Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 7, No.3, July 2025, pp. 02503021-01 ~ 02503021-18
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v7i3.2201

Predicting Habitat Suitability of Mahseer Fish (*Tor* spp.) in Tropical River Systems Using MaxEnt and Google Earth Engine: A Geospatial Modeling Approach

Jefri Permadi[1*], Nia Kurniawan [1], Diana Arfiati [2], Agung Pramana Warih Marhendra[1]

[1]Department of Biology, Faculty of Mathematics and Natural Sciences, Brawijaya University, Jl. Veteran, Malang 65145, East Java, Indonesia

[2]Department of Aquatic Resources and Management, Faculty of Fisheries and Marine Science, Brawijaya University, Jl. Veteran, Malang 65145, East Java, Indonesia

*jefri.pmm@gmail.com

Abstract. Rivers are vital freshwater habitats that face threats of degradation and climate change. Mahseer fish, a key species, is in decline. This study predicted Mahseer fish habitats in Central Java using the Google Earth Engine and the MaxEnt machine learning algorithm. Environmental predictors, including NDVI, elevation, slope, river order, temperature, and rainfall, were extracted from Sentinel, SRTM, MODIS, and CHIRPS data. The model identified river order as the most influential variable (73%), followed by elevation (18%) and rainfall (8%), with an AUC score of 0.7, indicating fair accuracy. Suitable habitats were located in upstream river orders (1–3), typically at higher elevations. These findings provide spatial guidance for conservation planning, such as identifying critical habitats, prioritizing upstream areas, and establishing seasonal fishing ban. This approach supports biodiversity protection and aligns with the Sustainable Development Goals by offering a scalable tool for freshwater ecosystem management. Using MaxEnt with GEE shows promise for rapid, and cost-effective species distribution modeling in data-limited tropical regions.

Keywords: Mahseer Fish, Habitat Prediction, Google Earth Engine, Remote Sensing, MaxEnt Machine Learning

(Received 2025-06-20, Revised 2025-06-22, Accepted 2025-06-22, Available Online by 2025-07-24)

1. Introduction

Mahseer fish is included in the Cyprinidae group which has important value from socio-economic and ecological aspects. Mahseer fish is an icon of local wisdom and a high-value aquaculture commodity because of its complete nutritional content [1,2]. From an ecological aspect, this fish as a flagship species and bioindicator of the ecological function of an aquatic environment that is still good, this is because of habitat of Mahseer fish requires a clear water with dissolved oxygen levels in the range of 8 to 9 ppm and a water temperature of 18 to 32 ºC. [1,3,4].

In addition to their importance, several Mahseer fish species, namely, *Tor putitora*, *T. malabaricus, T. rimadevii, and T. laterivivatus,* are currently vulnerable to extinction and have critical populations in

Asia [5–8]. In Indonesia, the populations of *T. tambroides* and *T. tambra* tend to decline due to the loss of Mahseer fish habitat, overfishing, anthropogenic stress, invasive species, and climate change [9–11]. In these cases, the habitat dynamics of Mahseer fish are crucial for understanding their distribution and accompanying factors. Predicting the habitat with species distribution modeling approach integrated with machine learning can help study the distribution of Mahseer fish habitat.

Machine learning has been applied in many biology studies, such as the detection of Grasserie disease in *Bombyx mori* using a combination of the Histogram of Oriented Gradients (HOG) for feature extraction, Kernel Principal Component Analysis (KPCA), and Support Vector Machine (SVM) as reduction dimension and classifier with an accuracy value of 93,16% and an AUC value of 0,94 [12]. Moreover, machine learning is also used in agriculture, such as the detection of corn crop health by classifying blight and leaf spots of corn leaves with data augmentation techniques in Convolutional Neural Network (CNN) architecture. This technique can detect symptoms of corn disease with a validity value until to 98,50% [13]. The effectiveness of these studies underlines the utility of deep learning and machine learning approaches for imagery analysis in ecological and conservation-related applications.

Maximum Entropy (MaxEnt) analysis is classified as a machine learning method because it randomly categorizes species presence data using R, Python, and JavaScript. MaxEnt is promising for habitat modeling because it is a user-friendly, effective machine learning algorithm that predicts species distributions based on presence-only data, making it suitable for various ecological applications. Its ability to handle complex relationships between species and environmental variables, along with its widespread validation and application in ecological studies [14,15]. In prior investigations, the MaxEnt analysis technique was able to forecast multiple species of freshwater fish beyond their normal environment in rivers in the Virginia area [16]. Furthermore, MaxEnt is capable of predicting potential loss in endemic fish populations and economically valuable species such as *Seminemacheilus lendlii, T. tor*, *T. putitora*, and green snapper in the subtropical region over the coming decades with good accuracy (AUC 0.8) with limited presence points than other modeling methods [17–21].

The R studio is commonly used for MaxEnt analysis, with environmental data as predictor accessed from *www.worldclim.org/data/bioclim.html*. R Studio is designed in the R programming language, offering the advantages of being user-friendly, effective in classifying data in both text and spatial formats, and able to reduce memory usage while processing massive volumes of data [22]. Despite these advantages, The R language still needs to be adjusted for a larger area with different latitudes and not integrated with the data catalog, which will require a computer with high specifications and reduce the efficiency of raster data processing. Moreover, applying merely weather data to depict environmental variables for understanding riverine fish habitats is insufficient, because biological factors cannot stand alone in influencing the environmental dynamics of river ecosystems, but are usually supplemented by geomorphic factors [23–25].

In the context of these shortcomings, the Google Earth Engine (GEE) platform and JavaScript language were used in this study for MaxEnt analysis. The advantages of GEE are that it is integrated with satellite imagery as a catalog, so that image acquisition and correction can be carried out on the same platform. GEE is based on Google's cloud infrastructure [26], and can be more effective for processing raster and shapefile data in MaxEnt analysis.

This research area covers most of Central Java because it has a complex topography in the form of mountainous areas and lowlands that form river basins with diverse characteristics, which can represent tropical rivers in Asia. Several of these river basins are known to be the habitat of Mahseer fish [27–29], which are currently experiencing critical conditions due to forest degradation and land conversion [30], [31], therefore, conservation efforts are needed to protect the Mahseer fish population from the threat of extinction.

The conservation of Mahseer fish will have a positive link to the subset of SDG's goals of zero hunger by increasing small-scale agriculture and fisheries productivity, including freshwater fisheries, clean water sanitation by protecting and restoring aquatic environments, including lakes and rivers, which are important habitats for freshwater species, climate action, and life on land by conserving terrestrial and aquatic ecosystems to stop biodiversity loss and controlling invasive species that threaten local species [32][33].

Therefore, the initial information on habitat suitability of Mahseer fish in tropical rivers such as Indonesia needs to be explored with a novel approach in the era of data openness. The combination of Satellite Imagery to generate environmental data and geomorphological conditions with cloud-based processing, such as the GEE platform, would propose a rapid and effective approach to analyzing rapid environmental changes. This study proposed a novel approach for obtaining initial information on habitat suitability of Mahseer fish using MaxEnt machine learning in the GEE platform, which does not depend entirely on computer specifications and has not previously been applied in this region. This study aimed to predict the distribution of flagship species of Mahseer fish in the rapidly changing and high-pressure environmental regions of Java, which could be an interpretation of another tropical region in Asia.

2. Methods

This research was conducted from August 2022 to July 2023 in the upstream river systems of Central Java. A purposive random sampling method was applied, and ensuring a minimum spatial distance of 5 km between points to reduce spatial autocorrelation. Ten GPS-recorded presence points were collected and saved as shapefiles using QGIS 3.22. Extracting and processing raster data was performed using JavaScript in the Google Earth Engine platform, while multicollinearity test was performed using Microsoft Excel. The methodology used in this study is an adaptation of prior research [34] and consists of six stages:

2.1. Sample collection and presence data

The presence of Mahseer fish (Tor spp.) was obtained from literature studies and interviews with local anglers and was used as a basis for determining the location of sample. Sampling gear was carried out using the casting fishing technique with a fishing rod in a rapid and rocky river flow. The coordinate points at the location of the Mahseer fish that were successfully caught were recorded using GPS and referred to as the presence point of Mahseer fish. The entire presence coordinates were tabulated using Microsoft Excel, saved in CSV format, and then converted into shapefile data using QGIS software version 3.22 for spatial analysis.

2.2. Merging sample presence and absence data

The sample presence and absence data were combined using polygon vector data from the research area in the form of a base map of Central Java Province and the Special Region of Yogyakarta, which was designated as a Region of Interest (ROI) (Figure 1), as well as river flow data downloaded in shapefile format from the page https://www.hydrosheds.org/products/hydrorivers. The river flow vector data were analyzed in QGIS software version 3.33 using the buffering methodology, resulting in a Strahler-based river order classification. In addition, the two vector datasets were overlaid with the presence point dataset to calculate the location of the absence points. Meanwhile, absent points were generated randomly in places outside reach of the ROI's using the randomPoints function. All presence and absence data were then combined and used as a training dataset for MaxEnt modeling on the Google Earth Engine (GEE) platform using the JavaScript programming language, as demonstrated in the following script:

```
var absence = ee.FeatureCollection.randomPoints({
region: vector_area_absence,
seed: 1,
points: presence.size().multiply(3)
}).map(function(x){ return x.set({presence: 0}) });
var merge = presence.merge(absence)
```

Figure 1. Research location in Central Java

2.3. Envrionmental predictor and prepocessing

The environmental predictor variables used were NDVI (Sentinel-2, 10 m), rainfall (CHIRPS Daily), temperature (MODIS MOD11A1, 1 km), elevation and slope (SRTM, 30 m), and river order (HydroSHEDS). All those imageries then generated into raster data were accessed via Google Earth Engine (GEE) and temporally averaged using `.mean ()` from August 2022 to July 2023 by these JavaScript below:

Sentinel Harmonized-2 10m and NDVI

```
var image = ee.ImageCollection("COPERNICUS/S2_SR_HARMONIZED").filterBounds(ROI).
    filterDate('time_start','rime_end').map(function(x){var scl = x.select('SCL')
var cloud = scl.eq(3) .or(scl.eq(7)) .or(scl.eq(8)).or(scl.eq(9)) .or(scl.eq(10))return
    x.updateMask(cloud.eq(0))}).mean().clip(ROI)
var NDVI = image.normalizedDifference(['B8','B4']).rename('NDVI')
```

Rainfall

```
var Rainfall = ee.ImageCollection("UCSB-CHG/CHIRPS/DAILY").filterBounds(ROI)
    .filterDate('time_start,'time_end').sum().clip(ROI).rename('Rainfall').convolve(ee.Kernel.circle({r
    adius: 4}))
```

Temperature

```
var Temperature = ee.ImageCollection("MODIS/061/MOD11A1").select('LST_Day_1km')
    .filterBounds(ROI).filterDate('time_start','time_end').map(function(x){ return x .multiply(0.02)
    .subtract(273.15).clip(batas1).copyProperties(x,['system:time_start','system:time_end'])})
    .Mean().clip(ROI).rename('Temperature').convolve(ee.Kernel.circle({radius: 4}))
```

Elevation and Slope

```
var Elevation=ee.Image("USGS/SRTMGL1_003").clip(ROI)
var slope = ee.Terrain.slope(Elevation)
```

2.4. Mutlicolinearity test

The multicollinearity test was used to identify the presence of predictors that has significant correlations with other predictors. A high correlation between independent variables can reduce the accuracy of coefficient estimation in the model and make it more difficult to comprehend each predictor's contribution to the dependent variable. Thus, detecting and addressing with multicollinearity are critical elements in the modeling process. In this stage, the multicollinearity test (VIF) was performed using QGIS to extract raster values and the Microsoft Excel program, and the predictors with VIF > 10 were excluded. All selected variables had VIF < 10, indicating acceptable levels of collinearity. The multicollinearity test is proposed by [28], with the following calculation formula:

$$Tolerance = 1 - r^2$$
$$VIF = \frac{1}{Tolerance}$$

2.5. MaxEnt modeling

Data processing commenced with the import of vector layers from Google Earth Engine (GEE) assets, including Mahseer presence points, river flow networks, and administrative base maps, into the code editor. The required raster datasets were acquired from the Earth Engine Data Catalogue and loaded into the same workspace. All preparation operations involving vector and raster data, along with MaxEnt modeling techniques, were implemented in JavaScript. The MaxEnt model generated species response curves, estimated contributions from each environmental predictor, model evaluation metrics, and a habitat suitability map for Mahseer fish in Central Java. Habitat suitability is determined using the Habitat Suitability Index (HSI) which is classified as binary, with values approaching 1 indicating extremely suitable habitat conditions and values near 0 indicating unsuitable locations [29]. The threshold in this study was 0.5. HSI ≥ 0.5 indicates suitable habitat; < 0.5 indicates unsuitable habitat. MaxEnt modeling was performed using the following JavaScript function:

```
var algorithm_maxent = ee.Classifier.amnhMaxent().setOutputMode('Probability')
.train({ features: sampling_maxent, classProperty: 'presence',inputProperties:
predictors.bandNames()});
```

2.6. Validation of MaxEnt Model

MaxEnt, which is executed on the Google Earth Engine platform, validates data using a statistical evaluation approach based on the Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. The table below provides an interpretation of AUC values.

Table 1. Interpretation of AUC Value

AUC Value	***Interpretation***
0.5	*Random*
0.6 – 0.7	*Low - Fair*
0.7 – 0.8	*Fair - Good*
0.8 – 0.9	*Good – Very Good*
1	*Accurate*

3. Results and Discussion

3.1. Presence-absence point

A presence-absence map of Mahseer fish in Central Java was generated by combining presence points and Region of Interest (ROI) vector data with the *reduceToVectors()* method and randomly sampling points inside the polygon area (Figure 2). Each point of absence of Mahseer fish was assigned a presence attribute of 0. The pseudo-absence points were generated randomly from the presence points by multiplying the number of presence points by three to produce a larger number of absence points to produce a sufficient presence-absence proportion model, allowing the model to recognize a greater number of non-habitat conditions and increasing prediction stability [35]. However, the number of presence points required to obtain the optimal absence points is determined by the modeling method utilized [36].

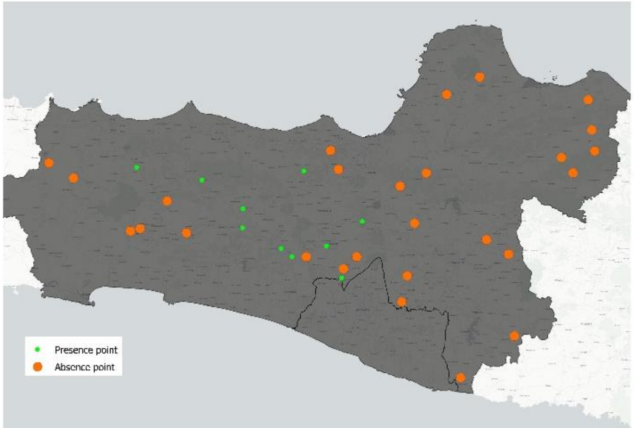

Figure 2. Map of presence-absence point of Mahseer fish in Central Java

Mahseer fish are most commonly found upstream of major streams that flow in hilly places at elevations ranging from 300 to 900 m above sea level. This is consistent with the fact that Mahseer fish prefer large, clear, fast-flowing rivers with rocky substrates [37], such as the upstream areas of the Serayu and Progo rivers, which flow through the Dieng, Sindoro, and Sumbing mountain ranges. Meanwhile, other rivers, such as the Comal, Bodri, Bogowonto, and Jali, have fewer meeting spots than Progo and Serayu. The frequency of Mahseer fish presence sites in these rivers suggests that the downstream regions of these rivers are less appropriate for Mahseer habitat, despite having habitat criteria for Mahseer fish at the sample site.

3.2. Vector data processing

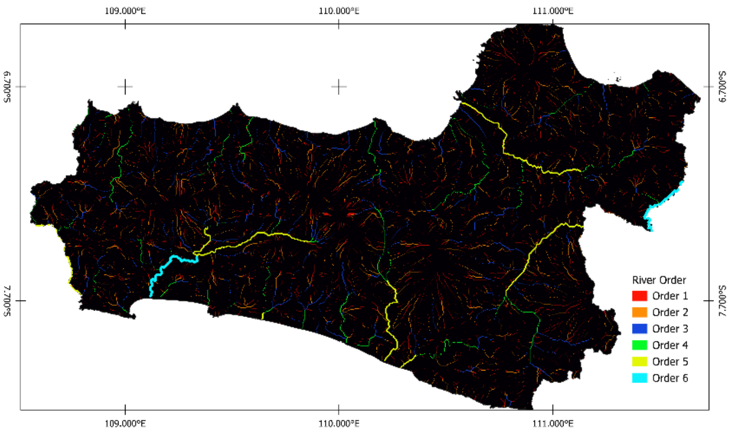

Figure 3. *Map of classification of river order*

The vector data downloaded from https://www.hydrosheds.org/products/hydrorivers then processed in QGIS with buffering method. The river vector data buffered acording to the river Strahler order classification produced six levels of river flow classification in Central Java (Figure 3). River order 1 is indicated by a small red river flow and is the starting point of the river flow. Order 2 was brown, longer, broader, and intersected with order 1. Order 3 is blue and wider than order 2 and is a river flow resulting from the intersection of order 2. The intersection of order 3 river flows will be an order 4 river, which is green and wider. Rivers with order 4 in the Progo and Elo rivers are the fourth-order rivers in the Progo Watershed.

Both rivers have the characteristics of fast-flowing rivers with rocky, gravel and sandy substrates, and greater erosion and sediment transport. The intersection of order 4 rivers will produce order 5 rivers. Order 5 rivers are wider, and have more complex characteristics than smaller orders, and the water flow is more stable because it receives flow input from smaller river orders and plays an important role in sediment deposition and forming deltas downstream of the river.

River order indirectly contributes to the biodiversity richness of the river ecosystem by forming the size of the channel and substrate types that support the habitat ranges of organisms. In the context of habitat modelling of endemic species in restricted habitats, such as Mahseer fish, river order, and other fluvial fundamental information, such as channel size and connectivity between river segments and river paths, are provided from the hydrology dataset in shapefile form [38,39]. Mahseer fish presence in the map (Figure 2) tend to inhabit in the upstream whereas characterized by cold and clean water, fast flowing and low temperatures [40], those characteristics are typically of upstream in lower order.

3.3. Raster data processing

The predictors used in species habitat modelling of Mahseer fish consist of NDVI (Sentinel-2, 10 m), rainfall (CHIRPS Daily), temperature (MODIS MOD11A1, 1 km), elevation, and slope (SRTM, 30 m) which are form in raster data format that is accessed and processed the cloud-based platform of Google Earth Engine with JavaScript language. The average NDVI value in Central Java between August 2022 and July 2023 was 0.62 in January 2023 and 0.51 in November 2022 (Figure 4.a).

The NDVI index was generated by comparing of the spectral reflection values of green light (NIR) and red light (Red) from Sentinel imagery. NDVI is used to monitor vegetation density around freshwater ecosystems, including rivers, because NDVI is related to the habitat of aquatic organisms as a food provider for organisms in rivers and their surroundings, natural filtration from land to rivers and at the same time a buffer against fluctuations in lake water temperature [41–45]. Meanwhile, daily rainfall fluctuations in Central Java from August 2022 to July 2023 ranged from 0 to 35 mm/d with significant peak rainfall in November 2022 and January 2023 (Figure 4.b).

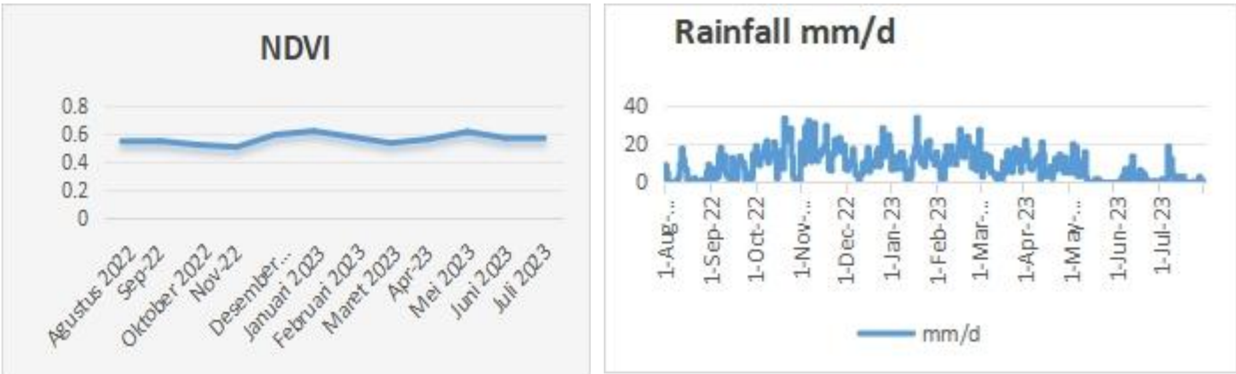

Figure 4. a.) NDVI value, b.) Average of rainfall within 12 months in Central Java

Fluctuations in rainfall can illustrate the presence of seasonal rivers upstream during the rainy season, as well as new inundation areas that will become seasonal migration destinations for freshwater fish to lay eggs and avoid anglers [46–48]. Rainfall can cause a decrease in SO_4- compounds and, sediment dissolution into water bodies, which increase N, P, NH4+ -N compounds, increasing the turbidity of river water. Such, conditions can be a place of shelter for fish against predator attacks [49,50]. Temperature values were obtained from MOD11A1.061 Terra Land Surface Temperature and Emissivity Daily Global 1 km in the period from August 2022 to July 2023 in the Central Java region. The temperature recorded from the image is between 7℃ which is interpreted in blue to red at a temperature of 36℃ (Figure 5).

Most of the wide rivers in the map (Figure 5) are green, while the small rivers in the highlands are blue, indicating low temperatures in the area, and the yellow and red colors in the lowlands indicate warmer temperatures. In general, the temperature distribution in Central Java over a period of 1 year was relatively similar, as indicated by the green color. Temperature fluctuations were detected in the Bengawan Solo River flow on its southeast side. The Bengawan Solo River flow upstream from August 2022 to March 2023 was red indicating an increase in temperature, and yellowish green from April to July 2023, indicating lower temperatures.

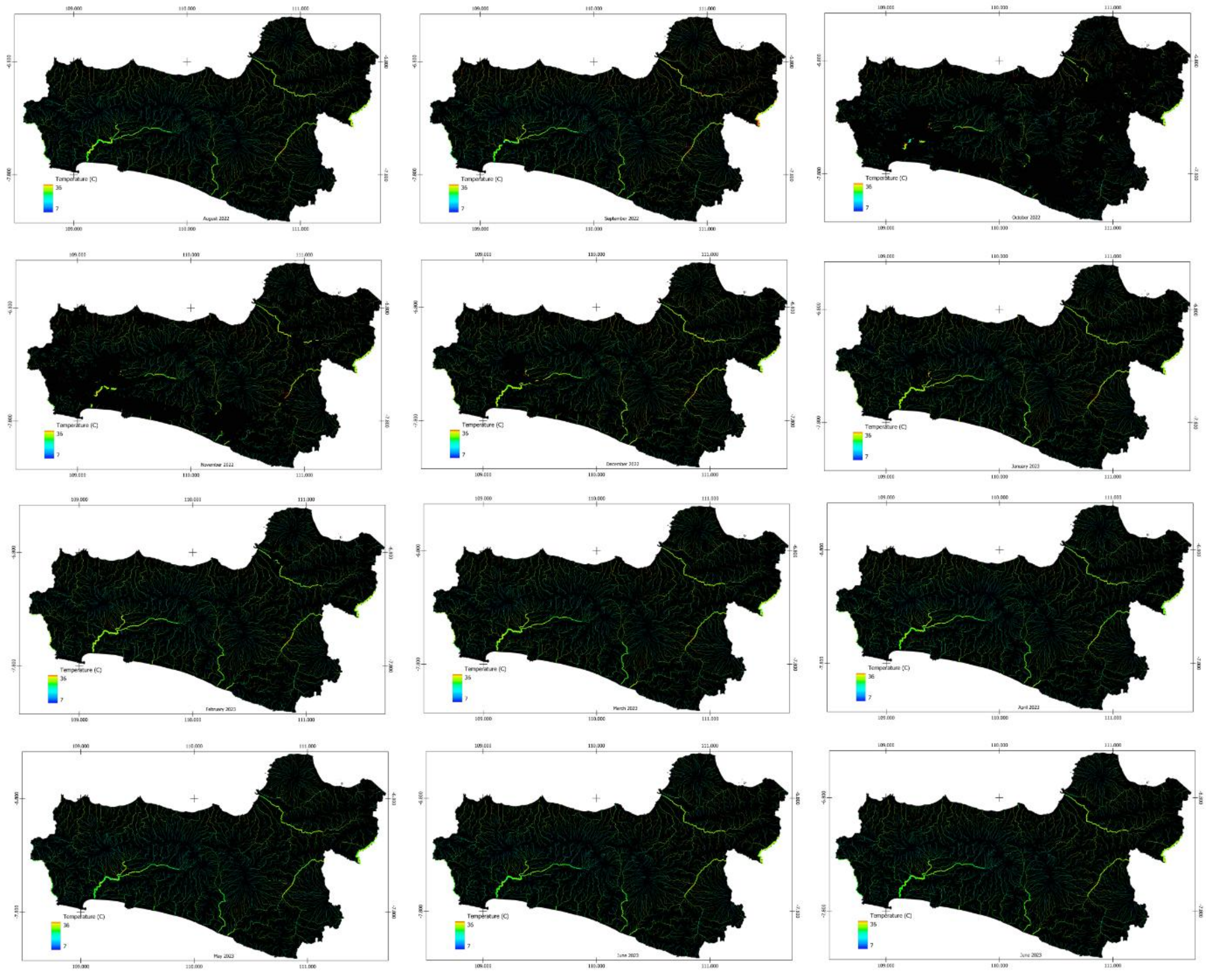

Figure 5. Average temperature over 12 months.

The river flow elevation map in most areas of Central Java was acquired from NASA SRTM Digital Elevation 30 m imagery showing an elevation range of 0 to 2000 m above sea level (Figure 6.a). The yellowish-brown river flow indicates an altitude between 0 and 500 m above sea level, whereas the dark brown to green river flow indicates an altitude between 500 m above sea level and 2000 m above sea level. Rivers at an altitude of more than 100 m above sea level are generally headwaters in mountainous areas. The slope map of Central Java in degrees (Figure 6.b.) was generated from an overlay of NASA SRTM Digital Elevation 30 m imagery with a vector map of river flow with a slope range of 0° (indicated by the yellow area) to 50° (indicated by the green area).

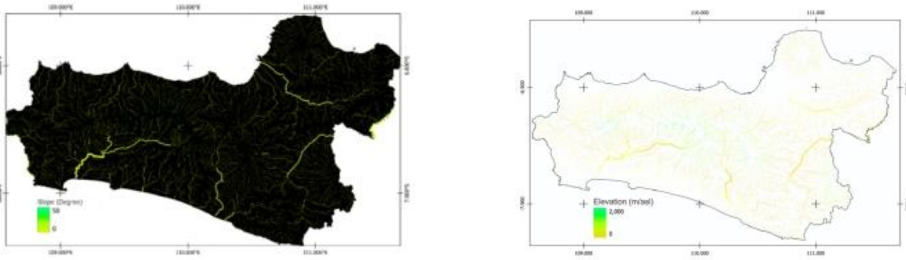

Figure 6. a.) Elevation, b.) Slope in Central Java.

The slope map shows smaller rivers that tend to be greenish located in mountainous areas, indicating a slope of more than 25˚, while rivers with a color approaching yellow indicate a slope value below 25˚ located in lower areas. River slope was positively correlated with species richness, and rivers with steep slopes were habitats for few fish species [51]. The Habitat Suitability Index (HSI) map of Mahseer fish in Central Java, analyzed using the Maximum Entropy approach, was interpreted in a range of values from 0 (red) to 1 (green) (Figure 7). The habitat suitability map of Mahseer fish was interpreted from August 2022 to July 2023. According to the habitat suitability map, Mahseer fish in Central Java have habitat suitability in the upstream areas of highland rivers in the Serayu and Lawu mountains. River flows with orders 1, 2, and 3 are suitable for Mahseer fish habitat, and 4 is suitable in the rainy season. In addition, the range of Mahseer fish habitat suitability changes throughout the year.

The Serayu River flow had a suitable habitat for Mahseer fish from August 2022 to October 2022, and January 2023 to March 2023, while June 2023 and July 2023 were only found in Order 4 near the branching of Order 3, which was marked with green to dark green with a habitat suitability score of 0.7. Meanwhile, in November 2022 and December 2022, almost the entire Serayu River flow was in green, which means the habitat suitability value for Mahseer fish was more than 0.5. In April 2023 and May 2023, the Serayu River flow of orders 5 and 6 and the Mawar Bedono River flow of order 4 were yellow, which means that the habitat suitability value for Mahseer fish in the area was approximately 0.5. Other areas in the flow tended to be reddish-yellow which indicating that the habitat was less suitable for Mahseer fish. The river basins around the Serayu River Basin such as Bogowonto, Jali, Progo, Comal, Lampir, and Bodri, also show the same indications. Meanwhile, the river basins outside the Serayu mountain area that have suitable habitat for Mahseer fish are the upstream streams of Bengawan Solo, Tuntang and Opak Rivers in Yogyakarta.

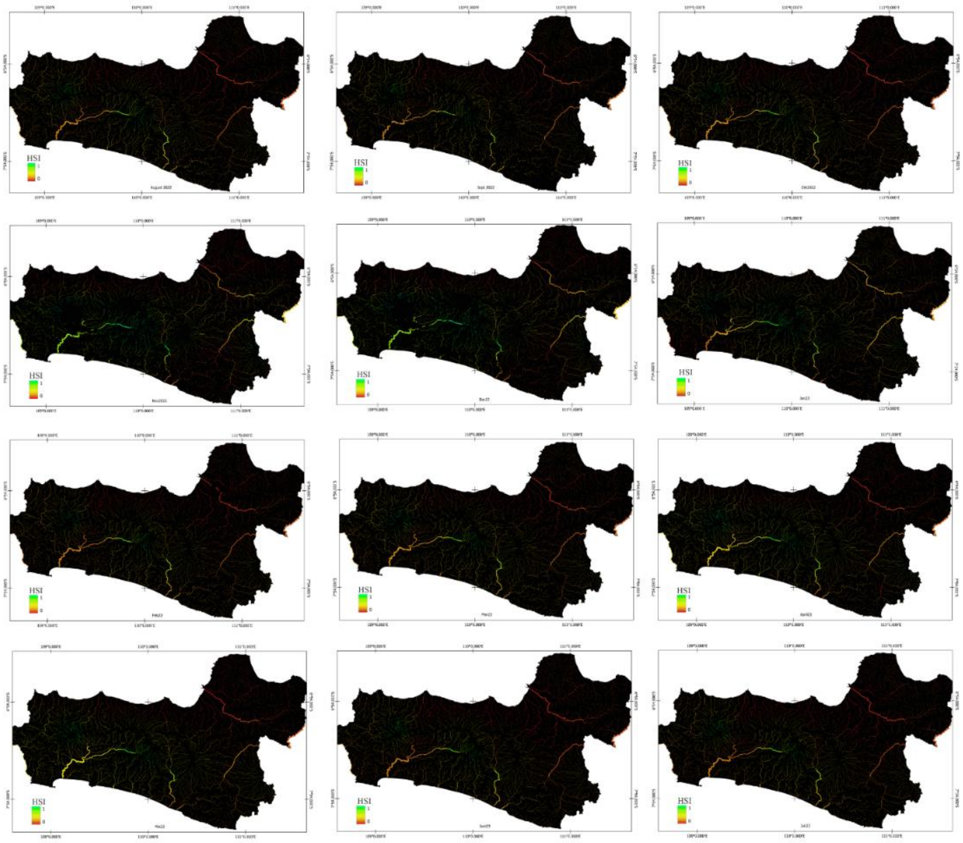

Figure 7. Map of habitat suitability of mahseer fish in Central Java

Based on the habitat suitability map of Mahseer fish in Central Java, the river basin in the upstream part of the Serayu Mountains is a stable habitat for Mahseer fish throughout the year. The predictor variables that had a large contribution to the suitability of the Mahseer fish habitat were River Order (73%), Elevation (18%) and Rainfall (8%), while NDVI, Temperature and Slope gradient did not contribute to the suitability of the Mahseer fish habitat. The accuracy value of the Mahseer fish habitat suitability map was 0.70, which means that the modeling of the Mahseer fish habitat suitability was quite good and close to accurate. This accuracy value is interpreted as the Area Under Curve (AUC) value on the Google Earth Engine platform console.

3.4. Multicolinearity test

The results of the multicollinearity test, calculated using the variance inflation factor (VIF) formula, produced a VIF value between 1 and 10. The VIF value reflects the extent to which the variation in a regression coefficient estimate can be attributed to a linear relationship with other variables. A lower VIF indicates reduced linearity among predictors, whereas a higher VIF indicates increased linearity between predictors. The threshold for the VIF value in this study was set at 7, indicating a moderate degree of correlation among the predictors. The results of the multicollinearity test (Table 2), indicate that the average correlation value between the predictors (r) was 0.034. Additionally, the coefficient of determination (r^2) was 0.176, the tolerance level was 0.8, and the variance inflation factor (VIF) is 1.21.

Table 2. Multicollinearity value between predictors

Predictors	**r**	**r^2**	**Tolerance**	**VIF**	**Conculsion**
Elevation vs NDVI	0.451	0.204	0.796	1.255587	<10
Elevation vs Rainfall	-0.054	0.003	0.997	1.002974	<10
Elevation vs Slope	0.684	0.468	0.532	1.879939	<10
Elevation vs Temperature	-0.708	0.501	0.499	2.00434	<10
Elevation vs River Order	0.655	0.429	0.571	1.751667	<10
NDVI vs Rainfall	-0.186	0.034	0.966	1.035678	<10
NDVI vs Slope	0.067	0.004	0.996	1.004487	<10
NDVI vs Temperature	-0.287	0.082	0.918	1.089623	<10
NDVI vs River Order	0.437	0.191	0.809	1.236663	<10
Rainfall vs Slope	-0.362	0.131	0.869	1.15127	<10
Rainfall vs Temperature	0.087	0.008	0.992	1.007626	<10
Rainfall vs River Order	0.108	0.012	0.988	1.011904	<10
Slope vs Temperature	-0.598	0.357	0.643	1.555988	<10
Slope vs River Order	0.414	0.171	0.829	1.206658	<10
Temperature vs River Order	-0.194	0.038	0.962	1.039184	<10
Average	0.034	0.176	0.824	1.213038	

The highest correlation value (r) between the predictors occurred with elevation and temperature, exhibiting the strongest correlation of -0.708, although in a negative direction. The coefficient of determination (r^2) was 0.501, suggesting that elevation accounted for 50% of the variance in temperature. The tolerance value was 0.499, and the variance inflation factor (VIF) was 2. Although the relationship between elevation and temperature is essential, the coefficient of determination (r^2), tolerance, and VIF values were all below the established thresholds of 0.501, 0.499, and 2 for VIF, respectively. The threshold for the Variance Inflation Factor (VIF) in habitat modeling is primarily dependent on the analysis method employed, resulting in the absence of a universal criterion for its determination (Kock & Lynn, 2012). In the context of the SEM-PLS analysis of variance, a VIF value greater than 3 is indicative of multicollinearity symptoms. Conversely, in habitat suitability analysis utilizing the MaxEnt machine learning, a VIF value exceeding 5 suggests the presence of multicollinearity symptoms [52,53].

4. Discussion

In this study, the habitat suitability index for mahseer (*Tor* spp) in Central Java was modeled using a suite of environmental predictors obtained from the Google Earth Engine (GEE) catalog. Specifically, geomorphological variables (elevation, river order, and slope) were employed alongside climatic variables (temperature and rainfall) and a vegetation index derived via the Normalized Difference Vegetation Index (NDVI) approach.

The Area Under the Curve (AUC) value in this study was 0.70, with river order contributing 73% and elevation 18%, indicating a reasonably good predictive performance, although relatively lower than that reported in several other MaxEnt studies in tropical river systems. For example, a study on Tor putitora in the Himalayan river system reported an AUC of 0.867, with upstream elevation contributing 35.3% and flow length 21.4% [54]. Additionally, the invasion model for *Oreochromis niloticus* in the Ganges River achieved a test AUC of 0.999, with elevation as the dominant variable (69.2%), followed by slope (10.1%) [55]. These comparisons demonstrate that, although the AUC value in this study is relatively moderate, geomorphological variables such as river order and elevation exhibit important patterns that are consistent with general ecological trends in tropical river systems.

In this study, river order emerged as the most influential variable, surpassing elevation, a finding that stands out compared to many tropical fish habitat models, which often emphasize elevation. From a geomorphological perspective, Strahler's river order reflects the size and complexity of stream networks, where a higher number of tributaries supports the availability of natural food resources and provides refuge from predation. Mahseer fish in this study were primarily associated with river orders 1, 2, and 3, with a minor presence in order 4. This confirms that Mahseer fish are habitat specialists that require environments with abundant food sources but limited competition from other fish. Consequently, river order can be considered a critical indicator of Mahseer habitat suitability.

The model's performance underscores the strong dependency of Mahseer of genus Tor populations in Central Java on interconnected river corridors, a finding that is occasionally explicitly addressed in Southeast Asian literature. Studies on macroinvertebrates, which constitute a major food source for *Tor* spp. in tropical rivers, have also associated higher river orders with greater ecological diversity. For example, Shannon index values for macroinvertebrate communities typically decrease with river order, whereas fish communities show an inverse trend, increasing in diversity as the river order increases [56]. River order and the number of tributaries in a watershed reflect the interconnectedness of the river network, which, in turn, influences species composition and richness [57]. River order is a natural determinant of biodiversity, often showing a positive correlation with fish species richness [58,59]. In this study, at locations where Mahseer fish were frequently encountered, the presence of *Osteochilus vitatus*, a species with similar feeding habits, was also noted.

In addition to river order, elevation significantly influences Mahseer habitat suitability in Central Java. Based on the habitat suitability index, upstream river segments with elevations ranging from 150 to 927 m above sea level were identified as suitable habitats. This elevation range is particularly distinctive as it encompasses the hydrological and thermal conditions characteristic of highland areas in Central Java, which have not been previously quantified for Tor species on the island of Java. Higher elevations are generally associated with lower water temperatures, higher dissolved oxygen levels, and rocky substrates, all of which support nutrient availability and provide shelter for larvae and juvenile fish. This finding aligns with a study by [60], which found that *T. tambra* is commonly found in highland rivers in Malaysia's Baleh River. Changes in elevation along river gradients lead to alterations in the physical and chemical properties of water, thereby influencing species distribution patterns [61]. According to [60] fish abundance is positively correlated with elevation and pH and negatively correlated with turbidity.

Rainfall as a climate variable contributed approximately 8% to the model. Although this is lower than the contributions from topographic variables, rainfall plays a vital role in the reproductive cycle and larval development of Mahseer. Rainfall influences light penetration, flow regimes, and discharge fluctuations, which affect food availability and habitat space. In the Mahseer habitat suitability map of Central Java (Figure 7), river order 4 streams were yellowish-green during December and January, when rainfall was relatively high. This suggests that Mahseer fish may also be present in order 4 streams during the rainy season. A study on *Gymnocypris chilianensis* and *Triplophysa hsutschouensis* in the Hexi River system of the Qinghai–Tibet Plateau reported that rainfall was more influential than temperature as a climatic variable [62]. These findings affirm that while climate variables may not always be the primary contributors compared to topographic factors, their role in tropical freshwater habitat modeling should not be overlooked.

Based on the findings of the above study, geomorphological variables such as elevation, valley width, and slope can influence the structure of fish communities along river systems [63]. In this study, river order, elevation, and rainfall were the most significant contributors to the Mahseer habitat suitability index. Strahler's river order, a classification system emphasizing river flow size, including width and length, proved to be an important ecological indicator. The predictive model for Mahseer habitat in this study may be applied to other tropical river systems; however, its transferability is limited because of substantial differences in water physics-chemical parameters, land use, and anthropogenic pressures. While the principle of "niche conservatism" [64] supports model transferability under similar environmental conditions, the rivers in Central Java are uniquely influenced by dams, irrigation systems, and pollution, requiring local calibration. Thus, prior to model application in other regions, field validation and multicollinearity analysis of local variables are essential to ensure ecological relevance [65,66].

Another study also highlighted the limited transferability of SDM models across environmental and geographic space [67]. A model with a high AUC in one location (e.g., the Himalayas) may experience reduced accuracy when applied to other watersheds without recalibration. From this perspective, the moderate AUC value (0.70) observed in this study can be interpreted as a realistic reflection of a highly fragmented and anthropogenically impacted river system rather than a purely technical shortcoming.

The use of the Google Earth Engine (GEE) platform to estimate species habitat suitability using an SDM approach represents a relatively novel methodology. The GEE is particularly advantageous because of its integration with a multi-petabyte data catalog that enables rapid analysis and supports collaborative research efforts [34,68]. However, GEE lacks several essential analytical features, such as background data selection, model replication, Jackknife analysis, multicollinearity diagnostics, and AUC curve visualization [64]. These limitations can be addressed externally using simpler tools, such as QGIS and Microsoft Excel, for multicollinearity testing. Similarly, correlation verification can also be conducted using the R programming environment prior to modeling with software such as ArcGIS [69,70].

Future habitat suitability modeling should incorporate physiological species data, such as tolerance to temperature, pH, oxygen levels, and biotic interactions (e.g., competition and predation), to improve habitat prediction under changing climate scenarios [71–73]. This approach would provide a more comprehensive prediction framework, as climate change is expected to alter the ideal elevation ranges and rainfall patterns, rendering models based solely on historical data insufficient for future projections. In relation to climate change, species distribution modeling supports SDG 13 (Climate Action) by documenting the ecological impacts of climate change on freshwater fish distribution patterns.

Moreover, habitat suitability modeling can contribute to other Sustainable Development Goals (SDGs), such as SDG 15 (Life on Land), by identifying critical reproduction and migration areas for riverine fish, aiding in the establishment of protected areas and sustainable fisheries management. It also supports SDG 6 (Clean Water and Sanitation) by identifying environmental variables (e.g., water quality and flow patterns) essential for freshwater ecosystem monitoring and conservation, which are crucial for ensuring a clean water supply for communities [74].

5. Conclusions

MaxEnt modeling on the Google Earth Engine (GEE) platform in this study offers a new, scalable, and replicable approach to identify habitat suitability for Mahseer fish (*Tor* spp.) in a data-limited and dynamic ecological region in Central Java, with findings emphasizing the key role of geomorphological factors such as river order and elevation that shape habitat preference patterns. River orders 1 and 3 found in the upper reaches of the river play an important role as year-round Mahseer habitats, and order 4 as seasonal habitats in Central Java. The upper reaches of the river that are Mahseer habitats can be applied as priority zones for the protection and recovery of Mahseer fish populations and their habitats. River orders reflect river connectivity, and upstream integration has important ecological value for maintaining Mahseer fish populations. The results of this study also underline the importance of integrated watershed management policies that conserve upstream areas and maintain upstream connectivity amid massive ecological stress. This approach also helps identify potential refuges and spawning sites, which can inform spatial planning through seasonal fishing bans or community-based conservation zones. The integration of habitat suitability modeling through the GEE platform with MaxEnt machine learning can serve as a reference for other tropical river systems across Southeast Asia, enabling rapid assessments in data-poor areas and informing adaptive management strategies, especially when addressing the impacts of climate change and land use through predictive and seasonally dynamic models in future research
.

Acknowledgements

The authors are grateful to the Center for The Higher Education Funding (BPPT) and The Education Fund Management Agency (LPDP) through The Indonesia Education Scholarship (BPI) by the Ministry of Higher Education, Science and Technology (Kemendiktisaintek).

References

[1] Everard M, Pinder AC, Claussen JE, Orr S. Assessing the societal benefits of mahseer (Tor spp.) fishes to strengthen the basis for their conservation. Aquat Conserv 2021;31:2979–86. https://doi.org/10.1002/aqc.3683.

[2] Muchlisin ZA, Nur FM, Maulida S, Syahfrida Handayani L, Riska Rahayu S. Mahseer, the history of the king of the river. E3S Web of Conferences, vol. 339, 2022, p. 03006. https://doi.org/10.1051/e3sconf/202233903006.

[3] Johnson JA, Dhawan B, Sivakumar K. Study on ecology and migratory patterns of golden mahseer (Tor putitora) in river Ganga using Radio telemetry techniques, Wildlife Institute of India Project Report. Wildlife Institute of India: 2021. https://doi.org/10.13140/RG.2.2.30906.06084.

[4] Larashati S, Sulastri, Ridwansyah I, Afandi AY, Novianti R. Conservation efforts of ikan Batak (Tor spp. and Neolissochilus spp.) and its prospects to support ecotourism in Samosir Regency, North Sumatra Indonesia. IOP Conf Ser Earth Environ Sci 2020;535:0–10. https://doi.org/10.1088/1755-1315/535/1/012041.

[5] Jha BR, Rayamajhi A, Dahunakar N, Harrison A, Pinder A. Tor putitora. The IUCN Red List of Threatened Species 2018. The IUCN Red List of Threatened Species 2018: ET126319882A126322226 2018. https://doi.org/https://dx.doi.org/10.2305/IUCN.UK.

[6] Kottelat M. Tor laterivittatus. The IUCN Red List of Threatened Species 2018. The IUCN Red List of Threatened Species 2018: ET187921A126323049 2018. https://doi.org/https://dx.doi.org/10.2305/IUCN.UK.

[7] Pinder AC, Katwate U, Dahanukar N, Harrison AJ. Tor remadevii. The IUCN Red List of Threatened Species 2018. Tor Remadevii The IUCN Red List of Thretened Species 2018: ET56096394A56717605 2018. https://doi.org/https://dx.doi.org/10.2305/IUCN.UK.

[8] Raghavan R, Ali A. Tor malabaricus. The IUCN Red List of Threatened Species. The IUCN Red List of Threatened SpeciesA6895822 2011. https://doi.org/dx.doi.org/10.2305/IUCN.UK.

[9] Kottelat M, Pinder AC, Harrison AJ. Tor tambroides. The IUCN Red List of Threatened Species 2018. The IUCN Red List of Threatened Species 2018: ET187939A91076554 2018. https://doi.org/https://dx.doi.org/10.2305/IUCN.UK.

[10] Abass Z, Shah, T H, Bhat, F A, Ramteke K, Magloo, A H, Hamid I, et al. The mahseer: The tiger of water-an angler’s delight in the Himalayas and the undisputed king of sport fishing. Fish Res 2024;279. https://doi.org/https://doi.org/10.1016/j.fishres.2024.107147 Get rights and content.

[11] Gupta N, Everard M, Nautiyal P, Kochhar I, Sivakumar K, Johnson JA, et al. Potential impacts of non-native fish on the threatened mahseer (Tor) species of the Indian Himalayan biodiversity hot spot. Aquat Conserv 2020;30:394–401. https://doi.org/10.1002/aqc.3275.

[12] Manju G, Binson VA. Automated Disease Detection in Silkworms Using Machine Learning Techniques. Advance Sustainable Science, Engineering and Technology 2024;6:1–10. https://doi.org/10.26877/asset.v6i4.965.

[13] Ariska R, Sari CA, Rachmawanto EH. Classification of Corn Leaf Disease Using Convolutional Neural Network. Advance Sustainable Science Engineering and Technology 2024;6:0240408. https://doi.org/10.26877/asset.v6i4.772.

[14] Lissovsky AA, Dudov S V. Species-Distribution Modeling: Advantages and Limitations of Its Application. 2. MaxEnt. Biology Bulletin Reviews 2021;11:265–75. https://doi.org/10.1134/s2079086421030087.

[15] Kiedrzynsky M, Zielinska KM, Rewicz A, Kiedrzynska E. Habitat and spatial thinning improve the Maxent models performed with incomplete data. J Geophys Res Biogeosci 2017;122:1359–70. https://doi.org/doi:10.1002/ 2016JG003629.
[16] McGarvey DJ, Brown AL, Chen EB, Viverette CB, Tuley PA, Latham, O C, et al. Do fishes enjoy the view? A MaxEnt assessment of fish habitat suitability within scenic rivers. Biol Conserv 2021;263:109357. https://doi.org/https://doi.org/10.1016/j.biocon.2021.109357.
[17] Aksu S. Current and future potential habitat suitability prediction of an endemic freshwater fish species Seminemacheilus lendlii (Hankó, 1925) using Maximum Entropy Modelling (MaxEnt) under climate change scenarios: implications for conservation. Journal of Limnology and Freshwater Fisheries Research 2021;7:83–91. https://doi.org/10.17216/limnofish.758649.
[18] Mamun M, Kim S, An KG. Distribution pattern prediction of an invasive alien species largemouth bass using a maximum entropy model (MaxEnt) in the Korean peninsula. J Asia Pac Biodivers 2018;11:516–24. https://doi.org/10.1016/j.japb.2018.09.007.
[19] Mahato R, Nimasow G, Nimasow OD, Abujam S. Habitat Suitability Modeling of Tor tor (Hamilton, 1822) in the Indian Drainage Systems Using MaxEnt. Ecosystem and Species Habitat Modeling for Conservation and Restoration, Singapore: Springer Nature Singapore; 2023, p. 323–37. https://doi.org/10.1007/978-981-99-0131-9_17.
[20] Mahato R, Abujam S, Bushi D, Nimasow OD, Nimasow G, Das DN. Distribution modelling of Tor putitora (Hamilton, 1822), an endangered cyprinid in the Himalayan river system using MaxEnt. Acta Ecologica Sinica 2023;4:343–51. https://doi.org/https://doi.org/10.1016/j.chnaes.2022.01.004.
[21] Phillips SJ, Dudik M. Modeling of Species Distributions with MaxEnt: new extensions and a comprehensive evaluation. Ecography 2008;31:161–75. https://doi.org/10.1111/j.2007.0906-7590.05203.x.
[22] Jurka TimothyP. maxent: An R Package for Low-memory Multinomial Logistic Regression with Support for Semi-automated Text Classification. R J 2012;4:56–9.
[23] Elosegi A, Díez J, Mutz M. Effects of hydromorphological integrity on biodiversity and functioning of river ecosystems. Hydrobiologia 2010;657:199–215. https://doi.org/10.1007/s10750-009-0083-4.
[24] Yu Z, Fu Y, Zhang Y, Liu Z, Liu Y. Quantifying the Impact of Changes in Sinuosity on River Ecosystems. Water (Switzerland) 2023;15. https://doi.org/10.3390/w15152751.
[25] Zamroni A, Faustino-Eslava Decibel v. Geological Factors Influencing River Morphological Changes: Implications in the Agricultural Sector. In: Chiang PC, editor. Environment and Renewable Energy, Springer Singapore; 2023, p. 117–126. https://doi.org/https://doi.org/10.1007/978-981-97-0056-1_10.
[26] Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ 2017;202:18–27. https://doi.org/10.1016/j.rse.2017.06.031.
[27] Listyarini DW, Sulmartiwi L, Hasan V, Andriyono S. Morphological Characteristics Of Two Species Mahseers Fish (Cyprinidae; Torinae) From East Java. Jurnal Kelautan Dan Perikanan Terapan 2022;5:171–8.
[28] Jailani AQ, Mujtahidah T. Study of Iktiofauna Biodiversity Based on Geographical Information Systems (GIS) in the Progo River, Magelang, Central Java. Journal of Aquaculture Science 2021;6:38–47. https://doi.org/10.31093/joas.v6i1.136.
[29] Haryono, Tjakrawidjaja AH. Morphological Study for Identification Improvement of Tambra Fish (Tor spp.: Cyprinidae) from Indonesia. Biodiversitas 2006;7:59–62. https://doi.org/10.13057/biodiv/d070115.
[30] Nahib I, Amhar F, Wahyudin Y, Ambarwulan W, Suwarno Y, Suwedi N, et al. Spatial-Temporal Changes in Water Supply and Demand in the Citarum Watershed, West Java, Indonesia Using a Geospatial Approach. Sustainability (Switzerland) 2023;15. https://doi.org/10.3390/su15010562.

[31] Husodo T, Ali Y, Mardiyah SR, Shanida SS, Abdoellah OS, Wulandari I. Perubahan lahan vegetasi berbasis citra satelit di DAS Citarum, Bandung, Jawa Barat. Majalah Geografi Indonesia 2021;35:54. https://doi.org/10.22146/mgi.61217.

[32] Ho LT, Goethals PLM. Opportunities and challenges for the sustainability of lakes and reservoirs in relation to the Sustainable Development Goals (SDGs). Water (Switzerland) 2019;11:1–19. https://doi.org/10.3390/w11071462.

[33] Stokes GL, Lynch AJ, Smidt SJ, Steel EA, Dowd S, Britton JR, et al. Life on land needs fresh water (SDG 15). In: Mukherjee A, Matters W, editors. Water Matters, Elsevier; 2024, p. 295–309. https://doi.org/10.1016/b978-0-443-15537-6.00024-0.

[34] Crego RD, Stabach JA, Connette G. Implementation of species distribution models in Google Earth Engine. Divers Distrib 2022;28:904–16. https://doi.org/10.1111/ddi.13491.

[35] Warton DI, Shepherd LC. Poisson point process models solve the “pseudo-absence problem” for presence-only data in ecology. Annals of Applied Statistics 2010;4:1383–402. https://doi.org/10.1214/10-AOAS331.

[36] Barbet-Massin M, Jiguet F ric, Albert CH le`ne, Thuiller W. Selecting pseudo-absences for species distribution models: how, where and how many? Morgane. Methods Ecol Evol 2012;3:327–38. https://doi.org/doi: 10.1111/j.2041-210X.2011.00172.x.

[37] Jaafar F, Na-Nakorn U, Srisapoome P, Amornsakun T, Duong TY, Gonzales-Plasus MM, et al. A Current Update on the Distribution, Morphological Features, and Genetic Identity of the Southeast Asian Mahseers, Tor Species. Biology (Basel) 2021;10. https://doi.org/10.3390/biology10040286.

[38] Lehner B, Grill G. Global river hydrography and network routing: Baseline data and new approaches to study the world’s large river systems. Hydrol Process 2013;27:2171–86. https://doi.org/10.1002/hyp.9740.

[39] Amatulli G, Garcia Marquez J, Sethi T, Kiesel J, Grigoropoulou A, Üblacker MM, et al. Hydrography90m: A new high-resolution global hydrographic dataset. Earth Syst Sci Data 2022;14:4525–50. https://doi.org/10.5194/essd-14-4525-2022.

[40] Muchlisin ZA, Fadli N, Batubara AS, Nur FM, Irham M, Muhammadar AA, et al. Taxonomic Diversity of the Genus Tor (Cyprinidae) From Aceh Waters in Indonesia Based on Cytochrome Oxidase Sub-Unit I (Coi) Gene. Zoodiversity 2022;56:195–202. https://doi.org/10.15407/zoo2022.03.195.

[41] Huang C, Yang Q, Huang W. Analysis of the spatial and temporal changes of NDVI and its driving factors in the wei and jing river basins. Int J Environ Res Public Health 2021;18. https://doi.org/10.3390/ijerph182211863.

[42] Ndehedehe CE, Burford MA, Stewart-Koster B, Bunn SE. Satellite-derived changes in floodplain productivity and freshwater habitats in northern Australia (1991–2019). Ecol Indic 2020;114:106320. https://doi.org/10.1016/j.ecolind.2020.106320.

[43] Sharma P, Varga M, Kerezsi G, Kajári B, Halasi-Kovács B, Békefi E, et al. Estimating Reed Bed Cover in Hungarian Fish Ponds Using NDVI-Based Remote Sensing Technique. Water (Switzerland) 2023;15. https://doi.org/10.3390/w15081554.

[44] Vieira TB, Dias-Silva K, Pacífico ES. Effects of riparian vegetation integrity on fish and Heteroptera communities. Appl Ecol Environ Res 2015;13:53–65. https://doi.org/10.15666/aeer/1301_053065.

[45] Wang F, Wang X, Zhao Y, Yang Z. Temporal variations of NDVI and correlations between NDVI and hydro-climatological variables at Lake Baiyangdian, China. Int J Biometeorol 2014;58:1531–43. https://doi.org/10.1007/s00484-013-0758-4.

[46] Patrick AES. Influence of rainfall and water level on inland fisheries production: A review. Scholars Research Library Archives of Applied Science Research 2016;8:44–51.

[47] Sanches PV, Gogola TM, Silva RO, Topan DA, dos Santos Picapedra PH, Piana PA. Arms as areas for larval development of migratory fish species in a Neotropical reservoir and the

influence of rainfall over abundances. J Fish Biol 2020;97:1306–16. https://doi.org/10.1111/jfb.14474.
[48] Wang F, Chen H, Lian J, Fu Z, Nie Y. Hydrological response of karst stream to precipitation variation recognized through the quantitative separation of runoff components. Science of the Total Environment 2020;748:142483. https://doi.org/10.1016/j.scitotenv.2020.142483.
[49] Ng'onga M, Kalaba FK, Mwitwa J, Nyimbiri B. The interactive effects of rainfall, temperature and water level on fish yield in Lake Bangweulu fishery, Zambia. J Therm Biol 2019;84:45–52. https://doi.org/10.1016/j.jtherbio.2019.06.001.
[50] Koushali PH, Mastouri R, Khaledian MR. Impact of Precipitation and Flow Rate Changes on the Water Quality of a Coastal River. Shock and Vibration 2021;2021. https://doi.org/10.1155/2021/6557689.
[51] Camana M, Dala-Corte RB, Becker FG. Relation between species richness and stream slope in riffle fish assemblages is dependent on spatial scale. Environ Biol Fishes 2016;99:603–12. https://doi.org/10.1007/s10641-016-0502-0.
[52] Stuart CE, Wedding LM, Pittman SJ, Green SJ. Habitat suitability modeling to inform seascape connectivity conservation and management. Diversity (Basel) 2021;13:1–20. https://doi.org/10.3390/d13100465.
[53] Hair JF, Hult GTM, Ringle CM, Sarstedt M, Danks NP, Ray S. Evaluation of Formative Measurement Models. 2021. https://doi.org/10.1007/978-3-030-80519-7_5.
[54] Mahato R, Abujam S, Bushi D, Nimasow OD, Nimasow G, Das DN. Distribution modelling of Tor putitora (Hamilton, 1822), an endangered cyprinid in the Himalayan river system using MaxEnt. Acta Ecologica Sinica 2022. https://doi.org/10.1016/j.chnaes.2022.01.004.
[55] Singh AK, Srivastava SC, Verma P. MaxEnt distribution modeling for predicting Oreochromis niloticus invasion into the Ganga river system , India and conservation concern of native fish biodiversity. Aquat Ecosyst Health Manag 2021;24:43–51. https://doi.org/10.14321/aehm.024.02.08.
[56] Vorste VR, McElmurray P, Bell S, Eliason KM, Brown BL. Does stream size really explain biodiversity patterns in lotic systems? A call for mechanistic explanations. Diversity (Basel) 2017;9:1–21. https://doi.org/10.3390/d9030026.
[57] Zhang Z, Gao J, Cai Y. Effects of Land Use Characteristics, Physiochemical Variables, and River Connectivity on Fish Assemblages in a Lowland Basin. Sustainability (Switzerland) 2023;15. https://doi.org/10.3390/su152215960.
[58] Cheng ST, Herricks EE, Tsai WP, Chang FJ. Assessing the natural and anthropogenic influences on basin-wide fish species richness. Science of the Total Environment 2016;572:825–36. https://doi.org/10.1016/j.scitotenv.2016.07.120.
[59] Vorste VR, McElmurray P, Bell S, Eliason KM, Brown BL. Does stream size really explain biodiversity patterns in lotic systems? A call for mechanistic explanations. Diversity (Basel) 2017;9:1–21. https://doi.org/10.3390/d9030026.
[60] Soo CL, Nyanti L, Idris NE, Ling TY, Sim SF, Grinang J, et al. Fish biodiversity and assemblages along the altitudinal gradients of tropical mountainous forest streams. Sci Rep 2021;11:1–11. https://doi.org/10.1038/s41598-021-96253-3.
[61] Ghimire S, Koju NP. Short communication: Fish diversity and its relationship with environmental variables in kamala river, nepal. Biodiversitas 2021;22:4865–71. https://doi.org/10.13057/biodiv/d221119.
[62] Chen Z, Chen L, Wang Z, He D. Understanding the Effects of Climate Change on the Distributional Range of Plateau Fish : A Case Study of Species Endemic to the Hexi River System in the Qinghai – Tibetan Plateau. Diversity (Basel) 2022;14:1–19.
[63] Shields R, Pyron M, Arsenault ER, Thorp JH, Minder M, Artz C, et al. Geomorphology variables predict fish assemblages for forested and endorheic rivers of two continents. Ecol Evol 2021;11:16745–62. https://doi.org/https://doi.org/10.1002/ece3.8300.

[64] Campos JC, Garcia N, Alírio J, Arenas-Castro S, Teodoro AC, Sillero N. Ecological Niche Models using MaxEnt in Google Earth Engine: Evaluation, guidelines and recommendations. Ecol Inform 2023;76:102147. https://doi.org/10.1016/j.ecoinf.2023.102147.
[65] Johnson S, Molano-flores B, Zaya D. Field validation as a tool for mitigating uncertainty in species distribution modeling for conservation planning. Conserv Sci Pract 2023:1–16. https://doi.org/10.1111/csp2.12978.
[66] Tian L, Li Y, Zhang M. A variable selection method based on multicollinearity reduction for food origin traceability identification. Vib Spectrosc 2025;138.
[67] Charney ND, Record S, Gerstner BE, Merow C. A Test of Species Distribution Model Transferability Across Environmental and Geographic Space for 108 Western North American Tree Species. Front Ecol Evol 2021;9:1–16. https://doi.org/10.3389/fevo.2021.689295.
[68] Tian Z, Huo D, Yi K, Que J, Lu Z, Hou J. Evaluation of Suitable Habitats for Birds Based on MaxEnt and Google Earth Engine—A Case Study of Baer's Pochard (Aythya baeri) in Baiyangdian, China. Remote Sens (Basel) 2024;16. https://doi.org/10.3390/rs16010064.
[69] Rosenkrantz L. Leveraging geographic information systems (GIS) for environmental public health practice. Environ Health Rev Downloaded 2022;65:31–6. https://doi.org/10.5864/d2022-013.
[70] Koo H, Chun Y, Griffith DA. Integrating spatial data analysis functionalities in a GIS environment: Spatial Analysis using ArcGIS Engine and R (SAAR). Transactions in GIS 2018;22:721–36.
[71] Shan B, Huang W, Zhang M, Wang L. Integrating population genetics and species distribution models to predict red seabream distribution under climate change. Glob Ecol Conserv 2025;60:e03589. https://doi.org/10.1016/j.gecco.2025.e03589.
[72] Newman JC, Riddell EA, Williams LA, Sears MW, Barrett K, Kearney MR. Integrating physiology into correlative models can alter projections of habitat suitability under climate change for a threatened amphibian. Ecography 2022;2022:1–15. https://doi.org/10.1111/ecog.06082.
[73] Citores L, Ibaibarriaga L, Brewer MJ, Santos M, Chust G. Modelling species presence – absence in the ecological niche theory framework using shape-constrained generalized additive models. Ecol Modell 2020;418.
[74] Lo M, Reed J, Castello L, Steel EA, Frimpong EA. The Influence of Forests on Freshwater Fish in the Tropics : A Systematic Review. Bioscience 2020;70:404–14. https://doi.org/10.1093/biosci/biaa021.