Int. J. Aquat. Biol. (2024) 12(6): 533-545
ISSN: 2322-5270; P-ISSN: 2383-0956
Journal homepage: www.ij-aquaticbiology.com

Original Article

Quantifying how spatial resolution affects fish distribution model performance and prediction: A case study of Caspian Kutum, *Rutilus frisii*

Fateh Moëzzi, Soheil Eagderi*

Department of Fisheries, Faculty of Natural Resources, College of Agriculture and Natural Resources, University of Tehran, Tehran, Iran.

Abstract: The present study aimed to investigate the effect of spatial resolution of data on distribution modelling performance for the Caspian Kutum, *Rutilus frisii.* A set of spatial resolutions (4, 8, 16, 32, and 64 km) were considered in the modelling analyses, using sea surface temperature, chlorophyll-*a* concentration, particulate organic carbon content, bottom slope, and depth as environmental predictors of fish catch-per-unit-of-effort (CPUE). The boosted regression trees (BRT) method was applied as the modelling technique. The results showed considerable reductions in data variability (coefficient of variation (%) and variance) with decreasing spatial resolution for most environmental variables and CPUEs. The model performance (adj-R^2) was improved with decreasing resolutions, but the best prediction ability of the models was obtained with the BRTs fitted on the lowest resolutions (i.e. 4 and 8 km). While sea surface temperature was the main influencing predictor in the fitted BRTs at all resolutions, resolution-dependence differences were observed in the significance and response curves of other predictors of the models across the spatial resolutions. Overall, our findings indicated that using different levels of spatial resolution highly affects the modelling process, with more relevant explanations and higher prediction power using finer resolutions.

Article history:
Received 20 May 2024
Accepted 20 December 2024
Available online 25 December 2024

Keywords:
Boosted regression trees
Environmental predictors
Habitat modelling
Spatial scale

Introduction

Understanding the spatial characteristics of ecological processes and their related research frameworks is a fundamental influencing context in recognizing the impacts of changes in environmental conditions on marine fish populations. In aquatic organisms, the spatial dependence of the species-specific responses to the fluctuations in their surrounding environment is highly important in assessing their population dynamics (Luoto et al., 2007; Nystrom Sandman et al., 2013). Properly understanding the spatial ecology of different fish species can help fishery managers obtain more detailed and reliable information about the spatial and temporal trends of species distribution patterns and better plan conservation and exploitation activities (Cooke et al., 2016). The spatial frameworks used in ecological studies could affect environmental data quality and informativeness, analysis method, and the overall findings about ecosystem processes (Lecours et al., 2015). Generally, differences in the spatial dependence of these processes can lead to considerable differences in our understanding of species-environment relationships (Hale et al., 2019).

Species distribution models (SDMs), also known as habitat suitability models (HSMs), used extensively in assessing fish distribution, are based on identifying the relationships between species presence or abundance and a set of environmental variables describing their habitat (Nonez-Riboni et al., 2021). These models can predict changes in species distribution in space and time due to fluctuations in environmental conditions (Elith and Graham, 2000). Defining the appropriate scale of analyses is a crucial issue in SDM studies (Rushton et al., 2004; McPherson and Jetz, 2007; Hale et al., 2019). Using different resolutions (or scales) of data in fitting SDM models could result in different

*Correspondence: Soheil Eagderi
E-mail: soheil.eagderi@ut.ac.ir

DOI: https://doi.org/10.22034/ijab.v12i6.2398

conclusions for the same ecological question, leading to misunderstandings about species-environment associations (Hale et al., 2019; Moëzzi et al., 2024). Also, the role and importance of environmental predictors in SDMs can be different in models fitted on various scales (Nonez-Riboni et al., 2021). Therefore, in modelling fish distribution, the selection of the proper resolution of the analyses must be made considering the study goal, investigated system characteristics, and qualitative and quantitative measures of the available data (Elith and Leathwick, 2009; Nonez-Riboni et al., 2021).

Identifying a level of resolution (or scale) at which the strongest correlative relationships of species presence and its habitat characteristics could be found can effectively improve the management plans and conservation activities for aquatic organisms (Hale et al., 2019). The prior studies that investigated the role of spatial resolution on model performance presented contradictory findings. In some of the studies, it has been reported that using low-resolution (i.e. coarse scale) data obtained better model fits and understanding of fish-environment relationships (Rahbek and Graves, 2001; Tabalske, 2002; Luoto et al., 2007; Guisan et al., 2017). In contrast, some of the researchers presented improvements in model prediction accuracy using high-resolution (i.e. fine-scale) data (Thuiller et al., 2005; Guisan et al., 2007; Redfern et al., 2008; Seo et al., 2009; Becker et al., 2010; Dyer et al., 2013; Guisan et al., 2017; Nonez-Riboni et al., 2019, 2021). However, some other studies suggested applying a multi-scale (resolution) approach in evaluating distribution modelling to obtain the best practical levels of performance and prediction ability of the SDMs (Pearson et al., 2004; Bellier et al., 2010; de Knegt et al., 2010; Austin and Van Niel, 2011; Garcia-Callejas and Araujo, 2016; Kärcher et al., 2019). Overall, most of these research works emphasized selecting the proper resolution in modelling analyses considering the addressed question, the variability in environmental conditions, and the biological characteristics of the studied species.

Caspian Kutum, *Rutilus frisii* (Norman, 1840), from the Leuciscidae family (Eagderi et al., 2022), is one of the important Caspian Sea bony fish species. This fish is of high commercial and conservation importance (Esmaeili et al., 2015) and has a wide distribution area over the southern waters of the Caspian Sea (Rabazanov et al., 2019). The commercial catch data of this fish and the environmental condition at fishing locations along the Iranian coast of the Caspian Sea were used as a case study to assess the effect of the spatial resolution of data on distribution modelling performance and predictions. The research objectives were to investigate the variability of environmental parameters, model fitting, prediction performance, and differences in main model variables and their influencing patterns across a set of spatial resolutions.

Materials and Methods

Fish catch data: In this study, the commercial catch data of the Caspian Kutum were used as an abundance index of the fish, including fish biomass (kg), fishing operation time (hours), and the number of dragged seine nets. These data cover a 10-year period, including catch seasons from 2002/03 to 2011/12 and more than 150 fishing points along the southern coast of the Caspian Sea (Fig. 1). The catch data were standardized as catch per unit of effort (CPUE) using the equation of $CPUE_{ij} = (Catch_{ij} / (\text{Number of seine nets}_{ij} \times \text{Fishing time}_{ij})$, where $CPUE_{ij}$ is the CPUE (kg net^{-1} h^{-1}) for the catch season i, and fishing point j.

Environmental predictors: Five environmental variables were selected as the potential drivers of the Caspian Kutum distribution over the fishing points, regarding their reported ecological relevance in prior studies (Moëzzi et al., 2022, 2024), including near-surface chlorophyll-*a* concentration (Chl-*a*; mg m^{-3}), day-time sea surface temperature (SST; °C), particulate organic carbon concentration (POC; mg m^{-3}), bottom slope (°) and depth (m). The monthly data of the Chl-*a*, SST, and POC were obtained from the MODIS database (MODIS: Moderate Resolution Imaging Spectroradiometer, United States National Aeronautics and Space Administration (NASA) Goddard Space Flight Center, Ocean Ecology

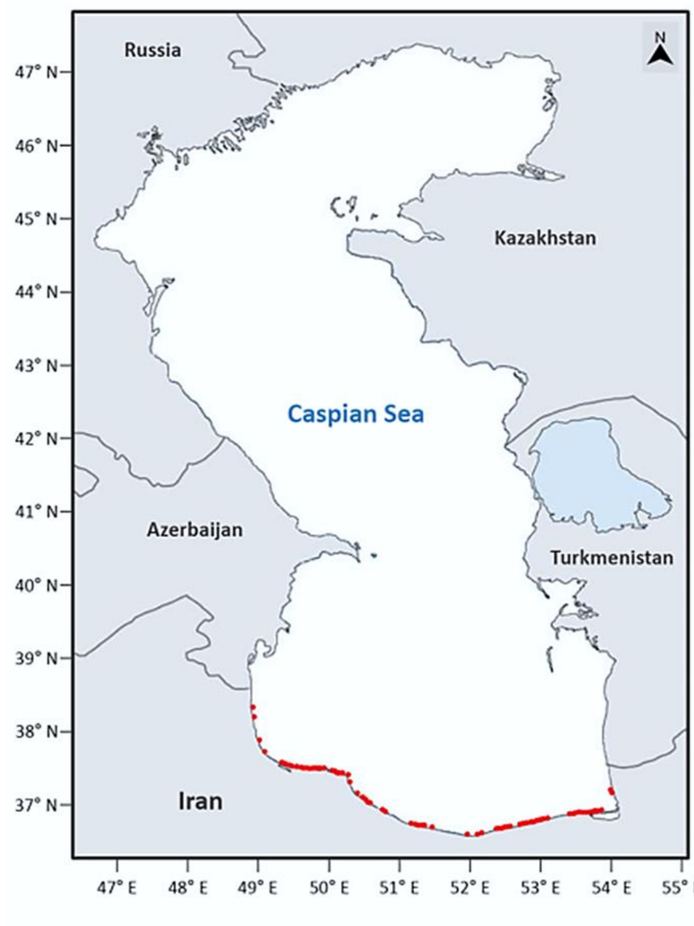

Figure 1. Geographical distribution of fishing points (●) for the Caspian Kutum, *Rutilus frisii*, over southern waters of the Caspian Sea.

Laboratory, 2021 (http://modis.gsfc.nasa.gov)). The depths of water at fishing points were obtained from the world bathymetry raster file (GEBCO: General Bathymetric Chart of the Oceans, 2021 (http://www.gebco.net)), following the approach used by Moëzzi et al. (2022). The levels of the bottom slope of the fishing locations were extracted from the slope map made from the bathymetry map. The extraction of environmental predictors' values from original remotely-sensed data formats and their conversion were conducted using the "raster" package (Version: 3.6-30; Hijmans, 2024) in R 4.4.2 (R Development Core Team, 2024).

Spatial resolutions: A set of spatial resolutions was selected based on the original resolution of the remotely sensed data: 4 km, including 4, 8, 16, 32, and 64 km. From here in the text, these are referred to as SR-A, SR-B, SR-C, SR-D, and SR-E, respectively. For each spatial resolution, the values of the environmental variables and CPUEs were averaged to obtain one single value of each variable for each spatial interval over each catch season.

Fitting and evaluation distribution models: The relationships between CPUEs and environmental predictors were assessed using boosted regression trees (BRT). This machine learning technique has been extensively used in distribution modelling studies of fishes (Anderson et al., 2016; Froeschke and Froeschke, 2016; Moëzzi et al., 2024), with high predictive performance and explanatory power (Elith et al., 2006; Froeschke and Froeschke, 2016). BRT models were fitted using the "gbm" package (Elith et al., 2008; Ridgeway, 2024). Model parameters were automatically tested until obtaining the best fit using an interaction depth of 3, a learning rate of 0.01-0.001, a bag fraction of 0.75, and maximum trees of 10,000 with a Gaussian error distribution (Moëzzi et al., 2024).

For each spatial resolution, 80% of the data (first eight catch seasons) were used for model training, and the fitted models were tested using the remaining 20% of data (the last two catch seasons). The fitting of the BRTs at each resolution was repeated using bootstrapped subsamples of datasets (10 iterations). The goodness-of-fit of the BRTs was evaluated using adjusted R squared (Adj-R^2), and their predictive ability was evaluated using testing datasets of all resolutions, based on normalized root mean squared error (nRMSE) scores, calculated as follows: $nRMSE = (\sqrt{\sum_{(i=1)}}\text{^}N(y_i - \hat{y}_i)\text{^}2 / N)) / (y_{max} - y_{min})$, where y_i and $\hat{y}_i$ are the raw and predicted values, N is the number of data points, and y_{max} and y_{min} represent the maximum and minimum values of y in each dataset, respectively.

Considering the five environmental predictors used for fitting the models, the parameters with relative importance (RI) scores higher than 20% (= 100/5) were recognized as significant predictors of the models (Thorn et al., 2016). Also, all predictors' partial dependency plots (PDPs) were compared between BRTs fitted on datasets with different resolutions.

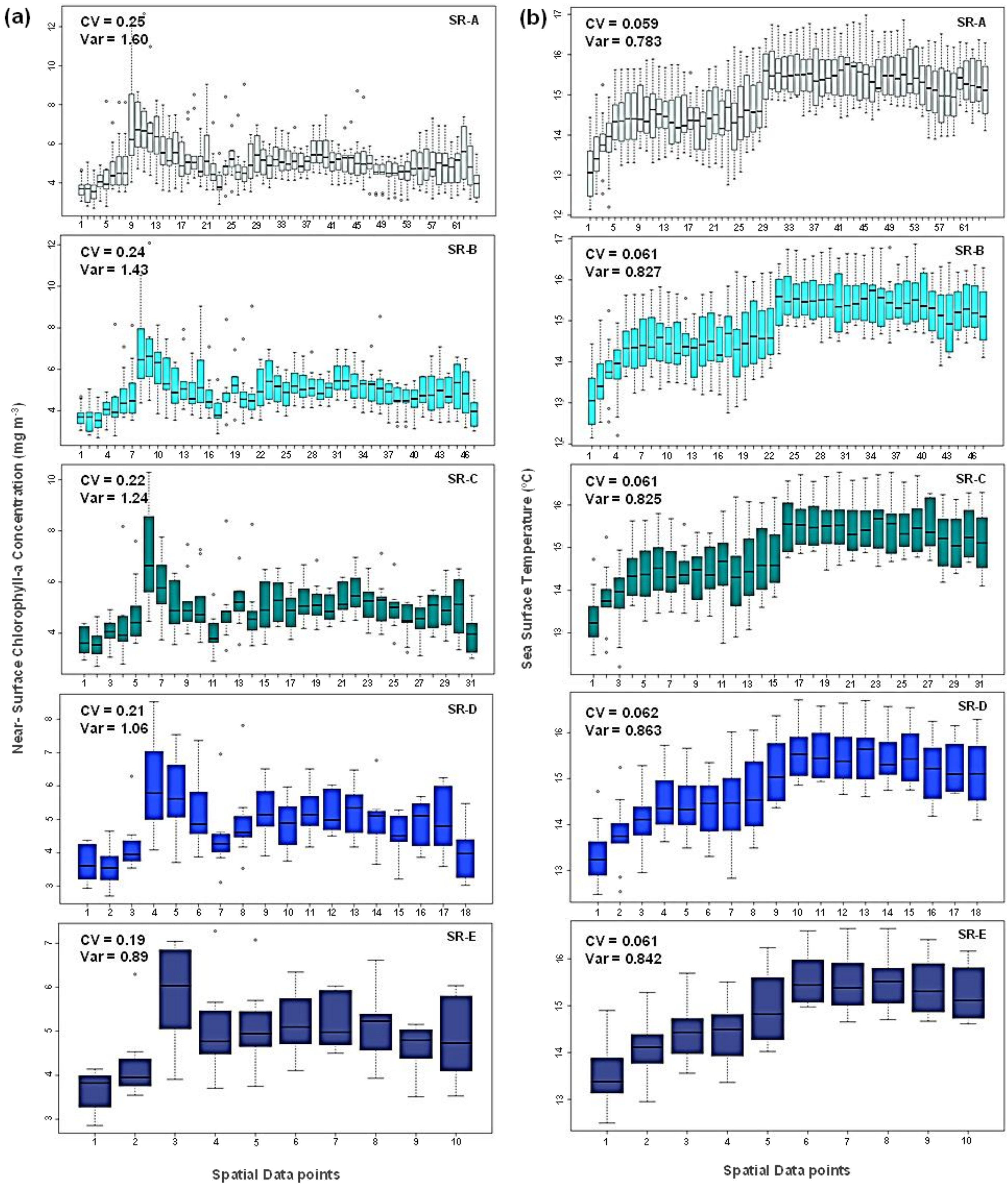

Figure 2. Boxplots of the environmental predictors and mean CPUEs over the studied fishing points at different spatial resolutions (SR-A: 4 km; SR-B: 8 km; SR-C: 16 km; SR-D: 32 km; and, SR-E: 64 km). (CV: coefficient of variation; Var: variance).

Results

Heterogeneity of environmental condition: The obtained distributions of environmental parameters' levels over the studied geographical extent with different resolutions (Fig. 2) showed that with decreases in spatial resolution (from SR-A to SR-E),

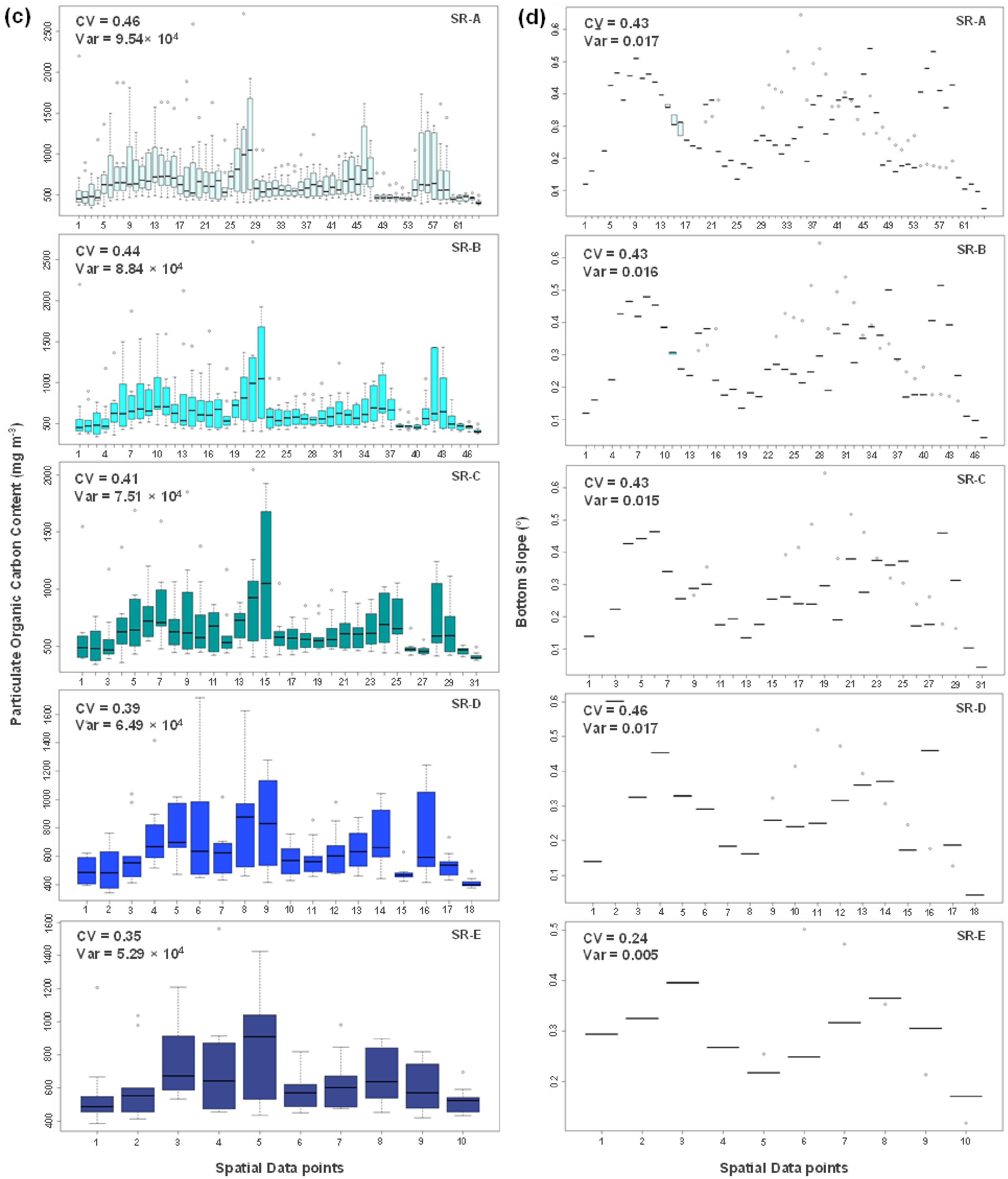

Figure 2. To be continued.

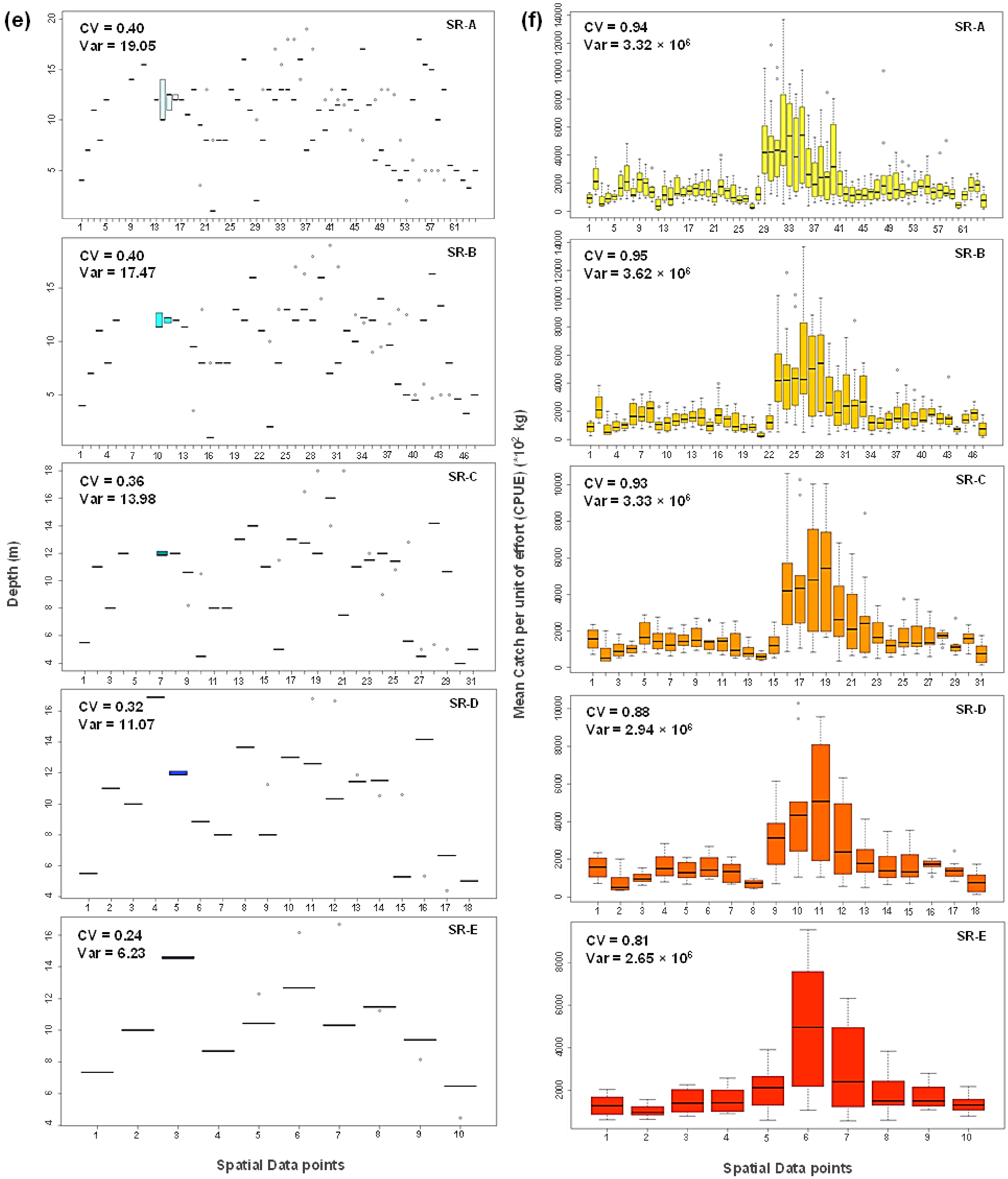

Figure 2. To be continued.

the coefficient of variation (CV) and variance of data were decreased for Chl-*a*, POC, bottom slope and depth, while there was no considerable difference in CV (6.2-5.9%) and variance (0.863-0.783) for SST

Table 1. Adjusted-R^2 (Adj-R^2) and normalized root mean squared error (nRMSE) scores of the fitted boosted regression trees (BRT) models using datasets at different spatial resolutions (SR-A: 4 km; SR-B: 8 km; SR-C: 16 km; SR-D: 32 km; and SR-E: 64 km).

Fitted model	Spatial resolution	Adj-R^2	nRMSE for Testing data				
			SR-A	SR-B	SR-C	SR-D	SR-E
BRT_{SR-A}	4 km	0.501	0.183	0.186	0.183	0.232	0.318
BRT_{SR-B}	8 km	0.587	0.185	0.175	0.182	0.221	0.304
BRT_{SR-C}	16 km	0.703	0.223	0.207	0.192	0.250	0.349
BRT_{SR-D}	32 km	0.851	0.246	0.223	0.240	0.284	0.326
BRT_{SR-E}	64 km	0.887	0.306	0.316	0.329	0.351	0.309

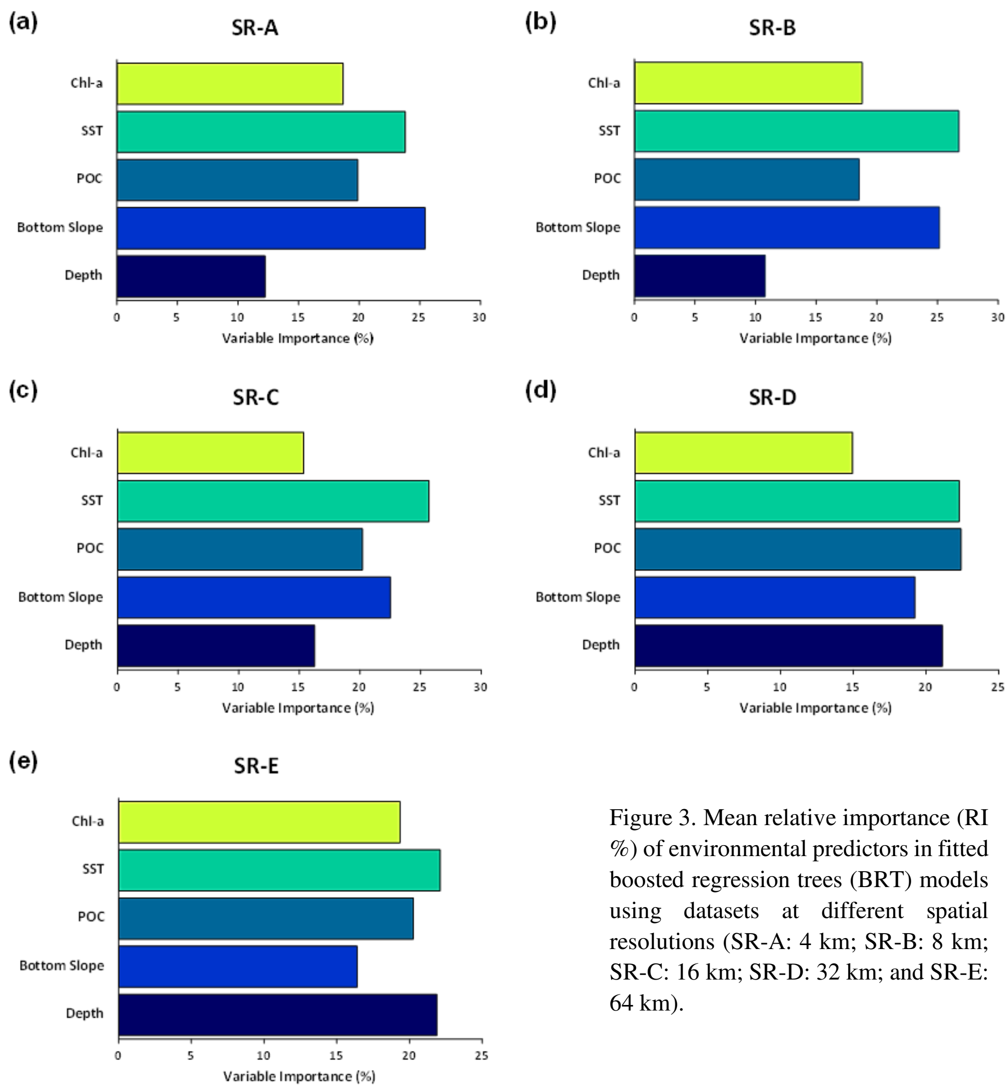

Figure 3. Mean relative importance (RI %) of environmental predictors in fitted boosted regression trees (BRT) models using datasets at different spatial resolutions (SR-A: 4 km; SR-B: 8 km; SR-C: 16 km; SR-D: 32 km; and SR-E: 64 km).

across different resolutions. The order of reductions in CV and Var for these parameters from SR-A to SR-E was: Depth (CV: 19%; Var: 70.5%) > bottom slope (CV: 16%; Var: 67.3%) > POC (CV: 11%; Var: 44.5%) > Chl-*a* (CV: 6%; Var: 44.3%). Also, averaged mean-CPUEs depicted decreases of 13% (from 94 to 85%) in CV and 20.8% (from 3.32×10^6 to 2.65×10^6) in variance with decreases in resolution from SR-A to SR-E.

Model performance and prediction evaluation: Fitting BRT models using training datasets with different resolutions (Table 1) showed higher adjusted-R^2 scores with decreased data resolution, whereas BRT_{SR-A} and BRT_{SR-E} models had the lowest and highest values of adj-R^2, respectively. Predictions of the fitted BRTs on testing datasets with different spatial resolutions indicated that BRT_{SR-B} had the best prediction performance on most testing datasets.

Other BRT models showed decreases in prediction ability with decreases in data resolution, even for models with the same resolutions for the training and testing datasets.

Environmental variables importance and response curves: The relative importance (RI) scores of BRT model parameters (Fig. 3) indicated differences in significant variables of the models fitted using the datasets with different spatial resolutions. SST (with RIs > 22.09%) was the main significant variable across all resolutions, and Chl-*a* was recognized as a non-significant variable in all the models (RIs < 19.35%). The bottom slope was significant in the BRT_{SR-A}, BRT_{SR-B}, and BRT_{SR-C} models, while it did not have significant effects in BRT_{SR-D} and BRT_{SR-E}. In contrast, depth was significant in low-resolution models (i.e., BRT_{SR-D} and BRT_{SR-E}), without significant effects in other BRTs fitted on higher-resolution data. The effect of POC was significant only in medium- to low-resolution models (i.e., BRT_{SR-C}, ERT_{SR-D}, and BRT_{SR-E}).

The response curves (i.e., partial dependency plots (PDPs)) of the environmental predictors were found with different patterns in BRTs fitted on different resolutions (Fig. 4). An overall decreasing trend was obtained for Chl-*a* across all of the resolutions but with higher differences in the maximum and minimum influencing levels at lower resolutions (Fig. 4a). The PDP graphs of SST depicted relatively similar increasing trends for all of the resolutions (Fig. 4b). The obtained PDPs for POC presented nearly opposite patterns between high to intermediate (BRT_{SR-A}, BRT_{SR-B}, and BRT_{SR-C}) and low-resolution (BRT_{SR-D} and BRT_{SR-E}) models (Fig. 4c). The influencing patterns of the bottom slope for SR-A to SR-D resolutions showed clear bell-shaped curves. For SR-E, the obtained curve had a decreasing trend over a shorter range of the parameter (Fig. 4d). The PDPs of depth for the BRT_{SR-A} and BRT_{SR-B} models with the higher influences at lower variable levels were different from the increasing curves of the models obtained using SR-C to SR-E resolutions, where the highest levels of influence in them belonged to the depths of more than 10 m. The highest difference between the maximum and minimum influencing levels was observed for the model of the lowest spatial resolution (Fig. 4e).

Discussions

The aggregation of data of the environmental variables with decreases in spatial resolution (i.e., increases of scale) changes the variability of data with considerable reductions in sampling units (or data points) for coarse resolutions that can affect fish distribution modelling (Scales et al., 2016). This study's results showed clear reductions in the variability of data for most of the environmental parameters (i.e., Chl-a, POC, bottom slope, and depth), as well as CPUEs, with decreases in resolution. Also, the levels of reduction were different between the investigated variables. The spatial trends of the values for these parameters nearly had the same fluctuations with the highest levels over the middle range of the fishing points, and the averaging of data obviously led to considerable decreases in CV and variance of datasets with decreases in the resolution. This condition is in accordance with the general outcome of the change in data resolution with the higher frequencies and variation in abundance metrics and environmental conditions in higher resolutions or finer-scale frameworks (Garcia-Charton et al., 2004; Nonez-Riboni et al., 2021). It has been reported that data aggregation can eliminate the noise in data and change patterns of environmental gradients and species abundance distribution (Tobalske, 2002; Redfern et al., 2008). However, in our study, the SST values had an increasing trend over the fishing points, and the results showed that aggregated averages of its values had no significant reduction in CV, nor in the variance across spatial resolutions, and the increasing pattern of SST was observed for all resolutions. Therefore, the geographical gradient of the environmental parameters could have different responses in the data aggregation process over spatial resolutions where the variability and variance of data could be changed or maintained in different spatial resolutions.

In the prior studies, the resolution (or scale) of the

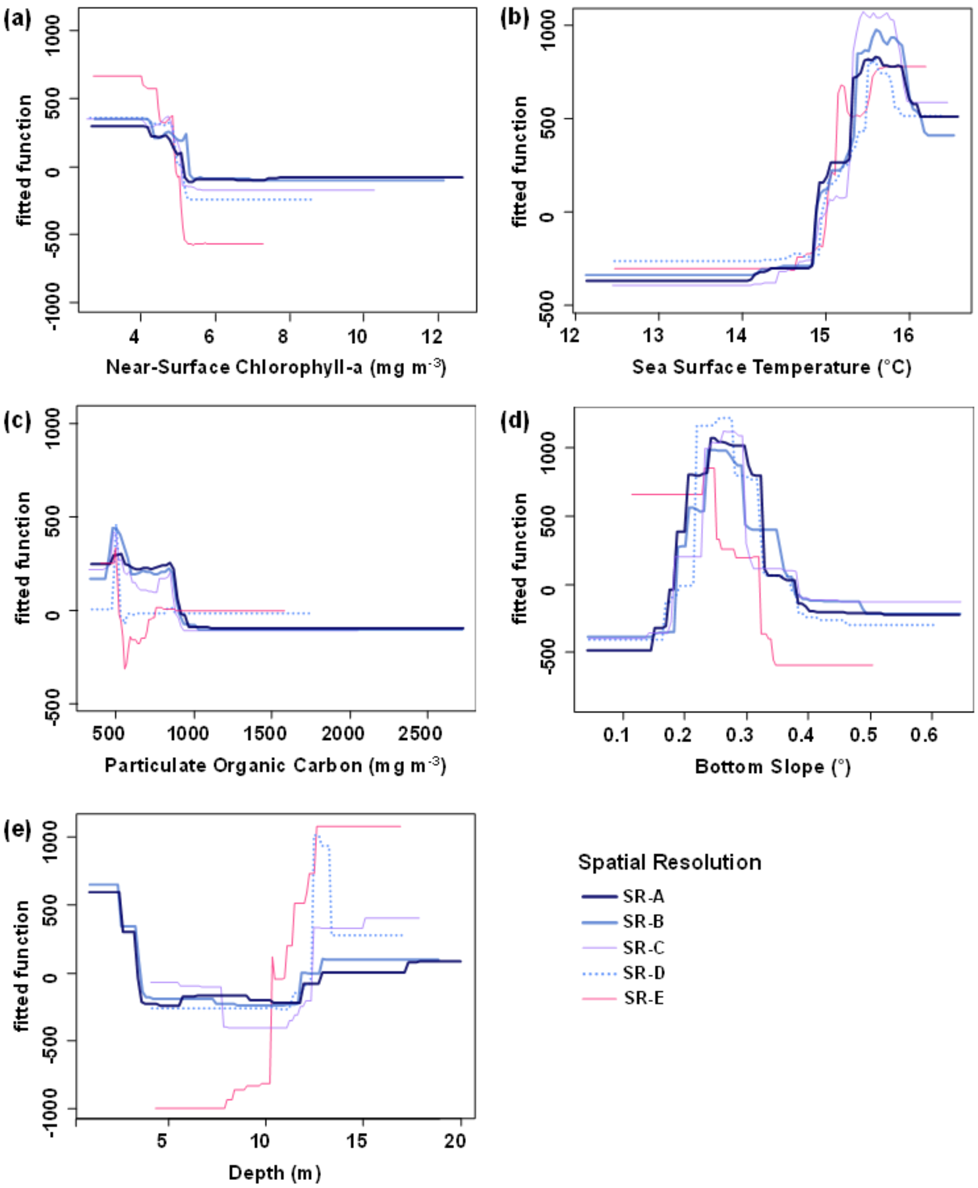

Figure 4. Mean partial dependency plots (PDPs) of the environmental predictors for the boosted regression trees (BRT) models fitted using datasets with different spatial resolutions (i.e., SR-A; 4 km; SR-B: 8 km; SR-C; 16 km; SR-D = 32 km; and SR-E: 64 km).

analyses has been mentioned as an important factor affecting the performance and predictive ability of SDMs (França and Cabral, 2016). In the present study, the goodness-of-fit of the models was improved by decreasing the spatial resolution of data where the best fish abundance – environment relationship was obtained at the lowest resolution (SR-E). Some of the research has claimed to obtain better model performance by fitting the models at lower spatial resolutions, or in other words, using larger spatial scales (Tabalske, 2002; Luoto et al., 2007; Hale et al., 2019; Nonez-Riboni et al., 2021). This condition is probably due to the lower noise and variation in data made by averaging the environmental variables and CPUEs, as well as fewer data points at lower-resolution data sets. It is difficult to capture the noise and variability in data at finer resolutions using statistical techniques and obtain a higher explaining ability of the models and strong fish-habitat associations compared to the situations using lower resolutions of data (Nonez-Riboni et al., 2021). Moreover, the decrease in sample counts at lower resolutions is reported as another factor that improved the performance of the models (Hale et al., 2019).

However, in our results, we found decreases in the predictive power of the models with decreases in spatial resolution. The $BRT_{SR\text{-}A}$ and $BRT_{SR\text{-}B}$ models, fitted at nearly high resolutions, had the best predictions, not only on the test data of the exact resolution but also at all of the studied spatial resolutions. Also, the fitted models' predictive power decreased with resolution decreases, where the lowest accuracy levels of predictions were obtained for the $BRT_{SR\text{-}E}$, fitted on the lowest resolution training data. This result is supported by some of the prior studies, which reported higher prediction accuracy with models fitted on finer scales (Collingham et al., 2000; Guisan et al., 2007; Seo et al., 2009; Dyer et al., 2013; Ross et al., 2015; França and Cabral, 2016). The inability to recognize reliable relationships between fish abundance and real environmental variables could lead to lower accuracy of model predictions (Fernandez et al., 2018). Overall, using intermediate to high spatial resolutions in SDM modelling analyses and considering more details in the data could lead to models with high predictive ability, even with nearly weak but realistic fish-environment relationships.

The estimates of the importance of the environmental predictors in BRT models in explaining CPUE variation showed two distinct conditions. In all models, SST was one of the main significant predictors with a considerable effect on fish distribution at all spatial resolutions. This means that this variable has an important effect on fish distribution without any specific dependency on the spatial scale of the analyses. Such a condition is previously reported by Hale et al. (2018). The response curves for this parameter also clearly showed the role of fish distribution, with nearly similar PDPs across all investigated resolutions and obvious increasing trends over the variable range. This condition is highly related to SST's constant variation metrics (CV and Var) in all resolutions. However, a resolution (or scale)-dependent pattern was observed for other environmental predictors, where there was a change in significance of variable from high-resolution (SR-A (bottom slope > SST) and SR-B (SST > bottom slope)) to intermediate (SR-C (SST > bottom slope > POC)) to low-resolution models (SR-D (POC > SST > depth) and SR-E (SST > depth > POC)). The response curves of these variables also depicted nearly similar and matching patterns for the resolutions. They are significant and somewhat contrasting with their curves at other resolutions. Differences in the contributions of environmental variables in SDMs and their differences across spatial scales are also reported by Pitman and Brown (2011), França and Cabral (2016), and Hale et al. (2018), indicating this fact that assessing fish distribution at different spatial resolutions could lead to different variables as the main habitat drivers of them with nearly different patterns of influence over their measured ranges. Based on this, we suggest using multiple spatial resolutions of data to assess model performance and prediction accuracy and consequently correctly determine the main drivers of fish distribution at the effective ecological scale.

Conclusion

In the present study, the effect of spatial resolution of data on distribution modelling of the Caspian Kutum was investigated across a set of resolutions including 4, 8, 16, 32 and 64 km. Our findings clearly showed decreases in the variability of data for most of the environmental predictors (except SST) as well as fish CPUEs, with decreases in spatial resolution. Fitting boosted regression trees (BRTs) presented clear improvements in model fits with an increase in data scale (i.e., a decrease of the resolution), but the BRT models fitted on the lowest resolutions (4 and 8 km) had the highest predictive power over all investigated resolutions. Also, based on the results, exceptionally for the SST that was significant in BRTs of all resolutions with nearly similar response curves, the other environmental parameters showed obvious resolution-dependent differences in their significance (i.e., relative importance) and influencing pattern in the fitted models across spatial resolutions. Considering all of the results, we suggest using a multi-resolution (or scale) approach in modelling analyses of fish distribution to detect the best spatial resolution of the modelling framework and obtain reliable fish-environment relationships and a proper understanding of the environmental drivers of the organism's distribution.

Acknowledgment
This study was funded by the Iran's National Foundation Elites and the University of Tehran.

References

Anderson O.F., Guinotte J.M., Rowden A.A., Clark M.R., Mormede S., Davies A.J., Bowden D.A. (2016). Field validation of habitat suitability models for vulnerable marine ecosystems in the South Pacific Ocean: Implications for the use of broad-scale models in fisheries management. Ocean and Coastal Management, 120: 110-126.

Austin M.P., Van Niel K.P. (2011). Improving species distribution models for climate change studies: Variable selection and scale. Journal of Biogeography, 38: 1-8.

Becker E.A., Forney K.A., Ferguson M.C., Foley D.G., Smith R.C., Barlow J., Redfern J.V. (2010). Comparing California Current cetacean–habitat models developed using in situ and remotely sensed sea surface temperature data. Marine Ecology Progress Series, 413: 163-183.

Bellier E., Certain G., Planque B., Monestiez P., Bretagnolle V. (2010). Modelling habitat selection at multiple scales with multivariate geostatistics: An application to seabirds in open sea. Oikos, 119: 988-999.

Collingham Y.C., Wadsworth R.A., Huntley B., Hulme P.E. (2000). Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. Journal of Applied Ecology, 37: 13-27.

Cooke S.J., Martins E.G., Struthers D.P., Gutowsky L.F.G., Power M., Doka S.E., Dettmers J.M., Crook D.A., Lucas M.C., Holbrook C.M., Krueger C.C. (2016). A moving target-incorporating knowledge of the spatial ecology of fish into the assessment and management of freshwater fish populations. Environmental Monitoring and Assessment, 188: 1-18.

de Knegt H.J., van Langevelde F., Coughenour M.B., Skidmore A.K., de Boer W.F., Heitkönig I.M.A., Knox N.M., Slotow R., van der Waal C., Prins H.H.T. (2010). Spatial autocorrelation and the scaling of species–environment relationships. Ecology, 91: 2455-2465.

Dyer J.J., Brewer S.K., Worthington T.A., Bergey E.A. (2013). The influence of coarse-scale environmental features on current and predicted future distributions of narrow-range endemic crayfish populations. Freshwater Biology, 6: 1071-1088.

Eagderi S., Mouludi-Saleh A., Esmaeili H.R., Sayyadzadeh G., Nasri M. (2022). Freshwater lamprey and fishes of Iran: a revised and updated annotated checklist-2022. Turkish Journal of Zoology, 46(6): 500-522.

Elith J., Leathwick J.R., Hastie T. (2008). A working guide to boosted regression trees. Journal of Animal Ecology, 77(4): 802-813.

Elith J., Leathwick J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics, 40: 677-697.

Elith J., Graham C.H. (2009). Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. Ecography, 32: 66-77.

Esmaeili H.R., Brian W.C., Mehraban H.R., Masoudi M., Khaefi R., Abbasi K., Mostafavi H., Vatandoust S. (2015). An updated checklist of fishes of the Caspian Sea basin of Iran with a note on their zoogeography. Iranian Journal of Ichthyology, 13: 152-184.

Fernandez M., Yesson C., Gannier A., Miller P.I., Azevedo J.M.N. (2018). A matter of timing: How temporal scale selection influences cetacean ecological niche modelling. Marine Ecology Progress Series, 595: 217-231.

França S., Cabral H. (2016). Predicting fish species distribution in estuaries: Influence of species' ecology in model accuracy. Estuarine, Coastal and Shelf Science, 180: 11-20.

Froeschke J.T., Froeschke B.F. (2016). Two-stage boosted regression tree model to characterize southern flounder distribution in Texas estuaries at varying population sizes. Marine and Coastal Fisheries, 8(1): 222-231.

García-Callejas D., Araújo M.B. (2016). The effects of model and data complexity on predictions from species distributions models. Ecological Modelling, 326: 4-12.

Garcia-Charton J.A., Petez-Ruzafa A., Sanchez-Jerez P., Bayle-Sempere J.T., Renones O., Moreno D. (2004). Multi-scale spatial heterogeneity, habitat structure, and the effect of marine reserves on Western Mediterranean rocky reef fish assemblages. Marine Biology, 144: 161-182.

GEBCO. (2021). General bathymetric chart of the oceans. https://www.gebconet/data_and_products/gridded_bat

hymetry_data
Guisan A., Graham C.H., Elith J., Huettmann F. (2007). Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions, 13: 332-340.
Guisan A., Thuiller W., Zimmermann N.E. (2017). Habitat suitability and distribution models. With applications in R. Cambridge University Press. 478 p.
Hale R., Colton M.A., Peng P., Swearer S.E. (2019). Do spatial scale and life history affect fish-habitat relationships? Journal of Animal Biology, 88: 439-449.
Hijmans R.J. (2024). Raster: Geographic data analysis and modeling. R package version 3.6-30. http://CRAN.R-project.org/package=raster
Kärcher O., Frank K., Walz A., Markovic D. (2019). Scale effects on the performance of niche-based models of freshwater fish distributions. Ecological Modelling, 405: 33-42.
Lecours V., Devillers R., Schneider D.C., Lucieer V.L., Brown C.J., Edinger E.N. (2015). Spatial scale and geographic context in benthic habitat mapping: a review and future directions. Marine Ecology Progress Series, 535: 259-284.
Luoto M., Virkkala R., Heikkinen R.K. (2007). The role of land cover in bioclimatic models depends on spatial resolution. Global Ecology and Biogeography, 1: 34-42.
McPherson J.M., Jetz W. (2007). Effects of species' ecology on the accuracy of distribution models. Ecography, 30: 135-151.
MODIS. (2021). Moderate Resolution Imaging Spectroradiometer, United States National Aeronautics and Space Administration (NASA) Goddard Space Flight Center. Ocean Laboratory Laboratory. http://modei.gsfc.nasa.gov/
Moëzzi F., Poorbagher H., Eagderi S., Feghhi J., Dormann C.F., Khorshidi Nergi S., Amiri K. (2022). Modelling habitat preference of Caspian Kutum, *Rutilus kutum,* using non-linear habitat suitability indices and generalized additive models. Regional Studies in Marine Science, 56: 102715.
Moëzzi F., Poorbagher H., Eagderi S., Feghhi J., Dormann C.F., Khorshidi Nergi S., Amiri K. (2022). The importance of temporal scale in distribution modeling of migratory Caspian Kutum, *Rutilus frisii*. Ecology ad Evolution, 14(9): e70259
Nunez-Riboni I., Akimova A., Sell A.F. (2021). Effect of data spatial scale on the performance of fish habitat models. Fish and Fisheries, 22: 955-973.
Núñez-Riboni I., Taylor M.H., Kempf A., Püts M., Mathis M. (2019). Spatially resolved past and projected changes of the suitable thermal habitat of North Sea cod (*Gadus morhua*) under climate change. ICES Journal of Marine Science, 76: 2389-2403.
Nyström Sandman A., Wikström S.A., Blomqvist M., Kautsky H., Isaeus M. (2013). Scale-dependent influence of environmental variables on species distribution: a case study on five coastal benthic species in the Baltic Sea. Ecography, 36(3): 354-363.
Pearson R.G., Dawson T.P., Liu C. (2004). Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. Ecography, 27: 285-298.
Pittman J.P., Brown A.B. (2011). Multi-scale approach for predicting fish species distributions across coral reef seascapes. PLoS ONE, 6(5): e20583.
R Development Core Team (2024). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Rabazanov N.I., Orlov A.M., Abdosamadov A.S., Barkhalov R.M. (2019). Caspian Kutum *Rutilus kutum*: A story of exploitation, survival, and revival. In: H.C. Krueger, W.W. Taylor (Eds.), From catastrophe to recovery: Stories of fishery management success. American Fisheries Society. pp: 485-508.
Rahbek C., Graves G.R. (2001). Multiscale assessment of patterns of avian species richness. Proceedings of the National Academy of Sciences, 8: 4534-4539.
Redfern J.V., Barlow J., Ballance L.T., Gerrodette T., Becker E.A. (2008). Absence of scale dependence in dolphin–habitat models for the eastern tropical Pacific Ocean. Marine Ecology Progress Series, 363: 1-14.
Ridgeway G., Developers G. (2024). gbm: Generalized boosted regression models. R package version 2.2.2. http://github.com/gbm-developers/gbm.
Ross L.K., Ross R.E., Stewart H.A., Howell K.L. (2015). The influence of data resolution on predicted distribution and estimates of extent of current protection of three "listed" deep-sea habitats. PLoS One, 10(10): e0140061.
Rushton S.P., Ormerod S.J., Kerby G. 2004. New paradigms for modelling species distributions? Journal of Applied Ecology, 41: 193-200.
Scales K.L., Hazen E.L., Jacox M.G., Edwards C.A., Boustany A.M., Oliver M.J., Bograd S.J. (2016). Scale of inference: on the sensitivity of habitat models for wide-ranging marine predators to the resolution of

environmental data. Ecography, 39: 001-011

Seo C., Thorne J.H., Hannah L., Thuiller W. (2009). Scale effects in species distribution models: Implications for conservation planning under climate change. Biology Letters, 1: 39-43.

Thuiller W., Lavorel A., Araújo M.B., Sykes M.T., Prentice I.C. (2005). Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Science USA, 23, 8245.

Tobalske C. (2002). Effects of spatial scale on the predictive ability of habitat models for the Green Woodpecker in Switzerland. In: J.M. Scott, P.J. Heglund, F. Samson, J. Haufler, M. Morrison, M. Raphael, B. Wall (Eds.), Predicting species occurrences: Issues of accuracy and scale. Island Press. pp: 197-205.