Research Article

Water-body Segmentation in Heterogeneous Hydrodynamic and Morphodynamic Structured Coastal Areas by Machine Learning

Irem Gumuscu[1] , **Furkan Altas[1]** , **Beril Turkekul[1]** , **Hasan Alper Kaya[1]** , **Firat Erdem[2,*]** , **Tolga Bakirman[3]** , **Bulent Bayram[3]**

[1] Coastal and Harbor Laboratory, Department of Civil Engineering, Yildiz Technical University, Istanbul, TURKIYE
[2] Institute of Earth and Space Sciences, Eskisehir Technical University, Eskisehir, TURKIYE
[3] Department of Geoinformatics, Yildiz Technical University, Istanbul, TURKIYE

Abstract

Coastal areas constitute the most important part of the world when considered in terms of their socio-economic and natural values. Measuring and monitoring the coastal areas accurately is an important issue for coastal management. Compared to ground-based studies, remote sensing applications enriched with machine learning algorithms such as Random Forest (RF) and Support Vector Machine (SVM) provide significant benefits in terms of cost, time, and size of the study area. Within the scope of this study, Sentinel-2 images for five coastal areas located in Turkey with different morphological and hydrodynamic properties were classified as land and water-bodies using SVM and RF algorithms. Water-body segmentation results of the SVM and RF classification for the different band combinations of Sentinel-2 images have been compared. The reasons affecting the results of the accuracy analysis were examined in accordance with the geography of each area. The accuracy analysis (Accuracy, IoU, F1 Score, Precision, Recall) results for all study areas, which consider different classification methods and band combinations are in the order of 99%. Among the band combinations, the best results were obtained in combinations of BRN for İzmir Aliağa and Hatay Samandağ regions, BGRN for Sakarya Karasu region, GRN for Rize Iyidere region and BGN for Samsun Bafra region. In addition, experimental results show that the utilized machine learning methods provide satisfactory results for combinations involving the NIR band in all study areas.

Keywords: Sentinel-2, shoreline extraction, Support Vector Machines, Random Forest, water-body classification

Introduction

The interface or transition zones between land and sea are generally referred to as coastal areas (Gazioglu et al., 1997; Burak et al., 2004; Cao and Wong, 2007). Coastal areas constitute the most important part of the world in terms of their socio-economic and natural values. More than half of the world's population lives in coastal areas (Haslett, 2008). Coastal areas are always popular settlement areas for many reasons such as transportation, industrial activities, agricultural purposes, tourism, and human settlement (Small and Nicholls, 2003). In addition to all these human-based activities on coastal regions, there are also natural impacts affecting coastal zones such as global warming, coastal erosion, sea-level rise, and sedimentation (Patil and Deo, 2020). Since they include a substantial amount of economic activities of the world and attractive areas for human migration, coastal management and planning are essential (Neumann, et al., 2015). Minimizing adverse effects on human and ecosystem health is difficult for natural resource managers and environmental decision-makers (Mahboubi, et al., 2015). Therefore, sustainable coastal zone monitoring and management have vital importance to protect the environment and biodiversity processes. Remote sensing provides spatially explicit information of large areas for this purpose (Osborne, et al., 2012; Zhang, et al., 2016; Demir, et al., 2019).

Various approaches have been proposed for shoreline extraction and water-body segmentation. Although traditional image segmentation methods have been utilized, these methods are use-case specific and cannot be generalized for all conditions. For example, water indices like Normalized Difference Water Index (NDWI) (McFeeters, 1996) can provide unsatisfactory results, especially in shallow water areas. Additionally, pixel-based supervised (Maglione, et al., 2014) and unsupervised (Sekovski, et al., 2014) algorithms can produce noisy results. To overcome these limitations, object-based methods can be used (Zhang, et al., 2013; Coiman, 2020; Dervisoglu, et al., 2020; Celik and Gazioglu, 2022). However, these methods are parameter-dependent and unsuitable parameters can produce poor results.

Recently, machine learning methods are widely used to avoid the mentioned problems and to create suitable solutions for water-body segmentation (Dixon and Candade, 2008). Various studies have proposed shoreline extraction using machine learning from satellite imagery. Demir et al., 2017 used the Random Forest (RF) method as part of an integrated shoreline extraction approach. Choung and Jo, 2017 compared the Support Vector Machines (SVM) algorithm and NDWI for shoreline segmentation performance. Balázs, et al., 2018 used SVM

and RF classifiers to extract water areas, and their studies showed that bands 4–5–7 of Landsat imagery are the most informative in this analysis. Paul et al., 2018 compared the SVM, Artificial Neural Network, K-Nearest Neighbor, discriminant analysis, and RF algorithms for water-body extraction from LISS-III, AWIFS, and Landsat data. Ai et al., 2019 examined the coastline change of the Pearl River Estuary using the SVM algorithm. Bangira et al., 2019 compared the thresholding methods and five machine learning algorithms (decision tree, K-Nearest Neighbor, RF, and two implementations of the SVM) for shoreline extraction in a complex water area. Minghelli, et al., 2020 extracted shoreline from WorldView-2 Satellite Data in the presence of foam using SVM. Kumar et al., 2020 mapped shoreline change using Artificial Neural Network, K-Nearest Neighbor, and SVM algorithms. Bamdadinejad, et al., 2021 used SVM to extract shorelines from Landsat 5 and Landsat 8 images. Similarly, Ngowo et al., 2021 used SVM to analyze shoreline change from 28-year Landsat 5 and Landsat 8 data. Although there are various studies focusing on shoreline extraction from satellite imagery, there is still a significant need for the assessment of the different types of shoreline segmentation methods (Toure et al., 2019). In this context, the main contributions of this study can be summarized as:

- To the best of our knowledge, there is no study to date analyzing the performance of machine learning methods for water-body segmentation in coastal areas with different hydrodynamic and morphological properties.
- Water-body segmentation was analyzed for coastal areas with three different seawater characteristics such as salinity, temperature, and density.
- Depending on the different precipitation rates, sediment dynamics, and the river regime, the study areas differing in terms of hydrology were selected.
- Sentinel-2 images for five coastal areas with different morphological and hydrodynamic properties are processed SVM and RF algorithms. Water-body segmentation results of the SVM and RF classification for the different band combinations of Sentinel-2 images are compared.
- The reasons affecting the results of the accuracy analysis of the study areas are examined by the geography of each field.

Study Area

In this study, Izmir-Aliaga, Hatay-Samandag, Rize-Iyidere, Samsun-Bafra, and Sakarya-Karasu coastal areas which are located in Turkey were selected as the study areas. Each study area has different hydrodynamics and morphological structures. The locations of the study areas are given in Figure-1. General characteristics of study areas have been given in Table-1.

Table-1. General characteristics of study areas. Temperature, tide, precipitation amount, and flood number data were provided from the Turkish State Meteorological Service (2021). Salinity data were provided from Çiftçi (2011).

| Study area | Features of the seaside of the shoreline | | Features of the land side of the shoreline | | | | Hydrodynamic effects | | Hydrological effects | |
	Salinity (PSU)	Temp (C) (Last 10 Years)	Region	Coastal Type	Coastal use	Flora	Wave effect	Tide effect (m)	Flood (quantity) (1940-2010)	Precipitation (mm/annual) (1927-2020)
Izmir Aliaga	~ 38	19.1	Aegean	Atlantic type	Harbour and port usage, commerce, tourism	Garrigue and maquis	Noneffective	< 0.5	~ 87	~ 710.5
Hatay Samandag	~ 39	22.2	Eastern Mediterranean	Pacific type	Beach, Settlement, Cultivated area	Maquis	Noneffective	~ 0.5	~ 45	~ 1163.5
Rize Iyidere	~ 18	15.2	Eastern Blacksea	Pacific type	Highway, Settlement, Protection structures	Forest	Effective	< 0.5	~ 98	~ 2301.5
Samsun Bafra	~ 18	15.2	Middle Blacksea	Pacific type	Dams, wetlands, groins, port, recreation, delta	Forest, Shrubbery	Effective	< 0.5	~ 38	~ 716.7
Sakarya Karasu	~ 18	15.2	Western Blacksea	Pacific type	The settlement, Commercial port	Forest	Effective	< 0.5	~ 18	~ 842.7
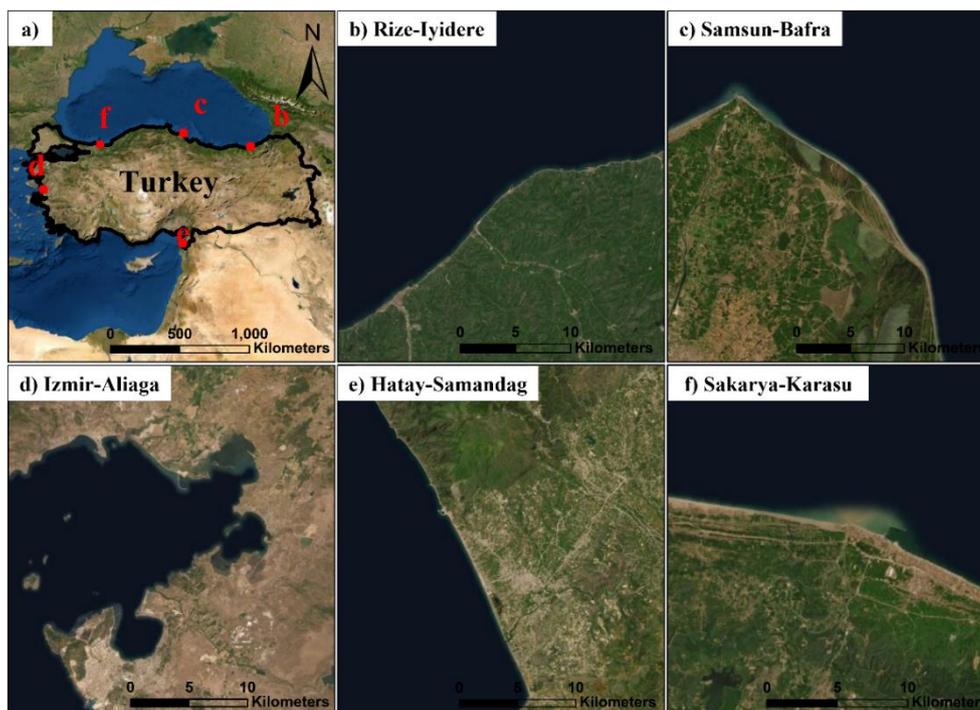
Fig. -1. Study Areas. a) Location of study areas. b) Rize-Iyidere. c) Samsun-Bafra. d) Izmir-Aliaga. e) Hatay-Samandag. f) Sakarya-Karasu.

Izmir-Aliaga

As a result of increasing economic activities such as industry, trade, tourism, and fishing, Aliaga and Candarlı coastal areas were severely affected. Major threats in the region include large port/pier constructions, coastal filling works, and unplanned housing construction.

With the filling of the sea during the gaining of industrial area and the construction of the port and pier, space was gained from the sea and the shoreline moved towards the sea (Guney and Polat, 2015). Dams built on streams prevent the deltas' sediment flow and also the removal of sand in delta areas causes coastal erosion.

Hatay-Samandag

The flood/delta plain formed by the Asi river, which is integrated with the coastal line, at the point where it flows into the sea is in a triangular shape. The Asi River filled the sea section here with the sediments it carried, and the coastline progressed regularly towards offshore. This coastal area is composed entirely of fine-medium and slightly coarse sands with no pebbles on the beach.

Although there is a balance between the river and marine processes in the development of the coastline and delta, today the human influence on natural processes has increased. Dams built on rivers, drying of wetlands, river bed structural arrangements, taking sand from the coastal area, pollution of the river and the shore negatively affect the natural development of the Asi delta (Öner, 2008).

Rize-Iyidere

Rize is located on the Eastern Black Sea which is the most mountainous part of the Black Sea Region. Due to the topographic structure and meteorological characteristics of the region, frequent floods and landslides occur due to

irregular construction. The Eastern Black Sea is the region that has the highest precipitation rate in Turkey (Turkish State Meteorological Service, 2021). The shoreline is not too indented and protruding. Therefore, strong waves caused the formation of cliffs on the shores.

The coastal zones of the region were extended to the sea to complete the international Black Sea Highway. Morphology of the coastal zones is affected directly or indirectly by highway, dam, and hydroelectric power plant constructions. Due to structures in the coastal zones and taking sediments from coasts, major changes have occurred in the sediment transportation regime (Berkün, et al. 2010).

Samsun-Bafra

Human-based activities are dense in the region such as agricultural purposes, energy generation via dams, drinking water supply, and irrigation (Samsun Investment Support Office, 2018). The study area also includes a large delta river and basin named Kizilirmak which is one of the largest and important wetlands having 73 km of coastline and also protected by the Ramsar Convention in Turkey (Ozturk et al. 2015).

Sakarya-Karasu

Karasu coast is in the Black Sea coastal zone of Turkey where the Sakarya province meets the Black sea. The coastline stretches out approximately 25 km each west and east side of the mouth of Sakarya River. With its tributaries, the Sakarya River has a length of 824 km and drains water and sediment to the Black Sea from the Sakarya Basin, which is one of the biggest basins in Turkey with a 56500 km2 watershed (TMMOB 2012). Three dams built on the river since 1965 caused a major decrease in the amount of water and sediment load carried by the river (Işık et al. 2006). Decreases in the sediment

load of the river lead to 7.5 m coastline erosion per year (Sahin 2020).

Materials and Methods
Materials

In this study, Sentinel-2 satellite images with a spatial resolution of 10 m have been used. Cloud-free images have been used which belong to different time periods. Satellite images were cropped to 2000 pixels x 2000 pixels (20 km x 20 km) for each study area. The same number of training polygons was used for all study areas. Blue (B), green (G), red (R), and near-infrared (N) bands of Sentinel-2 images were used. RGB, RGB-NIR, RG-NIR, RB-NIR, BR-NIR band combinations were selected for water-body segmentation. Since the study does not focus on change detection, the atmospheric correction is not applied, and the tide effect is not taken into account.

Machine Learning Methods
In this study, SVM and RF machine learning algorithms are used to extract water bodies from Sentinel-2 imagery.

Support Vector Machines (SVM)
SVM is a supervised non-parametric statistical learning method. SVM separates two classes with a hyperplane and assigns a class to samples from one of the two possible labels. The simplest form of SVMs is linear binary classifiers that predict the class of the given test sample from one of the two possible labels (Mountrakis et al., 2011).

The multi-spectral feature data are linearly separable in the input space according to the implementation of a linear SVM (Mountrakis et al., 2011). A kernel function is used to translate nonlinearly separable data into a higher dimensional space where it is presumed to be linearly separable (Jamil and Bayram, 2018). In order to be a valid kernel in SVMs, a kernel function typically needs to fulfill Mercer's Theorem in order to be a suitable kernel in SVMs (Scholkopf, et al., 2001). Some problems involve the identification of more than two classes. Using methods such as one-against-all, one-against-others, and directed acyclic graphs (Knerr et al., 1990) SVMs can be used in multi-class classification problems.

Random Forest (RF)
RF classification algorithm is a machine learning method based on decision trees (Classification and Regression Trees - CARTs). Decision trees predict the class of the test data according to the rules which were extracted from the training. These rules consist of many if-then conditions (Breiman, 2001). A subset of training samples is drawn using replacement to generate the trees (a bagging approach). As a result, the same sample can be chosen multiple times, while others may not be chosen at all (Gislason, et al., 2006). The Random Forest algorithm has two parameters. These are the number of decision trees to be generated (Ntree) and the number of variables to be selected and tested for the best split when growing the trees (Mtry) (Belgiu and Drăguţ, 2016).

The CART algorithm uses the GINI index to determine the best choice (Gislason et al., 2006). The GINI index measures the homogeneity of samples in each node. Each decision tree predicts the class membership of the test data therefore the test data is evaluated by all decision trees. Finally, the final membership class of the test data is selected by majority voting.

Results and Discussion

For each machine learning method, the training data with the equal quantity and same geometry properties were selected in each study area. All analyzes were carried out using the open-source software QGIS (3.10). Training and testing of machine learning algorithms were carried out in Orfeo Toolbox (7.1) integrated with QGIS software. For the training of machine learning algorithms, 100 polygons including 20200 pixels were used as training data. 60 of them were collected for the land class, 40 of them for the water class. The kernel type, gamma, and penalty parameters for SVM were Radial Basis Function, 0.250 and 100, respectively. The number of trees and random variables were set as 50 and 2 for the RF method. Parameters of SVM and RF algorithms were chosen empirically. Manual digitization was performed for obtaining ground truth data which was used as reference data in accuracy assessment. For evaluating the performance of water-body segmentation Precision, Recall, F1 score, IoU, and Accuracy metrics (Nowozin, 2014) have been used.
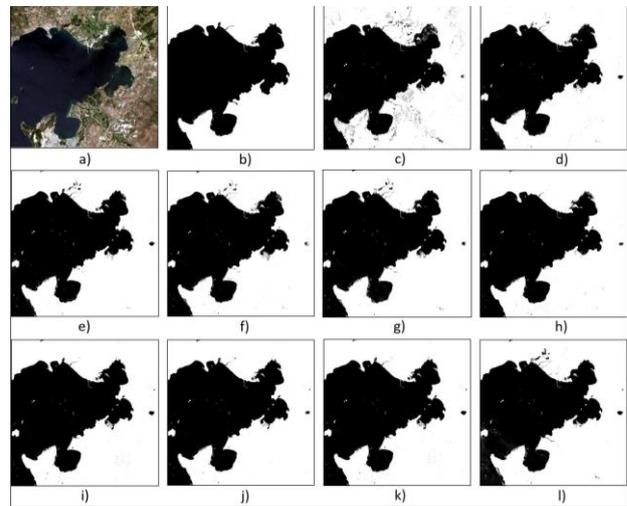

Fig. 2. Segmentation results for the Izmir-Aliaga region. a) Sentinel-2 Image, b) Manuel-digitization, c) BGR/RF, d) BGRN/RF, e) BGN/RF, f) BRN/RF, g) GRN/RF, h) BGR/SVM, i) BGRN/SVM, j) BGN/SVM, k) BRN/SVM, l) GRN/SVM

Izmir-Aliaga
The segmentation results for the Izmir-Aliaga study area are given in Figure-2. Similar results were obtained with all band combinations except BGR/RF. Noisy results were obtained in band combinations without the NIR band. In the accuracy analysis, BRN/SVM gave the highest result. The next best result was BGN/SVM. In this way, it can be said that the SVM method is more successful, although it is close to each other for this study

area. As it can be seen in Figure-3, there are marsh areas that have difficulty even manually interpretation of water or land areas. But there are still significant differences in comparing the best and worst results. Except for these marsh areas, higher segmentation accuracies have been obtained by SVM and RF methods as it can be seen in the Table-2. Both of the methods had no difficulty in classifying the breakwater and docking structure areas. In Table-2, accuracy analysis of Izmir-Aliaga results is

given. Accuracy, IoU, F1 Score, and Recall analyzes gave consistent results against each other. In all metrics except for recall, BRN combinations provided the highest success with the SVM method. According to the results of this analysis, BGR/RF tests gave the lowest accuracy. The SVM method has yielded more successful results than RF in this area, some of which are swamps, with an indented physical structure.
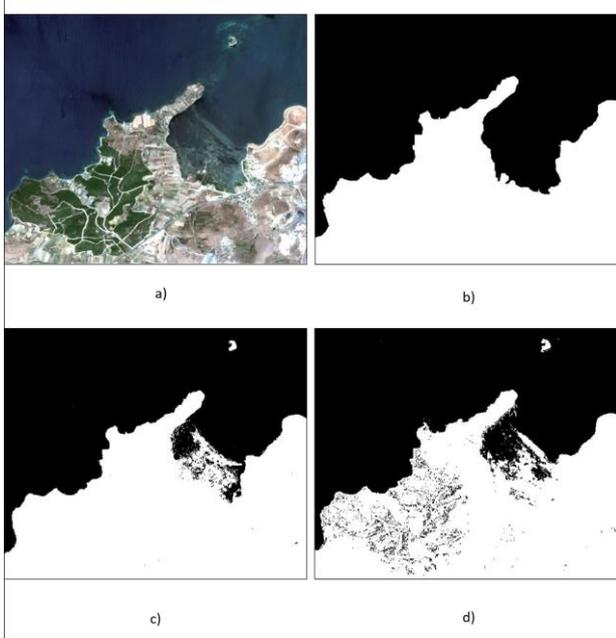

Fig. 3. Segmentation results for the marsh areas of the Izmir-Aliaga region. a) Sentinel-2 Image, b) manual-digitization, c) BRN/SVM (best result), d) BGR/RF (worst result).

Band Combination/ Method	Accuracy (%)	IoU (%)	F1 Score (%)	Precision (%)	Recall (%)
BGR/RF	97,638	95,918	97,916	97,810	98,023
BGR/SVM	98,515	97,439	98,703	97,632	**99,798**
BGRN/RF	98,687	97,725	98,849	98,076	99,635
BGRN/SVM	99,117	98,459	99,224	98,779	99,673
BGN/RF	98,734	97,806	98,891	98,150	99,644
BGN/SVM	99,122	98,470	99,229	98,676	99,788
BRN/RF	98,749	97,830	98,903	98,178	99,640
BRN/SVM	**99,135**	**98,490**	**99,239**	**98,797**	99,685
GRN/RF	98,427	97,289	98,626	97,624	99,648
GRN/SVM	98,530	97,457	98,712	97,976	99,459

Table-2. Accuracy analysis results for the Izmir-Aliaga region (best results are bold).

Hatay-Samandag

The segmentation results for the Hatay-Samandag region are given in Figure-4. According to the analysis results, the best approach was provided with the BRN band combination and RF method for Accuracy, IoU, F1 Score, and Recall metrics. The second-best highest accuracy was obtained with the BGN band combination and RF method. Noisy results were obtained with BGR band combinations for both methods. Extracting the river water with sediment supply is pretty hard in this band combination.

Large accuracy differences are observed especially in the river and estuarine areas between the most successful (BRN/SVM) and the least successful (BGR/SVM) band combination. At the estuarine and the point where it pours into the sea, the suspended solid material carried by the river has been identified as a land area with BGR bands combination (Figure-5). The accuracy assessment results of RF and SVM methods are given in Table-3. The BRN band combination with the RF method provided the highest accuracy values for all metrics except precision. The combination of the BGRN bands and the SVM method provides the highest precision. BGN combination with the RF method gave the second-highest accuracy for all metrics. BGR band combination with the SVM method gave the lowest accuracy. In this study area, the

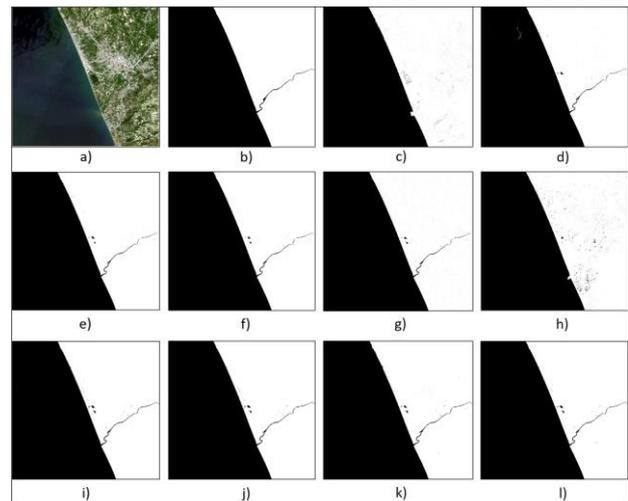performance of the RF method was better than the SVM method.

Fig. 4. Segmentation results for the Hatay-Samandag region. a) Sentinel-2 Image, b) Manuel-digitization, c) BGR RF, d) BGRN RF, e) BGN RF, f) BRN RF, g) GRN RF, h) BGR SVM, i) BGRN SVM, j) BGN SVM, k) BRN SVM, l) GRN SVM

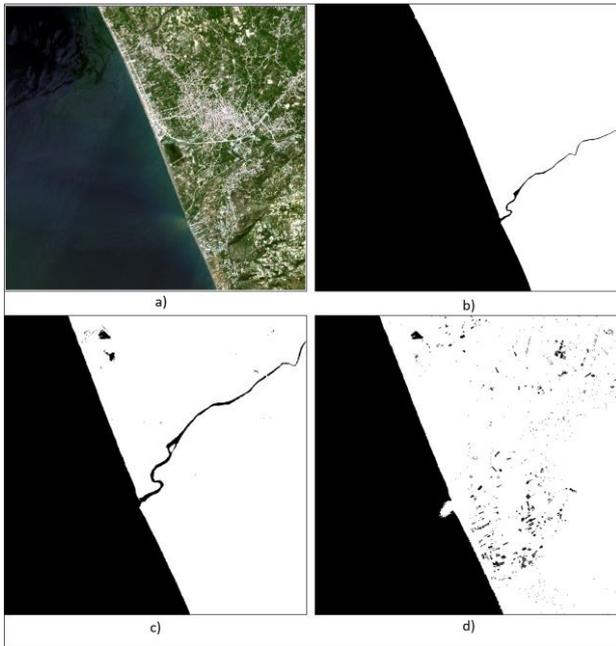| Band Combination/ Method	Accuracy (%)	IoU (%)	F1 Score (%)	Precision (%)	Recall (%)
BGR/RF	99,429	98,838	99,416	99,231	99,601
BGR/SVM	99,296	98,566	99,278	99,333	99,223
BGRN/RF	99,834	99,660	99,830	99,862	99,798
BGRN/SVM	99,819	99,629	99,814	**99,950**	99,678
BGN/RF	99,863	99,720	99,860	99,915	99,805
BGN/SVM	99,843	99,678	99,839	99,945	99,733
BRN/RF	**99,865**	**99,724**	**99,862**	99,917	**99,807**
BRN/SVM	99,743	99,473	99,736	99,942	99,531
GRN/RF	99,791	99,573	99,786	99,914	99,658
GRN/SVM	99,847	99,686	99,843	99,935	99,751

Fig. 5. Segmentation results for the river and estuarine area of the Hatay-Samandag region. a) Sentinel-2 Image, b) Manuel-digitization, c) BRN/RF (best result), d) BGR/SVM (worst result).

Table-3. Accuracy analysis results for the Hatay-Samandag region (best results are bold).

Rize-Iyidere

The segmentation results for the Rize - Iyidere study area have been given in Figure-6. It has been seen that reasonable results were obtained from all band combinations except BGR RF and BGR/SVM. Noisy results were obtained from these combinations.

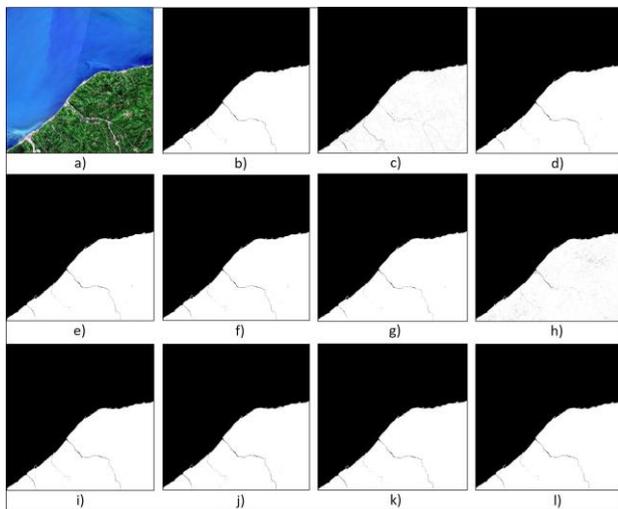high concentration of sediments in the area can cause miss-segmentation.

Fig. 6. The segmentation results for the Rize-Iyidere study area. a) Sentinel-2 Image, b) Manuel-digitization, c) BGR/RF, d) BGRN/RF, e) BGN/RF, f) BRN/RF, g) GRN/RF, h) BGR/SVM, i) BGRN/SVM, j) BGN/SVM, k) BRN/SVM, l) GRN/SVM

In all band combinations, despite the high segmentation accuracies of most parts of the coastal area, classification mistakes were observed in the section where T-head groins were located (Figure-7). In addition to T-head groins, it has been observed that the rivers which have a
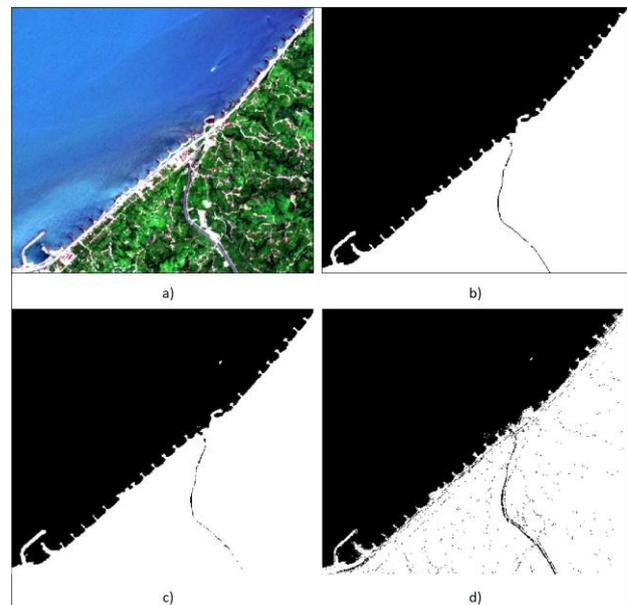

Fig. 7. The segmentation results for the section where T-head groins were located in the Rize-Iyidere study area. a) Sentinel-2 Image, b) Manuel-digitization, c) GRN/RF (best result), d) BGR/RF (worse result)

The accuracy analysis results for the Rize-Iyidere study area have been given in Table-4. When results of the accuracy analysis are examined, although all of the band combinations have given high accuracy, the highest accuracies of Accuracy, IoU, F1 Score, and Recall analyzes were obtained from the GRN/RF band combination. On the other hand, the highest accuracy of Precision analysis was obtained from the BGN SVM band

combination. Generally, the RF classifier gave higher accuracies than the SVM classifier in all band combinations except the BGR band combination. In addition, the lowest accuracies of all analyzes were obtained from the BGR RF combination. As a result, the GRN RF is the most suitable band combination for coastal areas such as Iyidere where floods happen frequently and there are many protection structures in there.

Table-4. Accuracy analysis results for the Rize-Iyidere region (best results are bold).

| Band Combination/ Method	Accuracy (%)	IoU (%)	F1 Score (%)	Precision (%)	Recall (%)
BGR/RF	99,371	98,374	99,180	99,662	98,703
BGR/SVM	99,523	98,765	99,379	99,771	98,989
BGRN/RF	99,935	99,831	99,915	99,939	99,892
BGRN/SVM	99,928	99,814	99,907	99,945	99,869
BGN/RF	99,935	99,831	99,915	99,939	99,892
BGN/SVM	99,914	99,777	99,888	**99,952**	99,824
BRN/RF	99,932	99,824	99,912	99,947	99,877
BRN/SVM	99,919	99,790	99,895	99,886	99,904
GRN/RF	**99,942**	**99,851**	**99,925**	99,913	**99,938**
GRN/SVM	99,929	99,816	99,908	99,914	99,903
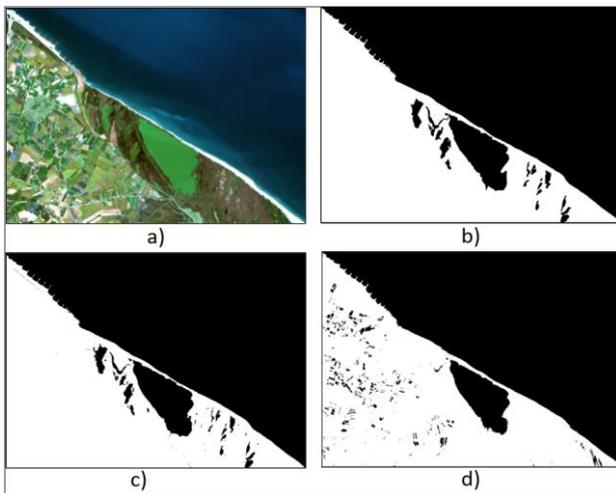
Fig. 9. Segmentation results for the Samsun-Bafra region. a) Sentinel-2 Image, b) manual-digitization, c) BGN/SVM (best result), d) BGR/SVM (worst result).

Sakarya-Karasu

Segmentation results of Sakarya-Karasu have been given in Figure-10. It was observed that combinations that included band-8 (NIR) yielded higher accurate results than combinations that did not involve NIR. Although the

Samsun-Bafra

Segmentation results of Samsun-Bafra have been given in Figure-8. The combinations with NIR band like BGN/SVM performed superior against combinations without NIR bands such as BGR/RF and BGR/SVM.
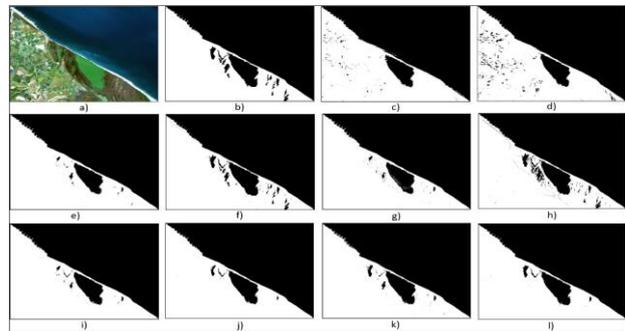

Fig. 8. The segmentation results for the Samsun-Bafra study area. a) Sentinel-2 Image, b) Manuel-digitization, c) BGR/RF, d) BGRN/RF, e) BGN/RF, f) BRN/RF, g) GRN/RF, h) BGR/SVM, i) BGRN/SVM, j) BGN/SVM, k) BRN/SVM, l) GRN/SVM

The region has complex hydrodynamic and morphodynamic aspects as mentioned in the previous sections. Figure-9 shows that the NIR band performed well in such a region but other combinations mostly failed to extract small water bodies like BGR/SVM. The accuracy analysis results for the Samsun-Bafra region are given in Table-5. The GRN band combination processed using the RF method performed the highest accuracy results for this region. In contrast, the BGR band combination processed with the SVM algorithm performed the worst in all band combinations.

Band Combination/ Method	Accuracy (%)	IoU (%)	F1 Score (%)	Precision (%)	Recall (%)
BGR/RF	97,456	94,363	97,100	95,218	99,057
BGR/SVM	96,837	92,936	96,339	95,909	96,772
BGRN/RF	98,834	97,352	98,658	97,634	99,704
BGRN/SVM	98,875	97,445	98,706	97,650	99,785
BGN/RF	98,360	96,326	98,129	96,351	**99,973**
BGN/SVM	**99,414**	**98,646**	**99,318**	**99,388**	99,248
BRN/RF	98,155	95,879	97,896	96,058	99,805
BRN/SVM	98,798	97,224	98,593	99,282	97,912
GRN/RF	98,517	96,660	98,302	96,814	99,836
GRN/SVM	98,593	96,827	98,388	96,941	99,879

Table-5. Accuracy analysis results for Samsun-Bafra region (best results are bold).

same training areas are used in all studies, several band combinations have produced water areas in urban areas where indeed there is no water area in those regions. Figure-11c and Figure-11d show that using the SVM algorithm, a combination of B, G, and R bands yields the

result of water area for highways, gardens, and even buildings, whereas including the NIR band in this combination produces the best result
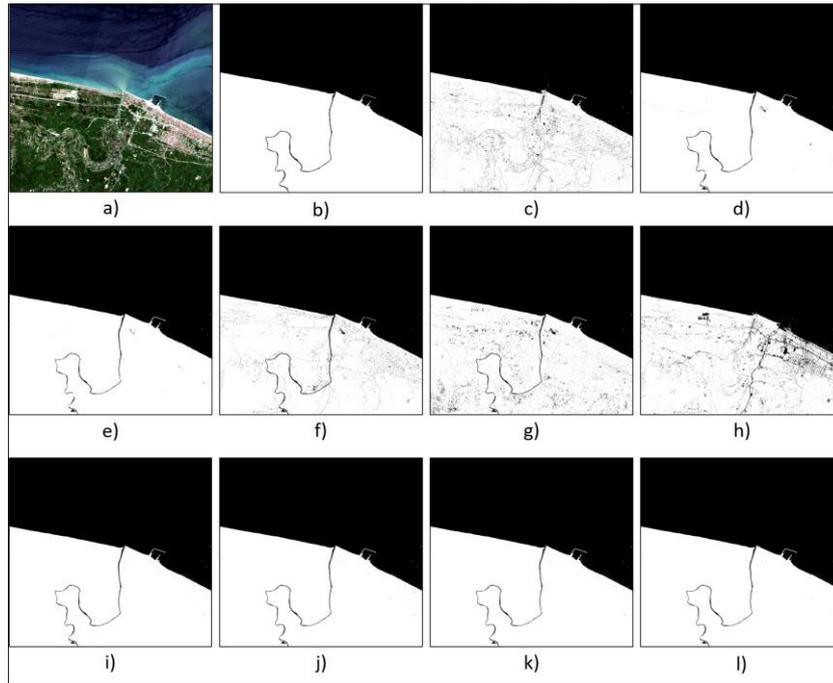

Fig. 10. The segmentation results for the Sakarya-Karasu study area. a) Sentinel-2 Image, b) Manuel-digitization, c) BGR/RF, d) BGRN/RF, e) BGN/RF, f) BRN/RF, g) GRN/RF, h) BGR/SVM, i) BGRN/SVM, j) BGN/SVM, k) BRN/SVM, l) GRN/SVM

Fig 11. Segmentation results for the Sakarya-Karasu region. a) Sentinel-2 Image, b) manual-digitization, c) BGRN/SVM (best result), d) BGR/SVM (worst result).

Band Combination/ Method	Accuracy (%)	IoU (%)	F1 Score (%)	Precision (%)	Recall (%)
BGR/RF	97,631	95,369	97,629	98,976	96,319
BGR/SVM	96,498	93,149	96,453	99,033	94,004
BGRN/RF	99,751	99,511	99,755	99,703	99,807
BGRN/SVM	**99,782**	**99,571**	**99,785**	99,690	99,880
BGN/RF	99,735	99,479	99,739	99,612	99,866
BGN/SVM	99,779	99,565	99,782	99,688	99,877
BRN/RF	98,775	97,587	98,779	**99,770**	97,807
BRN/SVM	99,780	99,568	99,783	99,682	**99,885**
GRN/RF	98,086	96,232	98,080	99,696	96,515
GRN/SVM	99,775	99,558	99,778	99,674	99,883

Table-6. Accuracy analysis results for the Sakarya-Karasu region (best results are bold).

Accuracy analysis results for the Sakarya-Karasu region are given in Table-6. Even though band combinations including the NIR band produce more accurate results, the results of the RF and SVM algorithms vary slightly. Combinations that included the NIR band generated more accurate results with the SVM algorithm than with the RF algorithm. On the other hand, the SVM algorithm produces the worst result in the combination that did not include NIR.

Accuracy Analysis for All Study Areas
Accuracy analysis results with the most successful results of each study area were given in Table-7. While the highest IoU and F1 Score results were obtained with BRN band combination and RF method in the Hatay-Samandag study area, the RF algorithm was similarly more successful in the Rize-Iyidere study area. On the other hand, the SVM algorithm was more successful in the other three study areas (Izmir-Aliaga, Samsun-Bafra, and

Sakarya-Karasu). For example, in Izmir-Aliaga, SVM's F1-score is 99.239 while RF's is 98.903. There was also a significant difference in the Samsun-Bafra area. In this area, SVM's F1-score is 99.785 while RF's is 98.658. Considering in general, it was observed that the study area where the highest F1-score was obtained with GRN SVM in the Rize-Iyidere region. The next best result is the BRN RF in the Hatay-Samandag region. The absence of any

sediment transport problem in these areas provided higher segmentation accuracy. Another remarkable result is that almost all of the best band combinations in each study area contain the NIR band. It is seen that the highest Recall value is obtained with BGR only in the Izmir-Aliaga region. However, the binary result of this band combination contains a lot of noise.

Table-7. The most successful results for each study area.

	Accuracy (%)		IoU (%)		F1 Score (%)		Precision (%)		Recall (%)	
	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF
Hatay-Samandag	99.847 (GRN)	99.865 (BRN)	99.686 (GRN)	99.724 (BRN)	99.843 (GRN)	99.862 (BRN)	99.950 (BGRN)	99.917 (BRN)	99.751 (GRN)	99.807 (BRN)
Izmir-Aliaga	99.135 (BRN)	98.749 (BRN)	98.490 (BRN)	97.830 (BRN)	99.239 (BRN)	98.903 (BRN)	98.797 (BRN)	98.178 (BRN)	99.798 (BGR)	99.648 (GRN)
Rize-Iyidere	99.929 (GRN)	99.942 (GRN)	99.816 (GRN)	99.851 (GRN)	99.908 (GRN)	99.925 (GRN)	99.952 (BGN)	99.947 (BRN)	99.904 (BRN)	99.938 (GRN)
Samsun-Bafra	99.414 (BGN)	98.834 (BGRN)	98.646 (BGN)	97.352 (BGRN)	99.318 (BGN)	98.658 (BGRN)	99.388 (BGN)	97.634 (BGRN)	99.879 (GRN)	99.973 (BGN)
Sakarya-Karasu	99.782 (BGRN)	99.751 (BGRN)	99.571 (BGRN)	99.511 (BGRN)	99.785 (BGRN)	99.755 (BGRN)	99.690 (BGRN)	99.770 (BRN)	99.885 (BRN)	99.866 (BGN)

Conclusion

The presented study has aimed to analyze the performance of RF and SVM machine learning methods for water-body segmentation in coastal areas with different hydrodynamic and morphological properties. For this purpose, water-body segmentation with SVM and RF methods have been carried out in five different study areas as Izmir-Aliaga, Hatay-Samandag, Rize-Iyidere, Samsun-Bafra, and Sakarya-Karasu which have different hydrodynamics, morphologic and hydrologic properties. For determining the performance of water-body extraction accuracy, IoU, F1 score, precision, and recall metrics have been used. The second aim of this study was to investigate the effects of band combinations on water-body segmentation performance. An optimal band combination was determined for each study area.

Approximately 99% F1 Score was obtained in the analysis of all study areas. The results of this investigation show that water-body segmentation can be realized with high accuracy with RF and SVM algorithms. There is embouchure sediment transport in Samsun-Bafra and Sakarya-Karasu study areas and this has resulted in low water-body segmentation accuracy. On the other hand, high accuracy results were obtained in the Rize-Iyidere study area, which does not have such a problem. The second major finding was that the optimal band combination was determined for each study area. The band combinations giving the highest accuracy were observed BRN for Izmir Aliaga and Hatay Samandag regions, BGRN for Sakarya Karasu region, GRN for Rize Iyidere region, and BGN for the Samsun Bafra region. The band combinations giving the lowest accuracy were observed BGR band combinations for all study areas. The common feature of band combinations giving the lowest accuracy results is that they do not contain the NIR band.

Thus, this study strengthens the idea that the NIR band should be used for water-body segmentation applications. The study contributes to our understanding of low accuracy results in water-body segmentation. The reasons affecting the accuracy analysis for each study area were analyzed separately according to the geography of the study areas. The issue of water-body segmentation in heterogeneous hydrodynamic and morphodynamic structured coastal areas is an intriguing one that could be usefully explored in further research.

In future work, the same band combinations for satellite images at different dates can be analyzed for coastal areas that have temporally different hydrodynamic processes. Higher-resolution satellite images can be analyzed with the same band combinations and classification method.

References

Ai, B., Zhang, R., Zhang, H., Ma, C., Gu, F. (2019). Dynamic process and artificial mechanism of coastline change in the Pearl River Estuary. *Regional Studies in Marine Science,* 30, 100715.

Balázs, B., Bíró, T., Dyke, G., Singh, S. K., Szabó, S. (2018). Extracting water-related features using reflectance data and principal component analysis of Landsat images. *Hydrological sciences journal,* 63(2), 269-284.

Bamdadinejad, M., Ketabdari, M. J., Chavooshi, S. M. H. (2021). Shoreline Extraction Using Image Processing of Satellite Imageries. Journal of the Indian Society of Remote Sensing, 1-11.

Bangira, T., Alfieri, S. M., Menenti, M., Van Niekerk, A. (2019). Comparing thresholding with machine learning classifiers for mapping complex water. *Remote Sensing,* 11(11), 1351.

Belgiu, M., Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing,* 114, 24-31.

Berkün, M., Anılan, T., Aras, E. (2010). Doğu Karadeniz Bölgesi'nde Sediment Taşınması ve Kıyı Erozyonu Etkileşimleri. *Türkiye Mühendislik Haberleri* (TMH), 3(4), 461-462.

Breiman, L. (2001). Random forests. *Machine Learning,* 45(1), 5-32.

Burak, S.; Doğan, E.; Gazioğlu, C. (2004) Impact of urbanization and tourism on coastal environment. *Ocean. Coast. Manag. 47*, 515–527.

Cao, W., Wong, M. H. (2007). Current status of coastal zone issues and management in China: *A review. Environment International,* 33(7), 985-992.

Choung, Y. J., Jo, M. H. (2017). Comparison between a machine-learning-based method and a water-index-based method for shoreline mapping using a high-resolution satellite image acquired in Hwado Island, South Korea. *Journal of Sensors,* 2017.

Coiman, A. (2020). Shoreline evolution of Valencia lake and land use and land cover changes in Zamora municipality, Aragua state, Venezuela, period 1986-2016. *International Journal of Environment and Geoinformatics,* 7(3), 305-318.

Celik, OI., Gazioglu, C. (2022). "Coast type based accuracy assessment for coastline extraction from satellite image with machine learning classifiers", *The Egyptian Journal of Remote Sensing and Space Science*, vol. 25, no. 1, pp. 289-299, 2022.

Çiftçi, N. S. (2011). *Türkiye Denizleri Açık Sularının Ekim 2000'deki Fitoplankton Kompozisyonu.* Süleyman Demirel Üniversitesi Eğirdir Su Ürünleri Fakültesi Dergisi, 7(2), 23-36.

Demir, N., Oy, S., Erdem, F., Şeker, D. Z., Bayram, B. (2017). Integrated shoreline extraction approach with use of Rasat MS and SENTINEL-1A SAR Images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* 4, 445.

Demir, N., Bayram, B., Şeker, D. Z., Oy, S., Erdem, F. (2019). A nonparametric fuzzy shoreline extraction approach from Sentinel-1A by integration of RASAT pan-sharpened imagery. *Geo-Marine Letters*, 39(5), 401-415.

Dervisoglu, A., Bilgilioğlu, B. B., Yağmur, N. (2020). Comparison of Pixel-Based and Object-Based Classification Methods in Determination of Wetland Coastline. *International Journal of Environment and Geoinformatics,* 7(2), 213-220.

Dixon, B., Candade, N. (2008). Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?. *International Journal of Remote Sensing*, 29(4), 1185-1206.

Gazioglu C, Yücel ZY, Burak S, Okus E, Alpar B. (1997) . Coastline changes and inadequate management between Kilyos and Karaburun shoreline. *Turkish J Mar Sci.* 3(2):111–122

Gislason, P. O., Benediktsson, J. A., Sveinsson, J. R. (2006). *Random forests for land cover classification. Pattern Recognition Letters,* 27(4), 294-300.

Guney, Y., Polat, S. (2015). Coastline Change Detection Using Remote Sensing Data: The Case of Aliağa And Çandarli. *Journal of Aeronautics and Space Technologies,* 8(1), 11-17.

Haslett, S. (2008). Coastal systems. Routledge.

Isik, S., Sasal, M., Dogan, E. (2006). *Sakarya Nehrinde Barajların Mansap Etkisinin Araştırılması.* Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 21(3), 401-408.

Jamil, A., Bayram, B. (2017). Tree species extraction and land use/cover classification from high-resolution digital orthophoto maps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11(1), 89-94.

Knerr, S., Personnaz, L., Dreyfus, G. (1990). Single-layer learning revisited: a stepwise procedure for building and training a neural network. *In Neurocomputing* (pp. 41-50). Springer, Berlin, Heidelberg.

Kumar, L., Afzal, M. S., Afzal, M. M. (2020). Mapping shoreline change using machine learning: a case study from the eastern Indian coast. *Acta Geophysica,* 68(4), 1127-1143.

Lin, Y. C., Cheng, Y. T., Zhou, T., Ravi, R., Hasheminasab, S. M., Flatt, J. E., … Habib, A. (2019). Evaluation of UAV LiDAR for mapping coastal environments. *Remote Sensing,* 11(24), 1–32. https://doi.org/10.3390/rs11242893.

Maglione, P., Parente, C., Vallario, A. (2014). Coastline extraction using high resolution WorldView-2 satellite imagery. *European Journal of Remote Sensing,* 47(1), 685-699.

Mahboubi, P., Parkes, M., Stephen, C., Chan, H. M. (2015). Using expert informed GIS to locate important marine social-ecological hotspots. *Journal of Environmental Management, 160*, 342-352.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, *International Journal of Remote Sensing,* 17(7): 1425-1432.

Minghelli, A., Spagnoli, J., Lei, M., Chami, M., Charmasson, S. (2020). Shoreline Extraction from WorldView2 Satellite Data in the Presence of Foam Pixels Using Multispectral Classification Method. *Remote Sensing*, 12(16), 2664.

Mountrakis, G., Im, J., Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS *Journal of Photogrammetry and Remote Sensing,* 66(3), 247-259.

Neumann, J. E., Emanuel, K., Ravela, S., Ludwig, L., Kirshen, P., Bosma, K., Martinich, J. (2015). Joint effects of storm surge and sea-level rise on US Coasts: new economic estimates of impacts, adaptation, and benefits of mitigation policy. *Climatic Change,* 129(1), 337-349.

Ngowo, R. G., Ribeiro, M. C., Pereira, M. J. (2021). Quantifying 28-year (1991–2019) shoreline change trends along the Mnazi Bay–Ruvuma Estuary Marine Park, Tanzania. *Remote Sensing Applications: Society and Environment,* 23, 100607.

Nowozin, S. (2014). Optimal decisions from probabilistic models: the intersection-over-union case. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 548-555).

Öner, E. (2008). Alluvial geomorphology and paleogeographical studieson the Asi (Orontes) delta plain (Antakya/HATAY). *Aegean Geograp. J,* 17, 1-25.

Osborne, B. P., Osborne, V. J., Kruger, M. L. (2012). Comparison of satellite surveying to traditional surveying methods for the resources industry. *Journal of the British Interplanetary Society,* 65(2), 98.

Ozturk, D., Beyazit, I., Kilic, F. (2015). Spatiotemporal analysis of shoreline changes of the Kizilirmak Delta. *Journal of Coastal Research,* 31(6), 1389-1402.

Patil, R. G., Deo, M. C. (2020). Sea Level Rise and Shoreline Change under Changing Climate Along the Indian Coastline. *Journal of Waterway, Port, Coastal, and Ocean Engineering,* 146(5), 05020002.

Paul, A., Tripathi, D., Dutta, D. (2018). Application and comparison of advanced supervised classifiers in extraction of water bodies from remote sensing images. *Sustainable Water Resources Management,* 4(4), 905-919.

Sahin, C. (2020). *Sakarya Nehri Deltası'nda Uzun Süreli Rüzgar ve Dalga İklimi Değişimleri.* Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 22(65), 353-368.

Samsun Investment Support Office, (2018). Bafra İlçesi Tarım Sektörü Raporu. Samsun, Türkiye.

Scholkopf, B., Herbrich, R., Smola, A. J. (2001). A generalized representer theorem. *In International conference on computational learning theory* (pp. 416-426). Springer, Berlin, Heidelberg.

Sekovski, I., Stecchi, F., Mancini, F. Del Rio, L. (2014). Image classification methods applied to shoreline extraction on very high-resolution multispectral imagery, *International Journal of Remote Sensing,* 35(10): 3556-3578.

Small, C., Nicholls, R. J. (2003). A global analysis of human settlement in coastal zones. *Journal of coastal research,* 584-599.

TMMOB (Union of Chambers of Turkish Engineers and Architects). (2012). *Karasu Kıyı Alanı Kıyı Daralması Raporu.*

Toure, S., Diop, O., Kpalma, K., Maiga, A. S. (2019). Shoreline detection using optical remote sensing: A review. *ISPRS International Journal of Geo-Information,* 8(2), 75.

Turkish State Meteorological Service (2021). The Latest Weather in | MiddleEast. https://www.mgm.gov.tr/eng/forecast-cities.aspx. Erişim Tarihi: 14.3.2021.

Zhang, F., Tiyip, T., Johnson, V. C., Wang, J., Nurmemet, I. (2016). Improved water extraction using Landsat TM/ETM+ images in Ebinur Lake, Xinjiang, China. Remote Sensing Applications: Society and Environment, 4, 109-118.

Zhang, T., Yang, X., Hu, S. Su, F. (2013). Extraction of coastline in aquaculture coast from multispectral remote sensing images: Object-based region growing integrating edge detection, *Remote Sensing,* 5(9): 4470-4487.