RESEARCH PAPER

Incorporating physiological knowledge into correlative species distribution models minimizes bias introduced by the choice of calibration area

Zhixin Zhang[1,2,3,4] · Jinxin Zhou[5] · Jorge García Molinos[6] · Stefano Mammola[7,8,9] · Ákos Bede-Fazekas[10,11] · Xiao Feng[12] · Daisuke Kitazawa[5] · Jorge Assis[13] · Tianlong Qiu[14] · Qiang Lin[1,3,15]

Abstract

Correlative species distribution models (SDMs) are important tools to estimate species' geographic distribution across space and time, but their reliability heavily relies on the availability and quality of occurrence data. Estimations can be biased when occurrences do not fully represent the environmental requirement of a species. We tested to what extent species' physiological knowledge might influence SDM estimations. Focusing on the Japanese sea cucumber *Apostichopus japonicus* within the coastal ocean of East Asia, we compiled a comprehensive dataset of occurrence records. We then explored the importance of incorporating physiological knowledge into SDMs by calibrating two types of correlative SDMs: a naïve model that solely depends on environmental correlates, and a physiologically informed model that further incorporates physiological information as priors. We further tested the models' sensitivity to calibration area choices by fitting them with different buffered areas around known presences. Compared with naïve models, the physiologically informed models successfully captured the negative influence of high temperature on *A. japonicus* and were less sensitive to the choice of calibration area. The naïve models resulted in more optimistic prediction of the changes of potential distributions under climate change (i.e., larger range expansion and less contraction) than the physiologically informed models. Our findings highlight benefits from incorporating physiological information into correlative SDMs, namely mitigating the uncertainties associated with the choice of calibration area. Given these promising features, we encourage future SDM studies to consider species physiological information where available.

Keywords Bayesian approach · Climate change · Habitat suitability · Physiological knowledge · Species distribution model

Introduction

Climate change affects all ecosystems on Earth—terrestrial (Chen et al. 2011; Mammola et al. 2019a), freshwater (Woodward et al. 2010), and marine alike (Brito-Morales et al. 2020; Levin and Le Bris 2015)—being recognized as one of the five direct drivers responsible for global biodiversity loss (IPBES 2019). As the largest ecosystems on Earth, oceans are now experiencing dramatic changes due to multiple drivers, such as warming, acidification, and deoxygenation (Doney et al. 2012; Poloczanska et al. 2013). These

changes have been reported to adversely affect organisms via diverse ecological mechanisms (Doney et al. 2012; Poloczanska et al. 2013, 2016), of which geographical range shifts of species are one of the most frequently observed outcomes (Doney et al. 2012; Dong et al. 2024; Lenoir et al. 2020; Pinsky et al. 2020). In response to climatic warming, marine organisms are shifting either poleward (Lenoir et al. 2020; Pinsky et al. 2020) or to deeper waters (Dulvy et al. 2008; Pinsky et al. 2013) in search of cooler environments. In a recent global study, Lenoir et al. (2020) demonstrated that, on average, marine species have shifted poleward over the course of the last century six times faster than species on land (5.92 versus 1.11 km per year, respectively). This quick global rearrangement of the mosaic of marine biodiversity can alter the composition and profitability of global fisheries catch (Cheung et al. 2013; Gaines et al. 2018) and affect population stability by creating novel species interactions

---

Zhixin Zhang and Jinxin Zhou contributed equally to this work.

Extended author information available on the last page of the article

---

(Potts et al. 2014; Vergés et al. 2016). For implementing effective management measures, it is critical to obtain a nuanced understanding of how species may respond to future scenarios of climate change.

Correlative species distribution models (SDMs) are tools to estimate species' habitat suitability by describing the statistical relationships between species distribution data and relevant predictors, which have become popular to investigate species response to climate change (Araújo et al. 2019; Feng et al. 2019; Guisan et al. 2017; Qiao et al. 2015; Taheri et al. 2021). Despite their growing popularity, correlative SDMs depend heavily on the quality of species distribution data, which are often susceptible to sampling bias and spatial uncertainty (Araújo et al. 2019; Feng et al. 2019; Hughes et al. 2021; Kramer-Schadt et al. 2013; Marcer et al. 2021). Furthermore, among the recognized limitations of correlative SDMs (e.g., Ryo et al. 2021; Smith et al. 2019; Zhang et al. 2024), their inability to account for species physiological limits is a major drawback (Hof 2021; Kearney and Porter 2009; Liao et al. 2021). Specific environmental preferences and tolerances mean that, although species may occupy a broad spatial extent, the spatial variability in the suitability of those habitats could be large (e.g., Champion et al. 2020). In other words, the habitat suitability of species occurrence records is not equal: some occurrences are from species optimal habitats, whereas others may be collected from marginally suitable regions. Given these biases in species distribution data, correlative SDMs without considering species' physiological knowledge are at high risk of failing to properly capture the environmental requirements of target species.

Compared with correlative SDMs, mechanistic SDMs directly use the physiological requirements of species to reasonably describe the realized niche (Dong et al. 2024; Kearney and Porter 2009; Peterson et al. 2015). However, mechanistic models require a number of physiological measurements for proper parameterization. These data are often costly and time-consuming to acquire, limiting the wide application of these models (Mammola et al. 2021; Peterson et al. 2015). Given the growing availability of physiological information, in addition to using species' physiological constraints to guide the selection of appropriate biologically relevant predictors, it is timely to explore the possibility of incorporating physiological knowledge directly into the fitting of correlative SDMs. In this regard, Bayesian SDMs, which are based on Bayes' theorem (Bayes 1764), make statistical inferences by incorporating species' physiological information as prior knowledge. Thus far, different Bayesian approaches have been developed and successfully applied in various studies, improving SDM performance and producing more reliable results (e.g., Brewer et al. 2016; Golding and Purse 2016; Talluto et al. 2016; Zhu et al. 2021). For example, Feng et al. (2020) tested the importance of integrating

physiological knowledge into SDMs for zebra mussel *Dreissena polymorpha* (Bivalvia: Dreissenidae) and reported that physiologically informed models improve model extrapolation ability. Similarly, Gamliel et al. (2020) compared the performance of physiologically informed and naïve SDMs for six Mediterranean marine species, demonstrating that models incorporating physiological information project less extreme range shifts. Despite the promising potential of integrating different data sources (i.e., distribution data and physiological knowledge), the significance of physiologically informed SDMs has not been fully appreciated yet, and the applications of incorporating physiological information into SDMs are still limited (e.g., Mammola et al. 2021).

The Japanese sea cucumber *Apostichopus japonicus* (Selenka 1867) (Echinodermata: Holothuroidea) (Fig. 1A) is widely distributed along the coastal regions of China, Japan, far eastern Russia, and Korea (Yang et al. 2015). This benthic species is an important aquaculture species of great ecological and economic importance (Yang et al. 2015). In 2019, the aquaculture area of *A. japonicus* in China reached approximately 246,745 hectares, and production of this species was 171,700 tonnes (China Fishery Statistical Yearbook 2020). The production of this species was equivalent to ~5.68 billion US dollars at a market price of 33.06 US dollars per kilogram (Yang et al. 2015). This species is particularly sensitive to high temperature and undergoes aestivation—a unique strategy in response to extreme conditions where species stop feeding, exhibit low activity and lose body weight (Yang et al. 2015).

Considering its sensitivity to high temperature regimes, the Japanese sea cucumber represents an ideal model species to explore the importance of integrating physiological information in SDMs. In this study, we developed two sets of SDMs for the Japanese sea cucumber to investigate the potential effects of climate change on habitat suitability for the species: (1) a SDM without considering physiological information (hereafter naïve model); and (2) a SDM that incorporates physiological information as priors to the model (hereafter physiologically informed model). We calibrate models within three buffer extents (1000 km, 1500 km, and 2000 km) and hypothesize that the incorporation of physiological knowledge into SDMs will reduce bias introduced by the choice of calibration areas.

Materials and methods

Model calibration area

The Japanese sea cucumber is a coastal water species; thereby, consistent with previous studies (Hu et al. 2021; Stephenson et al. 2020; Zhang et al. 2020), we restricted our analyses within the Exclusive Economic Zone, i.e., waters
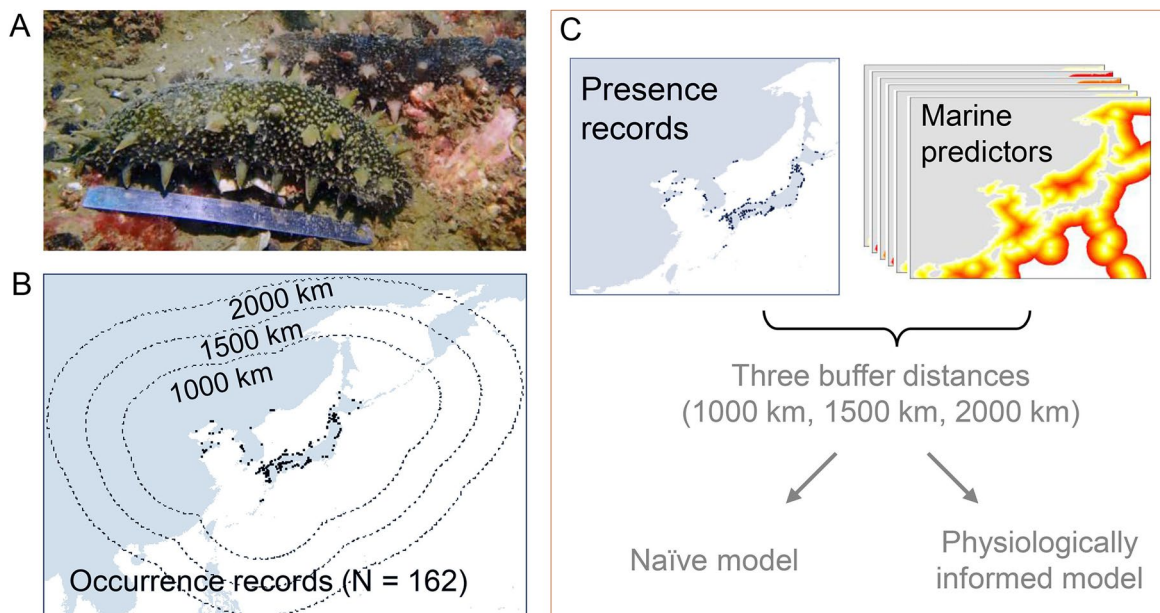
Fig. 1 Schematic diagram showing the study design. **A** Photograph of Japanese sea cucumber (*Apostichopus japonicus*), **B** species presence records (black dots) and three distinct buffer regions (equidistant cylindrical projection), and **C** schematic diagram of model design. We calibrated models within three buffer extents (1000 km, 1500 km, and 2000 km), as shown in Fig. 1B. For each calibration extent, we constructed two types of models: physiologically informed and naïve models. The photograph in Fig. 1A was taken on February 6, 2019 in Onagawa Bay (Japan) by Tohoku Ecosystem-Associated Marine Science, and reproduced with permission

extending up to 200 nautical miles from shore (Fig. 1B). The extent of the calibration area, a modeling area coherent with the biology and evolutionary history of the target species, is a critical feature to properly calibrate SDMs (Barve et al. 2011). It is always extraordinarily challenging to precisely define the calibration area (VanDerWal et al. 2009; Barve et al. 2011); therefore, following previous studies (e.g., Thuiller et al. 2019; Waldock et al. 2022), we considered the dispersal ability of the Japanese sea cucumber, and delineated calibration areas by creating buffers around species occurrence records (Fig. 1B). Sea cucumbers have planktonic larval stages and exhibit great dispersal potential (Yang et al. 2015). For example, the larvae of the Northern Sea Cucumber *Cucumaria frondosa* were reported to be transported over long distances (ranging from 800 to 1500 km) by ocean currents (So et al. 2011). Interestingly, the adults of sea cucumbers may actively modify their buoyancy and float with ocean currents, up to 90 km per day (Hamel et al. 2019). To reflect the maximal dispersal distance of sea cucumbers, we created buffers with three different radii using equidistant cylindrical projection: 1000 km, 1500 km, and 2000 km. In short, we calibrated both naïve and physiologically informed SDMs within three different buffers (Fig. 1C), and we used these SDMs to project habitat suitability of *A. japonicus* within the largest buffer (i.e., 2000-km buffer).

Species distribution data and marine predictors

We gathered distribution data for the Japanese sea cucumber through an extensive literature review (Supplementary Table S1). In particular, we treated the three-color variants (red, green, and black) of *A. japonicus* as belonging to the same species (Sun et al. 2010; Zhang et al. 2016). In addition to the literature review, we retrieved distribution data from online repositories including the Global Biodiversity Information Facility (GBIF; https://www.gbif.org) and the Ocean Biodiversity Information System (https://obis.org). Thus, we assembled a total of 458 presence records (Supplementary Fig. S1). Following previous recommendations (Kramer-Schadt et al. 2013), we filtered presence records by

1. removing records on land.
2. keeping only one presence record per 5 arcminute grid cell (the spatial resolution of marine predictors).
3. performing spatial thinning using a distance of 20 km (Hu et al. 2021) via the *thin* function in R package 'spThin' (version 0.2.0; Aiello-Lammens et al. 2015).

After the above cleaning procedures, we retained 162 presence records for subsequent analyses (Fig. 1B).

To map habitat suitability of the Japanese sea cucumber, we selected six environmental predictors out of an initial suite of 14 candidate predictors (Supplementary Table S2), i.e. water depth, distance to shore, annual mean benthic temperature, annual range of benthic temperature, annual mean benthic salinity, and annual mean benthic current velocity (Supplementary Fig. S2). Empirical evidence supports that these six predictors are ecologically relevant to the geographical distribution of the Japanese sea cucumber. In particular, it is widely accepted that water depth is associated with hydrostatic pressures, and marine benthic species typically will only tolerate certain depth ranges (Brown and Thatje 2014), i.e. less than 200 m for *A. japonicus* (Supplementary Table S3). Therefore, it is reasonable to include water depth into our models. Distance to shore correlates with production pathways and the supply of organic matter from the surface (Lapointe et al. 1994; Miller and Wheeler 2012). Therefore, the two geographical predictors have been commonly used in SDMs for marine organisms, including projecting species redistributions under climate change (e.g., Bosch et al. 2018; Hu et al. 2021; Peters et al. 2022). Temperature and salinity are among the most important environmental factors limiting the distribution of marine organisms (Castro and Huber 2005). Thus, we included annual mean temperature in our analyses instead of maximum temperature because although the Japanese sea cucumber is sensitive to high temperature, it survives high temperatures by aestivating (Wang et al. 2015; Yang et al. 2015). This is supported by the temperature information of species occurrence records: among the 162 records, annual mean temperature of all records was below 26 °C (the upper thermal limit of the Japanese sea cucumber; see Table S3 for details), whereas the maximum temperature of 54.3% (88 out of 162) records was higher than the upper thermal tolerance (Supplementary Fig. S3). We acknowledge that the seasonal variation of sea benthic temperature is large (e.g., Yang et al. 2015). Therefore, apart from annual mean temperature, we further considered the annual range of temperature to reflect the marked seasonal variation. Current velocity may influence the dispersal of marine species (Álvarez-Noriega et al. 2020) and interact with climate warming in driving range shifts (García Molinos et al. 2017). In addition, previous studies pointed out that current velocity may directly affect the movement activity and distribution of *A. japonicus* (e.g., Pan et al. 2015).

Multicollinearity across the six selected predictors was negligible, with variance inflation factor values < 10 (Naimi et al. 2014). We downloaded the two geographical predictors (i.e., water depth and distance to shore) with a spatial resolution of 5 arcminutes from the Global Marine Environment Datasets (http://gmed.auckland.ac.nz) (Basher et al. 2018),

and obtained the three marine environmental predictors under average present-day conditions (2000–2014) from the Bio-ORACLE version 2.1 dataset (https://www.bio-oracle.org) (Assis et al. 2018). We assumed that water depth and distance to shore remain constant across time. Future projections of the three marine environmental predictors under four distinct RCP scenarios in 2040–2050 were projected by three atmosphere–ocean general circulation models, including CCSM4, HadGEM2-ES, and MIROC5. We retrieved average outputs of the atmosphere–ocean general circulation models from the Bio-ORACLE version 2.1 dataset (Assis et al. 2018). The raster grids of the two databases did not align, and thus were unified to the same resolution (i.e., 5 arcminutes) using bilinear interpolation.

Naïve SDMs

In this study, we constructed SDMs via *plateau* approach, which is based on piecewise functions (Brewer et al. 2016; see detailed method in Supplementary Fig. S4). A *plateau* model consists of two intersecting linear functions and an additional constraint on the maximum value. The two linear functions were determined through their intersection point and corresponding slopes, then the maximum value was limited with an additional apex parameter. All parameters in a *plateau* model (see more details in Supplementary Fig. S4) were estimated based on Bayes' theorem, and numerically implemented via the Markov chain Monte Carlo (MCMC) method (Brewer et al. 2016). The MCMC setting remained identical for all calculations in this study: two parallel chains were run with the same initial values; in each chain a total of 5000 iterations were specified, where 4000 iterations were used for warm-up (also known as burn-in) and the rest 1000 interactions for posterior inference.

As with Feng et al. (2020), we fitted *plateau* models via presence/pseudo-absence data. To achieve this, we randomly generated the same number of pseudo-absences within the calibration areas as the presence data. To avoid the possible effects of pseudo-absence selections on *plateau* models, we developed 25 *plateau* models by randomly generating different pseudo-absence sets within each calibration area. The naïve SDMs were constructed with the default non-informative priors of the marine predictors (Brewer et al. 2016; see more details in Supplementary Fig. S4). Note that we normalized the values of each marine predictor to 0–1, following the recommended practice by Brewer et al. (2016).

Physiologically informed SDMs

To integrate the physiological information of *A. japonicus* into SDMs, we first collected physiological knowledge from literature of the Japanese sea cucumber related to our marine predictors. As it is commonly the case with marine species,

temperature is a parameter for which there is more abundant physiological information on *A. japonicus* (Yang et al. 2015). The optimal growth temperature for *A. japonicus* is generally between 15 and 18 °C (Supplementary Table S3). Regarding the upper thermal limit of the Japanese sea cucumber, field surveys reported that the density of this species becomes zero when sea water temperature reaches 24.7 ± 3.2 °C (mean ± standard deviation) (Minami et al. 2019). Furthermore, the threshold temperature inducing aestivation varies, but is normally reported at about 26 °C (Supplementary Table S3). After comprehensively considering the existing evidence, we denote the upper thermal limit of *A. japonicus* as 26 °C. The Japanese sea cucumber is a benthic species living mainly in coastal waters shallower than 40 m, and finds its maximum water depth limit at approximately 200 m (Supplementary Table S3). The species may also tolerate a wide range of salinities (25–45 psu), which covers the whole salinity range within our study region (29–35 psu) (Supplementary Table S3). Accordingly, we did not impose any constraints of salinity into our model. Water current velocity may be expected to exert an effect on both larval dispersal and adult habitat; however, the extent to which this parameter limits the range of *A. japonicus* remains unclear. Therefore, we did not consider any constraints of current velocity in our models. Similarly, no constraints were imposed associated with the distance to the coast as there is no available information about its (likely indirect) effect on this species.

With the given physiological information, we modified the prior distribution of annual mean temperature and water depth (see mathematical expressions in Supplementary Fig. S4). As for annual mean temperature, we constrained the apex value within the range of the optimal temperature, and calculated the slope of the linear function based on the upper thermal limit and apex value. For water depth, we determined the slope of the linear function based on the maximum depth limit (200 m) and optimal depth (0–40 m) (Supplementary Table S3). We normalized each constraint's values, as we did for each marine predictor (Brewer et al. 2016). We constructed *plateau* models via the R package 'R2WinBUGS' (version 2.1–21; Sturtz et al. 2005) that runs the WinBUGS (version 1.4) in batch mode using scripts (Brewer et al. 2016), and performed other analyses, including convergence tests, within R (version 4.0.3) (R Core Team 2020).

Model performance and projections

We adopted a fivefold spatial block cross-validation approach to evaluate predictive performances of each model. Thus, we split the distribution data into spatial blocks based on rows and columns (10 rows and 10 columns in our case),

and randomly allocated these blocks into five folds via the *cv_spatial* function in the R package 'blockCV' (version 3.1-3; Valavi et al. 2019). We calibrated models using four folds and validated model performance via the withheld fold. We repeated this procedure until all folds were used for model validation (Zhang et al. 2021 and reference therein). Previous studies highlighted that commonly used discrimination metrics such as AUC (area under the receiver operating characteristic curve) and TSS (true skill statistic) may be misleading when lacking high quality presence-absence data (Leroy et al. 2018; Lobo et al. 2008; Somodi et al. 2017). In our case, we only had species presence data and lacked species' true absence data (we used randomly generated pseudo-absences as alternatives). Therefore, instead of using AUC or TSS, we measured model performance via the Boyce index using the default setting of *ecospat.boyce* function in the R package 'ecospat' (version 3.1; Di Cola et al. 2017), which is specific for presence-only data and is regarded as a more robust metric for this type of data (Hirzel et al. 2006). We further quantified the variable importance for each predictor in the physiologically informed and naïve models determined by randomly permuting the values and then calculating the decrease in the goodness-of-fit measure, called "permutation importance" (Thuiller et al. 2009). Also, we determined the marginal response curves of the two constrained predictors (i.e., water depth and temperature) (Hu et al. 2021). The larvae of the sea cucumber are planktonic and may experience long-distance dispersal by ocean currents (So et al. 2011; Yang et al. 2015). Therefore, we adopted an unlimited dispersal assumption when projecting habitat suitability of the Japanese sea cucumber under future climates, i.e., sea cucumber will be able to inhabit newly suitable ranges in the future (Guisan et al. 2017). We binarized the continuous habitat suitability predictions via a 10% presence probability threshold (Hu et al. 2021; Zhang et al. 2021) and computed the range size changes of the Japanese sea cucumber under future climates.

For the physiologically informed and naïve models, we measured similarity between the suitable ranges predicted by different models via the Jaccard similarity index, which ranges from 0 (completely dissimilar) to 1 (completely similar). Take the models calibrated within 1000 km and 2000 km as an example, first we computed the Jaccard similarity index between all combinations of the suitable ranges predicted by the physiologically informed models (25 predictions of 1000 km buffer model × 25 predictions of 2000 km buffer model = 625 combinations), then conducted the same analyses for the naïve models. With the aid of this prior physiological knowledge, the physiologically informed models should always be able to properly delineate the species' tolerance of the constrained predictors. Thus, we reasonably hypothesize that the physiologically informed models should be less influenced by the choice

of calibration areas (i.e., the choice of pseudo-absences), resulting in more similar spatial distribution patterns (i.e., higher Jaccard similarity index). Also, we explored the range shifts (range expansion and range contraction) of *A. japonicus* under climate change projected by the different models.

Results

Model performance and predictor importance

All the models showed high predictive performance with Boyce values over 0.48 (Table 1). Regardless the calibration area, the physiologically informed SDMs showed a slight but insignificant advantage over the corresponding naïve models in terms of Boyce scores (two-sided paired Wilcoxon rank sum test, $p > 0.05$; see Supplementary Fig. S5 for all model predictive indices).

The physiologically informed and naïve models yielded different results regarding variable importance (Fig. 2). The physiologically informed models showed that water depth, distance to shore, and annual mean temperature represent the three most important predictors explaining the geographical distribution of the Japanese sea cucumber, whereas the naïve models identified the important roles of the distance to shore, annual mean and range of temperature (Fig. 2). All SDMs showed consistently that the contribution of salinity and current velocity was negligible (Fig. 2). The importance of the two constrained predictors (i.e., water depth and annual mean temperature) much higher improved in the physiologically informed models (Fig. 2).

Irrespective of the calibration regions, the two types of models produced different response curves with respect to annual mean temperature and water depth (Figs. 3, 4). On the one hand, the naïve SDMs failed to capture the intolerance of *A. japonicus* to high temperature and deep water with habitat suitability remaining relatively high even for high annual mean temperature (Fig. 3) or water depth

Table 1 The Boyce index (mean ± standard deviation) of species distribution models for the Japanese sea cucumber *Apostichopus japonicus*

Buffer (km)	Model type	
	Naïve model	Physiologically informed model
1000	0.519 ± 0.211	0.535 ± 0.178
1500	0.484 ± 0.232	0.537 ± 0.165
2000	0.496 ± 0.254	0.536 ± 0.169

Naïve model refers to the species distribution model without the physiological information, and Informed model is the physiologically informed species distribution model. Buffer refers to the different calibration extents for the species (see Fig. 1)
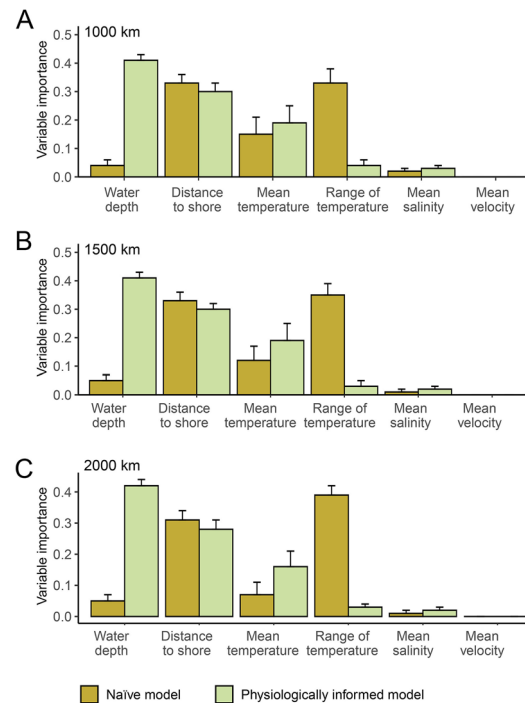
Fig. 2 Variable importance in the species distribution models of the Japanese sea cucumber *Apostichopus japonicus* regarding three buffer distances (**A–C**) with and without physiological constraints
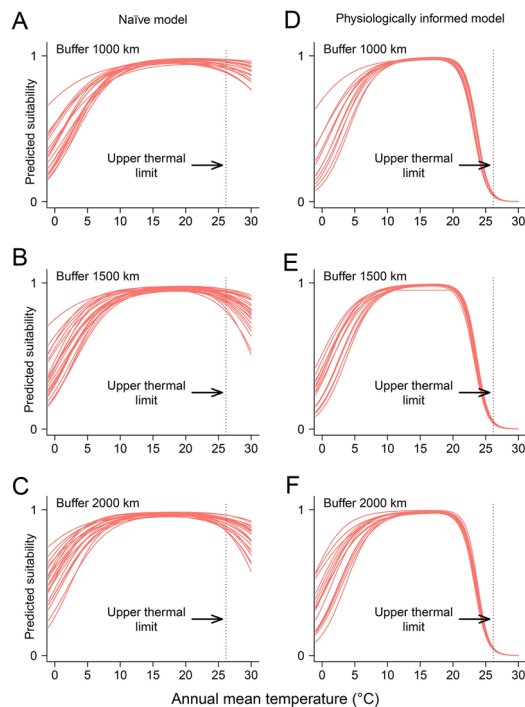

Fig. 3 Marginal response curves of the Japanese sea cucumber (*Apostichopus japonicus*) regarding annual mean temperature without (**A–C**) and with (**D–F**) physiological constraints. Vertical lines represent the upper thermal temperature of the focal species
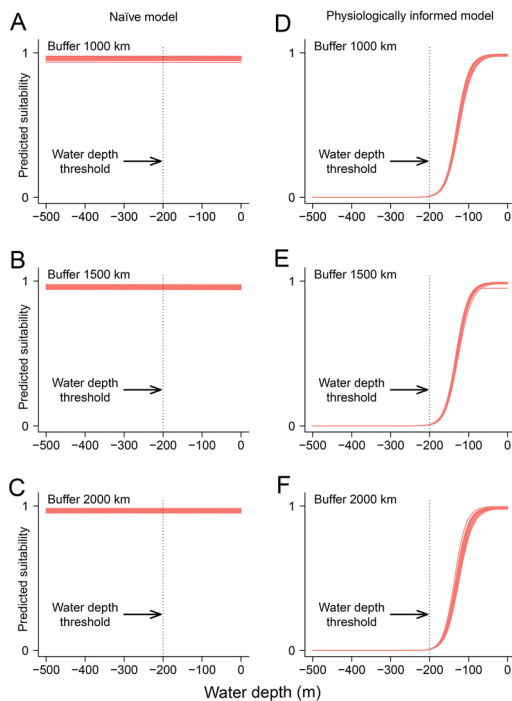
Fig. 4 Marginal response curves of the Japanese sea cucumber (*Apostichopus japonicus*) regarding water depth without (**A**–**C**) and with (**D**–**F**) physiological constraints. Vertical lines represent the maximum depth where the focal species can be found

(Fig. 4). Conversely, the physiologically informed models successfully detected that high temperature and deep water are unfavorable for the species (Figs. 3, 4). This indicates that the physiologically informed models incorporated successfully the constraints of annual mean temperature and water depth, and these models described better the thermal and bathymetric tolerance of the Japanese sea cucumber.

Present habitat suitability prediction

Both models showed consistently that a large proportion of coastal regions of China, Japan, and Korea are suitable for *A. japonicus* under present climatic conditions (Fig. 5). Moreover, suitable ranges for this species are predicted to be available in the south of the Yangtze River estuary in China (Fig. 5). The habitat suitability projected by each model are different in multiple aspects. First, the naïve models projected a more southward distribution limit than the corresponding physiologically informed models (two-sided paired Wilcoxon rank sum test, $p < 0.05$; Supplementary Fig. S6). Second, the physiologically informed models projected significantly larger suitable ranges compared with the naïve models (two-sided paired Wilcoxon rank sum test, $p < 0.05$; Supplementary Fig. S7). Third, and perhaps most

importantly, the suitable ranges projected by the physiologically informed models were less influenced by the choice of calibration areas. Our results showed that the Jaccard similarity indices between the suitable ranges predicted by the models fitted to three calibration areas were significantly higher for the physiologically informed models than that for the naïve models (one-sided paired Wilcoxon rank sum test, $p < 0.05$; see Supplementary Table S4 for statistical details), indicating that the physiologically informed models produced more similar predictions under distinct calibration areas (Fig. 5E).

Climate change effects on future habitat suitability

The physiologically informed and naïve SDMs projected distinct effects of climate change on the habitat suitability of *A. japonicus* (Fig. 6). We found that the naïve models projected significantly more range expansions but less range contractions under climate changes than the corresponding physiologically informed model (Figs. 6E, F; S8, S9). Taking the models calibrated within 2000-km buffer as an example, under representative concentration pathway (RCP) 8.5 in 2040–2050, compared with the naïve models (range expansion: $16.94\% \pm 2.59\%$; range contraction: $5.60\% \pm 1.89\%$), the corresponding physiologically informed models projected a significantly lower level of range expansion ($13.56\% \pm 2.08\%$) (two-sided paired Wilcoxon rank sum test, V statistic $= 297$; $p < 0.001$) but a significantly higher level of range contraction ($11.08\% \pm 0.88\%$) (two-sided paired Wilcoxon rank sum test, V statistic $= 0$; $p < 0.001$). With respect to range expansion, both types of models suggest that the future suitable ranges of *A. japonicus* will likely expand to northern regions, such as Sakhalin (Fig. 6C, D). As for range contraction, the naïve models identified only evident range contractions in the Bohai Sea, whereas the informed models detected severe range losses in the south Chinese coastal waters, Bohai Sea, Seto Inland, and Ariake seas (Fig. 6C, D).

Discussion

We used the Japanese sea cucumber as a model to explore the importance of incorporating physiological information into SDMs. Our results highlight that considering physiological information into correlative SDMs may greatly reduce the uncertainties associated with the choice of calibration areas. Compared with the naïve models, our physiologically informed models produced more stable habitat suitability projections with similar model predictive performance. Our findings reaffirm that we should not rely only on predictive
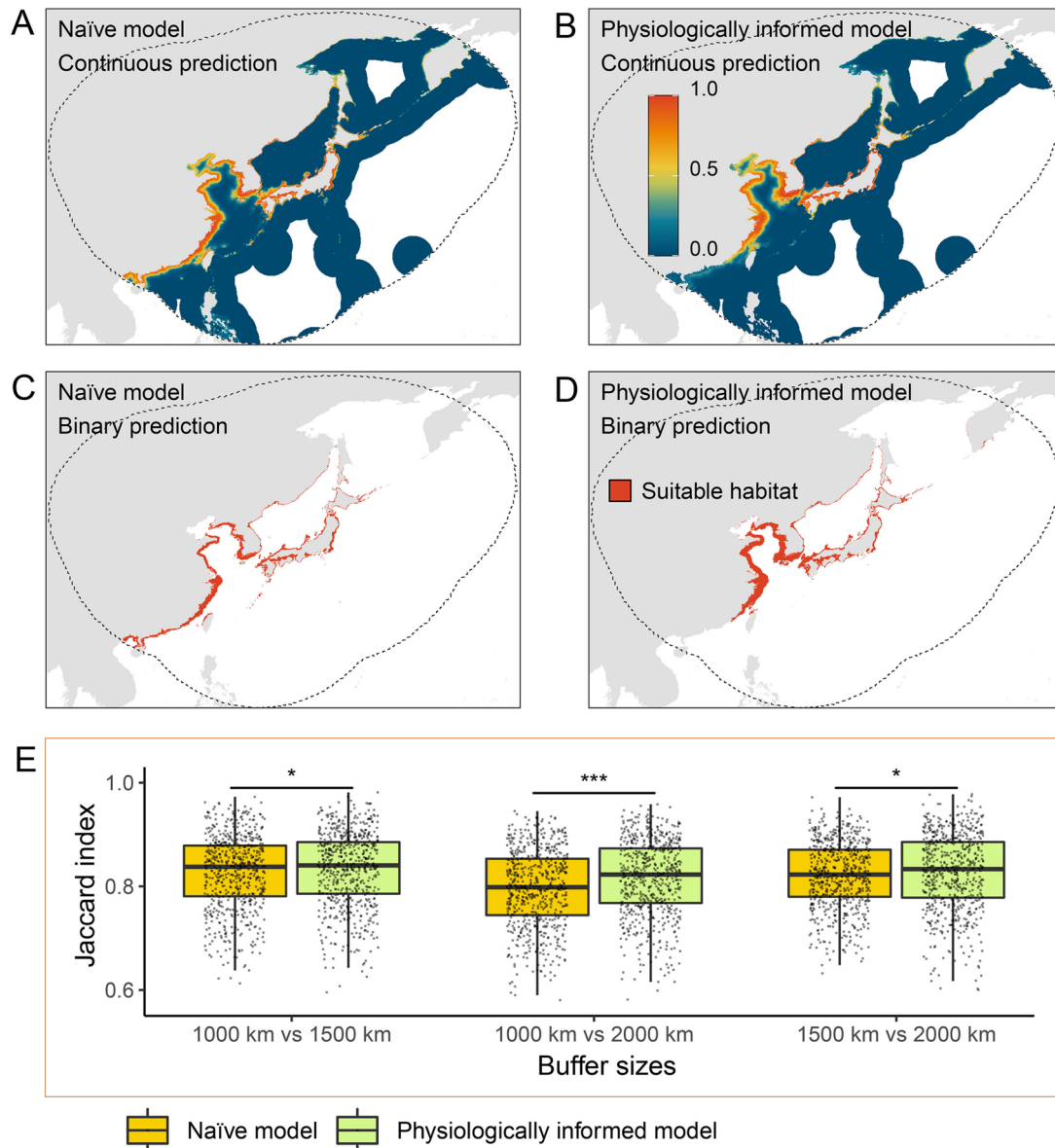
Fig. 5 Habitat suitability predictions of the Japanese sea cucumber (*Apostichopus japonicus*) by the two types of species distribution models under present-day climatic conditions using an example of 2000-km buffer. **A, B** Continuous habitat suitability predictions (equidistant cylindrical projection). **C, D** Binary habitat suitability predictions via a 10% presence probability threshold. **E** Similarity levels between suitable ranges predicted by models calibrated within differ-ent buffers. Dashed lines in a–d represent the 2000 km buffer zones. Data in **E** are expressed as mean ± standard error and the asterisk (*) indicates a significant difference ($p < 0.05$) with (***) for $p < 0.001$. **A–D** Illustrate one example among 25 models calibrated within 2000 km buffer, habitat suitability predictions for all 25 models and other buffers are available from Figshare (https://figshare.com/s/ec89e 8ca525b878fe5fc)

performance metrics in SDM studies (Lobo et al. 2008), but rather we should strive to balance statistical and biological considerations and examine the shapes of response curves to avoid any misleading predictions or potentially unreasonable biological trends. Furthermore, our results highlight the significance of integrating species' physiological knowledge into correlative SDMs, even in the form of a few, simple physiological constraints, as this practice may reduce greatly the uncertainties associated with the choice of model calibration extents, and better inform management and conservation strategies under future climate and environmental changes.
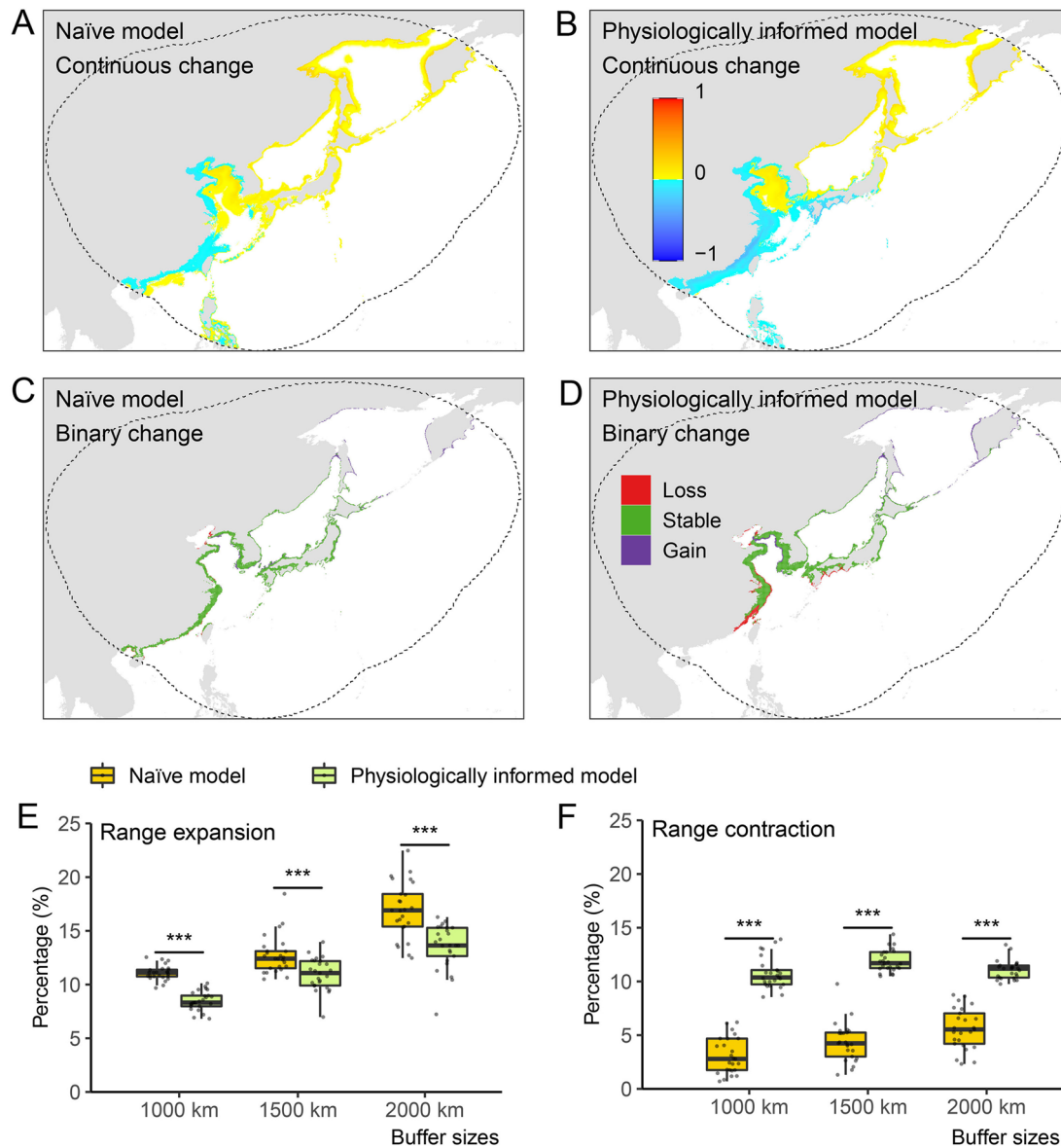
Fig. 6 Habitat suitability predictions and changes in projected suitable habitat of the Japanese sea cucumber (*Apostichopus japonicus*) under RCP 8.5 in 2040–2050 by two types of species distribution models using an example of 2000-km buffer. **A, B** Continuous habitat suitability predictions (equidistant cylindrical projection). **C, D** Binary habitat suitability predictions. Range **E** expansion and **F** contraction predicted by models calibrated within different buffers. Dashed lines in a–d represent the 2000 km buffer zones. Data are expressed as mean ± standard error in **E** and **F**, and the asterisk (***) indicates a significant difference ($p < 0.001$). **A–D** Illustrate one example among 25 models calibrated within 2000 km buffer. Habitat suitability predictions for all 25 models and other buffers are available from Figshare (https://figshare.com/s/ec89e8ca525b878fe5fc)

Critical role of calibration area in SDMs

Among the several critical assumptions of SDMs, a fundamental one is the model ability to capture the meaningful properties of the Hutchinson's realized niche of species (Guisan and Thuiller 2005; Guisan et al. 2017). This is often implicitly assumed as it is often unclear to what extent this assumption is met according to different research scenarios (Guisan and Thuiller 2005; Guisan et al. 2017), especially the calibration areas (Barbet-Massin et al. 2010; Guisan et al. 2017; Sánchez-Fernández et al. 2011; Thuiller et al. 2004). Previous studies demonstrated that SDMs based on regional distribution data, as opposed to whole distribution data, often produce truncated response curves, thus resulting in

misleading habitat suitability projections because of extrapolation (e.g., Barbet-Massin et al. 2010; Sánchez-Fernández et al. 2011; Thuiller et al. 2004). In our case, although we considered occurrence data from the whole known range of the Japanese sea cucumber, naïve SDMs failed to produce biologically meaningful response curves. Unlike the naïve models, our physiologically informed models were insensitive to the choice of calibration extents in this regard because the prior physiological knowledge had been successfully incorporated as prior information into models. Physiologically informed SDMs have been reported to have different advantages, such as improving extrapolation ability in novel conditions (Feng et al. 2020). Furthermore, our findings suggest that this type of model allows for an easy incorporation of physiological knowledge effectively mitigating the effects of the selection of calibration areas whilst calibrating SDMs. This finding may have broad implications as it is often challenging to correctly define calibration areas in SDM studies because of the frequent lack of data on dispersal potential and historical ranges (e.g., Barve et al. 2011). We encourage future studies to integrate the physiological information of target species to improve model reliability.

Statistical versus biological importance

In our study, the Boyce index suggests that the predictive capacities of the physiologically informed models were slightly higher than the corresponding naïve models although differences were statistically nonsignificant. Our results are consistent with those of Gamliel et al. (2020), who also reported a marginal performance improvement of physiologically informed models over naïve models. Regarding our results from a statistical point of view, it seems safe to conclude that both physiologically informed and naïve models are reliable given their high Boyce index. However, from a biological point of view, physiologically informed models are more meaningful in that they properly reflected the tolerance of the Japanese sea cucumber to temperature and water depth (Figs. 3, 4). Thus, the question remains whether we should favor biological realism versus blunt measures of statistical performance (Brewer et al. 2016; Mammola et al. 2019b). Previous studies demonstrated that incorporating physiological information into SDMs may make response curves more reasonable (e.g., Feng et al. 2020; Gamliel et al. 2020). Thus far in SDM studies, researchers rely heavily on evaluation criteria, such as AUC, TSS or the Boyce index to justify model reliability (Araújo et al. 2019; Feng et al. 2019; Guisan et al. 2017). We recommend that researchers should treat model predictive performance with caution, and not simply trust high goodness-of-fit values (Santini et al. 2021). In addition to following best-practice standards (Araújo et al. 2019) and reporting necessary information (Feng et al. 2019), these

results remind us of the critical role of examining response curves (Zurell et al. 2020).

Integration of physiological information into SDMs

Physiological information on depth distribution for the species was implemented into our model in terms of its absolute maximum depth limit and a relatively wide range of optimal depth. Without accurate information about the optimal depth range for the species, it is difficult to assess whether this is a sensible choice. The Japanese sea cucumber is associated mainly with shallow water environments, and drastic decreases in abundance have been reported with increases in depth even in very shallow environments. For example, Selin (2001) reported a 830-fold decrease in abundance of the Japanese sea cucumber in the Vostok Bay (Sea of Japan) between 0.5–1.5 m and 8–15 m depth ranges. Therefore, it is also possible that using the absolute maximum depth tolerance limit and a wide range of optimal depth for the species, we may have effectively biased the model towards overoptimistic predictions of the suitable habitat for the species in deeper waters. This highlights the need for the careful assessment of physiological information implemented into models and for the critical interpretation of results. Future studies are required to explore the extent to which incomplete physiological knowledge may bias predictions in this type of physiologically informed models.

Gamliel et al. (2020) reported that physiologically informed models predicted less drastic range shifts under climate change than naïve models. Contrary to the findings of Gamliel et al. (2020), our results show that physiologically informed models predict more severe impacts of climate change with greater range contraction and less range expansion. Nevertheless, the results of Gamliel et al. (2020) and our study indicate consistently that physiologically informed and naïve models projected different future distribution patterns. Incorporating physiological information into SDMs has promising benefits but is still in its infancy; thereby, we encourage future studies to further clarify the vital roles of physiology in predicting species redistribution under climate change.

Future perspectives

The IUCN expert-based range map shows clearly that the Japanese sea cucumber naturally occurs north of the Yangtze River in China (Hamel and Mercier 2013). This suggests that the Yangtze River outflow may well impose a geographical barrier defining the southern distributional limit of this species; something that seems to be corroborated by the lack of occurrence records for the species from these southern coastal waters. Both the physiologically informed and

naïve models predicted a widespread and suitable range for *A. japonicus* south of its known and actual distributional limits (Fig. 6). As any correlative SDMs, physiologically informed models predict where habitat conditions are suitable for species to be present (potential range), not where they actually occur (realized range). Therefore, care is needed when interpreting the output from these models in terms of species ranges particularly when informing those unfamiliar with the methods, such as managers and policy makers. Together with physiology, dispersal capacity, which may be limited by biogeographical barriers as in our present study, species interactions, physical habitat availability (e.g., seafloor substrate composition for benthic species) or human stressors other than climate changes, may all effectively condition the range of a species. There is clearly much to be gained from improving the reliability of SDMs through the combined inclusion of physiological information, species interactions (Zhang et al. 2024) or dispersal capacity and demographic factors (Miller and Holloway 2015). The development of better-resolved predictor layers, both in time and space, is also needed particularly for the marine realm. For example, as a deposit feeder, soft bottom sediments are critical for the habitat of *A. japonicus* (Yang et al. 2015), but high-resolution benthic substrate composition layers are still unavailable.

In summary, we integrated successfully physiological information of the Japanese sea cucumber into SDMs. Our results highlight that physiologically informed SDMs greatly reduce the uncertainties resulting from the choice of calibration extents. Given these promising features, we encourage future SDM studies to incorporate species' physiological information if available. However, despite their superiority, different types of challenges remain to be addressed, including disentangling the effects of macroclimates and microclimates, the link between field environments and laboratory-derived physiological information, and the lack of developmental stage-specific physiological knowledge [for a detailed discussion on this topic see Feng et al. (2020) and Gamliel et al. (2020)]. Furthermore, because physiological information is costly and time-consuming to acquire, the routine integration of physiology into species distribution modeling may still be a distant reality, especially for multi-species studies.

Author contributions ZZ and JZ designed the research. JZ, ZZ and TQ collected the data. JZ and ZZ performed the analyses. ZZ, JGM, SM, ÁB-F, XF, DK, QL and JA discussed and improved the results. JZ and ZZ prepared figures and tables and led the writing of the manuscript with critical suggestions from JGM, SM, and QL. ZZ and QL supervised this work. All authors revised and contributed to the paper.

Data availability The presence records of the Japanese sea cucumber are available from Figshare (https://figshare.com/s/ec89e8ca525b878fe5fc). Marine predictors can be retrieved from online repositories (https://www.bio-oracle.org and http://gmed.auckland.ac.nz). The present-day and future habitat suitability predictions of the Japanese sea cucumber are available from Figshare (https://figshare.com/s/ec89e8ca525b878fe5fc).

Declarations

Conflict of interest The authors declare no conflict of interest. Author Qiang Lin is one of the Editorial Board Members, but he was not involved in the journal's review of, or decision related to, this manuscript.

Animal and human rights statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP (2015) *spThin*: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38:541–545

Álvarez-Noriega M, Burgess SC, Byers JE, Pringle JM, Wares JP, Marshall DJ (2020) Global biogeography of marine dispersal potential. Nat Ecol Evol 4:1196–1203

Araújo MB, Anderson RP, Barbosa AM, Beale CM, Dormann CF, Early R, Garcia RA, Guisan A, Maiorano L, Naimi B, O'Hara RB, Zimmermann NE, Rahbek C (2019) Standards for distribution models in biodiversity assessments. Sci Adv 5:eaat4858

Assis J, Tyberghein L, Bosch S, Verbruggen H, Serrão EA, De Clerck O (2018) Bio-ORACLE v2. 0: Extending marine data layers for bioclimatic modelling. Glob Ecol Biogeogr 27:277–284

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol Evol 3:327–338

Barbet-Massin M, Thuiller W, Jiguet F (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? Ecography 33:878–886

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol Model 222:1810–1819

Basher Z, Bowden DA, Costello MJ (2018) Global marine environment datasets (GMED) version 2.0 (Rev.02.2018). http://gmed.auckland.ac.nz

Bayes T (1764) An essay toward solving a problem in the doctrine of chances. Biometrika 45:296–315

Bosch S, Tyberghein L, Deneudt K, Hernandez F, De Clerck O (2018) In search of relevant predictors for marine species distribution modelling using the MarineSPEED benchmark dataset. Divers Distrib 24:144–157

Brewer MJ, O'Hara RB, Anderson BJ, Ohlemüller R (2016) *Plateau*: a new method for ecologically plausible climate envelopes for species distribution modelling. Methods Ecol Evol 7:1489–1502

Brito-Morales I, Schoeman DS, García Molinos J, Burrows MT, Klein CJ, Arafeh-Dalmau N, Kaschner K, Garilao C, Kesner-Reyes K, Richardson AJ (2020) Climate velocity reveals increasing exposure of deep-ocean biodiversity to future warming. Nat Clim Change 10:576–581

Brown A, Thatje S (2014) Explaining bathymetric diversity patterns in marine benthic invertebrates and demersal fishes: physiological contributions to adaptation of life at depth. Bio Rev 89:406–426

Castro P, Huber ME (2005) Marine biology. McGraw-Hill, Boston, pp 64–84

Champion C, Hobday AJ, Pecl GT, Tracey SR (2020) Oceanographic habitat suitability is positively correlated with the body condition of a coastal-pelagic fish. Fish Oceanogr 29:100–110

Chen IC, Hill JK, Ohlemüller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. Science 333:1024–1026

Cheung WW, Watson R, Pauly D (2013) Signature of ocean warming in global fisheries catch. Nature 497:365–368

China Fishery Statistical Yearbook (2020) Ministry of Agriculture China Agriculture Press, Beijing, pp 22–50

Di Cola V, Broennimann O, Petitpierre B, Breiner FT, D'Amen M, Randin C, Engler R, Pottier J, Pio D, Dubuis A, Pellissier L, Mateo RG, Hordijk W, Salamin N, Guisan A (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography 40:774–787

Doney SC, Ruckelshaus M, Emmett Duffy J, Barry JP, Chan F, English CA, Galindo HM, Grebmeier JM, Hollowed AB, Knowlton N, Polovina J, Rabalais NN, Sydeman WJ, Talley LD (2012) Climate change impacts on marine ecosystems. Annu Rev Mar Sci 4:11–37

Dong Y, Bao M, Cheng J, Chen Y, Du J, Gao Y, Hu L, Li X, Liu C, Qin G, Sun J, Wang X, Yang G, Zhang C, Zhang X, Zhang Y,

Zhang Z, Zhan A, He Q, Sun J et al (2024) Advances of marine biogeography in China: species distribution model and its applications. Biodivers Sci. https://doi.org/10.17520/biods.2023453 (in Chinese with English abstract)

Dulvy NK, Rogers SI, Jennings S, Stelzenmüller V, Dye SR, Skjoldal HR (2008) Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. J Appl Ecol 45:1029–1039

Feng X, Liang Y, Gallardo B, Papeş M (2020) Physiology in ecological niche modeling: using zebra mussel's upper thermal tolerance to refine model predictions through Bayesian analysis. Ecography 43:270–282

Feng X, Park DS, Walker C, Peterson AT, Merow C, Papeş M (2019) A checklist for maximizing reproducibility of ecological niche models. Nat Ecol Evol 3:1382–1395

Gaines SD, Costello C, Owashi B, Mangin T, Bone J, García Molinos J, Burden M, Dennis H, Halpern BS, Kappel CV, Kleisner KM, Ovando D (2018) Improved fisheries management could offset many negative effects of climate change. Sci Adv 4:eaao1378

Gamliel I, Buba Y, Guy-Haim T, Garval T, Willette D, Rilov G, Belmaker J (2020) Incorporating physiology into species distribution models moderates the projected impact of warming on selected Mediterranean marine species. Ecography 43:1090–1106

García Molinos J, Burrows MT, Poloczanska ES (2017) Ocean currents modify the coupling between climate change and biogeographical shifts. Sci Rep 7:1332

Golding N, Purse BV (2016) Fast and flexible Bayesian species distribution modelling using Gaussian processes. Methods Ecol Evol 7:598–608

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009

Guisan A, Thuiller W, Zimmermann NE (2017) Habitat suitability and distribution models: with applications in R. Cambridge University Press, Cambridge, p 811

Hamel JF, Mercier A (2013) *Apostichopus japonicus*. The IUCN red list of threatened species 2013: e.T180424A1629389. https://doi.org/10.2305/IUCN.UK.2013-1.RLTS.T180424A1629389.en. Downloaded on 14 June 2021

Hamel JF, Sun J, Gianasi BL, Montgomery EM, Kenchington EL, Burel B, Rowe S, Winger PD, Mercier A (2019) Active buoyancy adjustment increases dispersal potential in benthic marine animals. J Anim Ecol 88:820–832

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecol Model 199:142–152

Hof C (2021) Towards more integration of physiology, dispersal and land-use change to understand the responses of species to climate change. J Exp Biol 224:jeb238352

Hu ZM, Zhang QS, Zhang J, Kass JM, Mammola S, Fresia P, Draisma SGA, Assis J, Jueterbock A, Yokota M, Zhang Z (2021) Intraspecific genetic variation matters when predicting seagrass distribution under climate change. Mol Ecol 30:3840–3855

Hughes AC, Orr MC, Ma K, Costello MJ, Waller J, Provoost P, Yang Q, Zhu C, Qiao H (2021) Sampling biases shape our view of the natural world. Ecography 44:1–11

IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services) (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES

Kearney M, Porter W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecol Lett 12:334–350

Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V, Stillfried M, Heckmann I, Scharf AK, Augeri DM, Cheyne SM, Hearn AJ, Ross J, Macdonald DW, Mathai J, Eaton

JA, Marshall AJ, Semiadi G, Rustam R, Bernard H et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. Divers Distrib 19:1366–1379

Lapointe BE, Tomasko DA, Matzie WR (1994) Eutrophication and trophic state classification of seagrass communities in the Florida Keys. B Mar Sci 54:696–717

Lenoir J, Bertrand R, Comte L, Bourgeaud L, Hattab T, Murienne J, Grenouillet G (2020) Species better track climate warming in the oceans than on land. Nat Ecol Evol 4:1044–1059

Leroy B, Delsol R, Hugueny B, Meynard CN, Barhoumi C, Barbet-Massin M, Bellard C (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. J Biogeogr 45:1994–2002

Levin LA, Le Bris N (2015) The deep ocean under climate change. Science 350:766–768

Liao M, Li G, Wang J, Marshall DJ, Hui TY, Ma S, Zhang Y, Helmuth B, Dong Y (2021) Physiological determinants of biogeography: the importance of metabolic depression to heat tolerance. Glob Change Biol 27:2561–2579

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Glob Ecol Biogeogr 17:145–151

Mammola S, Cardoso P, Culver DC, Deharveng L, Ferreira R, Fišer C, Galassi DM, Griebler C, Halse SA, Humphreys W, Isaia M, Malard F, Martínez A, Moldovan OT, Niemiller ML, Pavlek M, Reboleira AS, Souza-Silva M, Teeling EC, Wynne J et al (2019a) Scientists' warning on the conservation of subterranean ecosystems. Bioscience 69:641–650

Mammola S, Milano F, Vignal M, Andrieu J, Isaia M (2019b) Associations between habitat quality, body size and reproductive fitness in the alpine endemic spider *Vesubia jugorum*. Glob Ecol Biogeogr 28:1325–1335

Mammola S, Pétillon J, Hacala A, Monsimet J, Marti SL, Cardoso P, Lafage D (2021) Challenges and opportunities of species distribution modelling of terrestrial arthropod predators. Divers Distrib 27:2596–2614

Marcer A, Haston EM, Groom QJ, Ariño AH, Chapman A, Bakken T, Braun P, Dillen M, Ernst M, Escobar A, Fichtmüller D, Livermore L, Nicolson N, Paragamian K, Paul DL, Pettersson LB, Phillips S, Plummer J, Rainer H, Rey I et al (2021) Quality issues in georeferencing: from physical collections to digital data repositories for ecological research. Divers Distrib 27:564–567

Miller JA, Holloway P (2015) Incorporating movement in species distribution models. Prog Phys Geog 39:837–849

Miller CB, Wheeler PA (2012) Biological oceanography, 2nd edn. Wiley-Blackwell, Chichester, pp 588–637

Minami K, Masuda R, Takahashi K, Sawada H, Shirakawa H, Yamashita Y (2019) Seasonal and interannual variation in the density of visible *Apostichopus japonicus* (Japanese sea cucumber) in relation to sea water temperature. Estuar Coast Shelf Sci 229:106384

Naimi B, Hamm NA, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? Ecography 37:191–203

Pan Y, Zhang L, Lin C, Sun J, Kan R, Yang H (2015) Influence of flow velocity on motor behavior of sea cucumber *Apostichopus japonicus*. Physiol Behav 144:52–59

Peters KJ, Stockin KA, Saltré F (2022) On the rise: Climate change in New Zealand will cause sperm and blue whales to seek higher latitudes. Ecol Indic 142:109235

Peterson AT, Papeş M, Soberón J (2015) Mechanistic and correlative models of ecological niches. Eur J Ecol 1:28–38

Pinsky ML, Selden RL, Kitchel ZJ (2020) Climate-driven shifts in marine species ranges: scaling from organisms to communities. Annu Rev Mar Sci 12:153–179

Pinsky ML, Worm B, Fogarty MJ, Sarmiento JL, Levin SA (2013) Marine taxa track local climate velocities. Science 341:1239–1242

Poloczanska ES, Brown CJ, Sydeman WJ, Kiessling W, Schoeman DS, Moore PJ, Brander K, Bruno JF, Buckley LB, Burrows MT, Duarte CM, Halpern BS, Holding JM, Kappel CV, O'Connor MI, Pandolfi JM, Parmesan C, Schwing FB, Thompson SA, Richardson AJ (2013) Global imprint of climate change on marine life. Nat Clim Change 3:919–925

Poloczanska ES, Burrows MT, Brown CJ, García Molinos J, Halpern BS, Hoegh-Guldberg O, Kappel CV, Moore PJ, Richardson AJ, Schoeman DS, Sydeman WJ (2016) Responses of marine organisms to climate change across oceans. Front Mar Sci 3:62

Potts WM, Henriques R, Santos CV, Munnik K, Ansorge I, Dufois F, Booth AJ, Kirchner CH, Sauer WH, Shaw PW (2014) Ocean warming, a rapid distributional shift, and the hybridization of a coastal fish species. Glob Change Biol 20:2765–2777

Qiao H, Soberon J, Peterson AT (2015) No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. Methods Ecol Evol 6:1126–1136

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Ryo M, Angelov B, Mammola S, Kass JM, Benito BM, Hartig F (2021) Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models. Ecography 44:199–205

Sánchez-Fernández D, Lobo JM, Hernández-Manrique OL (2011) Species distribution models that do not incorporate global data misrepresent potential distributions: a case study using Iberian diving beetles. Divers Distrib 17:163–171

Santini L, Benítez-López A, Maiorano L, Čengić M, Huijbregts MA (2021) Assessing the reliability of species distribution projections in climate change research. Divers Distrib 27:1035–1050

Selin NI (2001) Vertical distribution of the far east trepang *Apostichopus japonicus* in Vostok Bay, Sea of Japan. Russ J Mar Biol 27:256–258

Smith AB, Santos MJ (2020) Testing the ability of species distribution models to infer variable importance. Ecography 43:1801–1813

Smith AB, Godsoe W, Rodríguez-Sánchez F, Wang HH, Warren D (2019) Niche estimation above and below the species level. Trends Ecol Evol 34:260–273

So JJ, Uthicke S, Hamel JF, Mercier A (2011) Genetic population structure in a commercial marine invertebrate with long-lived lecithotrophic larvae: *Cucumaria frondosa* (Echinodermata: Holothuroidea). Mar Biol 158:859–870

Somodi I, Lepesi N, Botta-Dukát Z (2017) Prevalence dependence in model goodness measures with special emphasis on true skill statistics. Ecol Evol 7:863–872

Stephenson F, Goetz K, Sharp BR, Mouton TL, Beets FL, Roberts J, MacDiarmid AB, Constantine R, Lundquist CJ (2020) Modelling the spatial distribution of cetaceans in New Zealand waters. Divers Distrib 26:495–516

Sturtz S, Ligges U, Gelman A (2005) R2WinBUGS: a package for running WinBUGS from R. J Stat Softw 12:1–16

Sun XJ, Li Q, Kong LF (2010) Comparative mitochondrial genomics within sea cucumber (*Apostichopus japonicus*): provide new insights into relationships among color variants. Aquaculture 309:280–285

Taheri S, Naimi B, Rahbek C, Araújo MB (2021) Improvements in reports of species redistribution under climate change are required. Sci Adv 7:eabe1110

Talluto MV, Boulangeat I, Améztegui A, Aubin I, Berteaux D, Butler A, Doyon F, Drever CR, Fortin M, Franceschini T, Liénard JF, McKenney DW, Solarik KA, Strigul NS, Thuiller W, Gravel D (2016) Cross-scale integration of knowledge for predicting species ranges: a metamodeling framework. Glob Ecol Biogeogr 25:238–249

Thuiller W, Brotons L, Araújo MB, Lavorel S (2004) Effects of restricting environmental range of data to project current and future species distributions. Ecography 27:165–172

Thuiller W, Guéguen M, Renaud J, Karger DN, Zimmermann NE (2019) Uncertainty in ensembles of global biodiversity scenarios. Nat Commun 10:1446

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD—a platform for ensemble forecasting of species distributions. Ecography 32:369–373

Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2019) BLOCKCV: an r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods Ecol Evol 10:225–232

VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecol Model 220:589–594

Vergés A, Doropoulos C, Malcolm HA, Skye M, Garcia-Pizá M, Marzinelli EM, Campbell AH, Ballesteros E, Hoey AS, Vila-Concejo A, Bozec Y, Steinberg PD (2016) Long-term empirical evidence of ocean warming leading to tropicalization of fish communities, increased herbivory, and loss of kelp. Proc Natl Acad Sci 113:13791–13796

Waldock C, Stuart-Smith RD, Albouy C, Cheung WW, Edgar GJ, Mouillot D, Tjiputra J, Pellissier L (2022) A quantitative review of abundance-based species distribution models. Ecography 2022:e05694

Wang QL, Yu SS, Dong YW (2015) Parental effect of long acclimatization on thermal tolerance of juvenile sea cucumber *Apostichopus japonicus*. PLoS ONE 10:e0143372

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868–2883

Woodward G, Perkins DM, Brown LE (2010) Climate change and freshwater ecosystems: impacts across multiple levels of organization. Philos Trans R Soc B 365:2093–2106

Yang G, Xia C, Xiong X, Feng Z, Chen Z, Yang Y, Ma D, Ju X, Zheng Q, Yuan Y (2022) The seafloor heat flux driven by bottom water temperature variation in the Yellow and Bohai Seas. Ocean Model 177:102073

Yang H, Hamel JF, Mercier A (eds) (2015) The sea cucumber *Apostichopus japonicus*: history, biology and aquaculture. Academic Press, New York, p 453

Zhang W, Cao Z, Li Y, Zhao H, Huang J, Liang Z, Huang L (2016) Taxonomic status of the three color variants in sea cucumber (*Apostichopus japonicus*): evidence from mitochondrial phylogenomic analyses. Mitochondr DNA A 27:2330–2333

Zhang Z, Capinha C, Karger DN, Turon X, MacIsaac HJ, Zhan A (2020) Impacts of climate change on geographical distributions of invasive ascidians. Mar Environ 159:104993

Zhang Z, Kass JM, Mammola S, Koizumi I, Li X, Tanaka K, Ikeda K, Suzuki T, Yokota M, Usio N (2021) Lineage-level distribution models lead to more realistic climate change predictions for a threatened crayfish. Divers Distrib 27:684–695

Zhang Z, Ma S, Bede-Fazekas Á, Mammola S, Qu M, Zhou J, Feng EY, Qin G, Lin Q (2024) Considering biotic interactions exacerbates the predicted impacts of climate change on coral-dwelling species. J Biogeogr 51:769–782

Zhu G, Gutierrez Illan J, Crowder DW (2021) The use of insect life tables in optimizing invasive pest distributional models. Ecography 44:1501–1510

Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J, Fandos G, Feng X, Guillera-Arroita G, Guisan A, Lahoz-Monfort JJ, Leitão PJ, Park DS, Peterson AT, Rapacciuolo G, Schmatz DR, Schröder B, Serra-Diaz JM, Thuiller W, Yates KL et al (2020) A standard protocol for reporting species distribution models. Ecography 43:1261–1277

Authors and Affiliations

Zhixin Zhang[1,2,3,4] · **Jinxin Zhou[5]** · **Jorge García Molinos[6]** · **Stefano Mammola[7,8,9]** · **Ákos Bede-Fazekas[10,11]** · **Xiao Feng[12]** · **Daisuke Kitazawa[5]** · **Jorge Assis[13]** · **Tianlong Qiu[14]** · **Qiang Lin[1,3,15]**

✉ Zhixin Zhang
zxzhang@scsio.ac.cn

✉ Qiang Lin
linqiang@scsio.ac.cn

1 CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

2 Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

3 Marine Biodiversity and Ecological Evolution Research Center, South China Sea Institute of Oceanology, Guangzhou 510301, China

4 Global Ocean and Climate Research Center, South China Sea Institute of Oceanology, Guangzhou 510301, China

5 Institute of Industrial Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8574, Japan

6 Arctic Research Center, Hokkaido University, Sapporo, Hokkaido 001-0021, Japan

7 Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

8 Molecular Ecology Group (MEG), Water Research Institute (IRSA), National Research Council of Italy (CNR), 28922 Verbania Pallanza, Italy

9 National Biodiversity Future Center (NBFC), Palermo, Italy

10 Institute of Ecology and Botany, HUN-REN Centre for Ecological Research, Vácrátót, Hungary

11 Department of Environmental and Landscape Geography, ELTE Eötvös Loránd University, Budapest, Hungary

12 Department of Biology, University of North Carolina, Chapel Hill, NC 27599, USA

13 Centre of Marine Sciences, University of Algarve, Campus de Gambelas, Faro, Portugal

14 CAS Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

15 University of Chinese Academy of Sciences, Beijing 100049, China