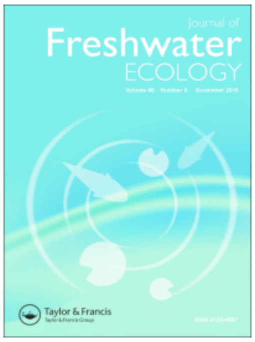
An application of maximum entropy model to evaluate the differential effect of cage aquaculture on the distribution of a native and an endemic fish species in Lake Maninjau, Indonesia

Yuniarti, Ivana; Glenk, Klaus; Yusri, Safran; Sutrisno

Published in:
Journal of Freshwater Ecology

DOI:
10.1080/02705060.2023.2192237

First published: 25/03/2023



Journal of Freshwater Ecology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjfe20

An application of maximum entropy model to evaluate the differential effect of cage aquaculture on the distribution of a native and an endemic fish species in Lake Maninjau, Indonesia

Ivana Yuniarti, Klaus Glenk, Safran Yusri & Sutrisno

To cite this article: Ivana Yuniarti, Klaus Glenk, Safran Yusri & Sutrisno (2023) An application of maximum entropy model to evaluate the differential effect of cage aquaculture on the distribution of a native and an endemic fish species in Lake Maninjau, Indonesia, Journal of Freshwater Ecology, 38:1, 2192237, DOI: 10.1080/02705060.2023.2192237

To link to this article: https://doi.org/10.1080/02705060.2023.2192237

Published online: 25 Mar 2023.

JOURNAL OF FRESHWATER ECOLOGY
2023, VOL. 38, NO. 1, 2192237
https://doi.org/10.1080/02705060.2023.2192237

RESEARCH ARTICLE

OPEN ACCESS

An application of maximum entropy model to evaluate the differential effect of cage aquaculture on the distribution of a native and an endemic fish species in Lake Maninjau, Indonesia

Ivana Yuniarti[a,b,c], Klaus Glenk[c], Safran Yusri[d] and Sutrisno[b]

[a]School of Geosciences, University of Edinburgh, Edinburgh, UK; [b]Research Centre for Limnology and Water Resources, National Research and Innovation Agency, Indonesian Institute of Sciences, Cibinong, Indonesia; [c]Department of Rural Economy, Environment and Society, SRUC, Edinburgh, UK; [d]Yayasan Terumbu Karang Indonesia, Depok, Indonesia

ABSTRACT

Physical cage aquaculture structure can attract native fish species in marine and freshwater ecosystems. Most studies on the effects of cage farms on native fish distribution have been undertaken in marine environments and outside of Asia as the main freshwater cage aquaculture producing region. Many studies have emphasised connections between native fish distribution and feeding time. Previous research also has shown the necessity to use modelling to monitor this effect to reduce data collection costs. Here we analyse the distribution of an endemic fish species, *Rasbora maninjau*, and a native fish, *Gobiopterus sp*., associated with tilapia cage aquaculture occurrence using a Maximum Entropy Model (MaxEnt). We find that the application of the MaxEnt model can produce reliable and accurate information on the impacts of cage aquaculture on the native fish species distribution aligning with the more expensive count data method. Our results also suggest that the species-specific interaction between the native fish and cage farms is mainly arising from an interaction between the ecological behaviour of the native fish with dimensions of the environmental condition such as turbidity. Our study therefore highlights the importance for improved appraisal of the ecology of native fish in the cage aquaculture risk assessment.

ARTICLE HISTORY
Received 16 January 2023
Accepted 13 March 2023

KEYWORDS
Aquaculture impacts; fish distribution; modelling; freshwater lake; tilapia farm

1. Introduction

Growing global demand for fish has led to a rapid increase in freshwater cage aquaculture (FAO 2016; Naylor et al. 2021). In Asia, freshwater bodies in populous fish-eating regions have become centres of intensive cage aquaculture development (Newton et al. 2021; Taskov et al. 2021). Freshwater cage aquaculture can contribute to nutritional security and its function as rural livelihood support has stimulated its growth in some

CONTACT Ivana Yuniarti ivan001@brin.go.id, ivanayuniarti02@gmail.com School of Geosciences, University of Edinburgh, Edinburgh, UK

countries (Rajee and Mun 2017; Shava and Gunhidzirai 2017; Njiru et al. 2018). However, there are increasing conflicts with other freshwater resource users, particularly where rapid expansion of cage aquaculture has resulted in eutrophication of reservoirs, declines in yields of native biota and increased occurrences of "fish kills" (Abery et al. 2005; Endah et al. 2017; Taskov et al. 2021).

The rapid expansion of freshwater cage aquaculture affects native fish communities in various ways. Firstly, the release of organic matter from feed remains results in increases in primary productivity which can lead to eutrophication, change in water quality (Abdel-Tawwab et al. 2002; Baccarin et al. 2005; Canonico et al. 2005; Chislock et al. 2013; Verdegem 2013) and decrease of benthic fauna richness (Tomassetti et al. 2016). Secondly, native fish communities may be affected by alterations of habitat and water quality and exchanges of parasites and diseases between the native and cultured fish (Barrett et al. 2018; Njiru et al. 2018). Thirdly, native fish populations may be negatively impacted by competition for resources, predation, and hybridization, whether intentionally or accidentally (Barrett et al. 2018).

Recent studies of cage aquaculture in freshwater systems have also investigated the effects of the cage structures on native fish populations (Demétrio et al. 2012; Ramos et al. 2013; Nobile et al. 2018). The physical cage structure may act to aggregate wild fish, offer shelter, and attract wild fish to the continuous food input (Sanchez-Jerez et al. 2011; Nobile et al. 2018). Studies in neotropical reservoirs have suggested that cage structures can increase habitat complexity in freshwater bodies and provide protection from predation for native fish communities (Nobile et al. 2018). However, studies in tropical lakes in other biogeographical regions, for example in Asia as the primary freshwater aquaculture fish producing region, are scarce. Further, previous research identifying species-specific relationships between marine cage farms and native fish species mainly emphasised the influence of feeding time and season (Uglem et al. 2009; Bacher et al. 2012; Ballester-Moltó et al. 2015). Research addressing which factors affect the species-specific relationship between cage farms and native fish distribution in tropical lakes is rare. This study aims at filling these gaps by analysing the effects of tilapia cage farms on a native and an endemic species in Lake Maninjau, Indonesia.

Most research to date aimed at monitoring the cage aquaculture impact on native fish distribution relied on count data collection (Šegvić Bubić et al. 2011; Bacher et al. 2015; Pereira et al. 2019). Collecting count data such as abundance is time consuming and requires specific sampling techniques (Marini et al. 2018; Siddiqui et al. 2018) making it less affordable for lake managers in the study area and in many other areas in Indonesia and Global South Nations where monitoring efforts are limited. Research by Yuniarti et al. (2021) revealed that the lake managers in the study area have voiced concern about limited monitoring resources such as time, funds, and human resources.

In this study, we employ Maximum Entropy Models (MaxEnt) to examine the distribution of two fish species (bada and rinuak) in relation to farm occurrence in Lake Maninjau. MaxEnt requires presence data only and environmental variables which can be collected with relative ease. MaxEnt models have been widely applied in conservation science to predict species distribution and habitat suitability (Sobek-Swant et al. 2012; Moore et al. 2016; Yi et al. 2016; Zhang et al. 2019; Zhang et al. 2018) and has been proven useful when only small sample sizes are available (Elith et al. 2006; Wisz et al. 2008; West et al. 2016). Considering these merits, we test whether the model application can provide a more affordable monitoring approach for cage impacts of tilapia on bada and rinuak, and reflect on the wider applicability of the MaxEnt model approach for other inland water bodies.

2. Materials and Methods

2.1. Site description

Lake Maninjau (Figure 1) is one of several large volcanic lakes in West Sumatra, Indonesia. The maximum depth in the lake is 165 m. It is 16.6 km from north to south, and 7.5 km at its widest point (Fakhrudin et al. 2002).

Cage aquaculture was introduced to the lake in 1992 with the installation of 16 cage units which produced 96 tons of fish in three months (Syandri and Azrita 2013). The currently farmed fish species on Lake Maninjau is Nile tilapia (*Oreochromis niloticus*). The number of cage units present in Lake Maninjau has increased rapidly, with 4,000 further units established in 2003, totalling 23,566 in 2016 (Syandri et al. 2014; Agam Regency Fisheries Department 2017). (The increasing presence of cage aquaculture in the lake has caused concern that aquaculture development has exceeded the lake's carrying capacity and accelerated eutrophication (Said et al. 2020).

Long term satellite data showed that cage aquaculture proliferation has caused significant decline of the lake's water transparency (Setiawan et al. 2019). Further, annual yet patchy data of the lake's water quality revealed that the expansion of cage aquaculture operation is correlated with an decrease of oxic layer, an increase of total phosphorus concentration, and an elevation of trophic state index (Sulastri et al. 2015).

2.2. The fish species

The main fish catch of the lake consists of two benthopelagic fish species: *Rasbora maninjau* (Lumbantobing 2014) (local name: bada) and *Gobiopterus sp*. (local name: rinuak). Bada is an important endemic fisheries species and a local culinary icon (Dina et al. 2019). Meanwhile, rinuak, is an important native fish for local communities around Lake Maninjau.

Studies on these two fish species, particularly on rinuak, are scarce and limited. Experts have not yet agreed on the classification and species name of rinuak. The species name for rinuak had not been assigned at the time of writing. A study by Roesma et al. (2020) using, NCBI database (GenBank) described rinuak as a close relative to *Gobiopterus*, but did not specifically classify its genus. Further, research by Larashati (2019, unpublished data) using the BOLD database showed that it belongs to *Gobiopterus*. Moreover, its biological and ecological characteristics remain understudied (Kottelat et al. 1993). More importantly, its catch rates remain largely unknown. To the best of our knowledge, rinuak was known to be susceptible to *tubo belerang* – a local name for the turnover of water

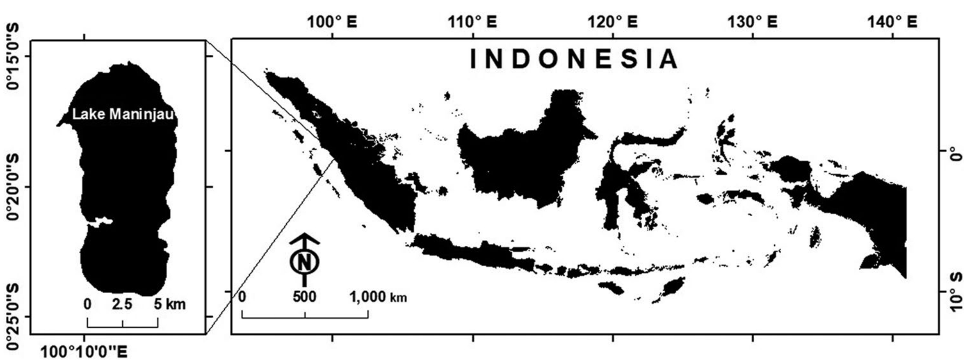


Figure 1. Study site (Image is processed in QGIS 3.2.1 based on map provided by, retrieved on 2 March 2021).

and toxic material (e.g. H_2S) from hypolimnion to epilimnion layer adding more severe impacts to anoxia condition due to eutrophication (Yuniarti et al. 2021b). It is also susceptible to the concentration of ammonia in the water (Yoga and Samir 2021).

On the contrary, bada shows higher survival capability during *tubo belerang* due to its agility and its strong connection with riverine waters, because they need to migrate when they are spawning (Hartoto and Mulyana, 1996 *cited in* Dina et al. 2019; Dina et al. 2019). Bada's diet, which primarily consists of zooplankton, aquatic insects, and small portions of phytoplankton, was known to overlap with the food of wild tilapia (Yuniarti and Sulastri 2010; Dina et al. 2019). It is reported that in recent years there has been a declining trend of bada production due to marble goby predation and overfishing (Dina et al. 2019; Linggi et al. 2020).

2.3. Selection of environmental variables

To select the environmental variables (EVs) for the models, we adopted the approach presented by Yiwen et al. (2016). Owing to the lack of ecological knowledge about bada and rinuak, we selected several EVs based on general fish ecological knowledge. The EVs that were selected *a priori* were cage presence, turbidity, presence of natural canopy, dissolved oxygen (DO), pH, conductivity, and water surface temperature (see Rosette et al. 2020)

2.4. Data collection and availability

Prior to conducting the habitat survey, a focus group discussion with local fishers was conducted to ensure that we also include the natural habitat of the fish in the survey. We also engaged one of the fishers to guide us during the survey.

A survey of the presence and absence of the fish and measurement of environmental variables was conducted in April 2019. An underwater camera (GoPro Hero 7) was used to survey fish presence. The survey was conducted from 9:00 am to 5:00 pm for four days for each species representing feeding and non-feeding time. Fish presence was recorded from the water surface (0 cm) and in different depths depending on habitat type. We selected four habitat types based on information gathered from interviews from eleven fishers. Surveys were conducted in: (1) leafy canopy covered lake margins (depth of underwater video: 0 cm, 30 cm, and the lake's bed); (2) the cage farm area (depth of the underwater video: 0, 30, 60, 90 and 130 cm – the average secchi depth in cage farm area at the time of sampling); (3) open waters 1 km from lake's edge (depth of the underwater video: 0, 30, 60, 90, and 175 cm – the average secchi depth in open water area, no canopies, no cages); (4) shallow water with no canopy (0 and 23 cm – the lake's bed). The water depth was measured using a secchi disk stick. However, we did not record any evidence fish presence below 50 cm. In total, we surveyed 36 sites for presence and absence of rinuak (Figure 2(a) and Table A), 23 sites for bada's presence (Figure 2(b) and Table B), and 33 sites for environmental variables (EVs) (Figure 2(c)).

We conducted a second field survey in March 2020 to collect data on presence and absence (Go Pro sampling) for independent cross validation referring to the work of West et al. (2016). The presence and absence data obtained from this second survey was used to validate models by providing ground data of the fish distribution. The survey was conducted randomly at 40 sites (Tables C and D). The random sampling sites were generated by using a research tool in QGIS by determining 1,000 metres as the minimum distance between points.

To verify that MaxEnt can be useful to assess the cage attraction to the fish compared to the usual method using count data, we conducted sampling of fish abundance in the

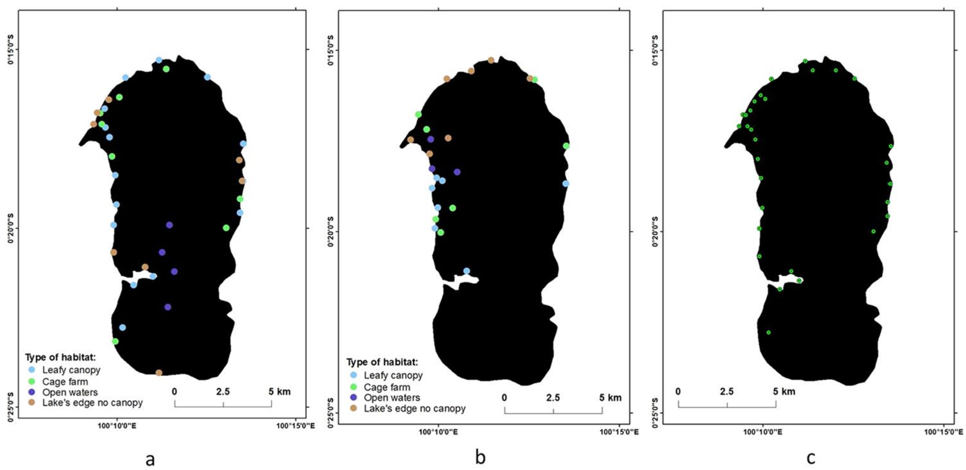


Figure 2. Sampling location (a) Rinuak, (b) Bada, (c) Environmental variables.

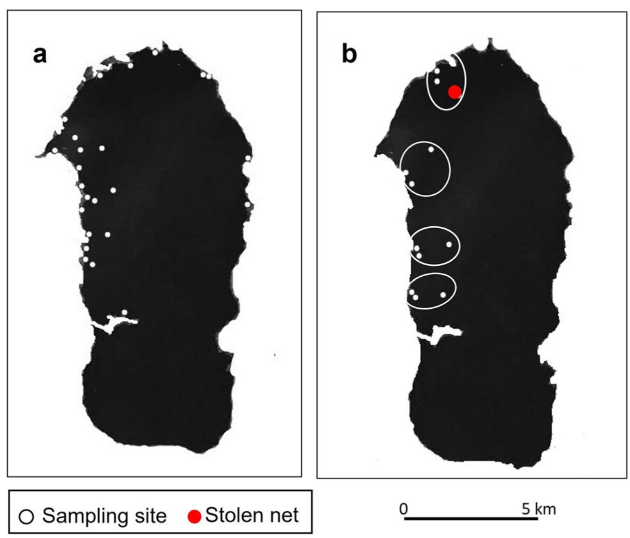


Figure 3. Relative abundance sampling locations (a) Rinuak, (b) Bada.

same month. The sampling was done following the work of Brandão et al. (2013). We used the same habitat criteria as for the presence/absence sampling to determine the abundance sampling sites (native/control habitat, the farm area, and open waters). Rinuak were caught using a scoop net with three repetitions at each site (Figure 3(a)). It was sampled in the morning between 6:00 am to 11:00 am to coincide with first light and during the period of maximum fish activity as informed by the local fishers.

Abundance data sampling of bada (Figure 3(b)) was conducted by experienced fishers using gill nets (size 100×3 m, mesh size 19, 25, and 38 mm) in April 2019. Multiple mesh size gill nets were recommended to sample fish from freshwater and estuarine areas (Gray et al. 2005). The nets were set from 2:00 am to 5:00 am following the guidance of the fishers and is the optimum period to catch this species, and not the farms' feeding time. The sampling was repeated for three consecutive nights. The data is provided in Tables C and D. All data are available in https://github.com/ivanayuniarti02/Maxent.

2.5. Data analysis

Before building the models, to test if the EVs data are spatially correlated, we calculated local Moran's I Index (I) using autoregressive STATA 15 (Anselin 1995; Kondo 2021). Local Moran's I index shows the indication of spatially autocorrelated environmental variable data, which may be caused by selecting sampling points that are too close to each other. In other words, it is used to test if the data is affected by the observation of a nearby sampling site (Rousset and Ferdy 2014). The I indices and their p-values were calculated and presented for DO, pH, turbidity, conductivity, and temperature.

To test the effects of the spatially autocorrelated EVs data on the model performance, we built two sets of models for each species. The first set (referred as total variables/ TVM) included all EVs. The second set (partial variables/PVM) was constructed by removing those EVs which are indicated to be spatially autocorrelated.

In the model building process, we ran the models using presence only data for both species, because we do not have absence data for bada. MaxEnt automatically generates predicted absence data using the environmental predictors' information to predict the absence of the species referred as pseudo-absence (Phillips and Dudı'k 2008). We run the model in default settings/auto features mode (linear, quadratic, product, hinge) to obtain the best fitting model following the work of Phillips and Dudı'k (2008), Jose et al. (2020), Yiwen et al. (2016).

Specifically, we set the regularization multiplier (RM) to avoid overfitting – a modelling error which happens when a function of the model relates too closely to a particular set of data looking the model too closely. We specified the RM at 0.5, 1.0, 1.5, 2.0, 4.0 for both model sets (Jose et al. 2020). RM is a tool in MaxEnt which can be used to regulate how focused or closely-fitted the output distribution is (Waszkowiak et al. 2002). Further, to obtain a stable model, we used a 5-fold cross-validation (CV) approach for both species in each model set (TVM or PVM). CV is a method to select the optimal model when the data is limited (De Bin et al. 2016)

We categorized the model settings as type 1 (CV: 5; RM:0.5), type 2 (5, RM: 1), Type 3 (5; RM:1.5); type 4 (5; RM:2), and type 5 (5; RM:4). Models were run with forty replications. Ten thousand random background data/grid points were selected to generate the model. In total, we constructed a total of 25 models for each species and each model set (5 models for each RM).

To compare models' performance and to select the best fitting model, we used the Area Under Curve presence only (AUC_{PO}) indicator. The AUC_{PO} is produced in MaxEnt using presence only and pseudo-absence data (Yackulic et al. 2013).

To validate the models using independent field data (obtained from the second presence data sampling), we calculated several indices, including sensitivity, specificity, and the True Skill Statistic (TSS) (Allouche et al. 2006; West et al. 2016) for the best fitting models (setting type 3 for both species). A cut-off point of 0.5 was used in this step to obtain predicted occurrence (probability ≥ 0.5 means presence, and otherwise).

To find the most influential environmental parameters, we used a stepwise removal method (Whittingham et al. 2006; Yiwen et al. 2016). The stepwise removal step was done on the best fitting models. We relied on the Jackknife analysis to generate indicators of the environmental predictors' importance. The process was repeated until two variables were obtained, because two variables are the minimum requirement for input data resulting in four level models for each species. Again, we used AUC_{PO} to evaluate the performance of the model performance (Phillips et al. 2016; Yang and Berdine 2017)

In the end, to test if the results of MaxEnt aligns with the results of the count data method, we used a maximum likelihood mixed effect model (ML) drawing on the relative abundance data – numbers of individual per species per square metre of area. We used STATA 15 to perform the test. ML was selected because the model can analyse the impact of repeated measurement (pseudo-replication). We tested the models against the null model. The Akaike Information Criteria (AIC) and Likelihood Ratio Test (LRT) were used to compare the models.

3. Results

3.1. Spatial autocorrelation of environmental variables

The local Moran's I indices (I) and the expected indices (E(I)) are generated to calculate the p-value (Table 1). The I values show that DO and pH are spatially autocorrelated, which means that the measurement of these parameters is influenced by the adjacent sampling sites. The p-values suggest that we can reject the null hypothesis that there is no spatial autocorrelation for both EVs. Further, the I values for both EVs are above zero indicating positive SAC, which means that the observation of DO and pH in one sampling point is affected by the results of observation in the nearby point. Meanwhile, SACs for the other three EVs are not detected. Considering the results, we built the models with and without these parameters to test whether removing the EVs with SAC would affect the models' performance.

3.2. Model selection

The estimated test statistics of Area Under ROC Curve (AUC) are quite similar for both model sets (Partial Variable Models/PVM and Total Variable Models/TVM). The values above 0.9 mean the models have good predicting capability. Therefore, that removing the spatially autocorrelated EVs does not dramatically change the model performance (Table 2).

The selected final models were partial variable models (PVM) to ensure that the explanatory variables were not spatially autocorrelated. Based on the optimum AUC values and the resulting maps, we selected PVM with model setting type 2 (5 replications with regularization multiplier 1) for rinuak. PVM with model setting type 1 (5 replications with regularization multiplier 0.5) is chosen for bada.

Drawing on the field validation using independent test data true statistic skill values (TSS) are estimated to be above zero indicating that the models performed better than random prediction for both species (Table 3) (Allouche et al. 2006; West et al. 2016).

The calculated TSS values reveal that the model is able to predict the actual occurrence of the fish as shown in Tables E and F. Although some predictions of the presence deviate from actual occurrence, most of the predicted absence aligns with the observed data.

Table 1. The local Moran's I index of the environmental variables presenting spatial autocorrelation in DO and pH.

Environmental variable	I	E(I)	Z	p-value
DO[a]	0.534	−0.029	3.899	$p < 0.05$
pH[a]	0.581	−0.029	4.224	$p < 0.05$
Turbidity	0.037	−0.029	0.519	0.302
Conductivity	0.009	−0.029	0.274	0.392
Temperature	0.040	−0.029	0.527	0.311

Note: H_0: spatial randomness
[a]statistically significant, indicating occurrence of spatial autocorrelation

Table 2. AUC values of the constructed models to choose the fittest models.

Species	Model settings	Total variable models (TVM)[c] Test AUC_{PO}[a]	Partial variable models (PVM)[d] Test AUC_{PO}[a]
Rinuak	Type 1 (Cross validation: 5, Regularization coefficient: 0.5)	0.928	0.928
	Type 2 (Cross validation: 5, Regularization coefficient: 1)[b]	0.943	0.930
	Type 3 (Cross validation: 5, Regularization coefficient: 1.5)	0.940	0.933
	Type 4 (Cross validation: 5, Regularization coefficient: 2)	0.937	0.933
	Type 5 (Cross validation: 5, Regularization coefficient: 4)	0.920	0.927
Bada	Type 1 (Cross validation: 5, Regularization coefficient: 0.5)[b]	0.952	0.977
	Type 2 (Cross validation: 5, Regularization coefficient: 1)	0.948	0.947
	Type 3 (Cross validation: 5, Regularization coefficient: 1.5)	0.952	0.943
	Type 4 (Cross validation: 5, Regularization coefficient: 2)	0.951	0.936
	Type 5 (Cross validation: 5, Regularization coefficient: 4)	0.935	0.919

Note: [a]AUC_{PO} (AUC Presence Only) as proposed by (Yackulic et al. 2013) to indicate that the AUC value is produced as a part of MaxEnt model building using presence only and pseudoabsence data.
[b]The selected model settings
[c]Total variable models (TVM): the models built with all included environmental variables. [d]Partial variable models (PVM): the models built with removing spatially correlated environmental variables (DO and pH).

Table 3. Results of independent data field validation showing the good prediction ability of the models.

Species	Sensitivity	Specificity	TSS
Rinuak	0.60	0.80	0.40
Bada	0.70	0.71	0.41

3.3. Species distribution maps

The maps (Figure 4(a,b)) present the predicted pattern of occurrence of both species in response of the environmental variables (cage, temperature, turbidity, conductivity, and canopies). Overall, predicted presence decreases with the distance from the lake's edge especially for rinuak. This species was mapped to be present mostly on the lake's edge compared to bada, and almost overlapped with the cage farms (Figure 4(a) vs. Figure 4(c)). Meanwhile, although there was overlap in some spots, bada was also predicted to occur in the more open waters (Figure 4(b,c)).

3.4. The important environmental variables affecting fish distribution

The results of the stepwise analysis shown in Table 4 are used to identify the principal environmental parameters influencing the occurrence of both species. Overall, temperature is the most influential variable for both rinuak and bada and turbidity ranks as the second most important variable for bada, and the third for rinuak. (Table 5). Meanwhile, cage presence, the second most influencing variable for rinuak distribution, is not closely associated to bada's presence.

3.5. Relative abundance in three habitat types

We could only find rinuak in leafy canopied and cage areas as the fish are absent in open water areas. Thus, we only use those two habitat types in the ML model for rinuak (Figure 5(a)). Meanwhile, the habitat types tested for bada are leafy canopied habitat, cage farms, and open waters as the fish were present in those three habitat types (Figure 5(b)).

The results of the mixed effects model (Table 5) show that only habitat types have significant influence on rinuak's relative abundance (Likelihood Ratio Test against null model/LRT[1] < 0.05 and LRT models with habitat and sampling location as variables and models

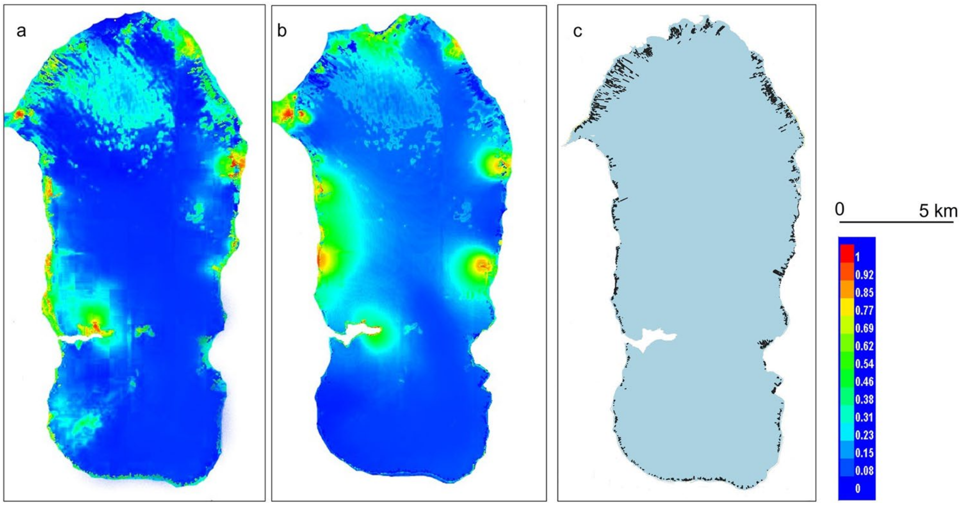


Figure 4. Maps of predicted species distribution in correlation with cage presence, red, yellow, and green colours show the predicted presence of the species (a) Rinuak, (b) Bada, (c) Cage farm.

Table 4. Results of stepwise analysis to obtain the most important environmental variables.

Species	Ranking of included variables	Test AUC_{PO}	Remark
Rinuak	Te, Ca, Tur, Con, Can	0.930	Excellent determination
	Te, Ca, Tur, Con	0.900	Excellent determination
	Te, Ca, Tur	0.921	Excellent determination
	Te, Ca	0.946[a]	Excellent determination
Bada	Te, Tur, Con, Can, Ca	0.949	Excellent determination
	Te, Tur, Con, Can	0.947	Excellent determination
	Te, Tur, Con	0.957[a]	Excellent determination
	Te,Tur	0.942	Excellent determination

Note: Te: temperature; Ca: cages; Tur: turbidity; Con: conductivity; and Can: natural canopy
[a]The best models

Table 5. Results of maximum likelihood mixed effect model (ML) for rinuak and bada.

Species	Response Variable	Mixed effects model	df	AIC	LL	LRT[b]	LRT[c]
Rinuak	Fish's relative abundance	Habitat + Sampling Locations	4	363.26	−177.63	$p = 0.04$	$p = 1.00$
		Habitat[a]	3	361.26	−177.63	$p = 0.01$	
		Null model	2	365.58	−180.79	–	
Bada	Fish's relative abundance	Habitat + Sampling Locations	4	52.81	−22.41	$p = 0.17$	$p = 0.10$
		Habitat	3	53.48	−23.74	$p = 0.34$	
		Null model[a]	2	52.39	−24.20	–	

Note: Df: degree of freedom; AIC: Akaike information criteria; LL: log likelihood; LRT[b]: likelihood ratio test against null model, LRT[c]: likelihood ratio test between models with habitat and sampling location as variables and models with habitat only.
[a]the best model

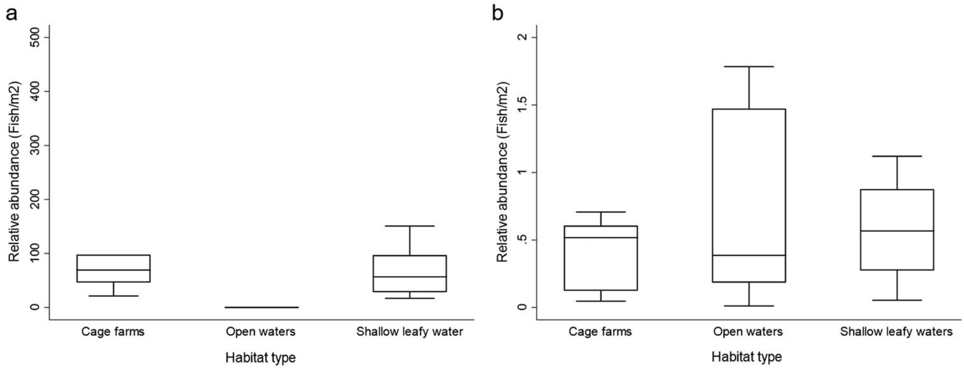


Figure 5. Relative abundance: (a) Rinuak, (b) Bada.

with habitat only variable/LRT^2 > 0.05). The results for bada indicate no statistically significant effect of habitat types and sampling locations on its relative abundance (LRT^1 < 0.05).

4. Discussion

Water temperature is the most influential variable affecting the distribution of both species. Few studies have addressed the impacts of warming temperature to fish distribution in tropical lakes (Cohen et al. 2016; Barbarossa et al. 2021). In general, rising temperature affects tropical freshwater fish distribution by decreasing their fitness due to increased stress physiology and required energy to reproduce and to survive, which forces the fish to be more actively find a suitable environment (Nivelle et al. 2019; Alfonso et al. 2020). The ability to find the suitable environment is constrained by limited habitat availability (Alfonso et al. 2020; Nivelle et al. 2019). Further, the warming temperature affects habitat availability by reducing water levels, jeopardising habitat connectivity, and disrupting riverine habitat (Ficke et al. 2007; Miranda et al. 2020; Barbarossa et al. 2021). The effect of rising temperature on habitat connectivity will especially affect bada, which migrates to riverine habitat for spawning. Additionally, the warming temperature will cause water retention changes, exacerbate eutrophication leading to anoxia condition, and increase the toxicity of pollutants (Ficke et al. 2007; Missaghi et al. 2017; Miranda et al. 2020). Thus, the impacts of increasing temperature on eutrophication will particularly affect rinuak, whose survival has been threatened by *the tubo belerang* event (Yuniarti et al. 2021a).

Moreover, our data corroborate that turbidity affects both species, and especially bada, which aligns with the study of Dina et al. (2019), which mentioned that bada prefers clean water. Turbidity may reduce bada's visual ability to spot zooplankton as their main diet (cf. Sulastri et al. 2010; Yuniarti and Sulastri 2010). Zooplankton can easily camouflage in the turbid water impeding bada's foraging behaviour, as indicated by Lunt and Smee (2015). On the contrary, turbid water around the cages could be beneficial for rinuak to avoid predators due to its tiny translucent body. Turbid water and cage structures, which increase habitat complexity, help rinuak to reduce its predation risk by decreasing encounter rates and foraging ability of its predators as explained by Ajemian et al. (2015) and McElroy et al. (2018). Therefore, it is not surprising that small fish like rinuak occupy complex littoral habitat or artificial structures and avoid open water habitat (Ajemian et al. 2015; Merz et al. 2021).

Turbidity is affected by the more frequent extreme weather predicted in Indonesia due to the warming temperature (Measey 2010; Supari et al. 2017). The anticipated longer wet season and increased rainfall will increase sediment run-off will add more debris avalanches to the lake, which is naturally high given the occurrence of earthquake and heavy rain in the study area (Wils et al. 2021). Turbidity was also predicted to be boosted by increasing human activities in the lake's watershed area (Antomi et al. 2016). In addition to that, a strong connection between Lake Maninjau's increased turbidity with the cage activities using long term satellite data has been identified (Setiawan et al. 2019).

Considering that natural and anthropogenic factors influence environmental variables and affect fish distribution, we underline the importance of continuing monitoring to help mitigation planning, as suggested by Callier et al. (2018) and Miranda et al. (2020). We suggest using MaxEnt to overcome monitoring constrains because it has shown good prediction ability of the fish distribution in our case study, making it useful to guide ecosystem managers to focus scientific effort as suggested by West et al. (2016) and Pearson et al. (2007). The MaxEnt models, which can be built based on species presence and environmental variables data, help reduce sampling cost. Using MaxEnt also requires lower specific skills to collect the data than common count data collection. Moreover, MaxEnt models can also be developed to forecast the dynamics of fish distribution related to changing environmental variables, such as changing temperature driven by future climate scenarios (Qin et al. 2017; Borzée et al. 2019; Hadgu et al. 2019).

Further, the Jackknife analysis feature in MaxEnt is beneficial to illustrate the significance of various environmental variables on fish distribution, including cage farming, one of the main anthropogenic activities in the study area. The results of the Jackknife analysis indicate that rinuak presence is highly related to the presence of the cage farms. However, the same does not apply for bada's distribution. We argue that the cage farms offer additional habitat to rinuak because the cage farms are not built under leafy canopies and expand up to 1 km from the lake's edge. The cage farms' attraction to rinuak is related to the habitat services provisioning, as suggested by TEEB (2010). The services offered by the farms are shelter and food source (Sanchez-Jerez et al. 2011; Goodbrand et al. 2013; Ramos et al. 2013; Uglem et al. 2014; Nobile et al. 2018). The farm's sheltering function is primarily associated with the provisioning of artificial physical structures acting like fish aggregating devices or artificial reefs (Dempster et al. 2011; Sanchez-Jerez et al. 2011; Uglem et al. 2014; Barrett et al. 2018; Callier et al. 2018).

The degree of association between wild fish distribution and farm cages is varied and species-specific (Sanchez-Jerez et al. 2011; Arechavala-Lopez et al. 2015; Ballester-Moltó et al. 2015). Our results support this finding by providing evidence that rinuak presence is strongly associated with cage farms, but not bada. As aforementioned, rinuak is likely to benefit from the complex structure of cage farms, waste feed, and turbid water, while bada does not rely on these aspects. Therefore, we suggest that the species-specific relationship between the cage farms and wild fish spatial distribution depends on the interaction between the ecology of the species and the environmental variables such as turbidity. This finding complements previous studies of cage farms' attraction to wild fish, which highlighted that the wild fish-cage farm relationship was affected by feeding time, fish' life stage, and seasonal event (e.g. reproductive cycle) (Sudirman et al. 2009; Uglem et al. 2009; Sanchez-Jerez et al. 2011; Šegvić Bubić et al. 2011; Bacher et al. 2012, 2015; Ballester-Moltó et al. 2015).

Our Data has limitations that should be addressed in future research. First, collecting data at various points in the year is advisable considering that cage aquaculture operates all year round in the lake. Second, future research on similar topics should include the effects of nutrient concentration in the lake's water, including total nitrogen, total phosphorus, and ammonia nitrogen, because these parameters are more sensitive to aquaculture activities. They may have impacts on native fish distribution through the occurrence of harmful algae bloom and DO depletion.

5. Conclusion

A species distribution model such as MaxEnt is shown to provide reliable information of the impacts of cage aquaculture on the distribution of native species compared to the more time consuming and expensive method using count data. Thus, we recommend using MaxEnt to help monitoring efforts in areas with limited monitoring capacities. The model can also be extended to advice the suitable area to move the cage aquaculture related to the future native fish distribution considering the warming temperature and increased turbidity. Further, we also emphasise the importance of an improved understanding of the ecology of native fish, including their interaction with café farms, for conducting environmental risk assessments of cage aquaculture.

Acknowledgements

We relay our gratitude to the local fishers for their kind help and assistance during the field study. We thank Dr. Neil Burns and Dr. Charlotte Hopkins for their valuable comments to improve the manuscript.

Disclosure statement

The authors declare that there is no conflict of interests.

Funding

This work was supported by Lembaga Pengelola Dana Pendidikan.

References

Abdel-Tawwab M, et al. 2002. Effect of different doses of inorganic fertilizer on water quality, primary productivity and production of Nile Tilapia (Oreochromis niloticus) in Earthen Ponds. Qatar Univ. Sci. J. 22:1–16.

Abery NW, et al. 2005. Fisheries and cage culture of three reservoirs in west Java, Indonesia; a case study of ambitious development and resulting interactions. Fish Manage Ecol. 12:315–330.

Agam Regency Fisheries Department. 2017. Data karamba dan perikanan Danau Maninjau tahun 2016.

Ajemian MJ, et al. 2015. Effects of turbidity and habitat complexity on antipredator behavior of three-spined sticklebacks (Gasterosteus aculeatus): antipredator behavior in sticklebacks. Environ Biol Fishes. 98(1):45–55.

Alfonso S, et al. 2020. Temperature increase and its effects on fish stress physiology in the context of global warming. J Fish Biol. 98(6):1496–1508.

Allouche O, et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol. 43:1223–1232.

Anselin L. 1995. Local indicators of spatial association—LISA. Geog Anal. 27(2):93–115.

Antomi Y, et al. 2016. Water quality index in Lake Maninjau as a parameter to determine the optimum economic growth of floating net cages and land-based livelihood. OIDA IJSD. 09(02):51–62.

Arechavala-Lopez P, et al. 2015. Aggregations of bluefish Pomatomus saltatrix (L.) at Mediterranean coastal fish farms: seasonal presence, daily patterns and influence of farming activity. Environ Biol Fishes. 98:499–510.

Baccarin AE, et al. 2005. Characterization and evaluation of the impact of feed management on the effluents of Nile tilapia (Oreochromis niloticus) culture. Braz Arch Biol Technol. 48(1):81–90.

Bacher K, et al. 2012. Spatial and temporal extension of wild fish aggregations at Sparus aurata and Thunnus thynnus farms in the north-western Mediterranean. Aquac Environ Interact. 2:239–252.

Bacher K, et al. 2015. Feeding activity strongly affects the variability of wild fish aggregations within fish farms: a sea bream farm as a case study. Aquacult Res. 46(3):552–564.

Ballester-Moltó M, et al. 2015. Husbandry and environmental conditions explain temporal variability of wild fish assemblages aggregated around a Mediterranean fish farm. Aquac Environ Interact. 7(3):193–203.

Barbarossa V, et al. 2021. Threats of global warming to the world's freshwater fishes. Nat Commun. 12(1):1–10.

Barrett LT, et al. 2018. Impacts of marine and freshwater aquaculture on wildlife: a global meta-analysis. Rev Aquacult. 11(4):1022–1044.

De Bin R, et al. 2016. Subsampling versus bootstrapping in resampling-based model selection for multivariable regression. Biometrics. 72(1):272–280.

Borzée A, et al. 2019. Climate change-based models predict range shifts in the distribution of the only Asian plethodontid salamander: karsenia koreana. Sci Rep. 9(1):1–9.

Brandão H, et al. 2013. Influence of a cage farming on the population of the fish species Apareiodon affinis (Steindachner, 1879) in the Chavantes reservoir, Paranapanema River SP/PR, Brazil. Acta Limnol Bras. 24(4):438–448.

Callier MD, et al. 2018. Attraction and repulsion of mobile wild organisms to finfish and shellfish aquaculture: a review. Rev Aquacult. 10:924–949.

Canonico GC, et al. 2005. The effects of introduced tilapias on native biodiversity. Aquat Conserv: Mar Freshw Ecosyst. 15:463–483.

Chislock MF, et al. 2013. Eutrophication: causes, consequences, and controls in aquatic ecosystems marine ecosystems in the world. Nat Edu Know. 4(4):10.

Cohen AS, et al. 2016. Climate warming reduces fish production and benthic habitat in Lake Tanganyika, one of the most biodiverse freshwater ecosystems. Proc Natl Acad Sci U S A. 113(34):9563–9568.

Demétrio JA, et al. 2012. Influence of net cage farming on the diet of associated wild fish in a Neotropical reservoir. Aquac. 330–333:172–178.

Dempster T, et al. 2011. Proxy measures of fitness suggest coastal fish farms can act as population sources and not ecological traps for wild gadoid fish. PLoS One. 6(1):e156546.

Dina R, et al. 2019. Fish and fisheries of Bada (Rasbora spp.) in Lake Maninjau, West Sumatra. IOP Conf Ser: Earth Environ Sci. 306(1):1–10.

Elith J, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 29(2):129–151.

Endah NH, et al. 2017. Pemanfaatan dan Peran Komunitas Lokal dalam Pelestarian Danau Maninjau. Jurnal Ekonomi Dan Pembangunan. 25(1):55–67.

Fakhrudin M, et al. 2002. 'Karakterisasi Hidrologi Danau Maninjau Sumatera Barat', in *Menuju Kesinambungan * Prosiding Seminar Nasional Limnologi : menuju Kesinambungan Pemanfaatan Sumberdaya Perairan*, 65–75.

FAO 2016. *The State of World Fisheries and Aquaculture: contributing to food security and nutrition for all, The State of World Fisheries and Aquaculture* 2016.

Ficke AD, et al. 2007. Potential impacts of global climate change on freshwater fisheries. Rev Fish Biol Fish. 17:581–613.

Goodbrand L, et al. 2013. Sea cage aquaculture affects distribution of wild fish at large spatial scales. Can J Fish AquatSci. 70(9):1289–1295.

Gray CA, et al. 2005. Utility and efficiency of multi-mesh gill nets and trammel nets for sampling assemblages and populations of estuarine fish. 1077–1088.

Hadgu M, et al. 2019. 'Modeling the potential climate change-induced impacts on future genus Rhipicephalus (Acari: ixodidae) tick distribution in semi-arid areas of Raya Azebo district, Northern Ethiopia. J Ecol Environ. 43(1):1–11.

Hartoto DI, Mulyana E. 1996. Hubungan parameter kualitas air dengan struktur ikhtiofauna perairan darat Pulau Siberut. Oseanologi dan Limnologi di Indonesia 29:41–55.

Jose VS, et al. 2020. The expanding distribution of the Indian Peafowl (Pavo cristatus) as an indicator of changing climate in Kerala, southern India: a modelling study using MaxEnt. Ecol Indic. 110(2019):105930.

Kondo K. 2021. Testing for global spatial autocorrelation in stata: Moransi version 1.21. Stat Softw Components S458473. 2021(2016):1–10. Available at: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/.

Kottelat M, et al. 1993. Gobiopterus brachypterus (Bleeker, 1855), Freshwater fishes of Western Indonesia and Sulawesi. Available at: https://www.fishbase.se/summary/Gobiopterus-brachypterus.html.

Larashati S. 2019. DNA sequence data for Gobiopterus sp. in Lake Maninjau. Unpublished.

Linggi GNT, et al. 2020. Pengelolaan Sumberdaya Ikan Bada (Rasbora sp) di Danau Maninjau, Sumatra Barat. Jurnal Pengelolaan Perairan. 3(1):1–15.

Lumbantobing DN. 2014. Four new species of Rasbora of the Sumatrana group (Teleostei: cyprinidae) from northern Sumatra, Indonesia. Zootaxa. 3764(1):1–25.

Lunt J, Smee DL. 2015. Turbidity interferes with foraging success of visual but not chemosensory predators. PeerJ. 3:1–12.
Marini S, et al. 2018. Tracking fish abundance by underwater image recognition. Sci Rep. 8(1):1–12.
McElroy KN, Beakes MP, Merz JE. 2018. Hide and seek: turbidity, cover, and ontogeny influence aggregation behavior in juvenile salmon. Ecosphere. 9(4):1–15.
Measey M. 2010. Indonesia: a vulnerable country in the face of climate change. Global Majority E-J. 1(1):31–45. Available at: http://unfccc.int/meetings/cop_13/items/4049txt.php.
Merz JE, et al. 2021. Comparison of three sampling methods for small-bodied fish in lentic nearshore and open water habitats. Environ Monit Assess. 193(5):1–20.
Miranda LE, Coppola G, Boxrucker J. 2020. Reservoir fish habitats: a perspective on coping with climate change. Rev Fish Sci Aquacu. 28(4):478–498.
Missaghi S, Hondzo M, Herb W. 2017. Prediction of lake water temperature, dissolved oxygen, and fish habitat under changing climate. Clim Change. 141(4):747–757.
Moore C, et al. 2016. Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. Marine Policy. 69:32–41.
Naylor RL, Hardy RW, Buschmann AH, Bush SR, Cao L, Klinger DH, Little DC, Lubchenco J, Shumway SE, Troell M, et al. 2021. A 20-year retrospective review of global aquaculture. Nature. 591(7851):551–563.
Newton R, Zhang W, Xian Z, McAdam B, Little DC. 2021. Intensification, regulation and diversification: the changing face of inland aquaculture in China. Ambio. 50(9):1739–1756.
Nivelle R, et al. 2019. Temperature preference of Nile tilapia (Oreochromis niloticus) juveniles induces spontaneous sex reversal. PLoS One. 14(2):1–19.
Njiru JM, Aura CM, Okechi JK. 2018. Cage fish culture in Lake Victoria: a boon or a disaster in waiting?. Fish Manage Ecol. 26(5):426–434.
Nobile AB, et al. 2018. Cage fish farm act as a source of changes in the fish community of a Neotropical reservoir. Aquacul. 495:780–785.
Pearson RG, et al. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr. 34(1):102–117.
Pereira LS, et al. 2019. Cage aquaculture in neotropical waters promotes attraction and aggregation of fish. Aquacult Res. 00:1–11.
Phillips SJ, et al. 2016. A comparison of GLM, GAM, and GWR modeling of fish distribution and abundance in Lake Ontario, Ecological modelling. Los Angeles, California, U.S.: University of Southern California.
Phillips SJ, Dudı'k M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation Steven. Ecography. 31:161–175.
Qin A, et al. 2017. Maxent modeling for predicting impacts of climate change on the potential distribution of Thuja sutchuenensis Franch, an extremely endangered conifer from southwestern China. Global Ecol Conserv. 10:139–146.
Rajee O, Mun ATK. 2017. Impact of aquaculture on the livelihoods and food security of rural communities. Int J Fish Aquat Stud. 5(2):278–283. Available at: https://psa.gov.ph/content/fisheries-statistics-philippines.
Ramos IP, et al. 2013. Interference of cage fish farm on diet, condition factor and numeric abundance on wild fish in a Neotropical reservoir. Aquacul. 414–415:56–62.
Roesma DI, Tjong DH, Aidil DR. 2020. Phylogenetic analysis of transparent gobies in three sumatran lakes, inferred from mitochondrial cytochrome oxidase I (COI) gene. Biodiversitas. 21(1):43–48.
Rosette ZL, et al. 2020. The influence of water quality parameters on fish species abundance and distribution near shoreline of Lake Victoria. AJENSR. 3(2):1–12. Available at: www.abjournals.org.
Rousset F, Ferdy JB. 2014. Testing environmental and genetic effects in the presence of spatial autocorrelation. Ecography. 37:781–790.
Said DSS, et al. 2020. Integrated multitrophic aquaculture in Maninjau Lake: converting eutrophic water into fish meal. IOP Conf Ser: earth Environ Sci. 535(1):1–11.
Sanchez-Jerez P, et al. 2011. Coastal fish farms as fish aggregation devices, artificial reefs in fisheries management. Florida, U.S.: CRC Press. Taylor & Francis Group.
Šegvić Bubić T, et al. 2011. Temporal and spatial variability of pelagic wild fish assemblages around Atlantic bluefin tuna Thunnus thynnus farms in the eastern Adriatic Sea. J Fish Biol. 78(1):78–97.
Setiawan F, et al. 2019. Long-term change of the secchi disk depth in Lake Maninjau, Indonesia shown by landsat TM and ETM + data. Remote Sens. 11(23):1–20.
Shava E, Gunhidzirai C. 2017. Fish farming as an innovative strategy for promoting food security in drought risk regions of Zimbabwe. Jamba: J Dis Risk Stud. 9(1):1–10.
Siddiqui SA, et al. 2018. Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. ICES J Mar Sci. 75(1):374–389.

Sobek-Swant S, et al. 2012. Potential distribution of emerald ash borer: what can we learn from ecological niche models using Maxent and GARP?. For Ecol Manage. 281:23–31.
Sudirman, et al. 2009. Wild fish associated with tropical sea cage aquaculture in South Sulawesi, Indonesia. Aquacul. 286:233–239.
Sulastri, et al. 2010. 'Karakteristik habitat, kebiasaan makan, dan sistem konservasi ikan bada Rasbora argyrotaenia di Danau Maninjau', in *Prosiding Seminar Nasional Ikan*, 487–497.
Sulastri S, Nomosatryo S, Sulawesty F. 2015. Long term monitoring of water quality and phytoplankton changes in Lake Maninjau, West Sumatra, Indonesia. Oseanologi Dan Limnologi di Indonesia. 1(3):341–355.
Supari, et al. 2017. Observed changes in extreme temperature and precipitation over Indonesia. Int J Climatol. 37(4):1979–1997.
Syandri H, et al. 2014. State of aquatic resources Maninjau Lake West Sumatra Province, Indonesia. Int J Ecol Environ Sci. 5(1):109–113. Available at: http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000261.
Syandri H, Azrita J. 2013. Loading and distribution of organic materials in Maninjau Lake. Proceed Int Sem Fish and Mar. 1:69–75.
Taskov DA, et al. 2021. Managing aquaculture in multi-use freshwater bodies: the case of Jatiluhur reservoir. Environ Res Lett. 16(4):1–13.
TEEB 2010. *Mainstreaming the economics of Natre, De Economist.*
Tomassetti P, et al. 2016. Benthic community response to sediment organic enrichment by Mediterranean fish farms: case studies. Aquacul. 450:262–272.
Uglem I, et al. 2009. High connectivity of salmon farms revealed by aggregation, residence and repeated movements of wild fish among farms. Mar Ecol Prog Ser. 384(May):251–260.
Uglem I, et al. 2014. Impacts of wild fishes attracted to open-cage salmonid farms in Norway. Aquacul Environ Inter. 6(1):91–103.
Verdegem MCJ. 2013. Nutrient discharge from aquaculture operations in function of system design and production environment. Rev Aquacult. 5:158–171.
West AM, et al. 2016. Field validation of an invasive species Maxent model. Ecol Inf. 36:126–134.
Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP. 2006. Why do we still use stepwise modelling in ecology and behaviour ?. J Anim Ecol. 75(5):1182–1189.
Wils K, et al. 2021. The sediments of Lake Singkarak and Lake Maninjau in West Sumatra reveal their earthquake, volcanic and rainfall history. Sediment Geol. 416:105863.
Wisz MS, et al. 2008. Effects of sample size on the performance of species distribution models. Divers Distrib. 14:763–773.
Yackulic CB, et al. 2013. Presence-only modelling using MAXENT: when can we trust the inferences?. Methods Ecol Evol. 4(3):236–243.
Yang S, Berdine G. 2017. The receiver operating characteristic (ROC) curve. The Southwest Respiratory and Critical Care Chronicles. 5(19):34–36.
Yi Y, et al. 2016. Maxent modeling for predicting the potential distribution of endangered medicinal plant (H. riparia Lour) in Yunnan, China. Ecol Eng. 92:260–269.
Yiwen Z, Wei LB, Yeo DCJ. 2016. Novel methods to select environmental variables in MaxEnt : a case study using invasive crayfish. Ecol Modell. 341:5–13.
Yoga GP, Samir O. 2021. Ammonia toxicity to Rinuak (Gobiopterus brachypterus) of Lake Maninjau. Indones J Limnol. 1(1):12–18.
Yuniarti I, Glenk K, et al. 2021a. An application of Bayesian Belief network to assess management scenarios for aquaculture in a complex tropical lake system in Indonesia. PLoS One. 16(4):1–23.
Yuniarti I, Glenk K, et al. 2021b. An application of Bayesian Belief Networks to assess management scenarios for aquaculture in a complex tropical lake system in Indonesia. PLoS One. 16:1–23.
Yuniarti I, Barnes C, et al. 2021. Challenges for the development of environmentally sustainable cage culture farming in Lake Maninjau, Indonesia: an institutional perspective. Ecosyst People. 17(1):248–263.
Yuniarti I, Sulastri S. 2010. Jaring-jaring makanan di Danau Maninjau. Prosiding Seminar Nasional Limnologi. V(2006):135–144.
Zhang J, Jiang F, Li G, Qin W, Li S, Gao H, Cai Z, Lin G, Zhang T. 2019. Maxent modeling for predicting the spatial distribution of three raptors in the Sanjiangyuan National Park, China. Ecol Evol. 9(11):6643–6654.
Zhang K, Yao L, Meng J, Tao J. 2018. Maxent modeling for predicting the potential geographical distribution of two peony species under climate change. Sci Total Environ. 634:1326–1334.

Supplementary materials

Table A. Presence-absence data for Rinuak.

Long	Lat	Presence = 1/ Absence = 0
100.1575	−0.27947	1
100.1588	−0.27964	1
100.1603	−0.65267	1
100.161	−0.27747	1
100.1629	−0.27333	1
100.1613	−0.28631	1
100.1678	−0.27219	1
100.1633	−0.29089	1
100.1706	−0.26308	1
100.1863	−0.25492	1
100.1896	−0.25917	1
100.209	−0.26281	1
100.2257	−0.29392	1
100.2237	−0.30147	1
100.2251	−0.31125	1
100.2255	−0.31125	1
100.2241	−0.31956	1
100.2241	−0.32603	1
100.2178	−0.33319	1
100.1596	−0.28472	1
100.1558	−0.28481	1
100.1643	−0.29983	1
100.1659	−0.30856	1
100.1665	−0.32233	1
100.1652	−0.33172	1
100.1652	−0.3445	1
100.1797	−0.35136	1
100.1834	−0.35569	1
100.1746	−0.35967	1
100.1694	−0.37956	1
100.1659	−0.38611	1
100.1862	−0.4006	0
100.1912	−0.33169	0
100.1878	−0.3445	0
100.1935	−0.35342	0
100.1905	−0.37	0

Table B. Presence-absence data for Bada.

Long	Lat	Presence = 1; Absence = 0
100.1655	−0.32767	1
100.1538	−0.29108	1
100.1819	−0.25942	1
100.1911	−0.25458	1
100.2111	−0.26339	1
100.1665	−0.32233	1
100.1659	−0.30856	1
100.1797	−0.35136	1
100.209	−0.26281	1
100.2255	−0.31125	1
100.1706	−0.26308	1
100.1575	−0.27947	1
100.1652	−0.33172	1
100.1628	−0.29753	1
100.1638	−0.30436	1
100.1638	−0.31325	1
100.1713	−0.29028	1
100.1755	−0.30594	1
100.1686	−0.30986	1
100.1678	−0.33364	1
100.1734	−0.32247	1
100.1633	−0.29089	1
100.2257	−0.29392	1
100.2255	−0.31125	1
100.1613	−0.28631	1

Table C. Abundance data of Rinuak.

Location	Repetition	Relative abundance (fish/m2)	Type of habitat
S 0.2847 E 100.1596	1	176.30	Cage area
	2	40.60	Cage area
	3	78.50	Cage area
S 0.2998 E 100.1643	1	456.00	Cage area
	2	69.20	Cage area
	3	64.60	Cage area
S 0.3445 E 100.1652	1	61.80	Cage area
	2	60.00	Cage area
	3	86.80	Cage area
S 0.3861 E 100.1659	1	46.20	Cage area
	2	21.20	Cage area
	3	27.70	Cage area
S 0.3858 E 100.1842	1	97.80	Cage area
	2	325.80	Cage area
	3	90.50	Cage area
S 0.2981 E 100.1558	1	97.04	Littoral zone
	2	28.40	Littoral zone
	3	56.80	Littoral zone
S 0.3086 E 100.1659	1	39.23	Littoral zone
	2	96.92	Littoral zone
	3	56.92	Littoral zone
S 0.4984 E 100.1652	1	69.23	Littoral zone
	2	21.54	Littoral zone
	3	43.08	Littoral zone
S 0.3317 E 100.1652	1	150.77	Littoral zone
	2	87.69	Littoral zone
	3	212.31	Littoral zone
S 0.3534 E 100.1839	1	16.92	Littoral zone
	2	27.69	Littoral zone
	3	55.38	Littoral zone
S 0.2859 E 100.1880	1	0.00	Open waters
	2	0.00	Open waters
	3	0.00	Open waters
S 100.1917 E 0.3329	1	0.00	Open waters
	2	0.00	Open waters
	3	0.00	Open waters
S 0.3348 E100.1889	1	0.00	Open waters
	2	0.00	Open waters
	3	0.00	Open waters
S 0.3564 E 100 1935	1	0.00	Open waters
	2	0.00	Open waters
	3	0.00	Open waters
S 0.3716 E 100.1917	1	0.00	Open waters
	2	0.00	Open waters
	3	0.00	Open waters

Table D. Relative abundance data of Bada.

			Relative abundance (fish/m2)		
No	Location	Repetition	Cage culture area	Littoral zone	Open waters
1	S 00 15.116′ E 100 11.759'	1 (26-02-19)	0.08	0.05	No data
		2 (27-02-19)	0.05	0.11	0.99
		3 (2-02-19)	0.10	1.10	0.25
2	S 00 18.603′ E 100 09.901'	1 (27/2/2019)	0.15	0.81	0.36
		2 (27-02-19)	0.57	0.36	0.55
		3 (1-02-1019)	0.64	0.56	0.23
3	S 00 19.729′ E 100 09.894'	1(28-02-2019)	0.56	0.94	1.78
		2(1-3-2019)	0.47	0.19	0.85
		3(1-3-2019)	0.56	0.38	1.69
4	S 00 15.778′ E 100 12.608'	1 (14-3-2019)	0.40	0.68	0.18
		2 (15-3-2019)	0.71	0.57	0.39
		3 (16-3-2019)	1.63	1.12	1.47

Table E. Expected and observed presence and absence for field data validation of Rinuak.

No	Long	Lat	Expected	Observed
1	100.2204	−0.3017	0	1
2	100.2204	−0.3017	0	1
3	110.2239	−0.2812	0	0
4	110.2239	−0.2812	0	0
5	100.2132	−0.2741	0	0
6	100.2132	−0.2741	0	0
7	100.1992	−0.2605	0	1
8	100.1992	−0.2605	0	1
9	100.1931	−0.2597	1	0
10	100.1931	−0.2597	1	0
11	100.1867	−0.2596	0	0
12	100.1867	−0.2596	0	0
13	100.1783	−0.2668	0	0
14	100.1783	−0.2668	0	0
15	100.1706	−0.2652	1	0
16	100.1706	−0.2652	1	0
17	100.1667	−0.273	1	0
18	100.1667	−0.273	1	0
19	100.1637	−0.2921	0	0
20	100.1637	−0.2921	0	0
21	100.1683	−0.3017	0	0
22	100.1683	−0.3017	0	0
23	100.1721	−0.3035	0	0
24	100.1721	−0.3035	0	0
25	100.1763	−0.3058	0	0
26	100.1763	−0.3058	0	0
27	100.1686	−0.3227	1	0
28	100.1686	−0.3227	1	0
29	100.1854	−0.3565	0	0
30	100.1854	−0.3565	0	0
31	100.1951	−0.3983	0	0
32	100.1951	−0.3983	0	0
33	100.1964	−0.3979	0	0
34	100.1964	−0.3979	0	0
35	100.1954	−0.3918	0	0
36	100.1954	−0.3918	0	0
37	100.2174	−0.3405	0	0
38	100.2174	−0.3405	0	0
39	100.2213	−0.3213	0	0
40	100.2213	−0.3213	0	0

Table F. Expected and observed presence and absence for field data validation of Bada.

No	Long	Lat	Expected	Observed
1	100.2241	−0.3028	1	1
2	100.2204	−0.3017	1	1
3	100.2222	−0.2997	1	1
4	100.2205	−0.2973	1	1
5	110.2239	−0.2812	1	1
6	100.2183	−0.2801	0	1
7	100.2151	−0.2789	0	0
8	100.2132	−0.2741	0	0
9	100.2095	−0.2719	0	0
10	100.1992	−0.2605	0	1
11	100.1978	−0.2611	1	1
12	100.1931	−0.2597	1	1
13	100.1867	−0.2596	1	1
14	100.183	−0.2644	1	1
15	100.1783	−0.2668	1	0
16	100.1706	−0.2652	1	1
17	100.1667	−0.273	0	0
18	100.176	−0.2688	1	1
19	100.1637	−0.2921	1	1
20	100.166	−0.2991	1	1
21	100.1683	−0.3017	1	1
22	100.1721	−0.3035	1	1
23	100.1763	−0.3058	1	1
24	100.1733	−0.3119	1	1
25	100.1686	−0.3227	1	1
26	100.1679	−0.3309	1	1
27	100.1713	−0.3336	1	1
28	100.1854	−0.3565	1	0
29	100.1693	−0.38	0	1
30	100.1695	−0.3844	0	1
31	100.1729	−0.3877	0	1
32	100.1815	−0.3955	0	1
33	100.1881	−0.3983	0	1
34	100.1951	−0.3983	0	0
35	100.1964	−0.3979	0	1
36	100.1954	−0.3918	0	1
37	100.2174	−0.3405	1	1
38	100.2196	−0.3403	1	1
39	100.2213	−0.3213	0	1
40	100.2255	−0.3019	1	1