International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 3, June 2024, pp. 2508~2521
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i3.pp2508-2521

A review on features and methods of potential fishing zone

Norsuzila Ya'acob[1,3], Nik Nur Shaadah Nik Dzulkefli[2], Mohd Azri Abdul Aziz[1], Azita Laily Yusof[1,3], Roslan Umar[4]

[1]School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Selangor, Malaysia
[2]School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Terengganu, Malaysia
[3]Wireless Communication Technology, School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Selangor, Malaysia
[4]Institut Penyelidikan Alam Sekitar Pantai Timur, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

Article Info

Article history:

Received Oct 5, 2023
Revised Feb 22, 2024
Accepted Feb 25, 2024

Keywords:

Classification methods
Features selection
Potential fishing zone
Sea surface chlorophyll a
Sea surface temperature
Support vector machine

ABSTRACT

This review focuses on the importance of identifying potential fishing zones in seawater for sustainable fishing practices. It explores features like sea surface temperature (SST) and sea surface height (SSH), along with classification methods such as classifiers. The features like SST, SSH, and different classifiers used to classify the data, have been figured out in this review study. This study underscores the importance of examining potential fishing zones using advanced analytical techniques. It thoroughly explores the methodologies employed by researchers, covering both past and current approaches. The examination centers on data characteristics and the application of classification algorithms for classification of potential fishing zones. Furthermore, the prediction of potential fishing zones relies significantly on the effectiveness of classification algorithms. Previous research has assessed the performance of models like support vector machines, naïve Bayes, and artificial neural networks (ANN). In the previous result, the results of support vector machine (SVM) were 97.6% more accurate than naive Bayes's 94.2% to classify test data for fisheries classification. By considering the recent works in this area, several recommendations for future works are presented to further improve the performance of the potential fishing zone models, which is important to the fisheries community.

Corresponding Author:

Nik Nur Shaadah Nik Dzulkefli
School of Electrical Engineering, College of Engineering Universiti Teknologi MARA
Cawangan Terengganu, Malaysia
Email: niknu5502@uitm.edu.my

1. INTRODUCTION

The contribution of fisheries production to human well-being is well known. Fish products are primarily used as food and a significant source of nutrition. Due to its high trading value, it also serves as an economic driver. Fisheries products play a very important role in ensuring human food security. Meanwhile, in terms of economics, fisheries products are among the most traded commodities, contributing to job creation, revenue generation, and regional economic growth and development [1]. Due to Malaysia's proximity to the sea and abundance of rivers and lakes, fisheries have also emerged as one of the country's most important economic sectors. These essential ecosystems provide natural fisheries and other aquatic resources for the people who live there. Despite the availability of numerous additional animal protein sources at competitive prices, the pattern has not changed much. Malaysians have a strong preference for fish, and it appears that there is no rival [2]. Malaysia ranked in the top 15 of world fisheries producers in

2018 with an estimated production of 1.45 million metric tons [1]. Fisheries resources, which are the backbone of the catch fisheries industry, are currently in jeopardy. Therefore, it is important to have a model for fisheries so that any fishing activities can be monitored and sustained. The suggested potential fishing zone (PFZ) model can make a significant contribution to understanding seasonal fishing activity and can help the fishing industry use its resources effectively, efficiently, and responsibly [3]. In recent years, researchers all over the world have become more aware of the need for modelling fisheries' productivity. Some research has been carried out to determine the accuracy of the potential fishing zone model. Each model utilizes a different set of environmental parameters that influence the potential fishing zone model, together with the application of geographic information system (GIS) and satellite remote sensing [4]. Satellite remote sensing is a useful tool for studying, managing, and harvesting fisheries because it lets us look at the oceans over time and compare different environmental factors that affect the amount and location of fisheries resources. It can also be used to do ecological studies at the community and ecosystem levels. The potential fishing zone prediction model is the major topic of this research proposal. There have been a few studies that have investigated this kind of research. Most of the researchers used sea surface temperature (SST) and chlorophyll-a (Chl-a) to determine the fishing zone, as in [5]–[11]. Since SST and Chl-a are correlated with fish distribution and abundance and other marine creatures, they are frequently used to identify possible fishing zones. The life and habitat of fisheries like skipjack tuna are significantly influenced by oceanographic factors, particularly those related to SST and chlorophyll-a content. SST values between 29 °C and 31 °C and chlorophyll-a concentrations between 0.15 and 0.35 mg/m^{-3} are ideal for skipjack tuna populations to congregate in. This is why determining the ideal fishing location requires an understanding of oceanic conditions and variations to ensure optimal fishing results [12]. The most important criteria in developing a prediction model are to determine the factors that will contribute to forecasting the potential fishing zone, since the factors will determine the accuracy and reliability of the model developed.

In this paper, the previous and present methodologies investigated by researchers were reviewed from a variety of angles, including the data, features, and methodologies used in determining the potential fishing zone. The main objectives of this review paper are to get insight into how to classify and determine the potential fishing zone by various features connected to the fisheries environment. Thus, fishing zone, model, remote sensing, and classifier were all the keywords that had been explored since all the keywords can be used to implement the new potential fishing zone model. This review paper will also include the techniques that previous researchers used to categorize the data for modeling the potential fishing zone. As a result, several suggestions are presented for the development and usage of the potential fishing zone in the future. The remainder of this paper is structured as follows. The features of the data for this research are discussed and reviewed in section 3. The various methods are covered in sections 4 and 5, and finally, the conclusion of this review paper is covered in section 7.

2. METHOD

The online search feature of a recognized publication database is a great tool for finding review structure. The review structure for modelling potential fishing zones is shown in Figure 1. These reviews were conducted to look for literature reviews, and most of the chosen journals and proceedings are from the IEEE Xplore database, which is indexed in Scopus and Science Direct. The research guide used a few keywords such as potential fishing zone, model, remote sensing, and classification method. All the review papers came from the six-year period that started in 2018 and ended in 2023. Most of the 2018 to 2023 journals and proceedings were picked. About 74 papers make up the total number of review papers. This research purposely wanted to answer a few questions on this research, such as the most influential features or data to determine fishing zones and techniques or methods used to determine fishing zones, whether using statistical methods or classification algorithms.

In the process of developing this review article, our primary objective is to delve into the practical applications of diverse data mining techniques. Specifically, we aim to explore the interplay between data features, feature selection methods, and machine learning algorithms employed in the classification of potential fishing zones within fisheries environments. While existing literature has touched upon these topics, a noteworthy observation is the limited use of datasets that have been widely applied by researchers. Notably, certain studies, such as those referenced in [5]–[7] have focused on determining potential fishing zones using a restricted set of data. In contrast, a study conducted in 2018 presented a unique approach by utilizing classification algorithms on underwater fish images. This innovative method while addressing challenges related to data volume, posed difficulties for researchers in acquiring a substantial amount of data [13]. Our contention is that the development of an accurate classification model hinges on the judicious application of preprocessing techniques, encompassing a well-defined set of data features, and the incorporation of an effective feature selection method. This review seeks to explore these aspects comprehensively, addressing the current landscape of research and identifying areas where advancements and optimizations can be made.

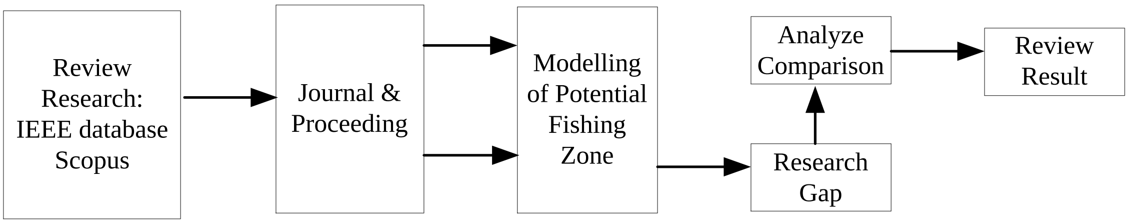

Figure 1. The review of modelling of potential fishing zone on features and method

3. FEATURES

The following is a review of the feature data that was used in the potential fishing zone study in several classifications of features. Researchers can find locations with high fish productivity and abundance and obtain a thorough grasp of the marine environment by examining these various feature classifications in the study of potential fishing zones. To maximize fishing operations while reducing detrimental effects on marine ecosystems, this knowledge is essential for fisheries management and conservation initiatives.

3.1. Oceanographic environmental based features

Oceanographic factors, which include water temperatures like the SST, sea vertical temperature distribution, salinity, and sea surface chlorophyll a (SSC), are major considerations in fisheries [14]. The SST and SSC are the most often referenced criteria for determining fishing locations [15]. Rupok *et al*. [16] also said that physical factors of water such as salinity, dissolved oxygen, and temperature also make up the basic components of the aquaculture system.

The physical factors of the water are very important to determine the potential fishing zone. Water temperature is a major factor in controlling the growth, physiology, and behavior of marine organisms. The SST is the most important oceanographic parameter influencing fisheries productivity. It directly affects the distribution of aquatic species and their populations, as well as their predator-prey interactions. Higher SSTs promote higher rates of photosynthesis by phytoplankton, which in turn increases the production of oxygen radicals and other toxic compounds. These toxins can damage fish tissue and hamper their growth. Too high of a temperature will kill the fish, while too low will not provide them with enough oxygen to survive. Physical factors also play a role in the migratory patterns of fish. For example, salmon migrate up rivers to spawn but will also travel downriver if the water temperature is too high or there is not enough oxygen available in the river.

Oceanographic factors, which include water temperatures like the SST, sea vertical temperature distribution, salinity, and SSC, are major considerations in fisheries. The SST and SSC are the most commonly used factors for determining fishing locations [15]. The primary production of phytoplankton, the base of the food chain for pelagic species like tuna, is strongly correlated with the SSC, a physical oceanographic measure that is frequently used to estimate fish abundance [17]. Oceanographic factors are used to forecast the potential fishing zones of SST and SSC [17]–[19]. When used to identify potential fishing areas, satellite images reveal the incredibly complex dynamics of marine waters, Tijani *et al*. [18] mapped the concentration of chlorophyll-a (Chl-a) and the temperature of the SST. The upwelling of cold, nutrient-rich waters (phytoplankton) creates these fronts. They are widespread in areas with a lot of pelagic fish and are identifiable by strong, antiparallel local gradients in SST and Chl-a. This technique significantly cuts down on search time, fuel use, and labor costs while simultaneously increasing catch per unit effort (CPUE). Sea surface temperature (SST), sea surface chlorophyll-a (SSC), sea surface salinity (SSS), mixed layer depth (MLD), sea surface height (SSH), and eddy kinetic energy (EKE) satellite-based oceanographic data were used to assess the effects of oceanographic conditions on the potential fishing zones for Albacore tuna and to investigate the spatial variability of these features in the South Indian Ocean (SIO) [20]. Most studies contend that SST and chlorophyll-a levels influence distribution patterns and tuna abundance variability. Lee *et al*. [20] said that the results of SSTs (18~21.5 °C) and relatively high chlorophyll-a concentrations (0.18~0.44 mg m^{-3}) were closest to the previous studies.

The diffuse attenuation coefficient is a widely used measure of turbidity in water. It denotes ocean turbidity and indicates turbid waters related to biological processes such as phytoplankton presence in the algal bloom zone by estimating light penetration in the sea [21]. Sea surface height determines the physical processes of oceans, such as upwelling. The phenomenon of upwelling pushes nutrient-rich marine life deeper. Phytoplankton absorbed the nutrients, thus increasing the amount of food in the sea. The frontal areas that are rich in nutrients are proven to be suitable fishing zones. Sea surface height anomaly (SSHa) is a characteristic that has been used to define potential fishing zones [22]. The factor will determine the heat stored in the ocean. Warm water is less dense than cold water. Consequently, higher locations are generally warmer than lower areas. Sea level is the level of the ocean when there is no sea level fluctuation due to tides, waves, surges, or other disturbances [21]. SST and SSC are also used in [23]–[25] to predict the hot spot of tuna fish. SSH is another important factor considered by researchers to determine the habitat of the fish. The study shows that the habitat of tuna species has positive results for SSH. While monsoon variability, which considers wind parameters such as wind speed and wind direction, affects ocean productivity activities and has various effects in different waters. Furthermore, variations in marine productivity have an impact on the fluctuation of fishery resources, making this parameter also important to consider [26]. Moreover, monitoring parameters like Chl-a concentration and other factors are very important in fisheries research, management, and harvesting because environmental factors influence the number and distribution of fish [27].

3.2. Image and video-based features

Besides environmental features, another kind of feature that can be used to predict potential fishing is by utilizing image and video features. One researcher used typical underwater video imagery to develop automated classification systems that can recognize fish underwater. Siddiqui *et al*. [13] used a state-of-the-art computer vision method based on deep learning techniques for fine-grained fish species classification and concluded that the accuracy for fish species from typical underwater video imagery captured off the coast of Western Australia after applying classification to test data performed by an support vector machine (SVM) on the features computed through the proposed method is about 94.3%. The study also suggested that developing automated classification systems that can recognize fish from underwater video images is a possible and cost-effective alternative to human identification. Another researcher, Konovalov *et al*. [28] also used underwater fish and non-fish images from 20 different habitats to develop a CNN-based fish detector. A method for efficiently labelling images was constructed to train a CNN-based binary image classifier (fish or nonfish) for fish recognition and localization. It is about 0.17% false positives and 0.61% false negatives on the project's 20,000 negative and 16,000 positive holdout test images, respectively. Almero *et al*. [29] said that fish detection via image and computer vision technology is thought to be a useful technique in fish monitoring to boost productivity and meet future worldwide demands. In underwater-based setups, object detection is a relatively recent area of research. Object detection is a computer vision task that identifies the type and location of an object in an image or video frame.

3.3. Biologically based features

Despite its importance in commerce and agriculture, fish recognition is now a highly challenging task. Distortion, noise, segmentation mistakes, overlap, and occlusion are some of the difficulties that can be encountered when trying to recognize fish accurately and reliably. Another researcher used a method based on the shape properties of fish that was divided into two subsets, the first of which consisted of 76 fish for training and the second of which consisted of 74 fish for testing. The body and five fin lengths were measured in centimeters: anal, caudal, dorsal, pelvic, and pectoral (cm) [30]. Fish recognition features typically involve the use of computer vision and image processing techniques to identify and categorize different fish species based on visual characteristics. Pitchai *et al*. [31] used datasets from the Fish-Pak website, which contains a number of images for distinguishing the two distinct fish species, Catla and Rohu, using the information on the head, body, and scales of each fish for fish classification. Not many studies use this kind of data since it requires the researchers to be present at the site, which can be hazardous, and the process takes too much time since this research requires big data. Engaging in direct capture for fish recognition data poses challenges in obtaining a large and diverse dataset, primarily due to factors such as the intricacies of underwater photography, the limited accessibility of certain species, and the time-intensive nature of fieldwork.

3.4. Vessel monitoring system

The vessel monitoring system (VMS) has provided more precise spatial and quantitative data about industrial fishing over the past ten years [32]. One of the technologies frequently utilized in several nations for monitoring and surveillance of fishing boats is the VMS. By giving precise information on fishing boat position and movements at any time, VMS can be utilized to effectively monitor and manage fishing activity. VMS technology makes use of satellite-based data transmission to provide information on the location of fishing boats automatically and in real-time. Additionally, data frequency can be changed to suit regulatory requirements. Fishing boats equipped with VMS transmitters will always be linked to the regulator's central monitor, allowing for continuous global monitoring [33]. Several researchers used the VMS as stated in [33] and [34]–[37]. The purpose of study [34] is to examine how fishing vessel activities can be tracked from space using trajectory data from the VMS. VMS data is used by study [37] to monitor illegal fishing activity by monitoring the density of light fishing fleets in the area of the fishing zone. These data from the visible infrared imaging radiometer suite for boat detection were examined, displayed, and compared with actual fishing data and environmental parameters of sea surface temperature and chlorophyll-a concentrations. Regular and frequent recording of vessel positions offers potential opportunities to enhance the estimation and management of fishing efforts but it is not made clear whether the ship is doing fishing activities or not [35]. The disadvantage of the VMS data is that all types of fishing boats operating in one country must be registered and have VMS technology installed so that the required data is always available. While the VMS data in study [36] was used to track fishing activity, whether fishing or non-fishing, and compare VMS speed data with the resultant speeds from logbooks.

Yugopuspito *et al*. [38] employed VMS data to determine the position of the tuna to lower labor and fuel costs in Indonesia by creating methods for predicting fish position utilizing a machine learning ensemble using support vector machine technology. It is concluded that the algorithm can anticipate the tuna's location with 97.6% accuracy [10]. Russo *et al*. [39], another researcher who has used vessels without monitoring equipment to predict annual fishing footprints using a model, namely a cascaded multilayer perceptron network (CMPN). The model exhibits strong predictive ability and permits the expansion of geographical assessments of fishing footprints to pertinent, though as-yet-undiscovered, fleet segments [39]. Torres-Irineo *et al*. [40] discover the role of environmental factors in the distribution of viable fishing grounds for this fleet using data from a pilot VMS, where a study was tested on a small fleet in the southeastern Gulf of Mexico (SGoM). For seven months, fishing vessels operating in four states provided tracking data for 1,608 daily trips. To locate possible fishing sites along the SGoM where this fleet operates, a researcher employed a correlative modelling strategy that took environmental factors into consideration. The outcomes showed that the principal factors influencing net primary output and sea surface temperature were the main factors that drive the spatiotemporal potential distribution of the study site's fishing grounds [40]. However, in looking for more information, Hery *et al*. [33] suggested that chlorophyll content, SST, and other environmental features should be considered in future developments in this particular research. It shows that environmental parameters are the most important factors that influenced the potential fishing zone and are the subject of the most potential research according to these features.

4. FEATURE SELECTION

Researchers incorporate a variety of techniques, including statistical analysis. Statistical analysis is one of the methods used by students, professors, workers, and users in all fields. Statistics are present in all disciplines of research involving collecting, managing, and sorting data; hence, some thought and perception are put into that occurrence, and from that information, will get new potential outcomes. The extraction of information from data to improve understanding of the situations it represents is one of the objectives of statistics. As a result, statistics can be viewed as the science of learning from data [41].

4.1. Habitat suitability index

To estimate habitat quality and species distributions, habitat suitability index (HSI) models are frequently employed. They are also used to develop biological surveys, evaluate reserve and management priorities, and foresee potential change under several scenarios for management or climate change [42]. Several researchers used the HSI to develop a model of potential fishing zones, as stated in studies [6], [25], [43]–[45]. Boitt and Aete [6] make use of surface temperature and chlorophyll-a remote sensing data by using a suitability index. They discovered that locations with sea surface temperatures between 23.0 to 28.3 ˚C and chlorophyll-a concentrations between 0.72 to 1.31 mg/m^3 accounted for more than 90% of the total catch. This study showed how a suitability index model, fish catch data, and satellite imagery components could be used to identify possible fishing zones for fish.

Hsu *et al*. [43] also used the HSI to determine the fishing grounds of skipjack tuna. The sea surface temperature and front, sea surface height, sea surface salinity, mixed layer depth, chlorophyll concentration, and finite-size Lyapunov exponents were among the environmental variables chosen to model the habitat. From that analysis, within 50 km of suitable habitat, about 94.9% of global data and 79.6% of Taiwanese data were detected. The findings demonstrated that the model did a good job of fitting both the actual fishing positions and daily prediction data. Zajac *et al*. [42] simulated hydrology input data from HSI models for two species of submerged aquatic vegetation (SAV) in Southwest Everglades National Park: Vallisneria Americana (tape grass) and Halodule wrightii, to demonstrate the GSA/UA architecture (shoal grass). Although distributions of HSI scores still permitted differentiation between sites with favorable versus poor circumstances, the researcher discovered significant regional variation in uncertainty for both species. When species distribution data are lacking, HSI models can be particularly helpful because they offer a way to utilize available environmental data to estimate past, present, and future habitat conditions [42].

4.2. Generalized additive model

Smooth nonlinear interactions between predictors and response variables are best described using generalized additive models (GAMs), which are gaining popularity. For modelling hierarchical functions for discrete responses that include intricate properties like zero-inflation, bounding, and uneven sampling, GAMs are particularly useful in ecology. However, because of their smooth functions, GAMs are less effective at making forecasts because they deliver unstable predictions outside the bounds of training data [46].

A few researchers such as in studies [5], [9], [11], and [47], [48] used the statistical model of the GAM to assess the influence of temperature change and Chl-a on fish distribution. Nurdin *et al*. [5] in their study set the goal of the study to find a connection between fish distribution and SST calculated by using the GAM. They used GAM to identify possible fishing zones and assess the influence of climate change on fish distribution using IPCC-AR5-RCP temperature projections. It is concluded that the preferred range of Chl-a at 0.30 to 0.40 mg/m^3 and SST at 30.00 to 31.00 °C were clearly connected (p 0.0001) with the distribution of Rastrelliger kanagurta. The areas with high potential catch were found near the coast to offshore (3–20 M), with an acceptable degree of map accuracy of 83.34% and a kappa value of 0.70, according to the potential fishing ground maps. An increase in temperature of 1.80 °C resulted in the relocation of potential fishing grounds to the southern portion of Makassar Straits, which leads to the archipelagic seas of Spermonde. Increased temperatures of 2.60 °C and 3.30 °C, on the other hand, led to a shift in the prospective fishing grounds area to the south. It was found that the findings of this study showed that remote sensing can help in determining the best fishing effort and making decisions for the long-term management of Rastrelliger kanagurta resources.

To investigate the relationship between yellowfin tuna, catch rates and oceanographic conditions using multispectral satellite images and to develop a habitat preference, Lan *et al*. [9] used GAMs fitted to two spatiotemporal fishery data sources, namely the 1° spatial grid and observer record longline fishery data from 2006 to 2010. According to the results, in the 1° spatial grid and observer record data, respectively, the cumulative deviances achieved using the chosen GAMs were 33.6% and 16.5%. Sea surface temperature was the environmental element in the study that contributed most to the deviation in the GAMs that were chosen. Zhang *et al.* [47] used 3 year (2000 to 2002) presence/absence data from squid fishing aggregations and environmental variables (nighttime sea surface temperature (SST), chlorophyll-a (Chl-a) concentration, Kd (490) (diffuse attenuation coefficients of downwelling irradiance at 490 nm), and bathymetry to develop a model to identify the presence and absence of Japanese common squid using statistical (generalized additive model (GAM) and generalized linear model (GLM)) and machine learning models (boosted regression tree (BRT)). It is found that the presence of squid was strongly influenced by SST and Chl-a concentration, but Kd (490), which is connected to water clarity, had a relatively less effect on the distribution of squid.

4.3. Statistical multiple regression model

Regression is a statistical technique used in engineering, business, finance, medicine, and other areas with the goal of determining the relationship between one dependent variable and several other independent variables [49]. Regression is a method to calculate the value of a single known variable as the dependent variable for a specific value or values of other variables or variables known as the independent variables while establishing a relationship between the two [50]. Many linear regression (MLR) techniques, which are part of the supervised learning group, can be used to make a model that shows how two or more interpretative factors (independent) and response variables (dependent) in observed data are related by using the right linear equations [51].

The statistical multiple regression model was used in [8] and [3] to analyze the effects of SST and Chl-a to forecast the distribution of fish. Nurdin *et al*. [8] set out to investigate the relation between SST and Chl-a of the marine environment's variability and productivity, which is crucial for exploring fishing resources. Monthly level 3 and daily level 1 images of SST and Chl-a derived by the moderate resolution imaging spectroradiometer satellite (MODIS) from July 2002 to June 2011 were used to investigate the relationship between SST and Chl-a and to forecast the potential fishing grounds of Rastrelliger kanagurta around the archipelagic waters of Spermonde, Indonesia. It is concluded that SST and Chl-a had a positive connection ($R = 0.3$, $p<0.05$), according to the findings. While SST and Chl-a were similarly shown to have a positive link with Rastrelliger kanagurta capture ($R=0.7$, $p<0.05$), Rastrelliger kanagurta's prospective fishing grounds were discovered along the coast (at an accuracy of 76.9%). It also found that by combining remote sensing technologies, statistical modelling, and GIS techniques, researchers were able to detect the association between SST and Chl-a, as well as predict Rastrelliger kanagurta aggregation. This could help with decision-making and reduce the time and cost of searching for fish during fishing activities.

A multi-linear regression model for mapping potential fishing zones (PFZ) was used in the research conducted by Daqamseh *et al*. [3] using parameters generated from MODIS satellite data, such as SSS, SST, and chlorophyll-a, where the study was conducted along the Saudi Arabian Red Sea beaches. With a high degree of consistency between mapped PFZ areas and fish capture data (R2=0.91), the proposed model shows great promise for use in the Red Sea region. It is recommended that, based on the findings of this study, the proposed PFZ model be utilized to assist fisheries in identifying high-potential fishing zones, permitting huge regions of the Red Sea to be used for a short period of time.

5. CLASSIFICATION ALGORITHMS

Researchers used a variety of techniques to determine the potential fishing zone. Most of the study used a statistical model to identify the most effective environmental parameters to identify potential fishing zones. But recently, machine learning (ML) approaches like SVM are effectively fitting with geospatial information science, remotely sensed data, and engineering fields. The details are explained below for both classifiers and statistical analysis methods. Scientific studies are increasingly focusing on the application of machine learning technologies to interpret and use ocean data. This strategy is vital for maintaining the ocean ecosystem, anticipating ocean elements, exploring the unknown, and responding to extreme weather. Most big data oceans are very complex, and the use of machine learning algorithms to examine the application of big data has become a hotly studied subject with the advancements in the field of machine learning and classifier technology. Researchers use various data collected in the water and integrate machine learning or classifier algorithms to conduct various experiments to better understand marine life [50]. There are two steps in the classification process. The model is developed during the learning phase, which comes before utilizing it to predict the class label in the next phases [52].

5.1. Support vector machine

One of the most significant and frequently applied problems in machine learning is classification, the goal of which is to develop a rule for classifying data into sets of pre-existing categories based on a collection of training sets. The SVM, which has been successfully used in many fields of science and engineering, is one of the most promising classification approaches in machine learning [53]. The classification and training model are based on the SVM algorithm. The method used for transforming the data and then searching for the best border between the outputs reduces the classification time and improves the accuracy of the system [54]. SVM can be used for regression and classification and can handle both continuous and categorical variables [55]–[57]. With its great accuracy and efficiency in identifying patterns in oceanic phenomena, the SVM algorithm has demonstrated itself to be a valuable tool for studying ocean data. Machine learning offers a wide variety of immediate applications in the fields of earth sciences and oceanography [56].

Image classification has gained a lot of interest lately to SVM in a variety of remote sensing. According to Sheykhmousa *et al*. [58], of the 251 relevant papers in the database that had been reviewed, 68% of the database applied SVM in the various studies regarding remote sensing respectively. Hery *et al*. [33] proposed the study to employ two machine learning techniques: naive Bayes and support vector machine. With an accuracy rate of 74.8%, the classifier analyzed 24,150 artificially tagged images featuring 15 common fish species in Taiwanese waters [59]. Siddiqui *et al*. [13] used SVM to classify test data and concluded that from typical underwater video imagery obtained off the coast of Western Australia, classification accuracy for fish species was 94.3%. Lukas *et al*. [33] concluded that the results of SVM are 97.6% more accurate than naive Bayes's 94.2%.

Fitrianah *et al*. [60] combined the ocean current and salinity parameters together with chlorophyll and SST parameters in exploring the features of the ocean to predict the potential tuna fishing zones using SVM. It is discovered that the combination of parameters said before yields the best classification performance as naïve Bayes reaches 69.03%, decision tree reaches 82.32%, and SVM reaches 68.30% of accuracy compared to the only two combinations of features of chlorophyll and SST when naïve Bayes reaches 57.44%, decision tree reaches 58.91%, and SVM reaches 56.74% of accuracy [60]. Andrews *et al*. [61] used SVM to classify the areas of the potential fishing zone at Kilakarai based on directions, where it had four directions: south, southeast, southwest, and east, and plot the support vectors to distinguish locations. The original data of Kilakarai from Indian National Centre for Ocean Information Services (INCOIS) is converted into fast Fourier transform (FFT) features so that it can be classified as a spot for potential fishing areas. Based on this research, the fishermen will be able to identify their safer borders according to the suggested accelerated SVM-based classification's clear identification of high-probability PFZ. This aids the fishermen in avoiding crossing international boundaries [61]. Ogunlana *et al* [30] also utilize the SVM-based technique to improve the classification of fish species. The method is based on the shape characteristics of fish, which were split into two subsets: the first set included 76 fish as training data, and the second set included 74 fish as testing data. The body was measured in centimeters along with the lengths of the anal, caudal, dorsal, pelvic, and pectoral fins (cm). Based on the results, the novel technique reveals a classification accuracy of 78.59%, which is much greater than that obtained by other techniques such as artificial neural networks (ANN), K-nearest neighbors (KNN), and K-mean clustering-based algorithms [30].

Mithoo *et al.* [62] proposed to apply support vector machines to images for classification, which apply generalized discriminant analysis (GDA) for parameter reduction, which requires less time and storage space [62]. Raj *et al.* [57] used historical data from various fish catchments to forecast future trends in fishing zones. To forecast the nature of these fish, the SVM technique is used. The system is put to the test using the supplied datasets, and the results demonstrate how effective the suggested system is [57]. Yugopuspito *et al.* [38] proposed a novel technique by using the VMS to determine the position of the tuna to minimize labor and fuel costs in Indonesia. The modelling from VMS data is built before comparing the results with weather data and earth conditions. Programming Python dash is used to create visualization techniques based on algorithms, and finally, support vector machine technology is utilized to build fish location prediction, where it is found that the system can forecast the tuna's location with about 97.6% accuracy [38]. After all, the SVM is the most efficient classifier out of all the tools available for data analysis, clustering, regression, and pattern identification [56].

5.2. Naïve Bayes

One method used in the classification process is the naive Bayes algorithm. The Bayesian network is a simple classifier but still plays an effective role when doing predictive modelling, as stated in [63], [64]. It is a probabilistic model that uses Bayes' theorem to calculate the probability of categories in each test set [63]–[65]. The Bayes principle is the foundation of the naïve Bayes method and classifies, using probability statistics knowledge, the test set of data. The Bayesian classification algorithm's error rate is low since it is due to its pure mathematical base. The Bayesian approaches are a combination of the prior and posterior probabilities. This does not rely on the preceding data, hence avoiding the subjective bias and staying away from the overfitting problem of probability when using information sampling data only [66].

Hery *et al.* [33] proposed the study to employ another machine learning technique, naive Bayes, besides support vector machines. Lukas *et al.* [33] concluded that the outcomes of this research are a website that will be used to determine the location of fishing tuna in Indonesian waters utilizing the naive Bayes and SVM-based on data VMS methods. It is also concluded that the results of SVM are 97.6% more accurate than naive Bayes (94.2%) in determining tuna, but naive Bayes is better in some areas. Zhou *et al.* [67] used eight possible options using SST and SSH data while comparing them to catch per unit area of effort, which represents the abundance of fish catches. The model is built by using the Bayes classifier model to predict and classify the fishing zone in the South China Sea [67]. Mustakim *et al.* [65] said that it is necessary to do research on the classification of aquatic species on Indonesian fishing boats to identify the impact that will take place on the fishing vessel. Thus, the classification technique in this study is carried out utilizing the nave Bayes classifier and algorithm for probabilistic neural networks (PNN). About 48% accuracy was obtained from the naive Bayes classifier algorithm using the RapidMiner tool [65].

5.3. Artificial neural network

Artificial neural networks often known as neural networks, are modern computing techniques for machine learning, knowledge demonstration, and, ultimately, the application of acquired information to maximize the output responses of complex systems [68]. Artificial neural networks are designed in the same way as the human brain, with neuron nodes interconnected in a web-like arrangement. The brain is made up of neuron cells, which number in the billions. Each neuron is made up of a cell body that transports information to and from the brain and processes it [69]. Such networks' concept is to process data and information to facilitate learning and knowledge creation. The creation of new structures for the information processing system is the main component of this concept [70]. Konovalov *et al.* [28] used 4,000 fish or no fish images from 20 different habitats. There are also 17,000 known negative (above-water general-domain (VOC2012)) pictures of missing) fish that were also used. An additional 27,000 positive/fish photos from above and below the water were provided from two publicly accessible fish-domain databases. In this paper, the researcher provides a labelling-efficient fish-detector training approach using CNN, resulting in an area under the ROC curve (AUC) of 99.94% [28]. Wang *et al.* [71] used an artificial neural network to explore PFZs of *O. Bartramii* from Chinese squid-jigging by using three important environmental variables: SST, SSH, and Chl-a concentration between 2003 and 2013. The well-developed back-propagation network model could predict the PFZ with an accuracy of 80% [71]. Nasir *et al.* [72] also used back-propagation of an artificial neural network to build a forecasting model of marine fish landings by using monthly marine landing data from East Johor and Pahang, which has 144 observations each. By using the root-mean-square error and mean absolute error values, the suggested model's performance is then compared to that of a traditional artificial neural network, showing that it can outperform the previous other [72]. Sivasankari *et al.* [73] realized the importance of having a better fishing community by developing a prediction model of the potential fishing zone using remote sensing images of chlorophyll, SST, and GPS location to predict the potential locations for fishing. Long short-term memory (FB-LSTM)-based recurrent neural networks and deep convolutional layers (RNN) are what the suggested architecture is made up of [73].

6. DISCUSSION

This section aims to provide a summary of prior research, with a particular focus on identifying notable knowledge gaps that warrant further investigation in the future. Table 1 offers an overview of the techniques employed in previous studies, their performance metrics, utilized datasets, and highlights any limitations or advantages found in each source. Through the systematic presentation of this information, we intend to bring to light discernible knowledge gaps within the existing body of literature, aligning with the central focus of our review paper. The insights gained from this summary table will guide the direction of our exploration, emphasizing areas where additional research can contribute meaningfully to the current understanding of the subject matter.

Table 1. Literature review summary of potential fishing zone

Author/Year	Performance	Technique	Dataset	Limitation
Nurdin *et al.* (2017) [5]	Chl-a at 0.30–0.40 mg/m^3 and SST at 30.00–31.00 °C were clearly connected ($p < 0.0001$) degree of map accuracy of 83.34% and a kappa value of 0.70	Generalized additive model (GAM)	Chlorophyll-a (Chl-a) and sea surface temperature (SST)	Limited data set
Boitt and Aete (2021) [6]	Sea surface temperature 23.0˚C - 28.3˚C and chlorophyll-a concentration 0.72 - 1.31 mg/m^3	Habitat suitability index	Surface temperature and chlorophyll-a levels	Limited data set
Mugo *et al.* (2020) [7]	AVI CVI CBI Mar 0.504 0.367 0.219 Apr 0.482 0.462 0.697 May 0.570 0.341 0.552 Jun 0.527 0.448 0.872 Jul 0.506 0.405 0.478 Aug 0.456 0.382 0.479 Sep 0.532 0.489 0.370 Oct 0.532 0.464 0.692 Nov 0.550 0.462 0.016	Ecological niche factor analysis (ENFA) models.	Sea surface temperature, chlorophyll-a, diffuse attenuation coefficient, sea surface heights and surface wind speed	Limited data set
Nurdin *et al.* (2015) [8]	$R=0.3$, $p<0.05$ $R=0.7$, $p<0.05$ at accuracy of 76.9 %	Statistical multiple regression model	SST and Chl-a	Limited data set
Siddiqui *et al.* (2018) [13]	Accuracy for fish species was 94.3%	Support vector machine (SVM)	Underwater videos	Hazardous to researches to get big amount of data
Lukas and Krisnadi (2021) [33]	Results of SVM is 97.6% more accurate than naive Bayes (94.2%)	Naive Bayes and support vector machine.	VMS data	Fishing boats operating in one country must be registered and have VMS technology installed
Huang *et al.* (2015) [59]	BEOTR method recognizes about 78% of the real, untrained valid fish images correctly.	Balance-enforced optimized tree with reject option (BEOTR)	Fish images	Hazardous to researches to get big amount of data
Lan, *et al.* (2017) [9]	($p < 0.01$) for SSHA, SST, and Chl-a	GAM	SST, Chl-a, Sea surface height anomaly (SSHa)	Limited data set
Yang, *et al.* (2020) [74]	Comparison of error between different models Grid search - 3.62 GA-SVM - 0.57 PSO-SVM - 0.4*	Grid search method particle swarm optimization (PSO-SVM) genetic algorithms (GA-SVM)	Catch fish data, SST	Limited data set
Fu *et al.* (2021) [10]	The results show that the proposed method is able to outperform the average fishers' ability by an average of 3%.	Faster R-CNN	sea water temperature and sea surface height (SSH)	Limited data set
Daqamseh *et al.* (2019) [3]	The optimal ranges of SST, SST front, SSH, SSS, and FSLE accounted for 60.00%, 62.66%, 51.26%, 57.40%, and 64.50% of total efforts, respectively	multi-linear regression model	Sea surface salinity (SSS), sea surface temperature (SST), and chlorophyll-a (Chl-a)	Limited data set
Mondal *et al.* (2021) [25]	AUC: BRT 0.72 GAM 0.71 GLM 0.64	Habitat suitability model (HSI)	sea surface ocean temperature (SST) and sea surface chlorophyll (SSC)	Limited data set
Putri *et al.* (2021) [48]	R^2 of MPN1 and MPN2 were 0.75 and 0.89, respectively.	GAM	SST and chlorophyll-a	Limited data set
Vayghan *et al.* (2020) [45]	$R^2 = 0.91$ 95.5%	AMM-HSI	SST and SSC concentration	Limited data set

Several studies, such as [5]-[8] have mostly used a small set of common environmental parameters, like sea surface temperature (SST) and chlorophyll-a, to find areas that might be good for fishing. Conversely, some researchers have ventured into more intricate and challenging datasets, notably utilizing hazardous data such as underwater fish images, as observed in [13], [33], [59], who advocates for the inclusion of additional environmental features, such as but not limited to chlorophyll content and SST, in future research endeavors, offers an important viewpoint in the search for additional insights. This suggestion underscores the significance of environmental parameters as pivotal factors influencing potential fishing zones. The emphasis on these features as potential subjects for extensive research highlights their critical role in shaping the understanding of the dynamics involved in determining optimal fishing areas.

The studies in [3], [5]-[9], [25], [48] used statistical approaches to predict the potential fishing zone. Nurdin *et al*. [5] also used GAM to identify possible fishing zones and assess the influence of climate change on fish distribution using IPCC-AR5-RCP temperature projections. It is found that a strong link (p 0.0001) between the preferred range of Chl-a at 0.30–0.40 mg/m^3 and SST at 30.00 °C to 31.00 °C and the distribution of Rastrelliger kanagurta. Areas with high potential catch were found near the coast to offshore (3–20 M), with an acceptable degree of map accuracy of 83.34% and a kappa value of 0.70, according to the potential fishing ground maps. SVM and naive Bayes are applied in [13] and [33]. The accuracy for fish species was 94.3%, and the results of SVM are 97.6% more accurate than naive Bayes (94.2%), respectively. They used underwater videos, which are a hazardous data set for researchers, and VMS data, whose fishing boats operating in one country must be registered and have VMS technology installed. Although the accuracy and performance of the technique used were good, the study did use a limited or hazardous data set, where to determine the potential fishing zone, researchers must have a large amount of data to make the classification more precise and effective.

This section highlights noticeable research gaps identified in the reviewed studies. These gaps underscore the opportunity for advancements, particularly in the integration of features of data, varied feature selection methods in conjunction with diverse classification algorithms. In section 5, an in-depth exploration of each classification algorithm was conducted, with their effectiveness in predicting potential fishing zones being evaluated. The performance of various classification algorithms was revealed, showcasing their respective strengths and limitations. Importantly, the incorporation of feature selection alongside these predictive models was found to be a critical factor, leading to enhanced outcomes in the uncovering of complex patterns related to factors influencing the determination of potential fishing zones. Throughout this section, it becomes evident that most of the studies solely utilized the common dataset, focusing on SST and chlorophyll-a. However, a notable observation is that some studies did not use other data set rather than common data set used before such as night sea surface temperature (NSST), wavelength of light (KD490), normalized fluorescence line height (nFLH), wind speed and rainfall. It is contended that the determination of potential fishing zones will be more accurate if the classification algorithms use a larger view of the dataset [33].

7. CONCLUSION

In this paper, a review of the modelling of the potential fishing zone is presented, focusing on the features of the data and the methods to classify the data and thus develop the model. The review addresses the numerous feature preferences used in accordance with the study and the variety of methods used in classifying the research's data. The papers that were keyword-related were gathered and reviewed in a digital library. This review paper only contains the most pertinent ones. Machine learning algorithms have become extensively employed across a variety of domains, with a notable emphasis on classification tasks. Despite the considerable insights revealed in numerous studies, there is ample room for additional exploration, encompassing diverse data types and preprocessing techniques. The choice of suitable algorithms frequently relies on factors such as the nature of the data, the duration of training, and the number of features. This research emphasizes its ongoing significance, particularly when contemplating the integration of new datasets such as NSST, wavelength of light (KD490), nFLH, wind speed, and rainfall, in conjunction with a variety of algorithms to get a precise model of a potential fishing zone. It was highlighted that the most used feature in the papers reviewed is environmental features. As mentioned earlier, SST and SSC are the most common environmental features used by researchers to determine the potential fishing zone. Apart from that, underwater video and imagery of fisheries were also used in previous research. Data acquisition for both types of data require the researchers to be at the site, which may pose a danger, and the process is very tedious. Thus, other environmental features should be considered in future development, in particular research, since the features are the most important factors that influenced the potential fishing zone and will be the most relevant research in this field. Numerous methods were reviewed that focused on locating the potential of a fishing zone. Different methods provide flexibility within the constraints of any particular application. As mentioned in the preceding section, the majority of studies have relied on public datasets, predominantly concentrating on parameters like SST and chlorophyll-a data. Hence, future research can incorporate additional variables, including NSST, wavelength of light (KD490), nFLH, wind speed, and rainfall. These variables, combined with previously employed common data, will contribute to a more comprehensive evaluation of potential fishing zones. Therefore, it is essential that exploration of potential fishing zones continues, and more innovative, low-cost, and effective ways are found to improve the future of modelling potential fishing zones.

ACKNOWLEDGEMENTS

The author would like to acknowledge the main sponsorship of this paper, Research Nexus UiTM which is under the incentive of "*Pembiayaan Yuran Penerbitan Artikel (PYPA)*".

REFERENCES

[1] FAO, "The State of World Fisheries and Aquaculture 2020," *Sustainability in action*, FAO, 2020.
[2] S. Fathi, A. N. Harun, S. Rambat, and N. A. Tukiran, "Current Issues in Aquaculture: Lessons from Malaysia," *Advanced Science Letters*, vol. 24, no. 1, pp. 503–505, Jan. 2018, doi: 10.1166/asl.2018.12051.
[3] S. T. Daqamseh, A. Al-Fugara, B. Pradhan, A. Al-Oraiqat, and M. Habib, "MODIS derived sea surface salinity, temperature, and chlorophyll-a data for potential fish zone mapping: West red sea coastal areas, Saudi Arabia," *Sensors (Switzerland)*, vol. 19, no. 9, p. 2069, May 2019, doi: 10.3390/s19092069.
[4] E. Chassot *et al.*, "Satellite remote sensing for an ecosystem approach to fisheries management," *ICES Journal of Marine Science*, vol. 68, no. 4, pp. 651–666, Jan. 2011, doi: 10.1093/icesjms/fsq195.
[5] S. Nurdin, M. A. Mustapha, T. Lihan, and M. Zainuddin, "Applicability of remote sensing oceanographic data in the detection of potential fishing grounds of Rastrelliger kanagurta in the archipelagic waters of Spermonde, Indonesia," *Fisheries Research*, vol. 196, pp. 1–12, Dec. 2017, doi: 10.1016/j.fishres.2017.07.029.
[6] M. K. Boitt and E. O. Aete, "Identification and Mapping of Essential Fish Habitats Using Remote Sensing and GIS on Lake Victoria, Kenya," *Journal of Geoscience and Environment Protection*, vol. 09, no. 10, pp. 91–109, 2021, doi: 10.4236/gep.2021.910007.
[7] R. Mugo *et al.*, "Identification of skipjack tuna (Katsuwonus pelamis) pelagic hotspots applying a satellite remote sensing-driven analysis of ecological niche factors: A short-term run," *PLoS ONE*, vol. 15, no. 8 August, p. e0237742, Aug. 2020, doi: 10.1371/journal.pone.0237742.
[8] S. Nurdin, M. A. Mustapha, T. Lihan, and M. A. Ghaffar, "Determination of potential fishing grounds of Rastrelliger kanagurta using satellite remote sensing and GIS technique," *Sains Malaysiana*, vol. 44, no. 2, pp. 225–232, Feb. 2015, doi: 10.17576/jsm-2015-4402-09.
[9] K. W. Lan, T. Shimada, M. A. Lee, N. J. Su, and Y. Chang, "Using remote-sensing environmental and fishery data to map potential yellowfin tuna habitats in the Tropical Pacific Ocean," *Remote Sensing*, vol. 9, no. 5, Art. no. 444, May 2017, doi: 10.3390/rs9050444.
[10] A. Fu, K. R. Patil, and M. Iiyama, "Region proposal and regression network for fishing spots detection from sea environment," *IEEE Access*, vol. 9, pp. 68366–68375, 2021, doi: 10.1109/ACCESS.2021.3077514.
[11] Nurholis, J. L. Gaol, A. F. Syah, and A. K. Dewi, "GIS-based spatiooral analysis on yellow fin tuna catch in Eastern Indian Ocean off Sumatera," *IOP Conference Series: Earth and Environmental Science*, vol. 429, no. 1, Art. no. 12041, Jan. 2020, doi: 10.1088/1755-1315/429/1/012041.
[12] S. N. Rifai, A. A. Ilham, and A. Achmad, "Temporal forecasting system of potential catching areas of skipjack tuna in bone sea using artificial neural network," in *Proceedings of the International Conference on Electrical Engineering and Informatics*, Sep. 2022, vol. 2022-Septe, pp. 71–77, doi: 10.1109/ICELTICs56128.2022.9932066.
[13] S. A. Siddiqui *et al.*, "Automatic fish species classification in underwater videos: Exploiting pre-trained deep neural network models to compensate for limited labelled data," *ICES Journal of Marine Science*, vol. 75, no. 1, pp. 374–389, Jul. 2018, doi: 10.1093/icesjms/fsx109.
[14] A. M. P. Santos, "Fisheries oceanography using satellite and airborne remote sensing methods: a review," *Fisheries Research*, vol. 49, no. 1, pp. 1–20, Nov. 2000, doi: 10.1016/S0165-7836(00)00201-0.
[15] P. Falco, A. Belardinelli, A. Santojanni, N. Cingolani, A. Russo, and E. Arneri, "An observing system for the collection of fishery and oceanographic data," *Ocean Science*, vol. 3, no. 2, pp. 189–203, May 2007, doi: 10.5194/os-3-189-2007.
[16] Q. S. S. Rupok, K. H. Suman, M. N. Sakib, and J. Agarwal, "Sensor data fusion for monitoring water quality toward sustainable freshwater fisheries," *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, Bangladesh, 2020, pp. 1-6, doi: 10.1109/ETCCE51779.2020.9350876.
[17] D. Fitrianah, A. N. Hidayanto, J. L. Gaol, H. Fahmi, and A. M. Arymurthy, "A spatio-temporal data-mining approach for identification of potential fishing zones based on oceanographic characteristics in the Eastern Indian Ocean," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 8, pp. 3720–3728, Aug. 2016, doi: 10.1109/JSTARS.2015.2492982.
[18] K. Tijani, M. T. Chiaradia, A. Morea, R. Nutricato, L. Guerriero, and G. Pasquariello, "Fishing forecasting system in Adriatic Sea - A model approach based on a normalized scalar product of the SST gradient and CHL gradient vectors," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2015, vol. 2015-Novem, pp. 2257–2260, doi: 10.1109/IGARSS.2015.7326256.
[19] K. Tijani, A. Morea, M. T. Chiaradia, R. Nutricato, and L. Guerriero, "Prototype of a multi-platform remote sensing service for fishing forecasting," Jun. 2016, doi: 10.1109/EESMS.2016.7504846.
[20] M. A. Lee, W. C. Yang, I. C. Hung, and S. Y. Teng, "Predicting winter potential fishing zones of Albacore tuna (Thunnus alalunga) using maximum entropy models and remotely sensed data in the South Indian Ocean," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2018, vol. 2018-July, pp. 957–959, doi: 10.1109/IGARSS.2018.8517359.
[21] Y. Xu, M. Lin, Q. Zheng, X. Ye, J. Li, and B. Zhu, "A study of long-term sea level variability in the East China Sea," *Acta Oceanologica Sinica*, vol. 34, no. 11, pp. 109–117, Nov. 2015, doi: 10.1007/s13131-015-0754-0.

[22] K. Nimit, M. N. Kumar, N. Swetha, J. Nayak, and T. S. Kumar, "Utility of sea surface height anomaly (SSHa) in determination of potential fishing zones," Technical Report, Indian National Centre for Ocean Information Services (INCOIS), 2015. Accessed: Jan. 08, 2024. [Online]. Available: https://incois.gov.in/documents/TechnicalReports/ESSO-INCOIS-ASG-TR-02-2015.pdf

[23] M. D. Setiawati, H. A. Rachman, A. R. As-Syakur, and A. Syahailatua, "Identification of commercial tuna hotspot in the Southern Waters of Java-Bali through satellite remote sensing data," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2021, vol. 2021-July, pp. 7434–7437, doi: 10.1109/IGARSS47720.2021.9553523.

[24] T. Y. Hsu and Y. Chang, "Modeling the habitat suitability index of skipjack tuna (katsuwonus pelamis) in the Western and Central Pacific Ocean," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2017, vol. 2017-July, pp. 2950–2953, doi: 10.1109/IGARSS.2017.8127617.

[25] S. Mondal, A. H. Vayghan, M. A. Lee, Y. C. Wang, and B. Semedi, "Habitat suitability modeling for the feeding ground of immature albacore in the southern indian ocean using satellite-derived sea surface temperature and chlorophyll data," *Remote Sensing*, vol. 13, no. 14, Art. no. 2669, Jul. 2021, doi: 10.3390/rs13142669.

[26] Kunarso, M. Zainuri, R. Ario, B. Munandar, and H. Prayogi, "Impact of monsoon to aquatic productivity and fish landing at Pesawaran Regency waters," *IOP Conference Series: Earth and Environmental Science*, vol. 116, no. 1, Art. no. 12021, Feb. 2018, doi: 10.1088/1755-1315/116/1/012021.

[27] N. Hassan and C. S. Woo, "Machine learning application in water quality using satellite data," *IOP Conference Series: Earth and Environmental Science*, vol. 842, no. 1, p. 12018, Aug. 2021, doi: 10.1088/1755-1315/842/1/012018.

[28] D. A. Konovalov, A. Saleh, M. Bradley, M. Sankupellay, S. Marini, and M. Sheaves, "Underwater fish detection with weak multi-domain supervision," in *Proceedings of the International Joint Conference on Neural Networks*, Jul. 2019, vol. 2019-July, doi: 10.1109/IJCNN.2019.8851907.

[29] V. J. D. Almero, R. S. Concepcion, E. Sybingco, and E. P. Dadios, "An image classifier for underwater fish detection using classification tree-artificial neural network hybrid," *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, Ho Chi Minh City, Vietnam, 2020, pp. 1-6, doi: 10.1109/RIVF48685.2020.9140795.

[30] S. O. Ogunlana, O. Olabode, S. A. Oluwadare, and G. B. Iwasokun, "Fish classification using support vector machine," *African Journal of Computing & ICT*, vol. 8, no. 2, pp. 75-82, 2015.

[31] P. Pitchai, T. Rajasekaran, V. K. Kaliappan, R. Sivaramakrishnan, and T. Thigarajan, "Fish species classification using SVM kernels," *Innovations in Information and Communication Technology Series*, pp. 466–470, Dec. 2020, doi: 10.46532/978-81-950008-1-4_102.

[32] A. N. Tassetti, A. Galdelli, J. Pulcinella, A. Mancini, and L. Bolognini, "Addressing gaps in small-scale fisheries: a low-cost tracking system," *Sensors*, vol. 22, no. 3, p. 839, Jan. 2022, doi: 10.3390/s22030839.

[33] Hery, S. Lukas, P. Yugopuspito, I. M. Murwantara, and D. Krisnadi, "Website design for locating tuna fishing spot using naïve Bayes and SVM based on VMS data on Indonesian Sea," in *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020*, Dec. 2020, pp. 89–93, doi: 10.1109/ISRITI51436.2020.9315338.

[34] M. I. Marzuki, P. Gaspar, R. Garello, V. Kerbaol, and R. Fablet, "Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories," *IEEE Journal of Oceanic Engineering*, vol. 43, no. 3, pp. 689–699, Jul. 2018, doi: 10.1109/JOE.2017.2723278.

[35] R. Joo, S. Bertrand, A. Chaigneau, and M. Ñiquen, "Optimization of an artificial neural network for identifying fishing set positions from VMS data: An example from the Peruvian anchovy purse seine fishery," *Ecological Modelling*, vol. 222, no. 4, pp. 1048–1059, Feb. 2011, doi: 10.1016/j.ecolmodel.2010.08.039.

[36] Y. Wang, Y. Wang, and J. Zheng, "Analyses of trawling track and fishing activity based on the data of vessel monitoring system (VMS): A case study of the single otter trawl vessels in the Zhoushan fishing ground," *Journal of Ocean University of China*, vol. 14, no. 1, pp. 89–96, Oct. 2015, doi: 10.1007/s11802-015-2467-6.

[37] Nurholis, J. Lumban-Gaol and F. S. Achmad, "Light fishing fleets monitoring by GIS-based spatiotemporal analysis in West Sumatera waters," *Pertanika Journal of Science and Technology*, vol. 28, pp. 327–351, Jan. 2020.

[38] P. Yugopuspito, S. Lukas, I. M. Murwantara, S. Albert, and Hery, "A novel fishing finder technique based on VMS data in Indonesia," in *ACM International Conference Proceeding Series*, Oct. 2019, pp. 107–111, doi: 10.1145/3369114.3369161.

[39] T. Russo, S. Franceschini, L. D'Andrea, M. Scardi, A. Parisi, and S. Cataudella, "Predicting fishing footprint of trawlers from environmental and fleet data: an application of artificial neural networks," *Frontiers in Marine Science*, vol. 6, Nov. 2019, doi: 10.3389/fmars.2019.00670.

[40] E. Torres-Irineo *et al.*, "Spatio-temporal determination of small-scale vessels' fishing grounds using a vessel monitoring system in the Southeastern Gulf of Mexico," *Frontiers in Marine Science*, vol. 8, May 2021, doi: 10.3389/fmars.2021.643318.

[41] R. P. Sarmento and V. Costa, "An Overview of Statistical Data Analysis," *arXiv:1908.07390*, Aug. 2019.

[42] Z. Zajac, B. Stith, A. C. Bowling, C. A. Langtimm, and E. D. Swain, "Evaluation of habitat suitability index models by global sensitivity and uncertainty analyses: A case study for submerged aquatic vegetation," *Ecology and Evolution*, vol. 5, no. 13, pp. 2503–2517, Jun. 2015, doi: 10.1002/ece3.1520.

[43] T. Y. Hsu, Y. Chang, M. A. Lee, R. F. Wu, and S. C. Hsiao, "Predicting skipjack tuna fishing grounds in the Western and Central Pacific Ocean based on high-spatial– temporal-resolution satellite data," *Remote Sensing*, vol. 13, no. 5, pp. 1–17, Feb. 2021, doi: 10.3390/rs13050861.

[44] M. A. Lee, J. S. Weng, K. W. Lan, A. H. Vayghan, Y. C. Wang, and J. W. Chan, "Empirical habitat suitability model for immature albacore tuna in the North Pacific Ocean obtained using multisatellite remote sensing data," *International Journal of Remote Sensing*, vol. 41, no. 15, pp. 5819–5837, Sep. 2020, doi: 10.1080/01431161.2019.1666317.

[45] A. H. Vayghan, M. A. Lee, J. S. Weng, S. Mondal, C. Te Lin, and Y. C. Wang, "Multisatellite-based feeding habitat suitability modeling of albacore tuna in the Southern Atlantic Ocean," *Remote Sensing*, vol. 12, no. 16, Art. no. 2515, Aug. 2020, doi: 10.3390/RS12162515.

[46] N. J. Clark and K. Wells, "Dynamic generalised additive models (DGAM) for forecasting discrete ecological time series," *bioRxiv*, 2022, doi: 10.1101/2022.02.22.481550.

[47] X. Zhang, S. I. Saitoh, and T. Hirawake, "Predicting potential fishing zones of Japanese common squid (*Todarodes pacificus*) using remotely sensed images in coastal waters of South-Western Hokkaido, Japan," *International Journal of Remote Sensing*, vol. 38, no. 21, pp. 6129–6146, Dec. 2017, doi: 10.1080/01431161.2016.1266114.

[48] R. S. Putri *et al.*, "Analysis of small pelagic fishing grounds using a generalized additive model in the Makassar Strait," *IOP Conference Series: Earth and Environmental Science*, vol. 763, no. 1, Art. no. 12042, May 2021, doi: 10.1088/1755-1315/763/1/012042.

[49] M. A. Iqbal, "Application of regression techniques with their advantages and disadvantages," *Elektron Mag*, vol. 4, pp. 11–17, Sep. 2021.

[50] S. Dasgupta, A. Baral, and A. Lahiri, "Estimation of electric stress in a vacuum interrupter using multiple linear regression," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 29, no. 3, pp. 859–865, Jun. 2022, doi: 10.1109/TDEI.2022.3173521.
[51] Y. Kittichotsatsawat, N. Tippayawong, and K. Y. Tippayawong, "Prediction of arabica coffee production using artificial neural network and multiple linear regression techniques," *Scientific Reports*, vol. 12, no. 1, Aug. 2022, doi: 10.1038/s41598-022-18635-5.
[52] A. Singh and R. Sathyaraj, "A comparison between classification algorithms on different datasets methodologies using RapidMiner," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 5, 2016, doi: 10.17148/IJARCCE.2016.55140.
[53] M. H. Almaspoor, A. Safaei, A. Salajegheh, B. Minaei-Bidgoli, and A. A. Safaei, "Support vector machines in big data classification: a systematic literature review support vector machines in big data classification: a systematic literature review," *Preprint, Research Square*, Aug. 2021, doi: 10.21203/rs.3.rs-663359/v1.
[54] L. K. Tolentino *et al.*, "Fish freshness determination through support vector machine," *Journal of Telecommunication*, vol. 9, pp. 139–143, Aug. 2017.
[55] D. M. C. Nieto, E. A. P. Quiroz, and M. A. Cano Lengua, "A systematic literature review on support vector machines applied to regression," *2021 IEEE Sciences and Humanities International Research Conference (SHIRCON)*, Lima, Peru, 2021, pp. 1-4, doi: 10.1109/SHIRCON53068.2021.9652268.
[56] H. Ahmad, "Applications of support vector machine learning in oceanography," *Preprint*, 2020. doi: 10.13140/RG.2.2.22561.68961.
[57] R. J. R. Raj, T. Sasipraba, M. Vasudev, S. Gupta, M. Rizwan, and P. Srivastava, "Predicting the impact of climate change on tidal zone fishes using SVM Approach," *Procedia Computer Science*, vol. 92, pp. 237–243, 2016, doi: 10.1016/j.procs.2016.07.351.
[58] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308–6325, 2020, doi: 10.1109/JSTARS.2020.3026724.
[59] P. X. Huang, B. J. Boom, and R. B. Fisher, "Hierarchical classification with reject option for live fish recognition," *Machine Vision and Applications*, vol. 26, no. 1, pp. 89–102, Sep. 2015, doi: 10.1007/s00138-014-0641-2.
[60] D. Fitrianah, "Feature exploration for prediction of potential tuna fishing zones," *International Journal of Information and Electronics Engineering*, 2015, doi: 10.7763/ijiee.2015.v5.543.
[61] M. R. Swaminathan, S. Andrews, P. Subash, and B. S. Varunbabu, "Finding the high probabilistic potential fishing zone by accelerated SVM classification," *International Journal of Information and Communication Technology*, vol. 11, no. 4, 2017, doi: 10.1504/ijict.2017.10008318.
[62] P. Mithoo and P. Balhara, "Fishing ground prediction using support vector machine," *International journal of engineering technology science and research*, 2017.
[63] S. Bhatia and J. Malhotra, "Naïve Bayes classifier for predicting the novel coronavirus," in *Proceedings of the 3rd International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, ICICV 2021*, Feb. 2021, pp. 880–883, doi: 10.1109/ICICV50876.2021.9388410.
[64] M. Mustakim, A. Hidayat, Z. Efendi, A. Aszani, R. Novita, and E. T. Lestari, "Algorithm comparison of naive bayes classifier and probabilistic neural network for water area classification of fishing vessel in Indonesia," *Journal of Theoretical and Applied Information Technology*, vol. 96, pp. 4114–4125, Jul. 2018.
[65] A. Kelly and M. A. Johnson, "Investigating the statistical assumptions of naïve Bayes classifiers," *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, 2021, pp. 1-6, doi: 10.1109/CISS50987.2021.9400215.
[66] C. Li, L. Chen, S. Wu, Y. Mo, and L. Chen, "Application research of naive Bayes algorithm based on DIKW in weather website," in *2021 IEEE 23rd International Conference on High Performance Computing and Communications, 7th International Conference on Data Science and Systems, 19th International Conference on Smart City and 7th International Conference on Dependability in Sensor, Cl*, Dec. 2022, pp. 2253–2257, doi: 10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00337.
[67] W. F. Zhou, A. Z. Li, S. J. Ji, and Y. S. Qiu, "Yellowfin Tuna (Thunnusalbacares) fishing ground forecasting model based on bayes classifier in the South China Sea," *Polish Maritime Research*, vol. 24, no. s2, pp. 140–146, Aug. 2017, doi: 10.1515/pomr-2017-0076.
[68] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: a tutorial," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019, doi: 10.1109/COMST.2019.2926625.
[69] M. van Gerven and S. Bohte, "Editorial: artificial neural networks as models of neural information processing," *Frontiers in Computational Neuroscience*, vol. 11, Dec. 2017, doi: 10.3389/fncom.2017.00114.
[70] R. Dastres and M. Soori, "Artificial neural network systems," *International Journal of Imaging and Robotics*, vol. 21, no. 2, pp. 13–25, 2021.
[71] J. Wang, W. Yu, X. Chen, L. Lei, and Y. Chen, "Detection of potential fishing zones for neon flying squid based on remote-sensing data in the Northwest Pacific Ocean using an artificial neural network," *International Journal of Remote Sensing*, vol. 36, no. 13, pp. 3317–3330, Jul. 2015, doi: 10.1080/01431161.2015.1042121.
[72] N. Nasir, R. Samsudin, and A. Shabri, "Forecasting of monthly marine fish landings using artificial neural network," *International Journal of Advances in Soft Computing and its Applications*, vol. 9, pp. 75–89, Jan. 2017.
[73] M. Sivasankari, R. Anandan, and F. A. Chamato, "HE-DFNETS: a novel hybrid deep learning architecture for the prediction of potential fishing zone areas in Indian Ocean using remote sensing images," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–10, Jun. 2022, doi: 10.1155/2022/5081541.
[74] S. Yang, Y. Dai, W. Fan, and H. Shi, "Standardizing catch per unit effort by machine learning techniques in longline fisheries: A case study of bigeye tuna in the Atlantic Ocean," *Ocean and Coastal Research*, vol. 68, pp. 83–93, 2020, doi: 10.1590/S2675-28242020068226.

BIOGRAPHIES OF AUTHORS

Norsuzila Ya'acob is an associate professor at the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UiTM). In 2010, she was awarded a Ph.D. degree in electrical, electronics and systems engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia. She received her MSc degree from the University of Putra Malaysia (UPM) in remote sensing and geographic information systems in 2000. She also obtained her B.Eng. degree from the University of Putra Malaysia (UPM), Malaysia in electronics and computer engineering in 1999. Since joining UiTM, she has been actively involved in research as a principal investigator for national research grants and has been awarded 40 research grants funded by the government and university. Currently, she is supervising and co-supervising 5 Masters and 8 PhD students. She has published over 200 journal papers and conference proceedings with 14 H-index and 574 citations. She can be contacted at email: norsuzila@uitm.edu.my.

Nik Nur Shaadah Nik Dzulkefli received the B.Eng. degree in electrical engineering from Universiti Teknologi MARA, in 2009. She received her M.Eng. from University Technology MARA (UiTM) in 2012 and continue works as full-time lecturer at Universiti Teknologi MARA (UiTM), Terengganu. Currently, she is a senior lecturer at the School of Electrical Engineering, Universiti Teknologi MARA. Currently she does part-time PhD at Universiti Teknologi MARA (UiTM) Shah Alam and do research focused in wireless communication technology and remote sensing. She also assists numerous student final year project based on embedded system and IoT application. Her research interests include remote sensing and satellite systems. She can be contacted at email: niknu5502@uitm.edu.my.

Mohd Azri Abdul Aziz is a senior lecturer at the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UiTM), Malaysia. He received the B.Eng. (Hons) degree in electronic engineering from University of Manchester Institute of Science and Technology (UMIST), UK. M.Sc. degree in electrical engineering and Ph.D. degree in electrical engineering, both from Universiti Teknologi MARA, Malaysia. He is also a fellow at Innovative Electromobility Research Lab (ITEM) UiTM, primarily researching on autonomous vehicles. His research area includes pattern classification, object detection, and IoT-based application. Currently, he is supervising and co-supervising 5 master's and 6 PhD students. He has authored and co-authored 8 journals and 16 conference proceedings with 5 H-index and 73 citations. He can be contacted at email: azriaziz@uitm.edu.my.

Azita Laily Yusof is an associate professor in the discipline of communications, School of Electrical Engineering, Universiti Teknologi MARA (UiTM). In 2011, she was awarded a Ph.D. degree in electrical, electronic and systems engineering from Universiti Kebangsaan Malaysia (UKM). She also obtained her M. Eng. degree from Universiti Kebangsaan Malaysia in communications and computer engineering in 2000. Previously, she obtained her first degree from Universiti Kebangsaan Malaysia with Honours in electrical, electronic and systems engineering in 1999. She is also an active researcher and has been awarded 17 research grants funded by the government and university. She has published over 70 journal papers and conference proceedings on various topics related to wireless communications technologies. She is a group member of the Wireless Communication Technology Group (WiCoT) at UiTM. Her current research interests include radio resource and interference management in mobile communications networks. She can be contacted at email: azita968@uitm.edu.my.

Roslan Umar earns his PhD from Universiti Malaya (UM), Malaysia in the field of radio astronomy. His interest in exploring universe begin from his bachelor's degree internship program when he attached to Malaysia National Planetarium. After finished MSc, he began his academic career at Universiti Sultan Zainal Abidin (UniSZA) and was appointed as Head of Astronomy Unit at UNISZA Observatory. He is holding the position of Director of East Coast Environmental Research Institute (ESERI) and lecturer at Faculty of Islamic Contemporary Studies teaching science subject (astronomy, cosmology, and research method). His research interest is in the field of electromagnetic, radio frequency interference, radio astronomy and astrophysics, geographical information system (GIS), Islamic astronomy and environmental study. Currently, he is supervising and co-supervising 7 masters and 7 PhD students and graduated 3 PhD and 6 masters students. He has authored and co-authored 83 journals and 12 conference proceedings with 12 H-index and 852 citations. His email: roslan@unisza.edu.my.