versions of the SDMs generated different baseline maps but still maintained original high occurrence probability areas as observed in the Best-fit model, and so the resulting evaluations of layouts should be robust.

Our analysis illustrated a general guide (i.e., Best-fit) for adding new ARs in systems like Bohai Strait, with its oceanography and ecology and centrally-located pre-deployment 24 ARs. Our results suggest that when planning a low number of ARs, they should ideally be spaced about 100–200 meters apart. This recommendation aligns with the concept of a 30-meter ecological 'halo' for individual ARs (Reeds et al., 2018), which assumed high species abundances. Our findings may differ somewhat due to differences in the areas and the sizes of the AR monomers between studies.

Another factor to consider when planning a deployment of ARs is the foraging activities of the reef-dependent species. The foraging behavior of reef-dependent species around reefs is closely linked to the distribution and productivity of the surrounding biota (Posey and Ambrose, 1994). The size and degree of isolation of neighboring reef structures are likely to influence the extent and intensity of changes in the prey base associated with a reef (Zalmon et al., 2014). Specifically, rockfish typically have foraging distances of 50–200 meters from their refuge (Tolimieri et al., 2009), which supports our findings that appropriate inter-AR distances within a layout of ARs should offer sufficient space and refuge for these active but highly localized foraging behavior species. However, inter-AR distances become less of a consideration when numerous ARs are deployed because most of the grid becomes available for foraging (as depicted in the good categories in Figures 9-13).

Our analysis also illustrated another important consideration of the expectations for the spatial scale of the responses to the ARs. While occurrence probabilities may be unaffected or even decrease with increasing ARs, this can be offset by the large area affected. Managers may be interested in the total biomass in an area like our grid and, under high interactions, would need to trade off lowered occurrence probabilities with a broader spatial scale response. The change from baseline (no added ARs) becomes important as a basis for estimating grid-wide changes in occurrence probabilities and presumably also in biomass. This consideration is supported by Campbell et al. (2011) using individual-based modeling of reef fish communities. They found that while increasing the number of ARs enhanced local fish biomass and abundance, the effect exhibited diminishing returns as increasing spatial area was considered.

However, if the purpose of the AR deployment focuses on localized responses of specific species, our results showed that under some conditions, there could be important differences in where best to put ARs based on a single species versus all of the species. Focus on a single species can be complicated; the elbow shaped bad ARs pattern for *S. schlegelii* creates overlap of areas of good and bad performing cells (second from right column in Figure 10). Our analysis further showed similar challenging situations when the target is one species versus all four species. A good example is deploying 50 ARs under little interactions (Best-fit models) for *S. schlegelii* versus the geometric mean across all four species (Figures 10 versus 13). A focus on *S. schlegelii* would result in avoiding the central area (Figure 10A), whereas considering all species would focus on the central area (Figure 13A). Another example is the general tradeoff between the desirable species and Sea stars (Figures 9–13). The contrasting distributions of good performing ARs means that selecting locations only based on the desirable species would also add habitat that is conducive to promoting Sea stars. In some situations, it may be better to sacrifice some benefits for the desirable species to avoid exceeding some threshold of habitat for the Sea stars. More detailed, site-specific analyses may be able to identify specific areas where the tradeoff ensures benefits to some of the ultra-desirable species or species critically needing habitat enhancement while only marginally increasing habitat for Sea stars.

Our analyses also showed the potential importance of knowing the degree of interaction effects between ARs. There are multiple examples when Best-fit to Extreme models modified the spatial distribution of cells that support good performing ARs. For example, with 50 ARs, *C. japonica* and *H. otakii* had good performing ARs shifted from the upper triangle to lower triangle due to an emerging bad central area (leftmost versus rightmost columns in Figures 9, 11). *S. schlegelii* illustrated how increasing interactions can cause the filling in of the areas of cells that support bad performing ARs; with 50 ARs, the Best-fit model predicted a central area that was bad for ARs that then switched to supporting good performing ARs with Extreme (rightmost column in Figure 10). Sea stars showed how increasing interaction effects with 50 ARs can generate a general spreading out of good performing ARs until the Extreme models predicted that good performing ARs were distributed throughout the entire grid (right column in Figure 12).

The variation of considering trade-off between desirable and undesirable species and different effects related to the degree of inter-AR interactions indicate that careful attention to specific arrangement of ARs in deployments is important. In scenarios involving a small number of ARs, the selection of original habitat and proximity to existing ARs or natural reefs play a critical role in influencing species occurrence probabilities, although good performing ARs showed D_k values that were lower than bad for the three desirable species across all three model versions. In

contrast, when a relatively large number of ARs are planned that result in saturation effects, the importance of inter-reef interactions becomes more prominent. Because we intentionally maintained some realism for the Intermediate and Extreme models, they continued to show the consistent responses as observed in Best-fit model for D_k, B_k, and habitat usage (e.g., use of less preferred mud) in good versus bad performing layouts (Figure 14 vs. Supplementary Figure S24 vs. Supplementary Figure S25 and Figure 15 vs. Supplementary Figure S26 vs. Supplementary Figure S27).

A notable converse pattern emerged, where well-performing layouts shifted from a mid-center area toward the grid edges, with increasing numbers of ARs (e.g., *S. schlegelii* in the Best-fit model, top row in Figure 10A, Good.5 to Good.50). This shift was particularly evident when comparing scenarios with 5 and 10 ARs to those with 25 and 50 ARs. With low numbers of ARs, representative layout results consistently indicated that high occurrence probability AR arrangements were concentrated near pre-existing reef areas, especially in substrates such as rubble and boulder. However, with 25 or 50 ARs, the good performing layouts shifted markedly toward areas originally dominated by mud substrate (e.g., Figure 10A for *S. schlegelii*). This apparent reversal is driven by the spatial saturation of spillover effects: once a sufficient number of ARs have been deployed to ensure high coverage across the grid, which particularly central and high-quality areas, the marginal benefit of placing additional ARs in already favorable habitats diminishes.

Additional simulations (not shown) showed that beyond a critical threshold (approximately 25 ARs varied by species and models), additional ARs began to contribute most effectively when deployed in mud-dominated cells. These regions, previously considered low-quality habitats, became leverage points for further increasing occurrence probability due to the absence of prior spillover effects. Deploying ARs in mud areas transformed them into locally favorable environments, elevating their contribution to occurrence probability across the whole grid. Therefore, what appears as a converse pattern in representative layouts is actually an outcome of shifting marginal utility, driven by spillover saturation and the spatial distribution of existing ARs. Future deployment strategies should therefore consider both the spatial extent of the AR deployment and the strategic expansion into lower-quality substrates to maximize overall ecological benefit.

Our approach could be expanded by incorporating physical (hydrodynamic) models into our modelling framework, especially when considering early life stages (transport of eggs and larvae) and generation of hydrodynamic-related variables (e.g., water velocities) as possible additional environmental variables. For reef-dependent or demersal species, spawning often occurs in structurally complex reef habitats that provide shelter and suitable microhabitats. After hatching, eggs and larvae are often transported passively by local hydrodynamic processes, including upwelling and eddy-driven circulation around reef structures (Félix-Hackradt et al., 2014; Gatch et al., 2020). Early life stages can be represented as passive particles undergoing two- or three-dimensional advection by hydrodynamic or physical models, such as particle tracking approaches (Paris et al., 2013; Garnier et al., 2025). Numerous studies have employed physical models to investigate how different AR configurations alter surrounding flow fields, and in turn, affect early-life survival and growth (Tang et al., 2019; Shu et al., 2021; Guo et al., 2025). Hydrodynamic related variables, including vertical, horizontal, and depth-averaged current speed, have been incorporated as environmental covariates in other applications of SDMs (e.g., Fabrizio et al., 2021). Adding physical models to our analysis could greatly broaden the suite of potential covariates we consider in model building.

We used a statistical approach because our study organisms were mostly all adults. This reflected both the purpose of deploying ARs to enhance populations and the sampling data (traps and visual) that was available. Adults exhibit strong swimming capabilities and actively use ARs for feeding, shelter, and spawning, with relatively limited dependence on passive current-driven transport (Nathan et al., 2008; Lett et al., 2009). When AR deployment is considered from a full life cycle perspective, integrating multiple modeling approaches may provide a more comprehensive basis for decision-making. For example, Gouezo et al. (2025) successfully assessed restoration outcomes for early life stages of coral larvae in coral reef ecosystems by combining physical and statistical models. Coupled modeling frameworks that integrate hydrodynamic, biogeochemical, and individual-based models have demonstrated strong potential for quantitatively linking environmental conditions with biological processes (Rose et al., 2010; Curchitser et al., 2015). We anticipate that such integrated, coupled models approaches, particularly those applied at finer spatial scales, will become increasingly important for designing AR layouts.

The spatial resolution of the study grid determines the ecological relevance of the analysis and influences the reliability of model predictions. Selection of the spatial resolution should therefore be consistent with the study objectives, the design of the fish survey, and how well environmental variables can be generated at different resolutions. In this study, a grid interval of 50 m was adopted for the following four reasons. First, the objective of this modeling study was to examine how fish occurrence responds to the number and spatial arrangement of ARs, rather than to resolve fine-scale habitat use around individual reef monomers. Second, the selected resolution is ecologically and operationally relevant. Empirical studies have shown that the influence of reef structures on abundance and biodiversity is strongest within approximately 0–50 m of the reef (Puckeridge et al., 2021; Becker et al., 2023). Also, AR deployment (often in clusters) in practice is typically done to achieve spacing at 50 to 100 m. Third, the 50-m spatial resolution is consistent with the design of the fish survey data. Biological and environmental surveys were conducted using a stratified random sampling design aimed at capturing broad spatial gradients in species occurrence probability around the ARs, rather than for obtaining fine-scale information on occurrence. As a result, the response data from the survey support inference at the 50-m scale but would have high uncertainty if used at finer resolutions. Finally, the simulations of layouts using the SDMs require values of environmental variables for all spatial cells in the grid. All

environmental variables were therefore interpolated to generate complete spatial coverage, but they were derived from limited *in situ* observations and would not support reliable interpolation at finer spatial scales.

The use of spatial resolutions finer than 50 m is an important consideration for future simulation analyses, especially for questions that require fine-scale species-reef interactions. Finer scale resolution (e.g., 10 m or 25 m) may facilitate closer integration of SDMs with physical, hydrodynamic, or bioenergetic models, and may better represent processes operating at early life stages or within structurally complex reef habitats. However, such approaches require survey designs and environmental data with sufficient spatial resolution to support reliable inference. Increasing spatial resolution also introduces challenges, including the need for measurement of fine-scale movement patterns and increased demands to generate high-resolution environmental and biological data (Lechene et al., 2019; Asner et al., 2021). Using a finer resolution (e.g., 10 m or 25 m) in our analysis is unlikely to qualitatively change the primary conclusions but would be expected to modify model-level details and improve the spatial precision of simulated AR layouts. A recent analysis using SDMs showed that increasing spatial resolution (1000, 100, and 10 m) affected which environmental variables were selected and the relative importance of different variables, and generated habitat suitability maps that had similar coarse-scale patterns and summed suitable areas but differed at the spatial cell level (Nisonson et al., 2024).

The cost of AR deployment was not incorporated into our analysis because it was the one-off and sunk cost (Lan and Hsui, 2006) and linearly proportional to the number of ARs. Its economic effect was not associated with the focal species, and the differences across the costs are minor compared to the long-term benefits. We focused on evaluating the performance of spatial layouts based on occurrence probabilities rather than also including cost differences.

4.5 General applicability of using Intermediate and Extreme conditions

We derived the Intermediate and Extreme SDMs from the field data with the intent to assess how a relatively higher level of interactions among ARs (connectivity, competition) than observed in the field data would affect the performance of different spatial layouts. We decided to do this because the SDMs estimated from the field survey data (Best-fit) included little interaction effect. While the field survey data used to estimate the SDMs was reflective of a specific area, the generality of the field survey data was less certain. The resulting survey data likely depends to some extent on the presence of the 24 existing ARs and to the locations of the sampling sites within the area. The generality of survey data has been extensively discussed (Jaiteh et al., 2017; MaChado et al., 2021; Zhou et al., 2011).

Model results with the Intermediate and Extreme versions of the SDMs confirmed that we had captured results under unobserved but realistic and likely near the limit of interactions. The Intermediate models still maintained the main characteristics of the Best-fit model, while the Extreme models started to produce nearly unrealistic results suggesting it represented an upper limit of interaction effects. With the Extreme SDMs, a doughnut-shaped area of high species occurrence probabilities began to emerge at around 100m from the grid center. Species occurrence probability was higher at a suitable distance (100–300 m) from the existing 24 ARs, while it was lower both too close to and farther away from the existing ARs (Figure 6C).

Interactions among artificial reefs (ARs) have been well-documented in field surveys and field experiments. When multiple ARs are combined, the resulting ecological overlap can lead to competition (Campbell et al., 2011; Davis et al., 2015; Leitão et al., 2008). Traditionally, habitats with high structural complexity support more species and individuals, providing competitive refuges or a greater variety of discrete resources (Almany, 2004).

Conversely, deploying ARs too far apart can lead to a loss of ecological connectivity (Keller et al., 2017; Logan and Lowe, 2018; Vega Fernández et al., 2008). dos Santos et al. (2010) observed that fish richness and abundance sharply decreased with increasing distance from the reef, particularly when distances between ARs exceeded 300 meters. Similarly, Smith et al. (2017) found a significantly negative correlation between fish abundance and distance from ARs, consistent with our consideration that lower ecological connectivity would result in a lower probability of species occurrence.

4.6 Next steps

We envision several steps for further analyses. There are several ways to improve the survey data. While additional years and stations are always helpful, monitoring studies on movement and collecting data to refine the actual responses of species to the AR location that capture connectivity and competition effects would be informative. In addition, estimation of SDMs would improve with the refinement of how we survey bottom habitat substrate types (e.g., using multibeam echosounder swath mapping) and also by the use of a coupled hydrodynamic-water quality model to generate consistent environmental variables that can be aggregated into different spatial and temporal scales. In principle, if the relevant environmental variables are available for a spatial grid, then the relevant ecological predictions could be produced over the same grid (Araújo et al., 2011).

Several aspects of the focal species approach could also be refined. Consideration of additional species, especially species of high economic value, or groups of species based on life history and how they use ARs (Nakamura, 1985), would ensure model results are robust. The analyses can be expanded to include other ecological considerations (e.g., bycatch effects), the economic costs of different layouts compared to the benefits expressed in economic terms, and the negative ecological and economic effects of the invasive or undesirable species. Our focal species overlap to some degree in their life history and use of habitat. Future effort could focus on developing joint SDMs that account for multiple species simultaneously (Zhang et al., 2018) and ways to incorporate

movement dynamics into the connectivity and interaction aspects of ARs (Campbell et al., 2011).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The animal study was approved by Institutional Animal Care and Use Committee of Ocean University of China. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

HY: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing – original draft. KR: Conceptualization, Methodology, Writing – review & editing. GF: Data curation, Investigation, Writing – review & editing. YT: Resources, Supervision, Writing – review & editing. AB: Writing – review & editing. JF: Data curation, Investigation, Writing – review & editing. TZ: Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the National Natural Science Foundation of China (Grant No. 42306151), Primary Research and Development Plan of Guangxi Province (AB21220064), and Shandong Postdoctoral Science Foundation (Grant No. SDCXZG202301009). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

Sampling surveys were all conducted by members in Fisheries Technology Laboratory in Ocean University of China. It was truly grateful for the "Changdao Jiayi precious marine product Co., Ltd" (local marine ranching managing company) supported.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2026.1684304/full#supplementary-material

References

Alevizon, W. S., and Gorham, J. C. (1989). Effects of artificial reef deployment on nearby resident fishes. *Bull. Mar. Sci.* 44, 646–661.

Almany, G. R. (2004). Does increased habitat complexity reduce predation and competition in coral reef fish assemblages? *Oikos* 106, 275–284. doi: 10.1111/j.0030-1299.2004.13193.x

Araújo, M. B., Diogo, A., Mar, C., David, N. B., and Wilfried, T. (2011). Climate change threatens European conservation areas. *Ecol. Lett.* 14, 484–492. doi: 10.1111/j.1461-0248.2011.01610.x

Araújo, M. B., and Luoto, M. (2010). The importance of biotic interactions for modelling species distributions under climate change. *Glob. Ecol. Biogeogr.* 16, 743–753. doi: 10.1111/j.1466-8238.2007.00359.x

Asner, G. P., Vaughn, N., Grady, B. W., Foo, S. A., Anand, H., Carlson, R. R., et al. (2021). Regional reef fish survey design and scaling using high-resolution mapping and analysis. *Front. Mar. Sci. Volume* 8. doi: 10.3389/fmars.2021.683184

Austin, M. P., Belbin, L., Meyers, J. A., Doherty, M. D., and Luoto, M. (2006). Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. *Ecol. Model.* 199, 197–216. doi: 10.1016/j.ecolmodel.2006.05.023

Baynes, T. W., and Szmant, A. M. (1989). Effect of current on the sessile benthic community structure of an artificial reef. *Bull. Mar. Sci.* 44, 545–566.

Becker, A., Lowry, M. B., Fowler, A. M., and Taylor, M. D. (2023). Hydroacoustic surveys reveal the distribution of mid-water fish around two artificial reef designs in temperate Australia. *Fish. Res.* 257, 106509. doi: 10.1016/j.fishres.2022.106509

Becker, A., Smith, J. A., Taylor, M. D., McLeod, J., and Lowry, M. B. (2019). Distribution of pelagic and epi-benthic fish around a multi-module artificial reef-field: Close module spacing supports a connected assemblage. *Fish. Res.* 209, 75–85. doi: 10.1016/j.fishres.2018.09.020

Becker, A., Taylor, M. D., and Lowry, M. B. (2017). Monitoring of reef associated and pelagic fish communities on Australia's first purpose built offshore artificial reef. *Ices J. Mar. Sci.* 74, 277–285. doi: 10.1093/icesjms/fsw133

Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity* (New York: John Wiley & Sons). doi: 10.1002/0471725153

Bohnsack, J. A., and Talbot, F. H. (1980). Species-packing by reef fishes on Australian and caribbean reefs: an experimental approach. *Bull. Mar. Sci.* 30, 710–723.

Byrne, M., Morrice, M. G., and Wolf, B. (1997). Introduction of the northern Pacific asteroid Asterias amurensis to Tasmania: reproduction and current distribution. *Mar. Biol.* 127, 673–685. doi: 10.1007/s002270050058

Caddy, J. F. (2011). *How artificial reefs could reduce the impacts of bottlenecks in reef fish productivity within natural fractal habitats* Vol. 332 (Boca Raton: Artificial reefs in fisheries management, CRC Press). doi: 10.1201/b10910-5

Campbell, M. D., Rose, K., Boswell, K., and Cowan, J. (2011). Individual-based modeling of an artificial reef fish community: Effects of habitat quantity and degree of refuge. *Ecol. Model.* 222, 3895–3909. doi: 10.1016/j.ecolmodel.2011.10.009

Cannizzo, Z. J., Lang, S. Q., Benitez-Nelson, B., and Griffen, B. D. (2020). An artificial habitat increases the reproductive fitness of a range-shifting species within a newly colonized ecosystem. *Sci. Rep.* 10, 554. doi: 10.1038/s41598-019-56228-x

Castège, I., Milon, E., Fourneau, G., and Tauzia, A. (2016). First results of fauna community structure and dynamics on two artificial reefs in the south of the Bay of Biscay (France). *Estuar. Coast. Shelf Sci.* 179, 172–180. doi: 10.1016/j.ecss.2016.02.015

Charbonnel, E., Serre, C., Ruitton, S., Harmelin, J. G., and Jensen, A. (2002). Effects of increased habitat complexity on fish assemblages associated with large artificial reef units (French Mediterranean coast). *ICES J. Mar. Sci.* 59, S208–S213. doi: 10.2307/1959434

Chittaro, P. M., Usseglio, P., and Sale, P. F. (2005). Variation in fish density, assemblage composition and relative rates of predation among mangrove, seagrass and coral reef habitats. *Environ. Biol. Fish.* 72, 175–187. doi: 10.1007/s10641-004-9077-2

Coll, J., Moranta, J., Renones, O., Garcia-Rubies, A., and Moreno, I. (1998). Influence of substrate and deployment time on fish assemblages on an artificial reef at Formentera Island (Balearic Islands, western Mediterranean). *Hydrobiologia* 385, 139–152. doi: 10.1023/A:1003457810293

Costa, B., Taylor, J. C., Kracker, L., Battista, T., and Pittman, S. (2014). Mapping reef fish and the seascape: using acoustics and spatial modeling to guide coastal management. *PloS One* 9, e85555. doi: 10.1371/journal.pone.0085555

Curchitser, E. N., Rose, K. A., Ito, S.-i., Peck, M. A., and Kishi, M. J. (2015). Combining modeling and observations to better understand marine ecosystem dynamics. *Prog. Oceanogr* 138, 327–330. doi: 10.1016/j.pocean.2015.11.001

Davis, W. T., Drymon, J. M., and Powers, S. P. (2015). Spatial and Dietary Overlap Creates Potential for Competition between Red Snapper (Lutjanus campechanus) and Vermilion snapper (Rhomboplites aurorubens). *PloS One* 10, e0144051. doi: 10.1371/journal.pone.0144051

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x

dos Santos, L. N., Brotto, D. S., and Zalmon, I. R. (2010). Fish responses to increasing distance from artificial reefs on the Southeastern Brazilian Coast. *J. Exp. Mar. Biol. Ecol.* 386, 54–60. doi: 10.1016/j.jembe.2010.01.018

Duffy-Anderson, J. T., Manderson, J. P., and Able, K. W. (2003). A characterization of juvenile fish assemblages around man-made structures in the New York; New Jersey Harbor estuary, U.S.A. *Bull. Mar. Sci.* 72, 877–889.

Elith, J., and Leathwick, J. R. (2007). Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Divers. Distrib.* 13, 265–275. doi: 10.1111/j.1472-4642.2007.00340.x

Elith, J., Leathwick, J. R., and Hastie, T. (2008). A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813. doi: 10.1111/j.1365-2656.2008.01390.x

Fabricius, K., De'ath, G., McCook, L., Turak, E., and Williams, D. M. (2005). Changes in algal, coral and fish assemblages along water quality gradients on the inshore Great Barrier Reef. *Mar. pollut. Bull.* 51, 384–398. doi: 10.1016/j.marpolbul.2004.10.041

Fabrizio, M. C., Tuckey, T. D., Bever, A. J., and MacWilliams, M. L. (2021). The extent of seasonally suitable habitats may limit forage fish production in a temperate estuary. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.706666

Félix-Hackradt, F. C., Hackradt, C. W., Treviño-Otón, J., Pérez-Ruzafa, A., and García-Charton, J. A. (2014). Habitat use and ontogenetic shifts of fish life stages at rocky reefs in South-western Mediterranean Sea. *J. Sea Res.* 88, 67–77. doi: 10.1016/j.seares.2013.12.018

Fielding, A. H., and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49. doi: 10.1017/S0376892997000088

Fowler, A. M., and Booth, D. J. (2012). How well do sunken vessels approximate fish assemblages on coral reefs? *Conserv. implications vessel-reef deployments. Mar. Biol.* 159, 2787–2796. doi: 10.1007/s00227-012-2039-x

Garner, S. B., Boswell, K. M., Lewis, J. P., Tarnecki, J. H., and Patterson, W. F. (2019). Effect of reef morphology and depth on fish community and trophic structure in the northcentral Gulf of Mexico. *Estuar. Coast. Shelf Sci.* 230, 106423. doi: 10.1016/j.ecss.2019.106423

Garnier, S., Murray, R. O. H., Gillibrand, P. A., Gallego, A., Robins, P., and Moriarty, M. (2025). Particle tracking modelling in coastal marine environments: Recommended practices and performance limitations. *Ecol. Model.* 501, 110999. doi: 10.1016/j.ecolmodel.2024.110999

Gatch, A. J., Koenigbauer, S. T., Roseman, E. F., and Höök, T. O. (2020). Assessment of two techniques for remediation of lacustrine rocky reef spawning habitat. *North Am. J. Fisheries Manage.* 41, 484–497. doi: 10.1002/nafm.10557

Gatts, P., Franco, M., dos Santos, L., Rocha, D., de Sá, F., Netto, E., et al. (2015). Impact of artificial patchy reef design on the ichthyofauna community of seasonally influenced shores at Southeastern Brazil. *Aquat. Ecol.* 49, 343–355. doi: 10.1007/s10452-015-9530-7

Gaudreau, J., Perez, L., and Harati, S. (2018). Towards modelling future trends of quebec's boreal birds' Species distribution under climate change. *ISPRS Int. J. Geo-Inf.* 7, 335. doi: 10.3390/ijgi7090335

Gouezo, M., Langlais, C., Beardsley, J., Roff, G., Harrison, P. L., Thomson, D. P., et al. (2025). Going with the flow: Leveraging reef-scale hydrodynamics for upscaling larval-based restoration. *Ecol. Appl.* 35, e70020. doi: 10.1002/eap.70020

Granneman, J. E., and Steele, M. A. (2015). Effects of reef attributes on fish assemblage similarity between artificial and natural reefs. *ICES J. Mar. Sci.* 72, 2385–2397. doi: 10.1093/icesjms/fsv094

Grove, R. S., Sonu, C. J., and Nakamura, M. (1991). "4 - Design and Engineering of Manufactured Habitats for Fisheries Enhancement," in *Artificial Habitats for Marine and Freshwater Fisheries*. Eds. W. Seaman and L. M. Sprague (Academic Press, San Diego), 109–152. doi: 10.1016/S0165-7836(97)00019-2

Guisan, A., and Thuiller, W. (2010). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993–1009. doi: 10.1111/j.1461-0248.2005.00792.x

Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186. doi: 10.1016/S0304-3800(00)00354-9

Guo, C., Tang, Y., Zhu, L., Liang, Z., Lv, S., Li, Z., et al. (2025). Numerical study of the effect of random longitudinal spacing on the flow field effect of unit reef. *Appl. Ocean Res.* 160, 104651. doi: 10.1016/j.apor.2025.104651

Huo, C. L., Wang, J. Y., Han, G. C., and Guan, D. M. (2002). Monitoring biological effects of contamination in fish Hexagrammos otakii along the Dalian Coasts by measurement of EROD activity. *Mar. Environ. Sci.* 21, 5–8.

Jaiteh, V. F., Hordyk, A. R., Braccini, M., Warren, C., and Loneragan, N. R. (2017). Shark finning in eastern Indonesia: assessing the sustainability of a data-poor fishery. *Ices J. Mar. Sci.* 74, 242–253. doi: 10.1093/icesjms/fsw170

Jaxion-Harm, J., and Szedlmayer, S. T. (2015). Depth and artificial reef type effects on size and distribution of red snapper in the northern gulf of Mexico. *N. Am. J. Fish. Manage.* 35, 86–96. doi: 10.1080/02755947.2014.982332

Keith, D. A., Resit, H., Wilfried, T., Midgley, G. F., Pearson, R. G., Phillips, S. J., et al. (2008). Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biol. Lett.* 4, 560–563. doi: 10.1098/rsbl.2008.0049

Keller, K., Smith, J. A., Lowry, M. B., Taylor, M. D., and Suthers, I. M. (2017). Multispecies presence and connectivity around a designed artificial reef. *Mar. Freshw. Res.* 68, 1489–1500. doi: 10.1071/MF16127_AC

Kilfoyle, A. K., Freeman, J., Jordan, L. K. B., Quinn, T. P., and Spieler, R. E. (2013). Fish assemblages on a mitigation boulder reef and neighboring hardbottom. *Ocean Coast. Manage.* 75, 53–62. doi: 10.1016/j.ocecoaman.2013.02.001

Kuklinski, P., Barnes, D. K. A., and Taylor, P. D. (2006). Latitudinal patterns of diversity and abundance in North Atlantic intertidal boulder-fields. *Mar. Biol.* 149, 1577–1583. doi: 10.1007/s00227-006-0311-7

Lan, C. H., and Hsui, C. Y. (2006). The deployment of artificial reef ecosystem: Modelling, simulation and application. *Simul. Model. Pract. Th.* 14, 663–675. doi: 10.1016/j.simpat.2005.10.011

Langhamer, O., Holand, H., and Rosenqvist, G. (2016). Effects of an offshore wind farm (OWF) on the common shore crab carcinus maenas: tagging pilot experiments in the lillgrund offshore wind farm (Sweden). *PloS One* 11, e0165096. doi: 10.1371/journal.pone.0165096

Lawesson, J. E. (2010). Effects of species partition on explanatory variables in direct gradient analysis: A case study from Senegal. *J. Veg. Sci.* 8, 409–414. doi: 10.2307/3237332

Lechene, M. A. A., Haberstroh, A. J., Byrne, M., Figueira, W., and Ferrari, R. (2019). Optimising sampling strategies in coral reefs using large-area mosaics. *Remote Sens.* 11, 2907. doi: 10.3390/rs11242907

Leitão, F., Santos, M. N., Erzini, K., and Monteiro, C. C. (2008). The effect of predation on artificial reef juvenile demersal fish species. *Mar. Biol.* 153, 1233–1244. doi: 10.1007/s00227-007-0898-3

Lett, C., Rose, K. A., and Megrey, B. A. (2009). "Biophysical models of small pelagic fish," in *Climate change and small pelagic fish*. Eds. C. Checkley, C. Roy, J. Alheit and Y. Oozeki (Cambridge University Press, Cambridge), 88–111.

Lima, J. S., Sanchez-Jerez, P., dos Santos, L. N., and Zalmon, I. R. (2020). Could artificial reefs increase access to estuarine fishery resources? Insights from a long-term assessment. *Estuar. Coast. Shelf Sci.* 242, 106858. doi: 10.1016/j.ecss.2020.106858

Lima, J. S., Zalmon, I. R., and Love, M. (2019). Overview and trends of ecological and socioeconomic research on artificial reefs. *Mar. Environ. Res.* 145, 81–96. doi: 10.1016/j.marenvres.2019.01.010

Liu, T. L., and Su, D. T. (2013). Numerical analysis of the influence of reef arrangements on artificial reef flow fields. *Ocean Eng.* 74, 81–89. doi: 10.1016/j.oceaneng.2013.09.006

Logan, R. K., and Lowe, C. G. (2018). Residency and inter-reef connectivity of three gamefishes between natural reefs and a large mitigation artificial reef. *Mar. Ecol. Prog. Ser.* 593, 111–126. doi: 10.3354/meps12527

Lowry, M. B., Glasby, T. M., Boys, C. A., Folpp, H., Suthers, I., and Gregson, M. (2014). Response of fish communities to the deployment of estuarine artificial reefs for fisheries enhancement. *Fisheries Manage. Ecol.* 21, 42–56. doi: 10.1111/fme.12048

MaChado, A. M. S., Giehl, E. L. H., Fernandes, L. P., Ingram, S. N., and Daura-Jorge, F. G. (2021). Alternative data sources can fill the gaps in data-poor fisheries. *Ices J. Mar. Sci.* 78, 1663–1671. doi: 10.1093/icesjms/fsab074

Macià, N., Bernadó-Mansilla, E., Orriols-Puig, A., and Kam Ho, T. (2013). Learner excellence biased by data set selection: A case for data characterisation and artificial data sets. *Pattern Recognit.* 46, 1054–1066. doi: 10.1016/j.patcog.2012.09.022

Mellin, C., Andréfouët, S., and Ponton, D. (2007). Spatial predictability of juvenile fish species richness and abundance in a coral reef environment. *Coral Reefs* 26, 895–907. doi: 10.1007/s00338-007-0281-3

Mesnil, B., Cotter, J., Fryer, R. J., Needle, C. L., and Trenkel, V. M. (2009). A review of fishery-independent assessment models, and initial evaluation based on simulated data. *Aquat. Living Resour.* 22, 207–216. doi: 10.1051/alr/2009003

Mi, C., Huettmann, F., Guo, Y., Han, X., and Wen, L. (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? *Three Asian crane species Models provide supporting evidence. Peerj* 5, e2849. doi: 10.7717/peerj.2849

Millers, K. (2015). Quantifying search and control performance during marine invasive surveys: a case study from Asterias amurensis. The University of Melbourne, Melbourne, VIC, Australia. PhD thesis.

Nakamura, M. (1985). Evolution of artificial fishing reef concepts in Japan. *Bull. Mar. Sci.* 37, 271–278.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., et al. (2008). A movement ecology paradigm for unifying organismal movement research. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19052–19059. doi: 10.1073/pnas.0800375105

Nisonson, H., Kiser, A., Gressler, B., Leight, A., and Young, J. (2024). *Pilot Framework for Fish Habitat Assessments Across Tidal and Non-Tidal Waters in the Patuxent River Basin* Vol. 332 (MD, United States: NOAA technical memorandum NOS NCCOS). doi: 10.25923/4jqw-mw29

Paris, C. B., Helgers, J., van Sebille, E., and Srinivasan, A. (2013). Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. *Environ. Model. Software* 42, 47–54. doi: 10.1016/j.envsoft.2012.12.006

Parris, K. M. (2004). Environmental and spatial variables influence the composition of frog assemblages in sub-tropical eastern Australia. *Ecography* 27, 392–400. doi: 10.1111/j.0906-7590.2004.03711.x

Parsons, D., Shears, N., Babcock, R., and Haggitt, T. (2004). Fine-scale habitat change in a marine reserve, mapped using radio-acoustically positioned video transects. *Mar. Freshw. Res.* 55, 257–265. doi: 10.1071/MF03190

Pearson, R. G., Raxworthy, C. J., Nakamura, M., and Peterson, A. T. (2010). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeography* 34, 102–117. doi: 10.1111/j.1365-2699.2006.01594.x

Perlich, C., and Swirszcz, G. (2011). On cross-validation and stacking: building seemingly predictive models on random data. *ACM Sigkdd Explor. Newslett.* 12, 11–15. doi: 10.1145/1964897.1964901

Posey, M. H., and Ambrose, W. G. (1994). Effects of proximity to an offshore hard-bottom reef on infaunal abundances. *Mar. Biol.* 118, 745–753. doi: 10.1007/Bf00347524

Puckeridge, A., Becker, A., Taylor, M. D., Lowry, M. B., McLeod, J., Schilling, H., et al. (2021). Foraging behaviour and movements of an ambush predator reveal benthopelagic coupling on artificial reefs. *Mar. Ecol. Prog. Ser.* 666, 171–182. doi: 10.3354/meps13691

Punt, M. J., Groeneveld, R. A., van Ierland, E. C., and Stel, J. H. (2009). Spatial planning of offshore wind farms: A windfall to marine environmental protection? *Ecol. Econ.* 69, 93–103. doi: 10.1016/j.ecolecon.2009.07.013

Reeds, K. A., Smith, J. A., Suthers, I. M., and Johnston, E. L. (2018). An ecological halo surrounding a large offshore artificial reef: Sediments, infauna, and fish foraging. *Mar. Environ. Res.* 141, 30–38. doi: 10.1016/j.marenvres.2018.07.011

Renwick, A. R., and Johnston, A. (2012). Modelling changes in species' abundance in response to projected climate change. *Divers. Distrib.* 18, 121–132. doi: 10.1111/j.1472-4642.2011.00827.x

Rose, K. A., Allen, J. I., Artioli, Y., Barange, M., Blackford, J., Carlotti, F., et al. (2010). End-to-end models for the analysis of marine ecosystems: challenges, issues, and next steps. *Mar. Coast. Fish.* 2, 115–130. doi: 10.1577/C09-059.1

Rosemond, R. C., Paxton, A. B., Lemoine, H. R., Fegley, S. R., and Peterson, C. H. (2018). Fish use of reef structures and adjacent sand flats: implications for selecting minimum buffer zones between new artificial reefs and existing reefs. *Mar. Ecol. Prog. Ser.* 587, 187–199. doi: 10.3354/meps12428

Schuh, E., Grilli, A. R., Groetsch, F., Grilli, S. T., Crowley, D., Ginis, I., et al. (2023). Assessing the morphodynamic response of a New England beach-barrier system to an artificial reef. *Coast. Eng.* 184, 104355. doi: 10.1016/j.coastaleng.2023.104355

Shu, A., Rubinato, M., Qin, J., Zhu, J., Sun, T., Yang, W., et al. (2021). The hydrodynamic characteristics induced by multiple layouts of typical artificial M-type reefs with sea currents typical of liaodong bay, bohai sea. *J. Mar. Sci. Eng.* 9, 1155. doi: 10.3390/jmse9111155

Smith, J. A., Cornwell, W. K., Lowry, M. B., and Suthers, I. M. (2017). Modelling the distribution of fish around an artificial reef. *Mar. Freshw. Res.* 68, 1955–1964. doi: 10.1071/MF16019

Smith, J. A., Lowry, M. B., and Suthers, I. M. (2015). Fish attraction to artificial reefs not always harmful: a simulation study. *Ecol. Evol.* 5, 4590–4602. doi: 10.1002/ece3.1730

Stallings, C. D. (2009). Predator identity and recruitment of coral-reef fishes: indirect effects of fishing. *Mar. Ecol. Prog. Ser.* 383, 251–259. doi: 10.3354/meps08004

Stephenson, F., Rowden, A. A., Anderson, O. F., Pitcher, C. R., Pinkerton, M. H., Petersen, G., et al. (2021). Presence-only habitat suitability models for vulnerable marine ecosystem indicator taxa in the South Pacific have reached their predictive limit. *ICES J. Mar. Sci.* 78, 2830–2843. doi: 10.1093/icesjms/fsab162

Strelcheck, A. J., Cowan, J. H., and Shah, A. (2005). Influence of reef location on artificial-reef fish assemblages in the northcentral Gulf of Mexico. *B. Mar. Sci.* 77, 425–440. doi: 10.1029/2004JC002816

Sungju, R., Jaekwon, C., Seonjae, K., Wookmin, P., Taisun, S., and Jaeho, H. (2014). Nutritional characteristics of black rockfish (Sebastes schlegeli) fed a diet of fish skin. *J. Aquac. Res. Dev.* 5. doi: 10.4172/2155-9546.1000239

Tang, Y., Hu, Q., Wang, X., Zhao, F., Huang, L., and Xie, T. (2019). "Evaluation of flow field in the layouts of cross-shaped artificial reefs," in *ASME 2019 38th International Conference on Ocean, Offshore and Arctic Engineering*. (Glasgow, Scotland, UK: American Society of Mechanical Engineers).

Tessier, A., Verdoit-Jarraya, M., Blouet, S., Dalias, N., and Lenfant, P. (2014). A case study of artificial reefs as a potential tool for maintaining artisanal fisheries in the French Mediterranean Sea. *Aquat. Biol.* 20, 255–272. doi: 10.3354/ab00563

Thanner, S. E., McIntosh, T. L., and Blair, S. M. (2006). Development of benthic and fish assemblages on artificial reef materials compared to adjacent natural reef assemblages in Miami-Dade County, Florida. *Bull. Mar. Sci.* 78, 57–70.

Thuiller, W., Lafourcade, B., Engler, R., and Araújo, M. B. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography* 32, 369–373. doi: 10.1111/j.1600-0587.2008.05742.x

Tolimieri, N., Andrews, K., Williams, G., Katz, S., and Levin, P. S. (2009). Home range size and patterns of space use by lingcod, copper rockfish and quillback rockfish in relation to diel and tidal cycles. *Mar. Ecol. Prog. Ser.* 380, 229–243. doi: 10.3354/meps07930

Ushiama, S., Smith, J. A., Suthers, I. M., Lowry, M., and Johnston, E. L. (2016). The effects of substratum material and surface orientation on the developing epibenthic community on a designed artificial reef. *Biofouling*. 32(9), 1049–1060. doi: 10.1080/08927014.2016.1224860

Vapnik, V. N. (2000). "The nature of statistical learning theory,". Ed. V. N. Vapnik (Springer Press, New York, NY, USA). doi: 10.1007/978-1-4757-2440-0

Vega Fernández, T., Anna, G., Badalamenti, F., and Pérez-Ruzafa, A. (2008). Habitat connectivity as a factor affecting fish assemblages in temperate reefs. *Aquat. Biol.* 1, 239–248. doi: 10.3354/ab00027

Velásquez-Tibatá, J., and Graham, C. H. (2013). Effects of climate change on species distribution, community structure, and conservation of birds in protected areas in Colombia. *Reg. Environ. Change* 13, 235–248. doi: 10.1007/s10113-012-0329-y

Vicente, M., Falcão, M., Santos, M. N., Caetano, M., Serpa, D., Vale, C., et al. (2008). Environmental assessment of two artificial reef systems off southern Portugal (Faro and Olhão): A question of location. *Cont. Shelf Res.* 28, 839–847. doi: 10.1016/j.csr.2007.12.009

Vijay, K., Neelamani, S., Nishad, C., and Sahoo, T. (2021). Gravity wave interaction with multiple submerged artificial reefs. *Proc. Institution Mechanical Engineers Part M: J. Eng. Maritime Environ.* 235, 607–622. doi: 10.1177/1475090220960848

Wang, H., Chen, P., Zhang, S., and Jia, X. (2009). Effect on fishery resources multiplication of artificial reefs. *Guangdong Agric. Sci.* 8, 18–21. doi: 10.1016/j.actao.2010.01.006

Wang, H., Yang, F., and Luo, Z. (2016). An experimental study of the intrinsic stability of random forest variable importance measures. *BMC Bioinform.* 17, 60. doi: 10.1186/s12859-016-0900-5

Weijerman, M., Gove, J. M., Williams, I. D., Walsh, W. J., Minton, D., and Polovina, J. J. (2018). Evaluating management strategies to optimise coral reef ecosystem services. *J. Appl. Ecol.* 55, 1823–1833. doi: 10.1111/1365-2664.13105

Wilborn, R., Rooper, C. N., Goddard, P., Li, L., Williams, K., and Towler, R. (2018). The potential effects of substrate type, currents, depth and fishing pressure on distribution, abundance, diversity, and height of cold-water corals and sponges in temperate, marine waters. *Hydrobiologia* 811, 251–268. doi: 10.1007/s10750-017-3492-9

Wilson, W. H. (1990). Competition and predation in marine soft-sediment communities. *Annu. Rev. Ecol. Evol.* S. 21, 221–241. doi: 10.1146/annurev.ecolsys.21.1.221

Yu, H., Fang, G., Rose, K. A., Tang, Y., and Song, X. F. (2022). Examining epibenthic assemblages associated with artificial reefs using a species archetype approach. *Mar. Coast. Fish.* 14, e10206. doi: 10.1002/mcf2.10206

Yu, C. G., Song, H. T., and Yao, G. Z. (2005). The quantity distribution and biological property of Charybdis japonica in the East China Sea. *J. Shanghai Fish. Univ.* 14, 40–45.

Yu, H., Yang, W., Liu, C., Tang, Y., Song, X., and Fang, G. (2020). Relationships between community structure and environmental factors in xixiakou artificial reef area. *J. Ocean Univ. China*. 19, 883–894. doi: 10.1007/s11802-020-4298-3

Zalmon, I. R., de Sa, F. S., Neto, E. J. D., De Rezende, C. E., Mota, P. M., and De Almeida, T. C. M. (2014). Impacts of artificial reef spatial configuration on infaunal community structure - southeastern Brazil. *J. Exp. Mar. Biol. Ecol.* 454, 9–17. doi: 10.1016/j.jembe.2014.01.015

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

Zhang, Y., Xu, Q., Alós, J., Liu, H., Xu, Q., and Yang, H. (2015). Short-term fidelity, habitat use and vertical movement behavior of the black rockfish Sebastes schlegelii as determined by acoustic telemetry. *PloS One* 10, e0134381. doi: 10.1371/journal.pone.0134381

Zhang, S., Xu, M., and Wang, Z. (2010). Review of artificial reef and stock enhancement. *Fish. Modernization* 37, 55–58.

Zhou, S., Smith, A. D., and Fuller, M. (2011). Quantitative ecological risk assessment for fishing effects on diverse data-poor non-target species in a multi-sector and multi-gear fishery. *Fish. Res.* 112, 168–178. doi: 10.1016/j.fishres.2010.09.028