Development of Practical Methods for Detecting Floating Marine Debris and Classification of Underwater Seaweed/Seagrass Beds in the Seto Inland Sea Using Satellite Data

衛星データによる瀬戸内海の海洋ごみ検出と藻場分類の実用的な手法開発に関する研究

March 2024

Shilin Song

Contents

Abstract

Since the last century, anthropogenic climate change, commonly referred to as 'global warming' has emerged as a pressing concern. The concentration of CO2 has surged from 280 parts per million (ppm) in 1750 to 368 ppm by 2000, reaching levels unprecedented in at least the past 420,000 years. Concurrently, global mean surface temperatures have witnessed a rise of approximately 0.6 degrees over the past century, contributing to the escalation of tropical cyclones and potentially amplifying their intensity. In light of these environmental challenges, the detection and classification of vegetation in a wide region have become crucial components in mitigating the greenhouse effect and facilitating post-disaster reconstruction. Specifically, the preservation and restoration of seaweed and seagrass beds have gained paramount significance due to their proven capacity for substantial carbon storage. Simultaneously, the imperative for immediate post-disaster reconstruction has intensified, driven by the escalating damages wrought by climate-related disasters.

However, regarding the detection and classification of vegetation, the existing algorithm in remote sensing fields, such as Leaf Area Index (LAF) and Normalized Vegetation Index (NDVI) have been proposed to solve the problem of detecting vegetation. And *Sargassum Index* (SI) has been proposed for *Sargassum* detection. Nevertheless, none of them can realize the detection and classifying seaweed and seagrass beds growing beneath the water surface. Simultaneously, concerning the detection of vegetation debris detection for post-disaster reconstruction, common methods involve the use of buoys equipped with Global Positioning System (GPS), aerial photographs, and visual observation. However, these methods face challenges in promptly locating marine debris, prompting the exploration of remote sensing

technology applications. Although there has been some algorithm that can detect floating vegetation, such as Floating Algae Index (FAI), the high turbidity marine environment after disaster obstacles to an accurate detection.

Based on the background mentioned, this study aims to propose practical methods for detecting and classifying vegetation in the Seto Inland Sea. Specially, this study initially utilized the corrected Floating Algae Index (cFAI) and the Otsu method to reproduce the marine debris distribution immediately after the July 2018 Heavy Rains in Western Japan, using multidate Landsat-8 Data. Secondly, an algorithm, named *Sargassum* and *Zostera* Distinguishing Index (SZDI) was proposed for distinguishing *Sargassum* and *Zostera* using Sentinel-2 Data. Thirdly, for reducing the algorithm development time and program creation time for distinguishing *Sargassum* and *Zostera*, an unsupervised learning model, K-Means has been attempted to be employed.

This dissertation is composed of six chapters.

Chapter 1 clarifies the background, the objectives, and the study areas of this study.

In Chapter 2, several existing plant detection algorithms and studies focused on distinguishing seaweed and seagrass beds are introduced.

In Chapter 3, a practical method for detecting marine debris immediately after the July 2018 heavy rains in western Japan using multidate Landsat-8 data is proposed and described. By abstracting the spectral reflectance of marine debris, it was confirmed that the marine debris consists of vegetation, meeting the prerequisite of the cFAI method. Subsequently, the turbid background caused by heavy rain was removed using the cFAI method, and effective automatic binarization of cFAI was

achieved through the application of the Otsu method. Ultimately, the marine debris distribution was reproduced.

In Chapter 4, an algorithm, named Sargassum and Zostera Distinguishing Index (SZDI) for distinguishing *Sargassum* and *Zostera* beds using Sentinel-2 data is proposed and described. The results of the field survey conducted in 2021 reveal the spectral reflectance characteristics of *Sargassum* and *Zostera*, indicating that both of them exhibit peak reflectance between the blue and red band, but *Zostera* exhibit relative high reflectance in the green band. Based on these characteristics, SZDI was proposed, and the distribution map of *Sargassum* and *Zostera* was generated.

In Chapter 5, an attempt to distinguish *Sargassum* and *Zostera* using machine learning is conducted and discussed. The K-means clustering is confirmed to be effective for classifying *Sargassum* and *Zostera* in this study area in some aspects. Besides, this model demonstrates prominent advantages, including an objective process and a shorter development time for algorithms and programming.

In Chapter 6, the conclusions of this study are described.

Chapter 1: General Introduction

1.1. Background

1.1.1. Heavy Rain disasters

Since the last century, anthropogenic climate change, commonly referred to as 'global warming', attributed to the increasing concentration of greenhouse gases, has become a serious concern. Fossil fuel emissions, particularly carbon dioxide (CO_2), constitute a major component of these gases. Between 1750 to 2000, the concentration of CO_2 has increased from 280 parts per million (ppm) to 368 ppm, reaching the highest level in at least the past 420,000 years. Furthermore, over the past century, global mean surface temperatures have risen by approximately 0.6 degrees, with Northern Hemisphere surface warming experiencing unprecedented levels over the past thousand years [1]. This global warming phenomenon provides more energy to fuel tropical cyclones, potentially making them more intense [1,2].

Tropical cyclones, as low-pressure weather systems that develop over the warm waters of the oceans, regularly impact populations and periodically result in devastating weather-related natural disasters. Over the past two centuries, tropical cyclones have caused an estimated 1.9 million deaths worldwide and are confirmed to be among the costliest of weather-related natural disasters [3–5]. Japan, in particular, ranks second globally in terms of "physical exposure to cyclones in percentage of population" [5]. In 2018, marked as the fourth-highest year for global land and ocean temperature annually in the recorded period from 1880 to 2018 (https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/201813/supplemental/page-1; Accessed on Dec. 7, 2023), Japan

experienced several large tropical cyclones. On Sep. 4, 2018, Typhoon Jebi made landfall first in Tokushima Prefecture, Japan, and rapidly moved to Osaka Bay, resulting in record-level storm surges and storm waves inside the bay. This event caused significant coastal flooding along the northern coasts of the inner part of the bay [6,7]. From Jun. 28 to Jul. 8, 2018, the Baiu front and Typhoon Prapiroon caused extremely heavy rain, resulting in an unprecedented amount of precipitation and tremendous damage. The event led to widespread and devastating floods, landslides, and mudflows [6,8,9].

Because of the heavy rain, a significant amount of sediments flowed from large and small rivers into the Seto Inland Sea adjacent to the affected area. This kind of inflow not only resulted in the generation of excessive marine debris but also caused a 68% reduction in the transparency of seawater off the Hiroshima shore [10]. According to the report by the Ministry of Land, Infrastructure, Transport, and Tourism, these marine debris primarily consisted of plant materials such as driftwood and plant fragments. However, the official marine debris collection conducted by the Ministry of Land, Infrastructure, Transport, and Tourism in the Seto Inland Sea off Hiroshima Prefecture took more than a month (7 July to 17 August, 2018, source: https://www.mlit.go.jp/kowan/kowan_tk3_000021.html; Accessed on Dec. 7, 2023)), posing a long-term hazard to the environment, fishing industry and navigation transportation. Although more than 3 years have passed since this disaster, the distribution of marine debris at that time remains unknown. Additionally, establishing a method for promptly obtaining information about marine debris after a disaster is crucial, especially considering the increasing frequency of torrential rains in recent years.

1.1.2. Coastal Blue Carbon Ecosystems in Global Warming

The term Blue Carbon (abbreviated as BC) was first coined a decade ago, and primarily refers to organic carbon captured and stored by the oceans and coastal ecosystems [11]. The ocean plays a crucial role in the Earth's global carbon cycle with just over 90% of the CO2 on the planet is found in the ocean [12]. Although coastal ecosystems, featuring plant foundation species such as seagrass beds, saltmarshes, and mangrove forests account for less than 0.5% of the ocean area, they account for 50-71% of the total carbon stock preserved in marine sediments. The annual carbon storage in these ecosystems, collectively known as blue carbon ecosystems, is roughly equivalent to the amount of carbon stored by all terrestrial plants on Earth. However, due to climate change and human activities, blue carbon ecosystems are rapidly disappearing, with a reported disappearance rate more than four times that of tropical rainforests [12]. Thus, the protection of blue carbon ecosystems has been an urgent task.

On Apr. 1, 2022, the "Act on Special Measures Concerning Conservation of the Environment of the Seto Inland Sea" was enacted at the 204th session of the Diet of Japan. The act aims to clarify the distribution of seaweed and seagrass beds, among other objectives (Source: https://www.env.go.jp/content/900542067.pdf; Accessed on Dec. 8, 2023). Through this kind of clarification, environmental conservation activities are expected to be encouraged, including the preservation of seaweed and seagrass beds. These actions are expected to enhance the ability of seaweed and seagrass beds for carbon capture and storage.

1.2. Objectives

This study aims to propose a practical method for detecting and classifying vegetation in the Seto Inland Sea, with a specific focus on the marine debris which are confirmed to consist of vegetation, and seagrass and seaweeds in the Seto Inland Sea. A variety of research has been conducted to develop methods for detecting the floating plants in sea areas, which are considered to be similar to marine debris studied in this research. Furthermore, there are existing research aimed at distinguishing seaweed and seagrass beds by clarifying their spectral reflectance. However, there is currently lack of methods for detecting marine debris in turbid water, a common occurrence after heavy rain disasters. Additionally, the method for distinguishing seaweed and seagrass beds is limited to reflectance analysis, lacking the development of classification algorithms and their application in specific areas. Finally, due to the extended time required for algorithm development and program creation, we aims to explore an unsupervised method for classifying seaweed and seagrass bed in the same area.

In order to address these gaps, several important works have been conducted. Each chapter focuses on specific research, and their objectives are as follows:

(1) To employing an automatic binarization method after clarifying the optical characteristics of marine debris. Finally, the goal is to reproduce the actual conditions of the marine debris distribution in the Seto Inland Sea off the coast of Hiroshima Prefecture after the heavy rain in July 2018 and verify the methods.

(2) To develop and verify a method able to spectroscopically distinguish *Sargassum* and *Zostera* in the sea using Sentinel-2 data, which exhibit the highest spatial resolution among currently available free data. Furthermore, this study aimed to understand changes in the distribution over time.

(3) To utilize an unsupervised learning model, K-Means to distinguish Sargassum and Zostera in the same area as discussed in the previous chapter, using Setinel-2 images.

Building upon the SZDI algorithm, the reflectances of band 2 (blue), band 3 (green), and band 4 (red) will be designated as the features for unlabeled data. Ultimately, the accuracy of this model will be evaluated using MSE and Kappa coeffitient.

1.3. The Study Area

1.3.1. Aki Nada in the Seto Inland Sea

As shown in Figure 1.1, the Seto Inland Sea is the largest inland sea in Japan. It is located in western Japan. The total length of the coastline is 500 km, and the average water depth is 30 m. More than 600 islands are located in this sea [13]. This inland sea has a complex structure in which a wide part called Nada is connected by a narrow waterway called Seto (Takeoka, 2001). Furthermore, the Seto Inland Sea has a large tidal range that exceeds 2 m around the innermost part of Nada. Hence, the tidal current in this sea is extremely strong, and there are many areas with rapid water flow at speeds of 1 m/s or more [13,14]. Additionally, different 643 rivers water originating in surrounding mountains flow into the sea (source: https://www.mlit.go.jp/kowan/kowan_tk3_000021.html; Accessed on Dec. 7, 2023)). These geometric factors create a complex marine environment. Simultaneously, this area is home to 28% of Japan's total population and plays an important role in country's fishing, aquaculture and industry. The fish catch in this area reached to 281,000 tons, aquaculture contributed 349,000 tons in 1993, and industrial production amounted to 94,500 billion in 1991 [15].This study focuses on the central waters of the Seto Inland Sea off the coast of Hiroshima Prefecture, named Aki Nada, where the heavy rains in July 2018 resulted in many casualties.

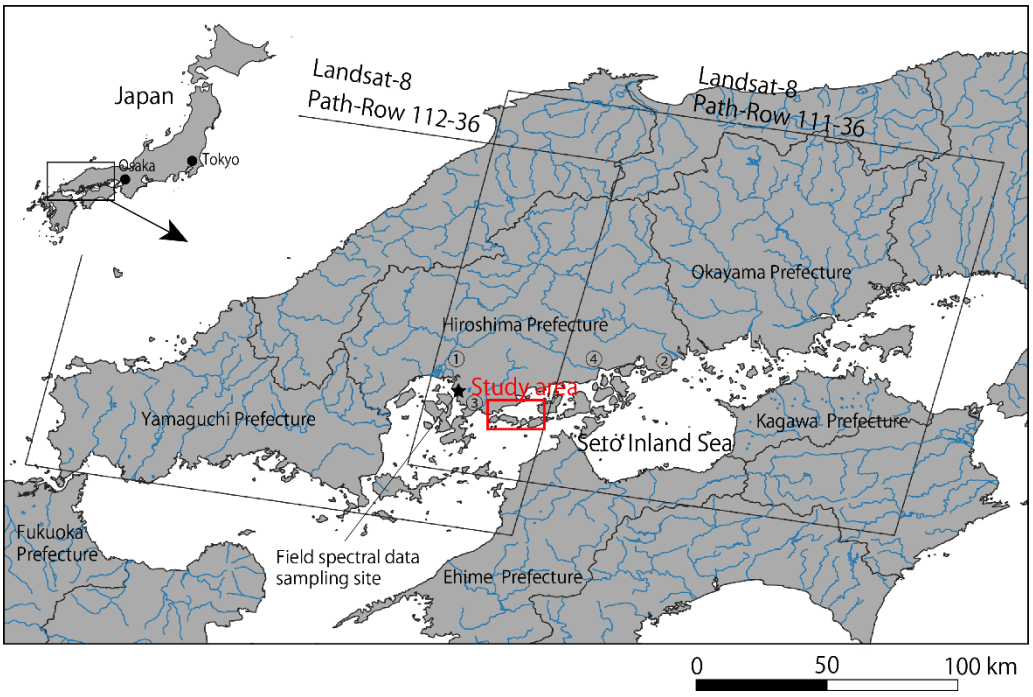


Figure 1.1. Location map of the study area with the orbital information of Landsat-8. Points labeled ①, ②, ③, and ④ indicate the positions of Ohta River, Ashida River, Kurose River, and Nuta River, respectively.

1.3.2. Takehara oki in the Seto Inland Sea

The study area was the Seto Inland Sea off the coast of Takehara City, located in the south-central part of Hiroshima Prefecture, as shown in Figure 1.2. According to Yoshida et al. [16], the SWSG beds area of the "Bingo–Geiyo–Seto" fishery area (conducted from 1989 to 1991) is 3636 ha. The total areas of seagrass, SWSG, and kelp beds are 2917 ha (80%), 1289 ha (35%), and 184 ha (5%), respectively (as there is mixed marine forest, the total does not add up to 100%). Thus, the *Sargassum* and *Zostera* beds are the main SWSG beds in the central Seto Inland Sea. Notably, *S. Confusum* and *S. Yezoense* are primarily found on the coast of Japan, extending to northeast coast of Honshu and the Seto Inland Sea, while *S. Trichophyllum* appears to be restricted to the Seto Inland Sea [17]. In addition, as for their habitat, *Zostera* beds are formed in calm sandy mud areas with little wind and waves, whereas *Sargassum* (formed by brown SWSG *Sargassum*) is formed in rocky reef areas. In the Seto Inland

Sea, there are many areas of water wherein sand and mud and rocky reef areas exist alternately; thus, *Sargassum* and *Zostera* beds coexist. These statistics are data for shallow waters below 10 m in depth, and as the marine forest in Figure 1 also exists mostly in shallow waters below 10 m, this study targeted waters below 10 m. In addition, the tidal level difference in the target water area is approximately 4 m at maximum.

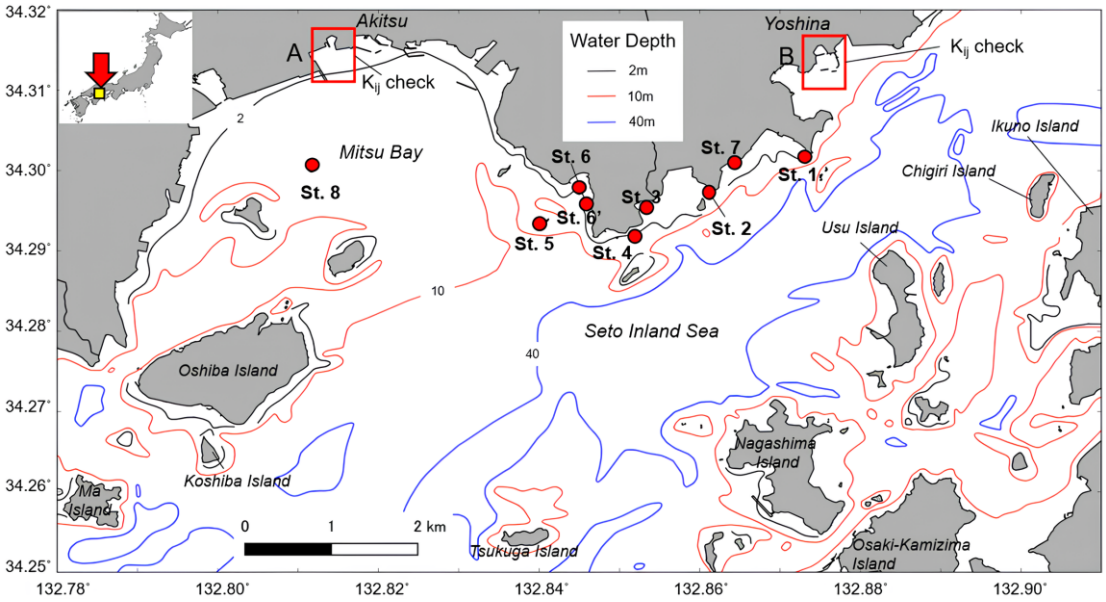


Figure 1.2. Study area depicting field survey points and water depth. This image is extracted from the yellow square in the top left corner. St. 1-St. 8 indicate the locations where we conducted field survey.

1.4. Outline of the Dissertation

The dissertation is composed of six chapters. Chapter 1 clarifies the background, the objectives, and the study areas of this study. In chapter 2, several existing plant detection algorithms and studies focused on distinguishing seaweed and seagrass beds are introduced. In chapter 3, a practical method for detecting marine debris immediately after the July 2018 heavy rains in western Japan using Multidate Landsat-8 data is proposed and described. And the distribution of marine debris is reproduced by this method. In Chapter 4, a practical algorithm for distinguishing *Sargassum* and *Zostera* beds in the Seto Inland Sea using Sentinel-2 data is proposed and described. And the

distribution of *Sargassum* and *Zostera* beds in the study area are estimated by this method. In Chapter 5, an attempt to distinguish *Sargassum* and *Zostera* using machine learning is conducted and discussed. This method is evaluated by Mean Square Error (MSE) and Kappa coefficient. In Chapter 6, the conclusions of this study are described.

Chapter 2: Remote Sensing Algorithms for Vegetation Detection and Distinguishing

2.1. Traditional Vegetation Detection Algorithm

2.1.1. Normalized Difference Vegetation Index (NDVI)

As we all know, more than 40 vegetation indices have been developed in the field of remote sensing for qualitatively and quantitatively evaluating vegetative covers using spectral measurement [18]. The inception of the vegetation detection algorithm can be traced back to 1972 when the first earth resources satellite, Landsat-MSS was launched. Subsequently, NASA's "Monitoring the Vernal Advancement and Retrogradation of Natural Vegetation" program was initiated [18–20], marking the beginning of vegetation detection. Another noteworthy American program, LACIE (Large Area Crop Inventory Experiment) contributed to vegetation detection studies. This program, including essential findings, highlighted the unique reflectance characteristics of vegetation in the red and near-infrared channels of sensors on satellites [18]. Tucker observed that the chlorophyll in vegetation absorbs visible radiation in the red channel (630-690nm), while the leaf cellular structures strongly reflect in the near-infrared channel (760-900nm) [21]. Subsequently, many mainline vegetation indices were developed based on these crucial characteristics.

Among them, Pearson and Miller are considered pioneers in the history of vegetation indices. They proposed the first two indice: the "Ratio Vegetation Index" (RVI) and the "Vegetation Index Number" (VIN), for the estimation and monitoring of vegetative covers.

$$RVI = R/NIR \tag{1}$$

$$VIN = NIR/R \quad (2)$$

where R represents the mean reflectance in the red channel, NIR represents the mean reflectance in the near-infrared channel [18,22].

However, these indices are sensitive to the optical properties of the ground [18,23], indicating that darker soil substrates lead to higher vegetation index values [24,25].

Additionally, a widely recognized index known as the "Normalized Difference Vegetation Index" (NDVI) has been proposed by Rouse [19,20]:

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad (3)$$

where the NIR represents the reflectance of MSS band 7 (Near-Infrared, 0.8 to 1.1 μm), and RED represents the reflectance of MSS band 5 Visible red (0.6 to 0.7 μm). Although it has a great promotion at mitigating the effect of sensor calibration degradation by approximately 6% of the overall index value [26], it shares a similar issue with RVI in being sensitive to the optical properties of the soil background [24,25]. Consequently, it is not well-suited for the detection of marine debris after heavy rain, which is one of the objectives in this study.

2.1.2. The Floating Algae Index (FAI)

Hu proposed a floating algae detection algorithm designed for identifying algae in open ocean environments using operational MODIS (Moderate Resolution Imaging Spectroradiometer) data. This algorithm is named as the Floating Algae Index (FAI), and is based on the ocean color. This algorithm aims to detect algae floating on the water surface, such as the brown macroalgae *Sargassum* spp., and reduce the sensitivity to varying environmental and observing conditions [27].

Relying on the reflectance characteristics that water strongly absorbs light in the RED-NIR-SWIR (short-waved infrared) wavelengths, whereas the algae floating on the water surface exhibit higher reflectance in the NIR compared to other wavelengths, the FAI was defined as the difference between Rayleigh-corrected reflectance in the NIR and a baseline formed by the red and SWIR bands. And the equation is defined as follows [27]:

$$\mathrm{FAI} = R_{rc}(NIR) - R'_{rc}(NIR) \quad (4)$$

$$R'_{rc}(NIR) = R_{rc}(RED) + (R_{rc}(SWIR) - R_{rc}(RED))\frac{(\lambda_{\mathrm{NIR}}-\lambda_{\mathrm{RED}})}{(\lambda_{\mathrm{SWIR}}-\lambda_{\mathrm{RED}})} \quad (5)$$

where $R'_{rc}(NIR)$ is the baseline reflectance in the NIR band derived from a linear interpolation between the red and SWIR bands.

This algorithm offers the advantage of mitigating the effects of the atmosphere. Whether in a turbid atmosphere or under significant sun glint, the counts of algae pixels and seawater pixels, as well as their differences, remain stable. Additionally, it has been demonstrated that the algorithm is also applicable to Landsat data. However, there are still challenges as researchers find it difficult to classify turbid water and floating plants [28]. Moreover, the effectiveness of the algorithm when applied to other species of floating plants has not been confirmed.

2.2. Traditional Detection and Classification Algorithm of Seaweed and Seagrass beds

2.2.1. The Bottom Index (BI)

To extract water depth and bottom-type information from passive multispectral scanner data, Lyzenga proposed a general algorithm known as the Bottom Index [29,30]. This algorithm is based on the assumption that a pair of wavelength bands can be found, making the ratio of the bottom reflectance in these two bands remain consistent across all the bottom types within a given scene. And it is defined as:

$$BI_{ij} = LN\left[\frac{(L_i - L_{si})}{(L_j - L_{ji})^{kij}}\right] = LN(L_i - L_{si}) - k_{ij}LN(L_j - L_{sj}) \quad (6)$$

where DNi means the digital value of the satellite image, i, j are arbitrary two bands. DNs represents dark pixels obtained from deep, clear water, while K_{ij} represents the ratio of dissipation coefficient of the two bands.

With a similar type of substrate, the Bottom Index value is expected to be consistent. Consequently, this algorithm is considered to be effective in identifying the seaweed and seagrass beds to some extent. Sochea et al. utilized Band 1 and Band 2 of ASTER satellite data to apply the Bottom Index and found it effective in detecting seaweed and seagrass beds below the water surface compared to the NDVI algorithm [31]. Matsunaga et al. indicated that the bottom index can be employed in classifying sand and coral / algae / seagrass habitat, with coral, algae, and seagrass commonly sharing the same Bottom Index value. Additionally, He also corrected the water depth using the tide level at the corresponding time [32]. However, due to potential inaccuracies arising from the combination of non-optimal wavelength bands and the reflectances not necessarily representing actual bottom types accurately, the classification of specific species is considered challenging using only this algorithm.

2.2.2. the *Sargassum* Index (SI)

In the study assessing the hyperspectral properties of the floating macroalgae *Sargassum* and aggregations of the seagrass *Syringodium filiforme* wrack in Greater Florida Bay, Dierssen et al. proposed a classification algorithm named the Sargassum Index (SI) [33]. This algorithm is developed based on the fact that *Sargassum* contains accessory pigments fucoxanthin and chlorophyll *c* [34], wherein the

chlorophyll *c* exhibits an absorption peak at 460, 485, and 635nm [35]. In contrast, the seagrass, which is indicated as *Syringodium* in Dierssen et al.'s study, contains chlorophyll *b*, exhibits absorption peaks in blue and red wavelengths at 470, 600, and 650nm.

Consequently, *Syringodium* wrack shows higher reflectance in the green wavelengths (500–600 nm) and low reflectance at 650 nm, coinciding with chlorophyll *b* pigment absorption. On the other hand, the *Sargassum* spectrum features a distinct dip at 630 nm, absent in *Syringodium*. The Sargassum Index (SI), used the difference reflectance values at 650nm and 630nm, and is defined as [33]:

$$Sargassum\ Index = \frac{R(650)}{R(630)} \tag{7}$$

where R(650) represents the reflectance at 650nm, and R(630) represents the reflectance at 630nm. Typically, The SI value for Sargassum is greater than 1, whereas the SI value for *Syringodium* is less than 1.

Because 630 nm and 650 nm are within the near-infrared red wavelength, which is typically not differentiated into two channels in a satellite sensor, this algorithm can only be applied in hyperspectral imagery through a spectrometer. For wider observation from satellites, an algorithm applicable to satellite data is expected to be developed.

Chapter 3: Reproduction of the Marine Debris Distribution Using Multidate Landsat-8 Data

3.1. Introduction

3.1.1. Backgrounds

In 2018, the western part of Japan experienced a torrential rain disaster (28 June to 8 July 2018, and heavy rain in July 2018) that caused more than 200 deaths [9,36]. Because of the heavy rain, much sediment flowed from large and small rivers into the Seto Inland Sea adjacent to the affected area. The heavy sediment inflow destroyed the special oyster farming raft and created excessive marine debris (mainly derived from plants such as driftwood and plant fragments). The official marine debris collection completed by the Ministry of Land, Infrastructure, Transport, and Tourism in the Seto Inland Sea off Hiroshima Prefecture took more than a month (7 July to 17 August 2018). Marine debris poses a significant obstacle to the fishing industry. Collecting the debris took a long time, mainly because it was difficult to locate due to the area's complicated topography and influence of ocean currents. Although more than 3 years have passed since this disaster, the distribution of marine debris at that time remains unknown. In addition, it is very important to establish a method for obtaining information about marine debris immediately after a disaster given the torrential rains that have become frequent in recent years.

3.1.2. Previous Studies

To grasp the distribution of marine debris, buoys with global positioning system (GPS) equipment, aerial photographs [37,38], and numerical simulation methods [39,40] used for exploring plastic waste can be considered. However, emergency observation during heavy rains is not easy, and the flow of the Seto Inland Sea is complex. It is not always possible to reproduce the position of marine debris. Therefore, the locations of marine debris can be identified by satellites that observe

the Earth periodically. Examples of exploration of suspended matter in water using satellites include the detection of farming rafts using synthetic aperture radar data [41,42] and the use of visible–near infrared data such as Terra/Aqua MODIS and Landsat-8 Operational Land Images (OLI) products for detecting floating algae using an external sensor [43,44]. Satellite remote sensing is further used to detect marine litter windrows for managing ocean plastic pollution [44]. In recent years, a scientific method for collecting word-of-mouth information using social network systems (SNS) or Internet information [45,46] has also attracted attention.

The floating algae index (FAI) algorithm is used to detect floating algae by employing conventional satellite remote sensing (RS). It realizes vegetation detection in a manner different from the method based on the normalized difference vegetation index (NDVI) by removing the influence of the seawater color in the background. In addition, a limitation of the traditional FAI algorithm [43] is that it erroneously recognizes floating plants and turbid water in water areas that are strongly affected by turbidity. The influence of turbidity in the sea is strong after a heavy rain disaster; hence, such misrecognition is considered likely when using the conventional FAI algorithm. Recently, Garcia et al. [47] were able to remove the influence of background water. Hu et al. [44] devised a new algorithm that improved detection performance in water areas affected by turbidity, and named this algorithm corrected FAI (cFAI). However, these algorithms are used to detect large living aquatic plants, and it is unknown whether they can be applied to detecting plant-derived disaster waste such as driftwood and plant pieces. Furthermore, the threshold for extracting suspended solids from cFAI is still mostly determined by a trial-and-error method. Hence, the results differ depending on the analyst.

3.1.3. Objectives

With the aforementioned research in mind, this study aims to reproduce the actual conditions of the marine debris distribution in the Seto Inland Sea off the coast of Hiroshima Prefecture after the heavy rain in July 2018. The data used for this purpose were the marine debris information acquired by cleaning boats and satellite images. First, information on the location and amount of marine debris was collected. This information, which was published on the Internet by public institutions on an

irregular basis, was mapped and graphed. Next, by applying the cFAI algorithm and objective binarization method to Landsat-8 data with a resolution of 30 m, plant-derived suspended matter (called marine debris in this study) is detected. The marine debris distribution map created using satellite data in this work is expected to aid the agency that manages the Seto Inland Sea in collecting marine debris immediately after a heavy rain disaster.

3.2. Materials and methods

3.2.1. Study Area and Landsat-8

As shown in Figure 1.1, the Seto Inland Sea is the largest inland sea in Japan. It is located in western Japan. The total length of the coastline is 500 km, and the average water depth is 30 m. More than 600 islands are located in this sea [13]. This inland sea has a complex structure in which a wide part called Nada is connected by a narrow waterway called Seto. Furthermore, the Seto Inland Sea has a large tidal range that exceeds 2 m around the innermost part of Nada. Hence, the tidal current in this sea is extremely strong, and there are many areas with rapid water flow at speeds of 1 m/s or more [13,14]. This study focuses on the central waters of the Seto Inland Sea off the coast of Hiroshima Prefecture, where the heavy rains in July 2018 resulted in many casualties.

According to the observation data of the Japan Meteorological Agency, the total rainfall from 0:00 on 28 June to 9:00 on 8 July exceeded 500 mm in the Chugoku region, including the Hiroshima Prefecture [18]. In particular, Higashi-Hiroshima City, Hiroshima Prefecture, experienced torrential rain of 426.5 mm in 48 h until 8:40 am on July 7. Because of this heavy rain, the major inflow rivers, such as the Ota River (point ① in Figure 3.1) and Ashida River (point ②), and small- and medium-sized rivers, such as Kurose River (point ③) and Nuta River (point ④), in Hiroshima Prefecture carried a large amount of sediment and disaster waste, such as plant species, to the sea.

Landsat-8 data are available for the Seto Inland Sea off Hiroshima Prefecture, which is the research area, where two satellite orbits (Path 111 and Path 112) overlap as shown in Figure 1. Therefore, although the Landsat-8 data usually have a 16-day

cycle, this overlap made it possible to obtain data with few clouds in a cycle of 7–9 days after the disaster. The Landsat-8 data used in the present study were the "Collection 2 Surface Reflectance" product of the OLI sensor distributed by the USGS Earth Explorer. This product is "atmospheric-corrected reflectance data" in which the effects of atmospheric gases, aerosols, and water vapor are removed. The downloaded Landsat-8 data were stored as 16-bit unsigned digital values (DN: digital number) and converted to ground surface reflectance (Rreal) using the following equation.

$$R_{real}(i) = DN(i) \times \mathrm{MULT}(i) + \mathrm{ADD}(i) \quad (8)$$

where i is the Landsat-8 level-2 band (Table 1) and MULT and ADD are the reflectance conversion coefficients stored in the Landsat-8 level-2 data metadata file. The values of MULT and ADD are 2.75×10^{-5} and -2, respectively, for all bands. In this study, in addition to the four images from 9 July, 16 July, 25 July, and 1 August immediately after the heavy rain, one image of the freshwater state before the heavy rain (acquired on 20 April) was used for cFAI processing as described later.

Table 3.1. Bands of landsat-8 operational land imager (OLI) level 2.

Band	Wavelength (μm)	Spatial Resolution (m)
1	0.43–0.45	30
2	0.45–0.51	30
3	0.53–0.59	30
4	0.64–0.67	30
5	0.85–0.88	30
6	1.57–1.65	30
7	2.11–2.29	30

3.2.2. Spectral Reflectance Data of Marine Debris

Almost no objective data other than satellite images are available to prove that large amounts of marine debris flowed into the Seto Inland Sea. In addition, during the studied heavy rainfall period, many of the roads to the port were cut off, and

marine debris was scattered even at sea, making it impossible to verify the debris distribution even by a research vessel in the sea immediately after the heavy rain. Therefore, in this study, we collected data on the approximate position and amount of debris collected by marine environment maintenance vessels (referred to as cleaning ships in this study) owned by the Chugoku Regional Development Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism. These data were published on the website of the Chugoku Regional Development Bureau. Table 3.2 lists the published materials on the status of marine debris collection. The authors created a compact map of the approximate location of marine debris by accumulating information from these published materials. In addition, the amount of marine debris collected was summarized as the time-series data of the debris accumulated every week.

Table 3.2. List of published reports (press releases; accessed on 1 November 2021) on marine debris collections by cleaning ships from the Ministry of Land, Infrastructure, Transport and Tourism after the heavy rain in July 2018.

Publication Date	Evidence Material
13 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180713-8top.pdf
16 July 2018	https://www.mlit.go.jp/common/001245276.pdf
17 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180717-5top.pdf
18 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180718-2top.pdf
19 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180719-4top.pdf
21 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180721-3top.pdf
22 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180722-2top.pdf
	https://www.cgr.mlit.go.jp/kisha/2018jul/180722-3top.pdf
24 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180724-1top.pdf
25 July 2018	https://www.cgr.mlit.go.jp/kisha/2018jul/180725-4top.pdf
1 August 2018	https://www.cgr.mlit.go.jp/kisha/2018aug/180801-2top.pdf

8 August 2018	https://www.cgr.mlit.go.jp/kisha/2018aug/180808-4top.pdf
15 August 2018	https://www.cgr.mlit.go.jp/kisha/2018aug/180815-1top.pdf

To understand the basic spectral characteristics of coastal garbage in Hiroshima Prefecture, samples of driftwood, polyvinyl chloride (PVC), plastic bottles (polyethylene terephthalate, or PET), and sand were acquired from Bayside Beach Saka (★ in Figure 1) in Mizushiri, Saka-cho, Aki-gun, Hiroshima Prefecture, on 4 May 2018. Note that although we collected spectral data of those materials in our study area, multiple spectral datasets of various material including plastics are now available online [19,20]. The spectrometer used was FieldSpec4 (ASD Inc., spectral range: 350–2500 nm; spectral resolution: 3 nm at 700 nm and 8 nm at 1400 and 2100 nm), and the radiance of objects was measured vertically downward. The reflectance of each object was derived from the radiance ratio of the white reference plate (SRT-99-050 from Labsphere, Inc., North Sutton, US) and the object.

3.2.3. Extraction of Marine Debris from Satellite Data

In this study, to detect plant-derived marine debris, we first calculated FAI using Hu's equation [9]:

$$\mathrm{FAI} = R_{rc}(NIR) - R'_{rc}(NIR) \quad (9)$$

$$R'_{rc}(NIR) = R_{rc}(RED) + (R_{rc}(SWIR) - R_{rc}(RED))\frac{(\lambda_{\mathrm{NIR}}\text{-}\lambda_{\mathrm{RED}})}{(\lambda_{\mathrm{SWIR}}\text{-}\lambda_{\mathrm{RED}})} \quad (10)$$

where R_{rc} is the Rayleigh-corrected reflectance; however, in this study, the level-2 aerosol-corrected reflectance was used instead of the Rayleigh-corrected one. RED indicates red (OLI Band 4); NIR indicates near-infrared (OLI Band 5); SWIR indicates mid-infrared (OLI Band 6); and λ indicates wavelength. This formula is based on the fact that the high NIR reflectance (A) of a plant-derived substance and the slope of the reflectance of background water with the same wavelength as the

baseline (B) formed by the NIR and SWIR reflectances are almost equal. We assumed that A–B can be used as an index of marine debris.

However, this assumption may not hold in the case of turbid background water, as is seen after heavy rain. To solve this problem, we used the method proposed by Hu et al. [44]. For the MODIS-derived FAI image of the Yellow Sea, Hu et al. [44] determined the FAI of background seawater (FAIsw) using two criteria and successfully detected drifting algae while suppressing false detection due to turbid water. Our FAI calculation performed according to the method of Hu et al. [44] is briefly described below.

For the first criterion, we calculated the difference between the gradient of the FAI image and the gradient of the R_{rc} (RED) image for each pixel (this difference is hereinafter referred to as the corrected gradient of FAI (cGFAI)) Subsequently, we considered the pixels with cGFAI below a certain threshold (TcG) as pixels of seawater without marine debris. For the gradient of FAI and R_{rc} (RED) images, the following equation given by Hu et al. [44] was used:

$$\nabla y_i = \sqrt{\frac{1}{8}\sum_{j=1}^{8}\left(\frac{y_i - y_j}{x_{i,j}}\right)^2}, \; j \neq i, \tag{11}$$

where ∇y_i is the magnitude of the gradient of the ith pixel of the image y (for example, FAI); y_j is the FAI value of the eight pixels adjecent to the ith pixel; and $x_{i,j}$ is the distance between the ith pixel and the eight adjacent pixels (i.e., 1 or 1.414 times the spatial resolution). The threshold TcG is the maximum cGFAI for the seawater that does not contain algae or marine debris. Hu et al. [44] calculated cGFAI for long-term data from several months without drifting algae and determined TcG as the long-term cGFAI with a cumulative frequency of 0.99 in all calculated frequency distributions. However, in this study, we could not obtain multiple good-quality Lansat-8 data from before the disaster. Hence, we calculated TcG using a Landsat image from April 20, 2018, which had the best image quality among images whose obtained dates are close to the heavy rain event. For the second criterion, we calculated the mean and standard deviation of the FAI in a 15 × 15 pixel kernel centered on the pixel of interest. If the FAI of the pixel of interest is less than that of the kernel (i.e., mean + 2σ), we judged that this pixel did not contain marine debris. Using these two criteria, we can divide an FAI image into two types of pixels—one

that may contain marine debris and one that corresponds to seawater. For a pixel regarded as seawater, the FAI of that pixel was used as FAI_{sw}. The FAI_{sw} of a pixel that may contain marine debris was taken as the average value calculated using only the pixels that do not contain marine debris in the 15 × 15 kernel centered on that pixel. Finally, the background effect was removed by subtracting FAI_{SW} from the FAI for each pixel. In this study, the FAI obtained by subtracting this FAI_{SW} was called the corrected FAI (cFAI), following Hu et al. [44].

Furthermore, it was necessary to determine the threshold value to create a marine debris map from the obtained cFAI. Many of the previous methods involved searching for the threshold value by trial and error, which lacks objectivity. Therefore, we adopted the Otsu method [48], which objectively determines the threshold value. This method is one of the main image-processing methods used for automatically determining a threshold value. The procedure is as follows.

1. Obtain the histogram of the cFAI image.
2. Calculate the minimum value (Imin), maximum value (Imax), and average value (μ_0) from the histogram.
3. Determine an appropriate threshold value T within the range of Imin and Imax.
4. Divide the histogram into two classes according to the threshold value T.
5. Obtain the variance (σ_1^2 and σ_2^2), average (μ_1 and μ_2), and number of pixels (n_1 and n_2).
6. Obtain the intraclass variance σ_w^2 and the interclass variance σ_b^2 from the following equations.

$$\sigma_w^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2} \tag{12}$$

$$\sigma_b^2 = \frac{n_1(\mu_1 - \mu_0)^2 + n_2(\mu_2 - \mu_0)^2}{n_1 + n_2} \tag{13}$$

7. From the two variances obtained in step 6, the degree of separation S of the following equation is obtained.

$$S = \frac{\sigma_b^2}{\sigma_w^2} \tag{14}$$

8. Repeat steps 4 to 6 to find all T values with a degree of separation S within the

range of minimum to maximum.

9. The T when the degree of separation S reaches its maximum is determined as the threshold value and is used for binarization processing.

Figure 3.2 shows the flow of marine debris extraction from Landsat-8 data and the conceptual diagram of cFAI threshold determination by the Otsu method.

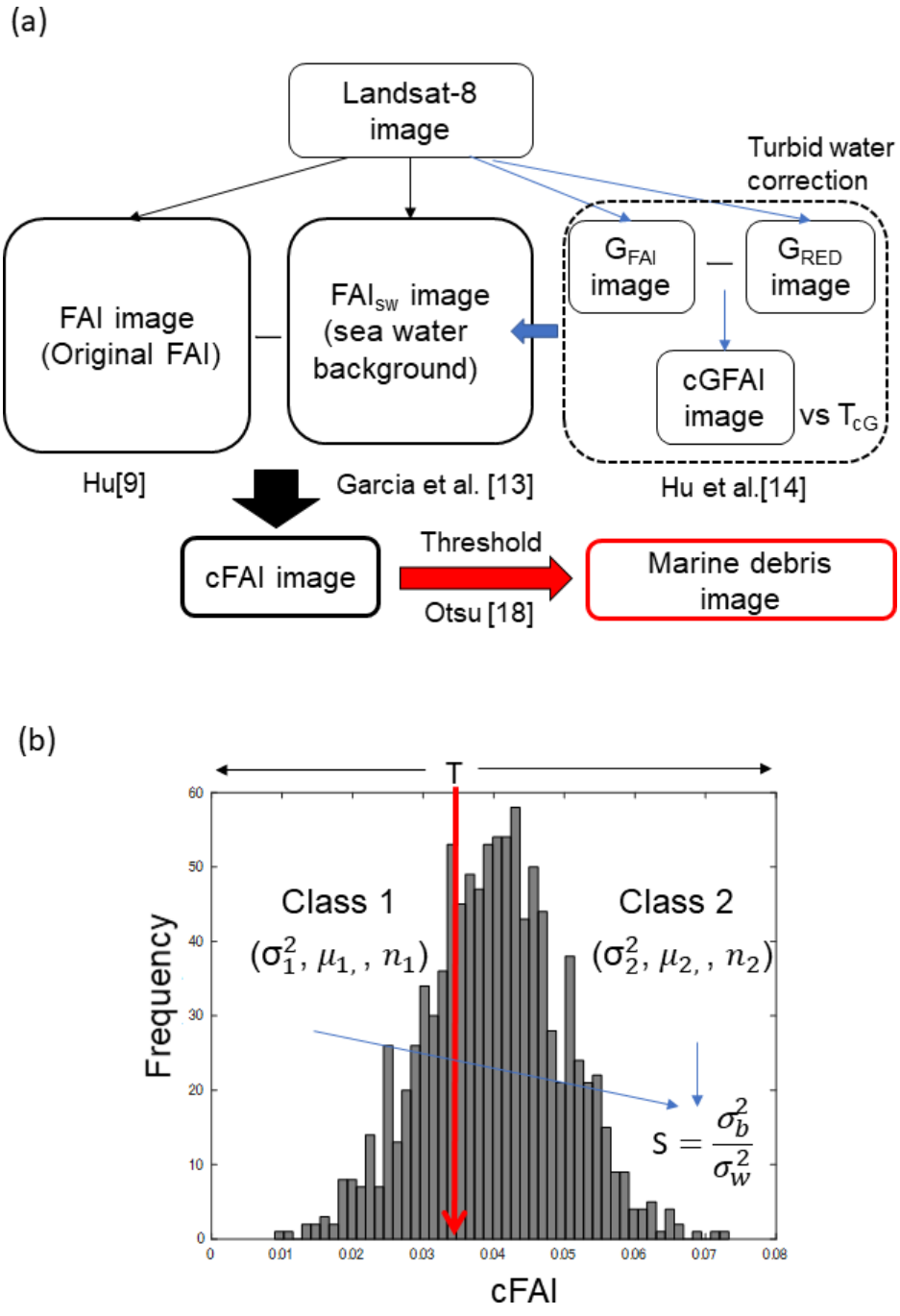


Figure 3.2. (**a**) Flow of marine debris extraction by Landsat-8 data and (**b**) conceptual diagram of corrected floating algae index (cFAI) threshold determination by the Otsu method.

3.3. Results

3.3.1. Marine Debris Collection Status by Cleaning Ships

Figure 3.3 shows the temporal change of marine debris distribution (over about 1 month) after the heavy rain event (see source information listed in Table 3.2). An overview of the marine debris acquisition locations determined by the cleaning ships is shown in the figure. Figure 3a,b shows the distribution of marine debris acquisition positions in the first half of the study period (i.e., 9 July to 23 July 2018), and Figure 3c shows the distribution of marine debris acquisition positions in the second half (i.e., 24 July to 14 August 2018). These figures show that, throughout this period, marine debris was floating in the same water area of the Seto Inland Sea off the coast of Hiroshima Prefecture. The debris frequently accumulated, in particular, in the area surrounded by the land and islands in the northern part of Aki Nada (the water area surrounded by the dotted line in Figure 3.3). Figure 3.4 summarizes the amount of garbage collected (mainly plant fragments) every week during this period. Assuming that 100% of the marine debris was collected by 14 August, the recovery rate was about 10% by 14 July, but more than 60% was recovered by 23 July. Thereafter, the amount of debris collected continued to decrease every week.

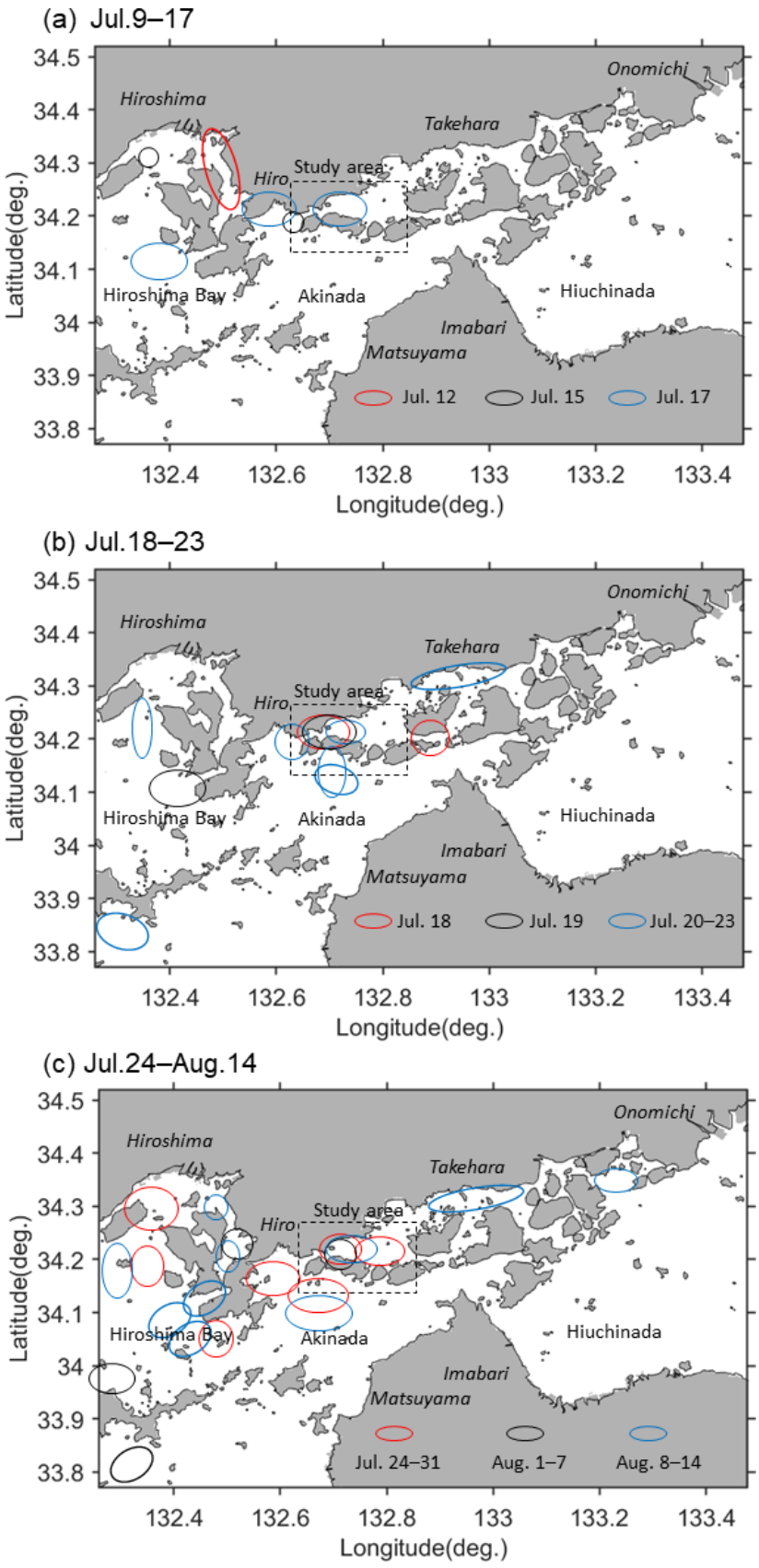


Figure 3.3. Summary of marine debris acquisition locations as determined by the cleaning ships.

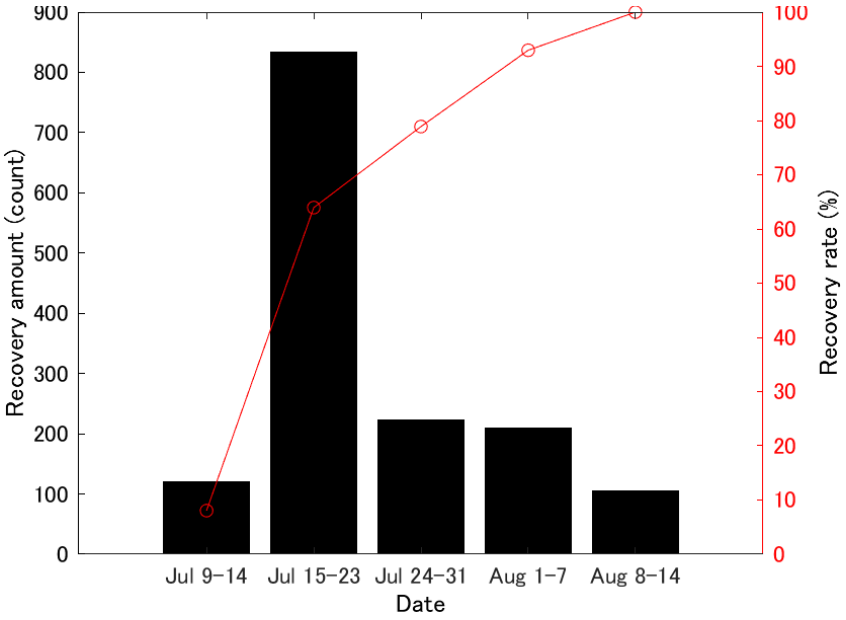


Figure 3.4. Time series of the amount of marine debris collected by the cleaning ships. The bar graph and line graph show the number of plant fragments and the recovery rate, respectively.

3.3.2. Spectral Characteristics of Marine Debris

Figure 3.5 shows an RGB color composite image of the Landsat-8 image taken on 9 July 2018, immediately after the heavy rain in the study area. This image shows floating objects that look red in color (examples: A, B, and C in Figure 3.5). The spectral reflectance characteristics extracted from the images at these three points are shown at the bottom of Figure 3.4. At each point, the spectral reflectance of three points of suspended matter (marked by red circles) and three points of background water (marked by blue circles) were obtained. From these spectral reflectance values, the spectral reflectance (red line) of the suspended matter was characterized by high reflectance in band 5 (central wavelength 865 nm). In addition, the reflectance gradually increased toward the long wavelength from band 1 to band 4, and the positive slopes of band 4 (central wavelength 655 nm) and band 5 were high. By contrast, the spectral reflectance of the background water was stable with a value lower than that of the suspended matter (blue lines in the bottom panels of Figure 3.5). With band 3 (central wavelength 560 nm) as the maximum value, the reflectance values at longer wavelengths (bands 4–7) were smaller than those at shorter

wavelengths (bands 1 and 2). For bands 4–7, the reflectance values increased with wavelength. The slopes of band 4 and band 6 (central wavelength 1610 nm) used as the baseline of FAI tended to be similar for suspended matter and background clear water.

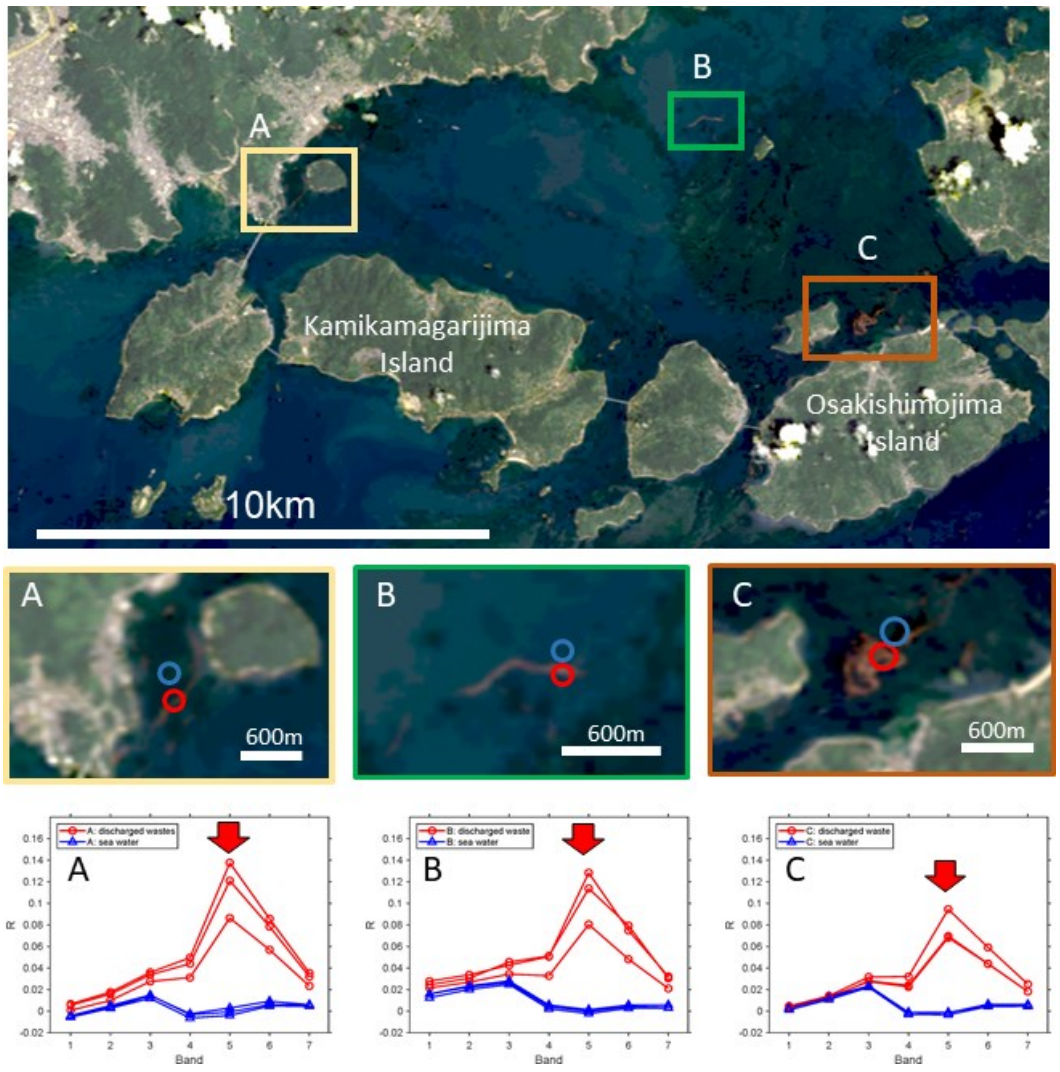


Figure 3.5. Landsat-8 RGB color composite image acquired on 9 July 2018 immediately after the heavy rainfall disaster. The areas labeled A, B, and C in the top panel are enlarged in the middle panels. The bottom panels show the spectral characteristics of the pixels in the red circle (suspended matter) and blue circle (background water) in each middle panel.

Figure 3.6 shows the measured spectral reflectance characteristics of marine debris along the coast of Hiroshima Prefecture. Figure 3.6a shows the measured original hyperspectral data, and Figure 3.6b shows the data obtained by extracting

only the Landsat-8 band data. Unlike the results shown in Figure 3.6, driftwood targeted in this study had the maximum value in band 6; however, the difference between bands 4 and 5, which is the basis of the height of FAI, is considerably larger than that between other substances.

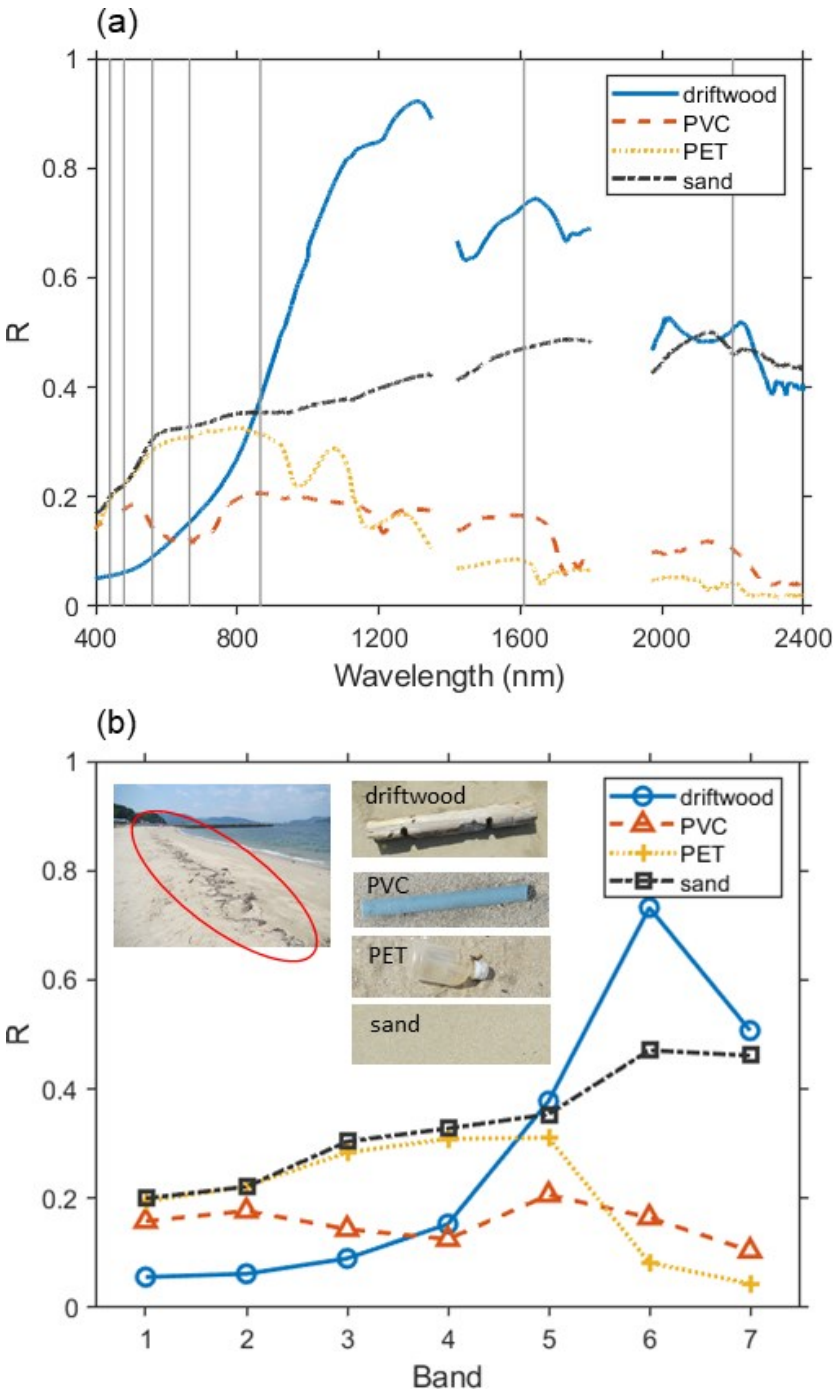


Figure 3.6. Example of spectral characteristics of the marine debris collected from the coast. (**a**) Hyperspectral data and (**b**) the data of Ladsat-8 band only.

3.3.3. Marine Debris Detection Results Obtained Using Satellite Data

Figure 3.7 shows (a) FAI, (b) cFAI and (c) marine debris images calculated from Landsat-8 data on 9 July 2018. Similarly, Figure A1 (in Appendix) shows the calculation results for the same wide area as that shown in Figure 3.3. From Figures 3.7 and A1, we can confirm the presence of innumerable linear marine debris that cannot be visually read in Figure 3.5 from the FAI image in (a). However, it is possible to confirm a high FAI water area that is difficult to distinguish from the background water, such as water area A. In the cFAI image in Figure 3.7b, only marine debris, which excludes the influence of background water, was detected. In the automatic binarized image of Figure 3.7c, the relatively high cFAI value detected by cFAI can be clearly observed.

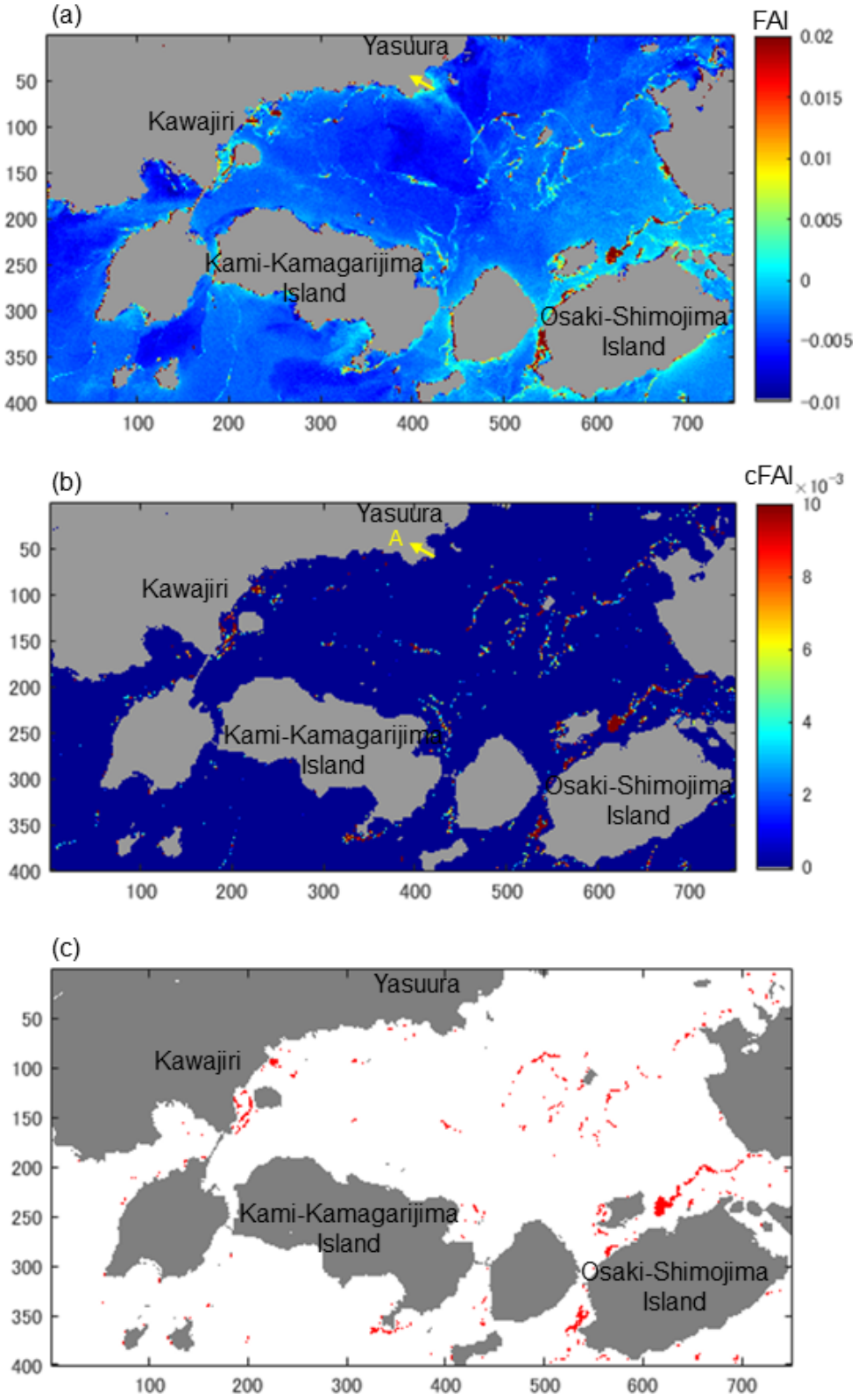


Figure 3.7. (**a**) FAI, (**b**) cFAI, and (**c**) marine debris images calculated from Landsat-8 data acquired on 9 July 2018.

Figure 3.8 shows the threshold comparison of cFAI when the Otsu method is applied to the acquired images of locations near Osaki-Shimojima at (a) 9 July, (b) 16 July, (c) 25 July and (d) 1 August 2018. The cFAI thresholds determined by using the Otsu method show similar values with small differences for each date (0.002 to 0.008). Without the use of the Otsu method, since the distribution of the cFAI histogram is not a bimodal distribution but a continuous one, it is difficult to determine the threshold value by trial and error. Furthermore, when we applied the Otsu method for FAI and cFAI on 9 July 2018, as shown in Figure 3.9, a considerably larger threshold value was obtained for FAI (0.06) than for cFAI (0.008). With such high threshold for FAI (0.06), most of the suspended matter could not be detected from the corresponding FAI image (see Figure 3.7a).

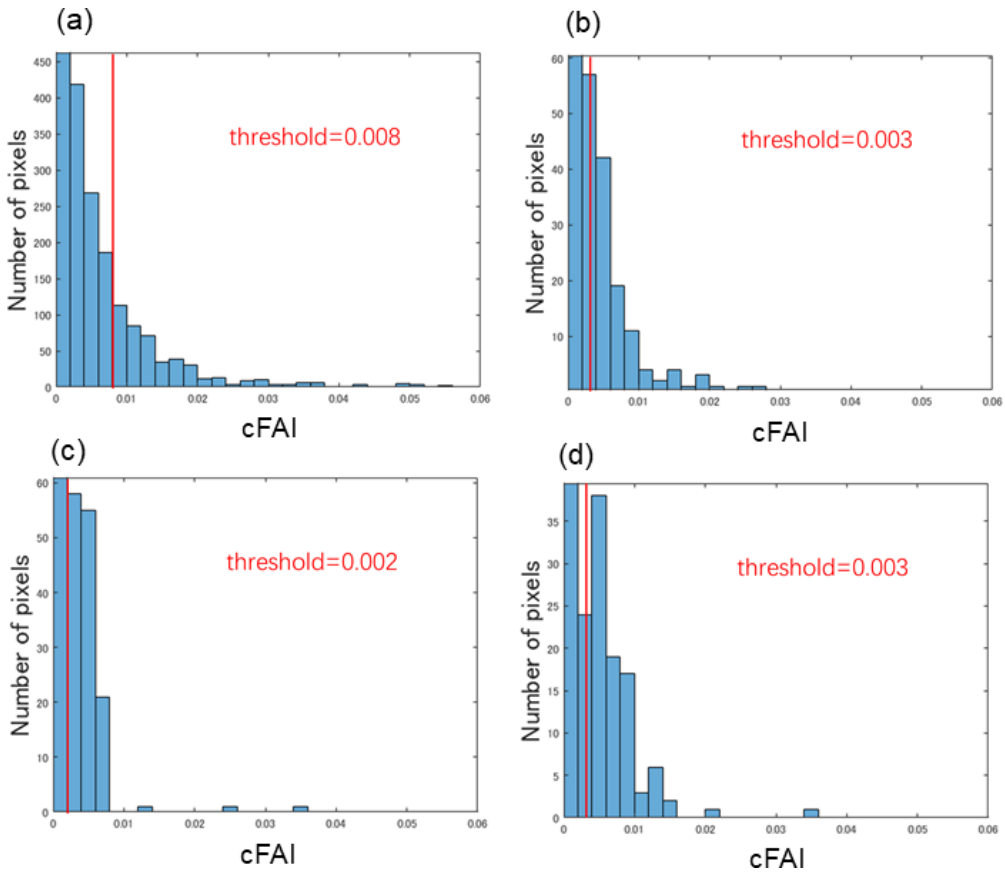


Figure 3.8. Comparison of the cFAI thresholds obtained by the Otsu method for four periods: (**a**) 9 July (**b**) 16 July (**c**) 25 July, and (**d**) 1 August 2018.

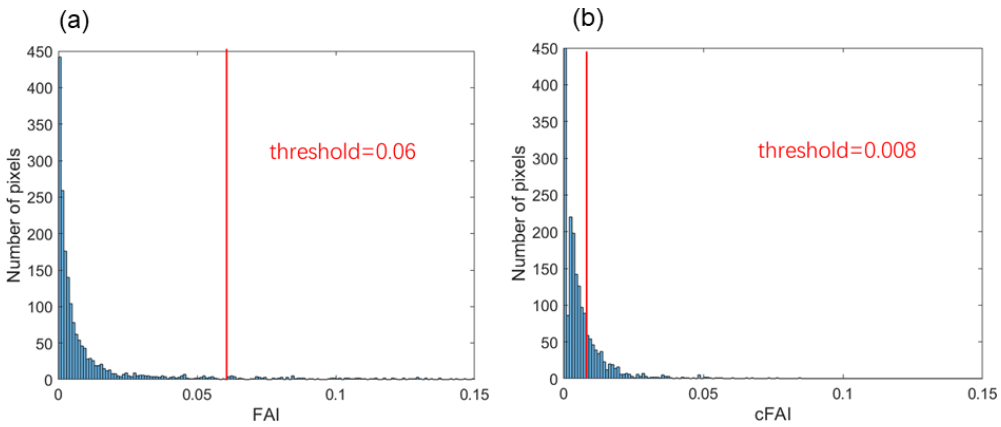


Figure 3.9. Comparison of the thresholds of FAI and cFAI data obtained by the Otsu method for 9 July 2018.

Figure 3.10 shows the change in marine debris distribution near Osaki-Shimojima detected using the cFAI threshold as shown in Figure 3.7. Figure 10 shows that within this range, on 9 July, the distribution extended 259,200 m^2 (about 0.26 km^2), and marine debris continued to decrease every week. About 3 weeks later, on 1 August, the debris extended only about 83,700 m^2 (0.08 km^2), which is about one-fourth of the extent on 9 July. In addition, the significant marine debris accumulation seen in the northern part of Osaki-Shimojima Island (point C in Figure 3.5) on 9 July was not observed 1 week later on 16 July.

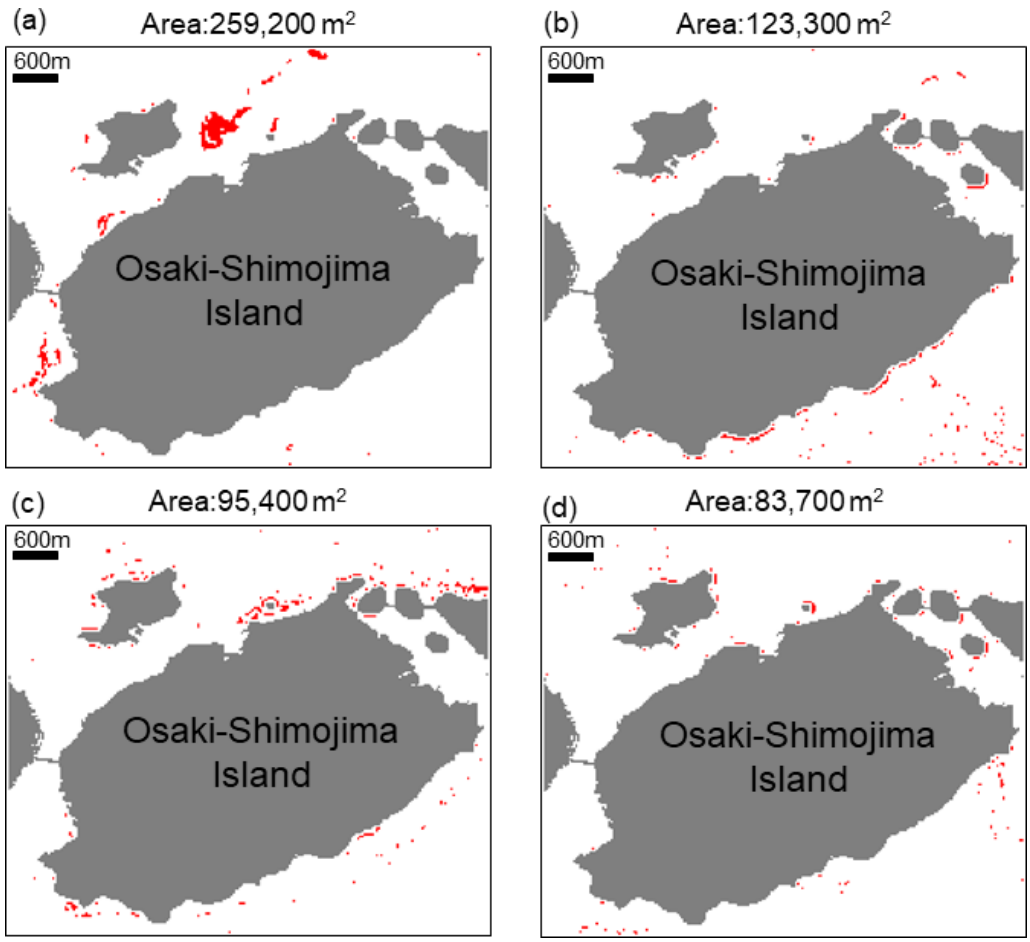


Figure 3.10. Changes in the marine debris distribution near Osaki-Shimojima Island using Landsat-8 data. The debris distributions on (**a**) 9 July, (**b**) 16 July, (**c**) 25 July, and (**d**) 1 August 2018 are shown.

3.4. Discussion

3.4.1. Validity of the FAI-Based Method for Marine Debris Detection

The spectral characteristics of marine debris as obtained from Landsat-8 images are characterized by high reflectance in band 5 (around 865 nm) as shown in Figure 3.5. These characteristics are almost identical to the spectral characteristics of the floating marine algae detected using the FAI [43]. Therefore, it can be concluded that marine debris is detected relatively reasonably by FAI, as shown in Figure 3.7a. Conversely, the spectral characteristics of driftwood measured along the coast reach their maximum value in band 6 (around 1620 nm) as evident in Figure 3.6. This means that if driftwood floats in the sea without getting wet, it is better to consider a method for detecting marine debris using band 6. However, it is probable that most

of the marine debris that flowed out after the heavy rain in the studied event was pieces of wet wood or thin pieces of wood such as reeds (see the images provided in the sources listed in Table 3.3). Oyama et al. [49] analyzed the measured spectral reflectance from 400 to 2400 nm and Landsat-5 data of large aquatic plants to observe the maximum reflectance near 1600 nm (Landsat-5 band 5). The FAI showed the effectiveness of separation of the target from the background water. These facts indicate that the FAI-based method is appropriate for marine debris detection. Note that, in this study, we did not intend precise classification of the types of debris. The detected debris (in Figures 3.7 and 3.10) included multiple types of debris that have red edge characteristics. For the purpose of classification of marine debris, Biermann et al. [50] suggested that the use of NDVI leads to better classification results than the use of FAI (FDI; floating debris index in their paper).

3.4.2. Validity of Marine Debris Detection by cFAI and Otsu Methods

Figure 3.7a shows that the detection of marine debris by the original FAI method [43] yields many places where the values of marine debris and background water are similar. In particular, when the background of the marine debris was seawater that has become turbid because of heavy rain, it was easier to extract marine debris using cFAI, as seen from Figures 3.7b and 3.9. This indicates that the methods of Hu et al. [44] are effective for marine debris extraction. When setting cGFAI that calculates the spatial gradient of reflectance, images with a relatively low resolution, such as MODIS data, may be more effective for detecting large-scale marine debris. This is because if the spatial resolution is too high, the gradient will be zero. However, most of the marine debris treated in this study was linearly distributed, and the scale was relatively small (Figure 3.7a); hence, it is considered that cGFAI treatment was not very difficult. Moreover, the histogram of cFAI in the target water area (Figure 3.8) clearly shows that it is difficult to determine the threshold value objectively from the shape. The Otsu method automatically determines the threshold value and provides reliable results such as those shown in Figures 3.7c and 3.10. Thus, its use as an automatic threshold value determination in the event of such a rainfall disaster is validated.

3.4.3. Reason for Change in the Amount of Marine Debris Detected

On the premise that the marine debris collected was the total amount present, the rate of recovery of the marine debris by the cleaning ships by 14 July was about 10%; however, by the end of July, about 80% of the marine debris was recovered, as shown in Figure 3.4. Figure 10 shows that the amount of debris recovered decreased. A large amount was recovered on 9 July , while one-fourth of that amount was recovered on 1 August. This result is assumed to be the effect of marine debris collection by the cleaning ships and the result of some of the debris being washed away to the open ocean or settling on the seabed because of the strong tidal current [13,14] in the Seto Inland Sea.

3.4.4. Limitations of Marine Debris Detection Using Landsat-8 Data

In the event of a disaster, it is important to grasp the current situation of debris distribution over a wide area, and, as shown in this study, the Landsat-8 data are useful for this purpose. A limitation is its temporal resolution; the observation cycle of Landsat-8 is 16 days, and even if overlapping paths are present, as in the present case, the observation frequency is only about once a week. Hence, marine debris locations detected by Landsat-8 data are approximate information that cannot be completely relied upon for practical recovery works. This is because the tidal current is strong in the Seto Inland Sea, and the marine debris is quickly displaced from the locations detected by Landsat-8 data. As shown in Figure 3.10, the large amount of marine debris identified on 9 July is not found in the data after a week. While we presume that the cleaning ship recovered the marine debris during this period, it is possible that it was dispersed from that area by the strong tidal currents. The fate of this marine debris (in Figure 3.10a) was not provided by the weekly Landsat-8 data in this study. In the Seto Inland Sea, which has strong tidal currents, we believe that at least hourly images are required to follow marine debris. Providing realtime-processed debris locations on an hourly basis helps aids in efficient recovery.

There are satellites and sensors that provides hourly images. For example, the Geostationary Ocean Color Imager (GOCI) onboard the COMS satellite is a satellite sensor with a high repeat cycle and can capture images at 1-h intervals. Although GOCI does not have a band corresponding to band 6 of Landsat-8, Son et al. [51] and

Qiu et al. [52] have made it possible to employ GOCI data to detect floating algae using FAI. Therefore, it can be assumed that marine debris can be detected at an early stage. However, because the spatial resolution of GOCI is 500 m, only large marine debris can be detected. Landsat-8 has a spatial resolution of 30 m, enabling it to detect marine debris better than GOCI, but it is still difficult to detect small-scale marine debris. Using sensors with insufficient spatial resolution would lead to the underestimation of amounts of marine debris. However, satellites or sensors with both sufficient temporal and spatial resolution are not yet available. Therefore, research that combines Sentinel-2 mission with a spatial resolution of 10 m and satellites with a spatial resolution of about 50 cm to 3 m such as DOVE, WorldView, and Pleiades [53] will be important.

3.5. Conclusions

In this study, an attempt was made to reproduce the actual state of marine debris distribution in the Seto Inland Sea off the coast of Hiroshima Prefecture during the heavy rains in July 2018. For this purpose, marine debris acquisition information from cleaning ships and Landsat-8 images was used. The following conclusions were drawn from our work.

1. From the data acquired by the cleaning ships on multiple days, the distribution of marine debris immediately after the heavy rain was approximated. In particular, the debris was concentrated in the water area surrounded by land and islands in the northern part of Aki Nada.
2. From the spectral reflectance data of Landsat-8 level-2, we confirmed that the marine debris had a high peak reflectance in band 5 (central wavelength 865 nm).
3. Unlike the original FAI method, the cFAI method enabled us to remove the background water signals from the Landsat-8 images.
4. The Otsu method for the automatic binarization of cFAI was effective in detecting marine debris from Landsat-8 images because it set an appropriate threshold value.

In the future, we would like to use data such as GOCI data, which obtains data more frequently than Landsat, and WorldView-2, which has a high spatial resolution. Using such data, we will aim to track the detailed movement of marine debris

(movement vector) and achieve higher spatial resolution. In recent years, heavy rainfall disasters have occurred frequently in Japan, and we would like to validate the effectiveness of this method in water areas other than the Seto Inland Sea.

Chapter 4: Distinguishing Sargassum and Zostera Using Sentinel-2 Data

4.1. Introduction

4.1.1. Backgrounds

Seaweed and seagrass (collectively referred to as SWSG) beds play a crucial role in the proliferation of aquatic resources owing to their functioning as spawning grounds for aquatic organisms, such as young and larval fish. Moreover, recently, marine forests have received renewed attention as sites of blue carbon, which absorbs inorganic carbon from water and supplies oxygen [54–57]. However, in the Seto Inland Sea, which is the study area, regulations on the discharge of nutrients and other salts have been implemented for many years to reduce the eutrophic state represented by red tide [58], which became a problem in the 1960s and 1970s. Consequently, water quality has improved considerably, except in certain areas; however, there have been problems, such as sluggish fish catches and discolored SWSG bets, owing to oligotrophication [59]. Under these circumstances, the "Act on Special Measures concerning Conservation of the Environment of the Seto Inland Sea" [60], which was amended and enforced in April 2022, aims to eliminate the lack of nutrients in the sea and promote "blue carbon". First, the current status of SWSG bed distribution throughout the year must be understood accurately to achieve this goal. However, the Seto Inland Sea has a complex topography, containing 600 islands at an average depth of 30 m [61]. Furthermore, it exhibits large regional and seasonal tidal differences (approximately 1–3 m and 3–4 m in the eastern and western sea areas, respectively). Conducting quick field surveys over a wide area is difficult and incurs enormous financial costs. The Ministry of the Environment, Government of Japan (MOEJ), compiles the date of field surveys that are reconducted once every few years and publishes a SWSG beds distribution map on the Internet. However, the seasonal changes over a wide area cannot be determined. Therefore, satellite remote sensing technology that can repeatedly grasp the distribution of SWSG beds over a wide area is considered promising for providing accurate estimates of blue

carbon. The difficulty in understanding the distribution of SWSG beds using satellite remote sensing is that the influence of light attenuation caused by water is extremely large when determining the presence of SWSG compared to that of the land cover classification on land. Additionally, as the attenuation rate of light underwater is greater for long wavelengths, above-water and information at wavelengths longer than that of near-infrared regions cannot be used and almost only information from light in the visible range can be used. Furthermore, because different types of SWSG often exhibit similar colors, distinguishing species based on their spectral characteristics is challenging. However, these weaknesses have been overcome and SWSG bed distribution maps are now being created using several satellite remote sensing-based several methods.

4.1.2. Previous Studies

There are broadly two types of methods for extracting SWSG beds using remote sensing. To extract SWSG, one method involves using existing water depth data, performing water depth correction, and then applying terrestrial land cover classification methods, such as unsupervised or supervised classification, LAI (leaf area index), which can measure leaf area based on optical assumptions regarding light attenuation, and NDVI (Normalized Vegetation Index), which measures plant activity based on red absorption and near-infrared scattering from the leaf, and so on [33,62]. The other method eliminates the influence of water depth and extracts SWSG only from image calculations without using water depth data [29,30]. The former method requires accurate water depth data that considers tidal level changes on the day the satellite image is captured. However, obtaining detailed water depth data in shallow waters is difficult. Thus, water depth estimation using images is often attempted [63,64]. However, this method exhibits poor accuracy in estimating water depth, particularly under conditions of a water depth of 4–5 m or greater. Moreover, it results in excessive water depth correction [unpublished]. The latter method is convenient because it facilitates the extraction of SWSG from satellite images without water depth data. However, it is generally difficult to identify SWSG species because it uses minute differences in the log reflectance ratio of the bottom sediment. Therefore, the latter method is used to extract the SWSG bed area in sea areas, such

as the Seto Inland Sea, wherein fairly accurate water depth and water level data can be obtained. In addition, in shallow sea areas of 10 m or less, the water depth data are corrected to determine the spectroscopic SWSG area. A hybrid method that performs species discrimination is considered to be effective. Currently, the MOEJ creates a marine forest map along the coast of Japan using field surveys and high-resolution satellite data, and the maps are classified into *Sargassum* bed, *Zostera* bed, and kelp bed [65]. There are also studies focused on the optical detection of *Sargassum* and *Zostera*. Gower initially identified *Sargassum* using the MERIS Maximum Chlorophyll Index [66]. Hu utilized the Floating Algae Index (FAI) with MODIS image data to detect *Sargassum* [27]. Gower and King employed the red-edge concept for *Sargassum* detection, acknowledging that other marine organisms can also cause elevated red-edge reflectance [67]. Taking into account the characteristic absorption feature of *Sargassum* at 630 nm, distinct from other floating vegetation, Dierssen et al. developed the *Sargassum* Index (SI), which uses the bands of 650 and 630 nm [68]. Field measurements by Hu indicated that *Sargassum* exhibits a distinctive reflectance curvature around 630 nm due to its chlorophyll c pigments [68]. While these methods have effectively diagnosed *Sargassum* slicks using pre-determined thresholds, a method for distinguishing between *Sargassum* and *Zostera* in the Seto Inland Sea using Sentinel-2 data has not been proposed. Yoshida's study on *Sargassum* and *Zostera* in the Seto Inland Sea reveals that *Zostera* thrives in shallow sandy areas, while *Sargassum* prefers deeper, rocky, and intertidal zones. In the Seto Inland Sea, *Sargassum* is mainly composed of three species: *S. condusum*, *S. yezoense*, and *S. trichophyllum*, each with distinct maturity periods ranging from March to June, June to August, and October to December, respectively [69].

4.1.3. Objectives

Against this background, this study developed and verified a method able to spectroscopically distinguish *Sargassum* and *Zostera* in the sea using Sentinel-2 data, which exhibit the highest spatial resolution among currently available free data. This study aimed to understand changes in the distribution over time. To achieve this goal, we first considered detecting marine forests from Sentinel-2 data using a method that did not require water depth data. Subsequently, we proposed a method to separate

Sargassum and *Zostera* based on their spectral reflectance measured in the field. Finally, the distributions of *Sargassum* and *Zostera* mapped using this method were integrated and verified through comparisons with existing marine forest maps. If such a method is valid, it will be possible to repeatedly update existing marine forest maps using satellites.

4.2. Materials and methods

4.2.1. Study area

The study area was the Seto Inland Sea off the coast of Takehara City, located in the south-central part of Hiroshima Prefecture, as shown in Figure 4.1. According to Yoshida et al. [16], the SWSG beds area of the "Bingo–Geiyo–Seto" fishery area (conducted from 1989 to 1991) is 3636 ha. The total areas of seagrass, SWSG, and kelp beds are 2917 ha (80%), 1289 ha (35%), and 184 ha (5%), respectively (as there is mixed marine forest, the total does not add up to 100%). Thus, the *Sargassum* and *Zostera* beds are the main SWSG beds in the central Seto Inland Sea. In addition, as for their habitat, *Zostera* beds are formed in calm sandy mud areas with little wind and waves, whereas *Sargassum* (formed by brown SWSG *Sargassum*) is formed in rocky reef areas. In the Seto Inland Sea, there are many areas of water wherein sand and mud and rocky reef areas exist alternately; thus, *Sargassum* and *Zostera* beds coexist. These statistics are data for shallow waters below 10 m in depth, and as the marine forest in Figure 1 also exists mostly in shallow waters below 10 m, this study targeted waters below 10 m. In addition, the tidal level difference in the target water area is approximately 4 m at maximum.

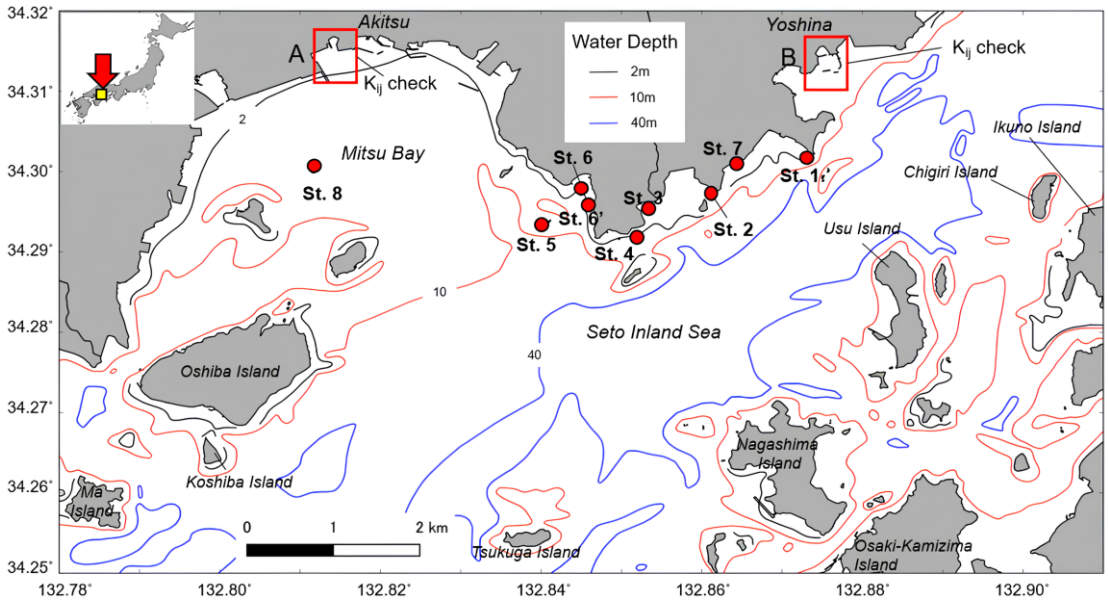


Figure 4.1. Study area depicting field survey points and water depth. This image is extracted from the yellow square in the top left corner. St. 1-St. 8 indicate the locations where we conducted field survey.

4.2.2. Distinguishing Method of Sargassum and Zostera Using Satellite Data

Figure 4.2 shows the procedure for creating a marine forest distribution map used in this study. To briefly explain the overall procedure, first, the acquired satellite data were processed as a land mask using near-infrared threshold processing and a cloud mask using Quality flag data. After correcting the water depth, the SWSG bed area was extracted. Furthermore, the *Sargassum* and *Zostera* Distinguishing Index (SZDI) method was proposed and applied to the extracted seagrass beds to separate them into *Zostera* and *Sargassum* beds. The most important water depth correction in this process is explained later. The specific method of SZDI is discussed later based on the results of the study.

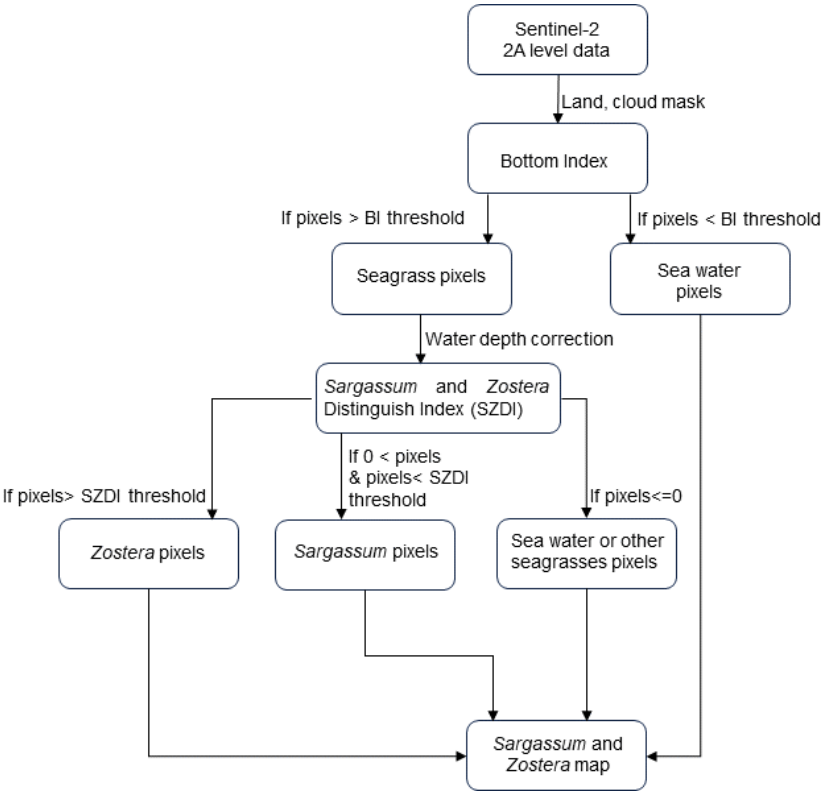


Figure 4.2. Flow chart of the study method employed.

In general, in underwater analysis, the influence of water dissipation is considerably greater than in land analysis; thus, in waters with a complex seafloor topography (the Seto Inland Sea), stable SWSG beds cannot be extracted without depth correction. Therefore, in remote sensing, there exist two methods: one involves using existing seafloor topography data, converting the underwater reflectance to be equivalent to the reflectance on land, and then performing analysis. The other method utilizes the surface reflectance ratio of two different bands to eliminate the influence of water depth and distinguish between various bottom sediments. Studies have attempted to remove the influence of water depth and extract the bottom sediment (in this case, SWSG beds). This method is referred to as the bottom index (BI) or depth invariant index (DII) [29,30].

First, we explain the former method, which uses seafloor topography data. If seafloor topography data exist, the most basic relationship between the reflectance R immediately above the water surface (equivalent to the atmospherically corrected reflectance of the satellite) and the seafloor reflectance information can be theoretically obtained by using the following formula [70]:

$$R_i = R_{\infty i}[1 - \exp(-2k_i z_B)] + R_{Bi}\exp(-2k_i z_B), \quad (15)$$

where R_∞ is the reflectance at an infinite water depth, k is the extinction coefficient of the target seawater, R_B is the underwater extinction coefficient of the band, Z_B is the seafloor depth, and the subscript i is the band. Furthermore, the first term on the right-hand side of Equation (1) accepts an extremely small value compared to the second term on the right-hand side. Therefore, R_B at a water depth of 0 m (equivalent to ground level) can be calculated approximately using the following formula. Hereinafter, such a correction is referred to as "water depth correction".

$$R_{Bi} \fallingdotseq \frac{R_i}{\exp(-2k_i z_B)}. \quad (16)$$

However, Z_B and k on the day of satellite observation must be known when using this formula. There exists a method to calculate k from the relation between Z_B and the brightness value of the image (the slope when approximated by a logarithmic function); however, if Z_B exceeds several meters, it results in overcorrection. It is extremely difficult to correct the water depth even when targeting topography below 10 m (as in this case). Therefore, we assumed that the extinction coefficients for the three periods were constant. Moreover, a correction coefficient m was introduced to prevent the value of each band from falling below 0, and corrections were made for the three periods using the following formula:

$$R_{Bi} \fallingdotseq \frac{R_i}{\exp\left(-2m_i k_{sw_i} Z_B\right)}, \quad (17)$$

where k_{sw} is the general extinction coefficient of seawater.

Next, we explain the latter method, which does not use bathymetric data. The BI method is based on the theory proposed by Lyzenga [29,30]. The bottom sediment index is a model developed based on the assumption that "if the bottom sediment is the same, the bottom sediment reflectance ratio (natural logarithm type) of two different bands will be constant". A brief explanation of the theory of sediment

indicators is given below. Lyzenga expressed the radiance ($L_{s\lambda}$) detected by a sensor mounted on a satellite in shallow water at wavelength λ using the following formula:

$$L_{s\lambda} = L_{B\lambda} - E_{\lambda} R_{B\lambda} \exp(-2k_{\lambda} z_B), \tag{18}$$

where $L_{B\lambda}$ is the radiance observed at a deep point in the water, E_{λ} is the solar irradiance reaching the ground, $R_{B\lambda}$ is the spectral reflectance of the bottom sediment, k_{λ} is the spectral extinction coefficient of water, and Z_B is the water depth. Equation (4) is extended to two wavelengths and the effect of water depth is removed by calculating the ratio. Therefore, the radiance observed by the satellite in two different bands *i* and *j* is expressed by the following equations using Equation (4).

$$L_{si} = L_{Bi} - E_i r R_{Bi} \exp(-2k_i z_B) \tag{19}$$

$$L_{sj} = L_{Bj} - E_j R_{Bj} \exp(-2k_j z_B). \tag{20}$$

Rearranging Equations (5) and (6), we obtain the following equation with water depth *z* removed:

$$LN\left[\frac{R_{Bi}E_i}{(R_{Bj}E_j)^{kij}}\right] = LN\left[\frac{(L_i - L_{si})}{(L_j - L_{ji})^{kij}}\right], \tag{21}$$

where most of the fluctuations on the left side are controlled by the "ratio of seafloor reflectance between the two bands" and contain information related to the bottom sediment. Therefore, this is defined as the BI and is set as the portion on the right side. Consequently, the bottom sediment index can be calculated using the following formula:

$$BI_{ij} = LN\left[\frac{(L_i - L_{si})}{(L_j - L_{ji})^{kij}}\right] = LN(L_i - L_{si}) - k_{ij} LN(L_j - L_{sj}). \tag{22}$$

Using Equation (8), the extinction coefficient ratio k_{ij} can be defined as the band ratio (the slope when the logarithm of the brightness values of the different bands is plotted) from the relation between the satellite data of two different bands. Theoretically, the influence of water depth can be calculated from the image alone. Consequently, the sediment (in this case, SWSG beds), from which the 200 nm has been removed (in the case of sediments with different reflectance ratios), can be distinguished. The actual extinction coefficient ratio k_{ij} is calculated at locations with the same bottom material but at different water depths (often using several pixels offshore from a sandy beach). If the slope is derived from the scattering value of the

logarithm of the luminance ratio of the two wavelengths (L_s corrected), it is equivalent to the theoretical extinction coefficient ratio. We assumed that the bottom sediment in areas A and B in Figure 1 (estuary area outlined in red) was sand, and four points at different water depths were selected. Equation (8) is at times used as an alternative to reflectance, as shown in the following equation [71]. Equation (9) was adopted herein as well.

$$BI_{ij} = LN(R_i - R_{si}) - k_{ij}LN\left(R_j - R_{sj}\right). \quad (23)$$

4.2.3. Field Spectral Reflectance Data Sampling

To develop an algorithm to separate *Sargassum* and *Zostera*, we primarily used an Ekman–Birge grab sampler and a boat hook to pull *Sargassum* or *Zostera* from the seabed onto the ship and measured their spectral reflectance. Figure 4.3 shows photographs of the *Sargassum* and *Zostera* captured underwater before being pulled up (bottom) and on board the ship (top). These were obtained at St. 2 on 19 March 2021 and 17 June 2021, respectively. The reflectance of SWSG was measured using a visible/near-infrared spectroradiometer "MS720" manufactured by EKO Ltd. (wavelength range: 350–1050 nm, wavelength interval: 3.3 nm) and a device manufactured by Japan Color Research Institute. It is a standard white board (ZB6010). Specifically, the reflectance was calculated by measuring the spectral irradiance of the SWSG and the spectral irradiance of the sunlight reflected by the white board three times from a height of approximately 20 cm directly above the surface and dividing the average value. Furthermore, by gathering SWSG of the same species as close together as possible, the effects of reflection from the background board were reduced. The resulting field survey dates were 19 March, 17 June, 19 July, and 21 September 2021, and the fundamental measurement points were St. 1, St. 2, St. 3 St. 4, St. 5, St. 6, St. 6', St. 7, and St. 8. However, St. 8, where the measured water depth exceeded 10 m and no SWSG could be confirmed, was not used in the analysis.

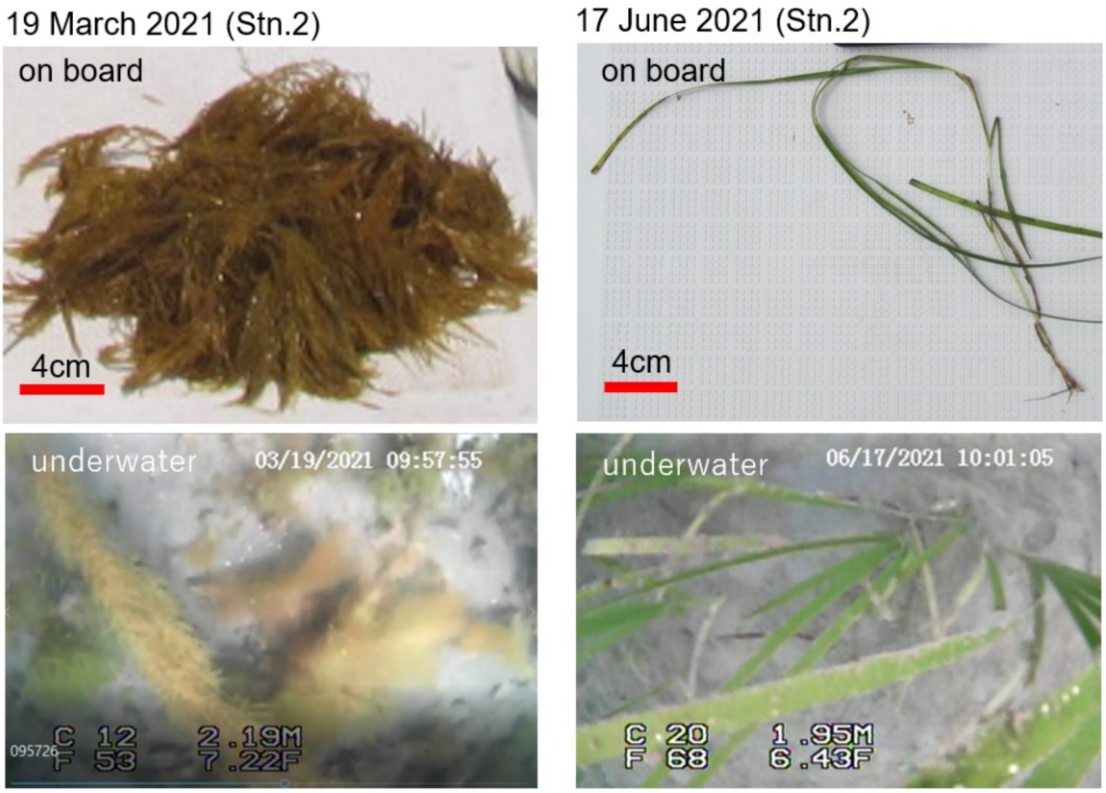


Figure 4.3. *Sargassum* (**left**) and *Zostera* (**right**) were obtained locally. The bottom photograph is a snapshot of an underwater video image of the seagrass, and the top photograph is of the seagrass, hoisted on board.

4.2.4. Satellite Data Used and Preprocessing Method

The satellite data used in this study were atmospherically corrected reflectance (level 2) data from the Sentinel-2 multispectral instrument (MSI), which can be downloaded free of charge from the Copernicus Open Access Hub. As shown in Table 4.1, MSI originally had 12 bands; however, in this study, data with a higher resolution were desired. Thus, only Band 2 (493 nm), Band 3 (560 nm), and Band 4 (665 nm) with a 10 m resolution were used. The satellite data used were that of 19 March, 23 May, and 27 July, which were close to the on-site observation dates in 2021, as shown in Table 4.2. The tide levels for satellite data acquisition were 2.37, 0.71, and 3.19 m, respectively. The k_{sw} employed for water depth correction using Equation (17) for the preprocessing of satellite images was the actual value of Moreland Prieur [72] (if there was no value for the Sentinel-2 wavelength, the values

for the previous and subsequent wavelengths were linearly interpolated); *m* was used at a water depth of 14 m, which is the maximum water depth to be corrected at 10 m, and the maximum tide level of this area was 4 m. Further, the limit value at which the reflectance of each band did not exceed 1 was used. Table 4.3 shows the values of k_{sw} and m of the water depth correction parameters used in this study. Furthermore, Rs in Equation (23) is the reflectance (equivalent to aerosol) in a clean body of water (the open ocean); however, in this case, the open ocean area was not near the target water area. Thus, the minimum value of each band in the screen was set [73] (Table 4.4).

Table 4.1. Characteristics of Sentinel-2 bands (https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/bands/. accessed on 9 November 2023).

Band	Central Wavelength	Spatial Resolution	Band	Central Wavelength	Spatial Resolution
Band 1	443 nm	60 m	Band 8	842 nm	10 m
Band 2	493 nm	10 m	Band 8a	865 nm	20 m
Band 3	560 nm	10 m	Band 9	945 nm	60 m
Band 4	665 nm	10 m	Band 10	1375 nm	60 m
Band 5	704 nm	20 m	Band 11	1610 nm	20 m
Band 6	740 nm	20 m	Band 12	2190 nm	20 m
Band 7	783 nm	20 m			

Table 4.2. Satellite data used with tidal level.

No.	Sentinel-2 (S-2)				Tidal Level * (m)
	Date	Time	Tile Number	Platform	
1	19 March 2021	10:47JST	T53SKU	Sentinel-2B	2.37
2	23 May 2021	10:47JST	T53SKU	Sentinel-2A	0.71
3	27 July 2021	10:47JST	T54SKU	Sentinel-2B	3.19

Note: * Takehara station of Japan Meteorological Agency at 11 am.

Table 4.3. Water depth correction parameters of Equation.

Sentinel-2 Band	Wavelength (nm)	k_{sw} (m^{-1})	m
2	493	0.0238	1.37
3	560	0.0720	0.45
4	665	0.4200	0.07

Table 4.4. R_s values on 19 March, 23 May and 27 July 2021.

Date	R_{s_Band2}	R_{s_Band3}	R_{s_Band4}
19 March 2021	0.018	0.165	0008
23 May 2021	0.030	0.050	0.040
27 July 2021	0.020	0.020	0.010

4.2.5. Digital Water Depth Data

In this study, existing digital water depth data (water depth equivalent to the lowest tide level) were used for water depth correction. Here, the M7000 series, a high-resolution submarine topography dataset along the coast of Japan created by the Japan Hydrographic Association, was employed in the M7018 region of the western Seto Inland Sea. The spacing of the contour lines of M7000 differs depending on the sea area; however, at depths below 100 m, the contour lines are densely spaced at intervals of 1–2 m. We used this irregular point cloud data by resampling it into 10 m mesh data, which is the same as Sentinel-2 data. The bathymetric map obtained in this manner is shown in Figure 4. In addition, when correcting the water depth using Equation (17), these tide levels were added to the water depth data described later.

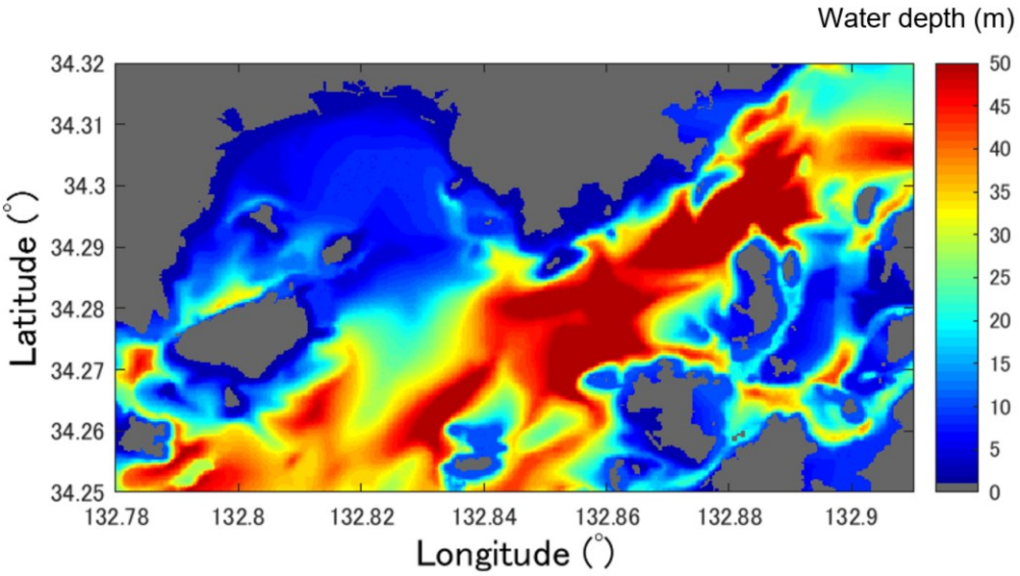


Figure 4.4. Reshaped Bottom topographic distribution map based on the M7000 series (https://www.jha.or.jp/en/jha/. accessed on 9 November 2023).

4.3. Results

4.3.1. Extraction Coefficient of Study Area

Equation (23) was used to extract the *Sargassum* and *Zostera* areas. First, to determine the extinction coefficient (k_{ij}) of Equation (21) for the three target periods, we calculated Sentinel-2 Band 2–3 (k23) and Band 3–4 (k23) in the two estuary areas, as shown in Figure 1. The relation (k34) was confirmed (Figure 4.5). Generally, if the bottom material is the same (here assumed to be sandy soil), a high correlation coefficient should be obtained for the natural logarithm relation between the two. Here, to aim for automation in the future, combinations that have a high correlation coefficient for any date must be selected. Consequently, as shown in Figure 4.5, all data exhibited almost the same correlation coefficient; however, overall, the correlation coefficient of k34 was higher than that of k23. Thus, k34 was adopted. It is notable that the value of Kij varies for each pixel in principle, but it is assumed to be consistent across the entire study area in this study. The Rs in Equation (9) was set as the value shown in Table 4.3. Figure 4.6 shows the results of calculating BI by applying this k34 to Equation (23) ((a–c) on the left) and the threshold value that

results in a water body with a water depth of approximately 10 m or less. The SWSG beds distribution map obtained when setting (March: −1, May: −0.95, July: −0.60) is shown in red. Hereinafter, this study assumed that this red area was a SWSG bed containing only *Sargassum* and *Zostera*.

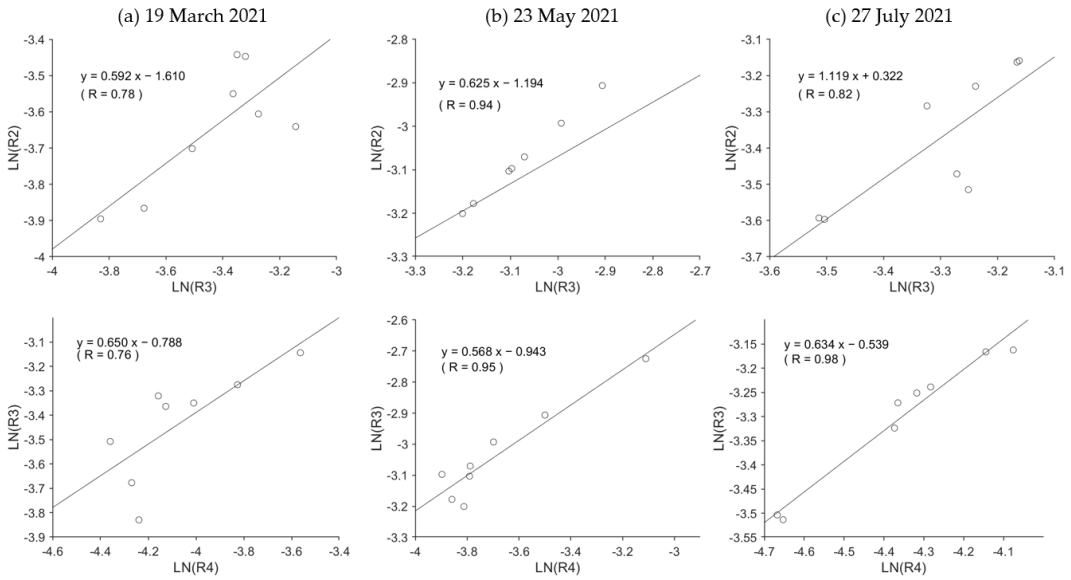


Figure 4.5. Scatterplots of Sentinel-2 Band 2–3 and Band 3–4 at the sand area of two river mouth sites for determining the extinction coefficient ratio. In (**a**), the upper plot illustrates the natural logarithm ratio between band 2–3, while the lower plot shows the natural logarithm ratio between band 3–4 on 19 March 2021. (**b**) shows the natural logarithm ratio on 23 May 2021, and (**c**) shows the ratio on 27 July 2021.

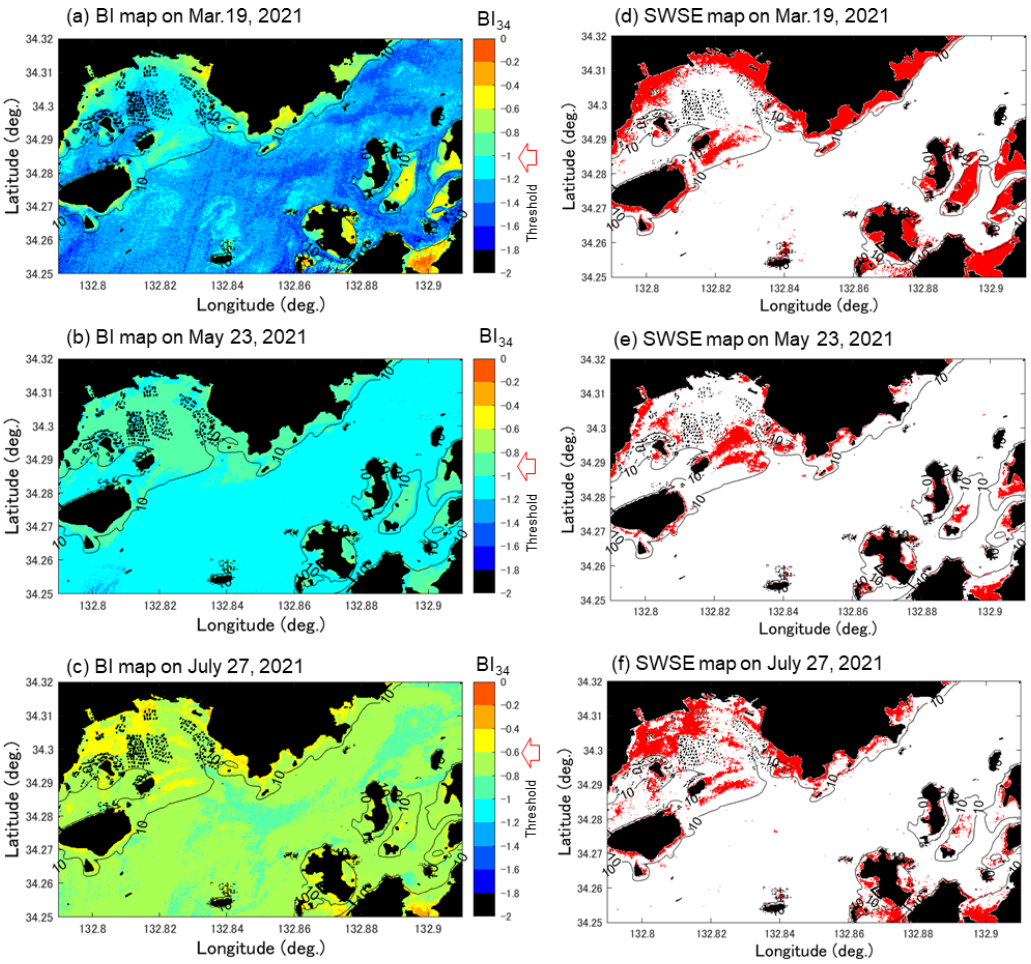


Figure 4.6. Bottom index (BI) map (**a-c**) and SWSG map (**d-f**) extracted from BI map on 19 March, 23 May, and 27 June 2021, as estimated from Sentinel-2 data.

4.3.2. Overview of Sargassum and Zostera Appearance

An occurrence table of SWSG types was created based on the observation results obtained at St. 1–St. 7 on 19 March, 17 June, 19 July, and 21 September 2021 (Table 4.5). In this table, "other" indicates the SWSG beds other than *Sargassum* and *Zostera*, "unknown" indicates a station wherein no survey was conducted on the day, and "none" indicates the absence of SWSG beds. In March 2021, *Sargassum* appeared at all four stations (St. 1, St. 4, St. 5). In June of the same year, *Zostera* appeared at all four stations (St. 2, St. 3, St. 4, St. 7) except St. 5, whereas *Sargassum* only appeared at St. 5. Furthermore, in July, *Zostera* newly appeared in St. 1. However, no SWSG beds were observed at any of the survey points in September.

These facts suggest that the SWSG beds in this region changed seasonally, even at the same station.

Table 4.5. Seagrass species observed at each location. "Other" indicates the SWSG beds other than *Sargassum* and *Zostera*. "Unknown" indicates a station wherein no survey was conducted on the day, and "None" indicates the absence of SWSG beds.

Date	Stations						
	St. 1	**St. 2**	**St. 3**	**St. 4**	**St. 5**	**St. 6**	**St. 7**
Mar 19	*Sargassum*	Other	Other	*Sargassum*	*Sargassum*	*Sargassum*	Unknown
June 17	None	*Zostera*	*Zostera*	*Zostera*	*Sargassum*	None	*Zostera*
July 19	*Zostera*	Unknown	Unknown	Other	*Sargassum*	Unknown	Unknown
September 21	None	None	None	None	None	None	None

4.3.3. Proposal of a Method Distinguishing between Sargassum and Zostera

Figure 4.7a shows the spectral reflectance characteristics of *Sargassum* and *Zostera* measured in 2021. The *Sargassum* at St. 4 and St. 6 on March 19, as shown in Table 4.5, was confirmed in the underwater video. However, spectral reflectance data could not be obtained as sampling was not possible. The common spectral characteristics of *Sargassum* and *Zostera* are that they both exhibit relatively low reflectance in the 400–500 nm range, relatively high reflectance at approximately 560 nm, and extremely strong absorption above 670 nm. The difference in the peak reflectance between the two was noticeable in the wavelength range of 500–670 nm. Thus, *Sargassum* and *Zostera* exhibit different reflectance maxima in the red region at ~600–670 nm (particularly ~610–650 nm) and the green region at 550–570 nm (particularly at ~560 nm). Figure 4.7b shows a diagram simulating the reflectance characteristics of the wavelengths observed by Sentinel-2. Only at the wavelength observed by Sentinel-2 was the difference in the reflectance between *Sargassum* and *Zostera* remarkable. The peak shift in the green and red areas was somewhat unclear in the low reflectance region (in particular, at the red of *Sargassum*, the peak disappeared).

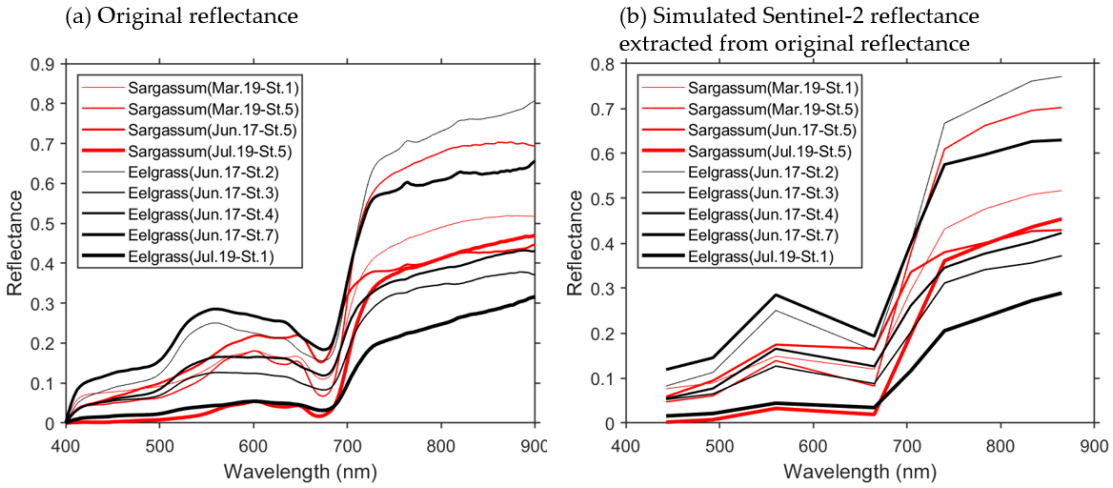


Figure 4.7. Spectral reflectance of *Sargassum* and *Zostera* obtained through field survey. Red and black indicate *Sargassum* and *Zostera*, respectively. (**a**) indicates the original reflectance and (**b**) indicates the simulated Sentinel-2 reflectance extracted from original reflectance.

Figure 4.8 shows the difference in spectral reflectance when the water depth is changed using Equation (17) for the data with the highest spectral reflectance for *Sargassum* and *Zostera*. As evident, the change in reflectance at 493 nm between *Sargassum* and *Zostera* was minimal. Moreover, even for red objects, such as *Sargassum*, the reflection was considerably weakened below a 1 m depth. Furthermore, the red reflectance was ~0 at a ~4-m depth and remained constant at approximately 0 at deeper water depths. However, the green band of *Zostera* tended to exhibit a considerably higher reflectance than that of *Sargassum*, at least at depths of ≥5 m.

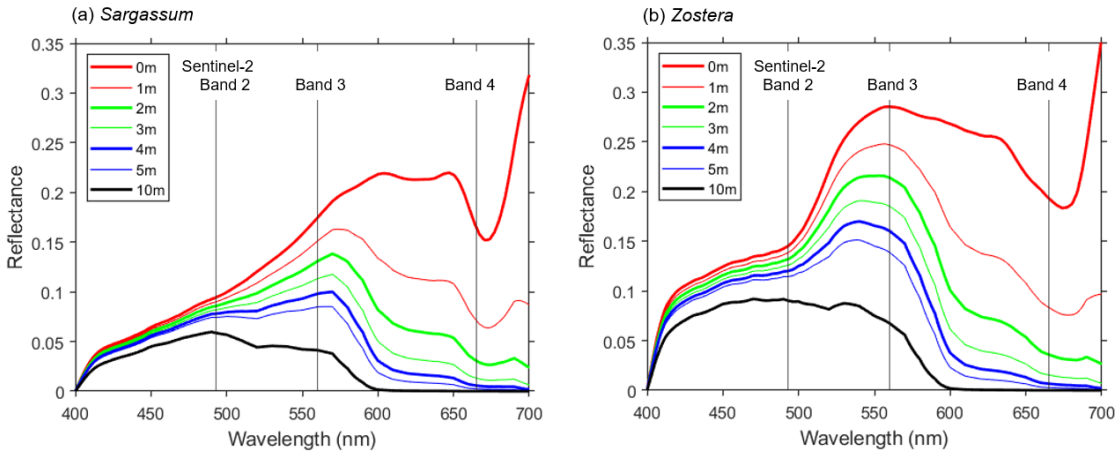


Figure 4.8. Simulated spectral reflectance of *Sargassum* and *Zostera* depending on water depth. (**a**) *Sargassum* at St. 5 on 17 June 2021. (**b**) *Zostera* at St. 7 on 17 June.

Based on the above results, Sentinel-2 can distinguish *Sargassum* and *Zostera* by changing the green area (560 nm: Sentinel-2 Band 2) reflectance at a depth of approximately 5 m or less, regardless of the water depth. In addition, the relative size of the green band is one indicator. Furthermore, following the floating algae index (FAI) [27,74], which is used to detect floating algae, the background can be detected by obtaining the difference between the line drawn by the reflectance of 493 and 665 nm (baseline) and the green reflectance. It is considered that the contrast between the reflectance of water and SWSG beds would become clearer. Therefore, we defined the SZDI from the relationship shown in Figure 4.9 and proposed the following calculation formula:

$$SZDI = R_{Green} - R_{Blue} - R'_{Green} \quad (24)$$

$$R'_{Green} = \frac{\lambda_{Green} - \lambda_{Blue}}{\lambda_{Red} - \lambda_{Blue}} (R_{Red} - R_{Blue}), \quad (25)$$

where R is the reflectance, λ is the wavelength, and the subscripts *Blue*, *Green*, and *Red* indicate *blue*, *green*, and *red* bands, respectively (in the case of Sentinel-2, they are 493, 560, and 655 nm, respectively). Hence, when Equations (24) and (25) are applied to satellite data, such as Sentinel-2, the larger the SZDI, the higher the possibility of *Sargassum*; meanwhile, the smaller the SZDI, the higher the possibility of *Zostera*.

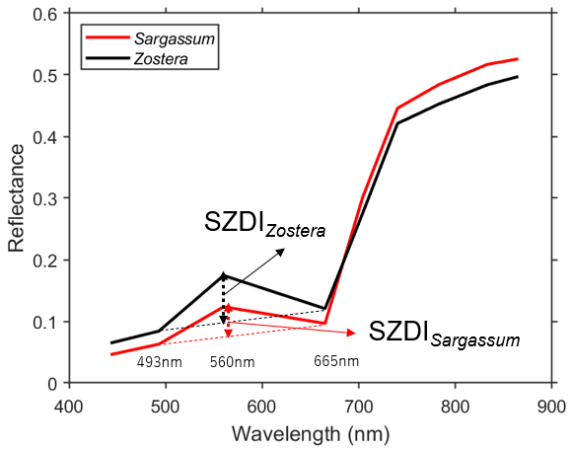


Figure 4.9. Schematic of the proposed *Sargassum* and *Zostera* distinguishing index (SZDI) of Sentinel-2 bands.

4.3.4. Validation of the Sargassum and Zostera Distinguish Index

Figure 4.10 shows the SZDI maps for 19 March, 23 May, and 27 June 2021 (a–c) and the SZDI thresholds for *Sargassum* and *Zostera*. The discrimination maps (d–f) are shown as well. The threshold values adopted are described below. First, referring to Table 4.5, based on actual measurements, we assumed that the images of St. 1 changed from March to July, and that the images of St. 4 changed from March to May. As shown in Figure 4.11, the threshold value was set at 0.015; this was the value near the boundary, wherein the SZDIs at both stations calculated from Sentinel-2 data were commonly divided. Consequently, it was determined that anything over 0.015 was *Zostera*, and anything less than that was *Sargassum*. Figure 4.10 (d–f), obtained in this manner, shows that *Zostera* increases from March to May and almost disappears in July. In contrast, *Sargassum* increases from March to May, but its area reduces in July, despite occupying large amounts of area in the study area. These temporal changes were synchronous with the overall measured results presented in Table 4.5. Furthermore, Figure 4.12 shows a comparison between the marine forest map obtained from Sentinel-2 and that previously conducted by the MOEJ [75] to verify the effectiveness of the SZDI method. The characteristics of the distribution of water area A, wherein *Sargassum* and *Zostera* tended to interchange easily, and water area B, wherein *Zostera* was abundant throughout the year, were observed. However, in Mitsu Bay on the left side of the screen, the wide distribution area of *Zostera* appeared to be considerably different. In a recent field survey of the western half of the bay by the MOEJ [76], as shown in Figure 4.13, *Zostera* was observed along the shore and around the islands and appeared mature from March to June but exhibited decrease in July, observed by Sentinel-2. It is notable that the absence of data on the right part of Mitsu Bay is attributed to the lack of field work conducted by Ministry of the Environment, Government of Japan. Figure 4.14 shows the changes in the area of *Sargassum* and *Zostera* throughout the screen during the three periods. In March, *Sargassum* and *Zostera* each occupied approximately 50% of the image. However, as May and July progressed, the proportion of *Zostera* decreased. Consequently, by July, the proportion had reached 15% each.

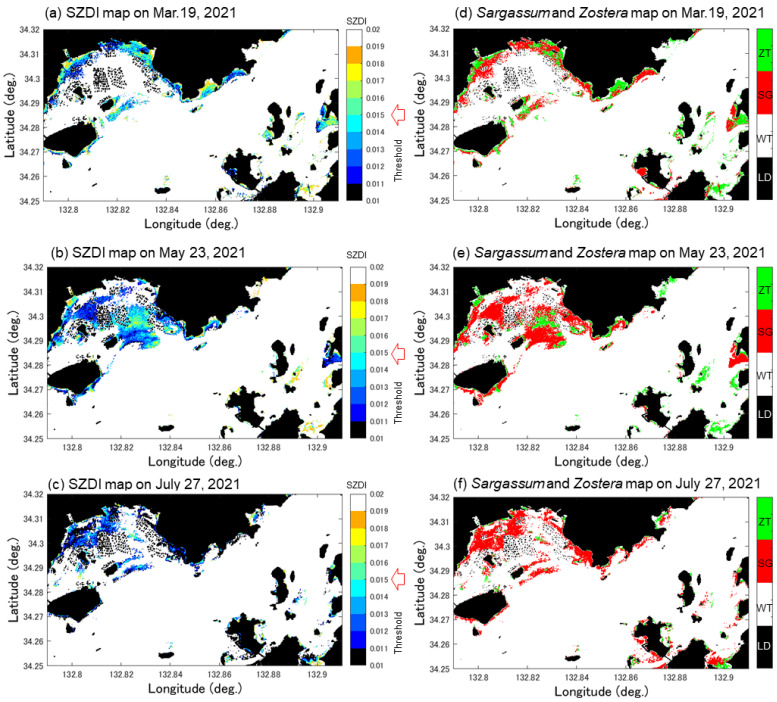


Figure 4.10. Distribution maps of *Sargassum* and *Zostera* on 19 March, 23 May, and 27 July 2021. The green and red colors represent *Zostera* (EL) and *Sargassum* (SG), respectively. The white and black colors represent sea water (WT) and land (LD), respectively. (**a**–**c**) represents SZDI maps on 19 March, 23 May, and 27 July 2021. (**d**–**f**) represents *Sargassum* and *Zostera* maps obtained after applying the determined threshold of SZDI.

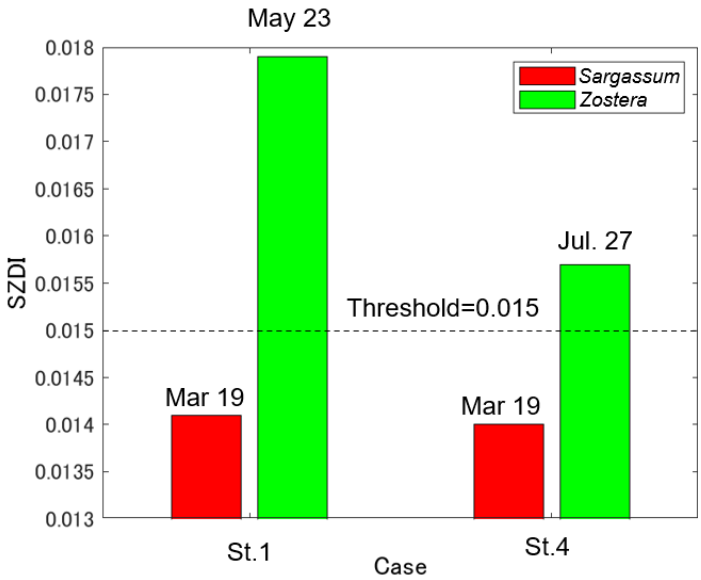


Figure 4.11. SZDI (*Sargassum* and *Zostera* Distinguish Index) threshold setting results.

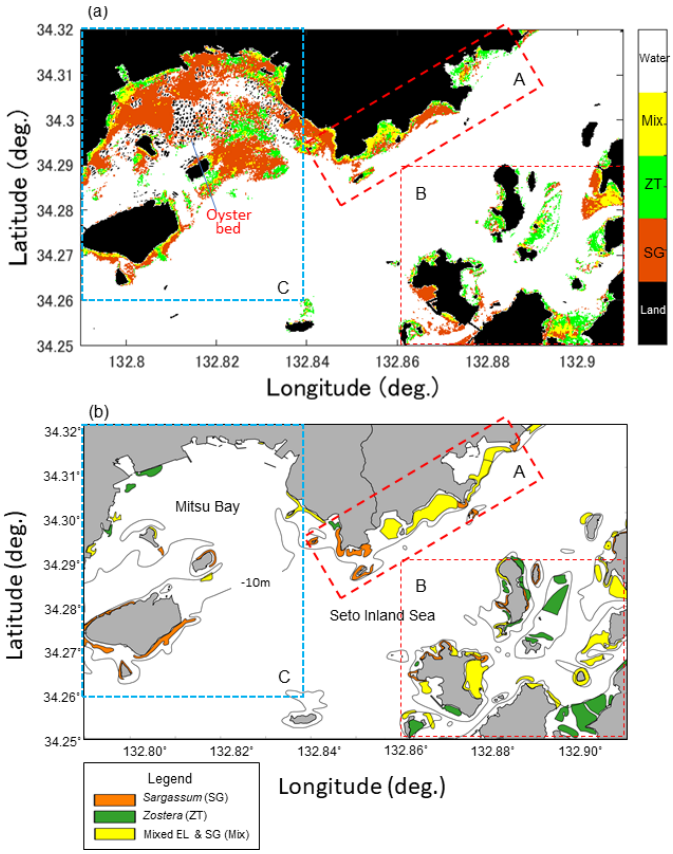


Figure 4.12. Comparison of *Sargassum* and *Zostera* distribution maps and Sentinel-2 map in this study. (**a**) and the redrawn reference map based on Ministry of the Environment, Government of Japan (1991) data. Black isoline in (**b**) indicates a 10 m water depth. A, B, and C are three water areas marked for convenient verification discussion.

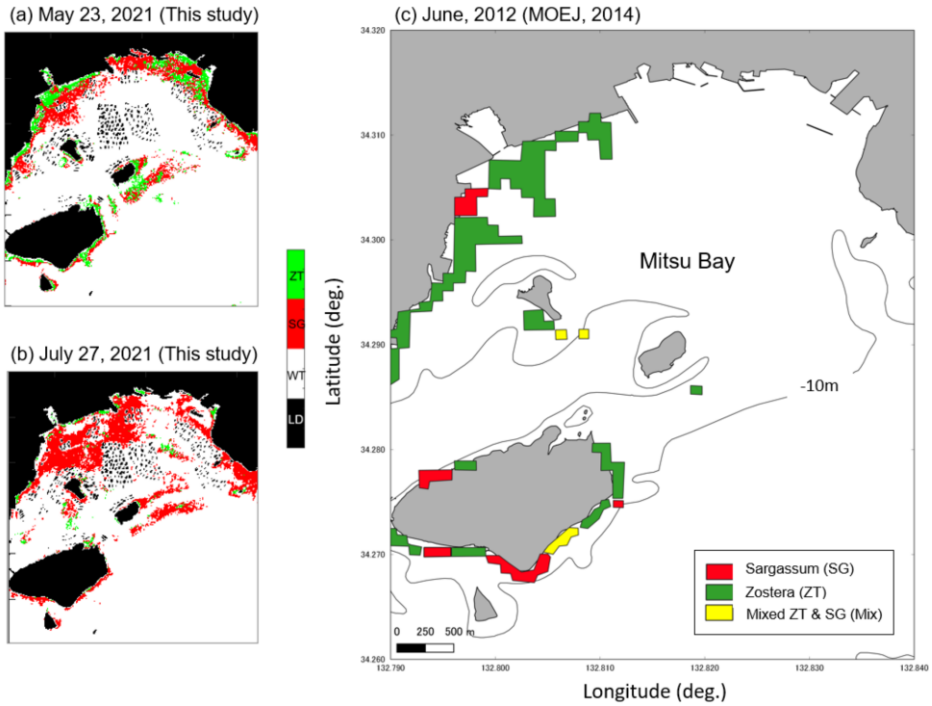


Figure 4.13. Comparison of *Sargassum* and *Zostera* distribution maps and Sentinel-2 maps both in May (a) and July (b) in the Mitsu Bay and redrawn reference map (c) based on the field work by Ministry of the Environment, Government of Japan (2014). Black isoline in (**c**) indicates 10 m water depth.

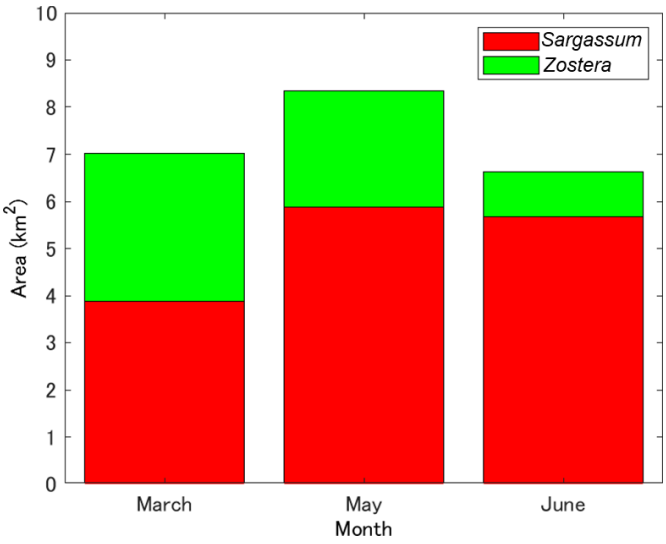


Figure 4.14. Temporal change in marine forest areas in this study area, derived from Sentinel-2 data.

4.4. Discussion

4.4.1. Validity of Underwater Forest Area Selection Using Bottom Index

This section discusses the validity of the marine forest distribution map of the study area based on the BI method shown in Figure 4.6a,b. As mentioned previously, marine forests in the Seto Inland Sea occur at depths of less than 10 m. The BI method used in this study included the SWSG bed distribution maps presented in Figures 4.12b and 4.13c, which were measured SWSG bed distribution maps over almost the entire area. Thus, it was judged to be appropriate in that respect. However, in general, the setting of k_{ij} (or the combination of bands) in Equation (23) is crucial when applying the BI method. Herein, the combination of Band 3 (green band) and Band 4 (red band) of Sentinel-2, which had a high correlation coefficient, was adopted. However, depending on the environment, a combination of blue and green bands is also effective [77]. This is because short wavelengths, which are more transparent, are effective in clean waters, such as coral reefs. However, in eutrophic waters, such as the Seto Inland Sea, absorption by color-dissolved organic matter (CDOM) in the blue band is relatively large owing to the strong absorption of CDOM and chlorophyll derived from rivers and sediments. Consequently, their absorption effects were weaker and the green and red bands appeared to be more effective. Moreover, upon the application of Terra ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)'s green and red bands to BI in the same region, they were effective in mapping *Zostera* beds [31], confirming the validity of the present results. Moreover, the potential monthly carbon storage amounts were estimated. According to the UNEP report [12], the average storage rate per unit area of seagrass is 0.83 carbon t/ha/year, with a range of 0.56 to 1.82 carbon t/ha/year. Applying these rates, the estimated carbon sequestration in our study area was found to be 53.08 t in March, 62.43 t in May, and 48.64 t in July [12,78]. It is important to note that the area of seagrass beds in the Seto Inland Sea decreased between 1960 and 1991, but has recently experienced revival due to improvements in water quality [12]. This highlights the significance of water quality improvement in enhancing the carbon sequestration capacity of seagrasses.

4.4.2. Spectral Characteristics of Zostera and Sargassum

This section discusses the spectral reflectance characteristics of *Sargassum* and *Zostera* in the target area, as shown in Figure 4.7. In this study, the differences in the spectral characteristics of *Sargassum* and *Zostera* can be approximately summarized as blue (400–500 nm), with low reflectance for both, red band (610–650 nm) for *Zostera*, and green band (550–560 nm) for *Sargassum*. This feature is also evident from the fact that the photographs of *Sargassum* and *Zostera* in Figure 4.3 appeared red and green, respectively. Upon further examining the reflectance characteristics, the present results showed that the reflectance for *Sargassum* was low in blue (400–500 nm) and in green (500–600 nm), and was high in near red (610 and 650 nm), compared to *Zostera*.

Previous studies on the spectral reflectance characteristics of these species include those on seagrass *Zostera capricorni*, *Posidonia australis*, and Halophila *ovalis* in eastern Australian waters [26] and the seagrass *Thalassia* in southern Florida [79]. Furthermore, studies on the green algae *Ulva* on the coast of China by Shin et al. [80] and *Sargassum* species in the Gulf of Mexico and the Atlantic Ocean by Hu et al. [68] represent other such research. For example, from past spectroscopic studies on *Sargassum*, *Sargassum* is known to be a type of brown macroalgae that contains the accessory pigments fucoxanthin and chlorophyll c [33]. The bio-absorption peaks of fucoxanthin are at blue and green wavelengths (480 and 520 nm). Additionally, chlorophyll c exhibits absorption peaks at blue (460 and 485 nm) and red (635 nm) [33,35] wavelengths. Conversely, the *Zostera* leaf and syringodium wrack contain chlorophyll b, with absorption peaks at blue and red wavelengths (470, 600, and 650 nm) [33]. To summarize these past studies, the spectral reflectance characteristics of *Sargassum* and *Zostera* are low for blue (460, 470, and 480 nm) wavelengths. A low reflectance at green (520 nm) wavelengths for *Sargassum* and a low reflectance for *Zostera* at red (600, 635, and 650 nm) wavelengths have also been confirmed. This indicates that the fundamental characteristics of *Sargassum* and *Zostera* in this study area are consistent with the characteristics of *Zostera* around the world.

4.4.3. Validity and Limitations of the SZDI Method

This section discusses the validity and limitations of the SZDI method, as shown in Figure 4.9 and Equations (24) and (25). As mentioned in Section 4.4.2, the definitive spectral characteristics of *Zostera* and *Sargassum* were characterized by a maximum near red (610 nm and 650 nm) wavelength for *Sargassum* and a minimum at ~630 nm for *Zostera*. Thus, it is desirable to use an index, such as the ratio of green to red, to separate *Zostera* and *Sargassum*. Dierssen et al. [33] proposed the ratio of reflectance values at 650 and 630 nm as the *Sargassum* index (SI). However, Sentinel-2 does not observe these bands (particularly the 630 nm band), and when considering only the Sentinel-2 bands, as shown in Figure 4.7b, the sentinel-2 reflectance does not necessarily reflect the green reflectance of *Sargassum*. The red area is not necessarily higher than the green area. Rather, the difference in the green reflectance appeared to be more useful for separation. Furthermore, by emphasizing the contrast with the reflectance of the background water, the difference in the green reflectance became clearer and more helpful when employed to distinguish *Sargassum* and *Zostera*. Considering all these ideas, Equations (24) and (25) (applicable to the FAI) [20,33] were considered to be effective methods for separating *Sargassum* and *Zostera*. This method, as summarized in Figure 4.9, calculated the baseline reflectance obtained using Band 2 (480 nm) and Band 4 (655 nm) from Sentinel-2 Band 3 (560 nm) reflectance. The higher the value, the more likely it was to be *Zostera*, whereas the lower the value, the more likely it was to be *Sargassum*. When the SZDI was calculated using Equations (24) and (25) and mapped using the thresholds set, as shown in Figure 4.10, seasonal distribution changes were evident. Furthermore, the distribution indicates that *Zostera* thrives in shallow sandy areas, while *Sargassum* prefers comparatively deeper, rocky, and intertidal zones, aligning well with their biological characteristics. In the Seto Inland Sea, *Sargassum* is predominantly composed of three species: *S. condusum*, *S. yezoense*, and *S. trichophyllum*, with maturity periods spanning from March to June, June to August, and October to December, respectively [69]. These findings contribute to explaining the rationale for the presence of Sargassum in our study during the period from March to July. Some studies also suggest that *Zostera* matures in June and experiences a decline thereafter [81], supporting our result that *Zostera*

occupied the largest area in the May image but decreased in the July image. In conclusion, the distribution not only exhibits a high level of consistency with the optical characteristics of *Sargassum* and *Zostera*, but also aligns closely with the findings from our field survey, as presented in Table 4.5 and the data provided by the MOEJ in Figure 4.12. Additionally, it resonates with the biological attributes of both *Sargassum* and *Zostera*.

However, at present, this method has certain limitations. One limitation is the influence of water depth z and k on Equation (17) within the SWSG bed area extracted using the BI method. Nevertheless, the reason why the seasonal changes in *Zostera* and *Sargassum* were valid (Figure 4.11) can be attributed to the assumptions or water depth correction method being effective to a certain extent during the analysis period. However, in reality, the spectrum of SWSG considerably changes depending on chlorophyll a and turbidity, which greatly affects k. Thus, to improve the accuracy of SWSG classification, the water quality information at the time of observation must be considered and the SWSG must be reflected. Furthermore, the rate may require correction. Meanwhile, in situations where the seagrass develops to maturation and releases floating seaweed, the upper seagrass may cover the lower seagrass, rendering the lower seagrass undetectable. Another challenge lies in the difficulty of achieving precise mathematical verification at present. In fact, there has only been one existing field survey conducted by the Ministry of the Environment, Government of Japan, primarily due to economic and time constraints; this was in 2014. Thus, there appears to be a lack of precise mathematical verification at present. To address these issues, the use of a hyperspectral sensor or the strategic design of a field survey tailored to match the satellite image of the specific area beforehand could be considered.

4.4.4. Application of the SZDI Method in the Study Area During Other Period

The field survey at Stn. 1-7 in the same region has been conducted in May, June and July of 2023. And the process of the algae detection and mearsurement remained the same. It is found that in the situation that the threshold of BI value was

appropriately set, the application of the SZDI method with a SZDI threshold of 0.015 can leading to a result correlated with the foundings from the field survey. The distribution maps of *Sargassum* and *Zostera* obtained through the SZDI method are shown in Figure 4.15. The observation results from the field survey are shown in Table 4.6. However, it is notable that the SZDI method may not able to accurately depict the distribution of *Sargassuma* and *Zostera* if a appropriate thereshold for BI value can not be chosen. Regarding the decision method for the theshold of the BI value, Matsunaga et al. mentionedthe bimodal characteristics of BI value to classify the "Benthic Community." which including coral, seagrass and algae, from sand. Matsunaga et al. finally chosed the combination of the blue and the green band of Landsat TM because the difference in means between "Sand" and "Benthic Community" of BI_{12} is smaller than in that of BI_{23} [32]. However, when applying this method to Sentinel-2 data of 2023, the BI value does not absorutely exhibit the bimodal characteristics. Therefore, objective observations from experienced individuals and the references to the result of field sureys are still necessary in deciding the threshold value of BI, posing an challenge to the accurate classification of *Sargassum* and *Zostera* to some extent.

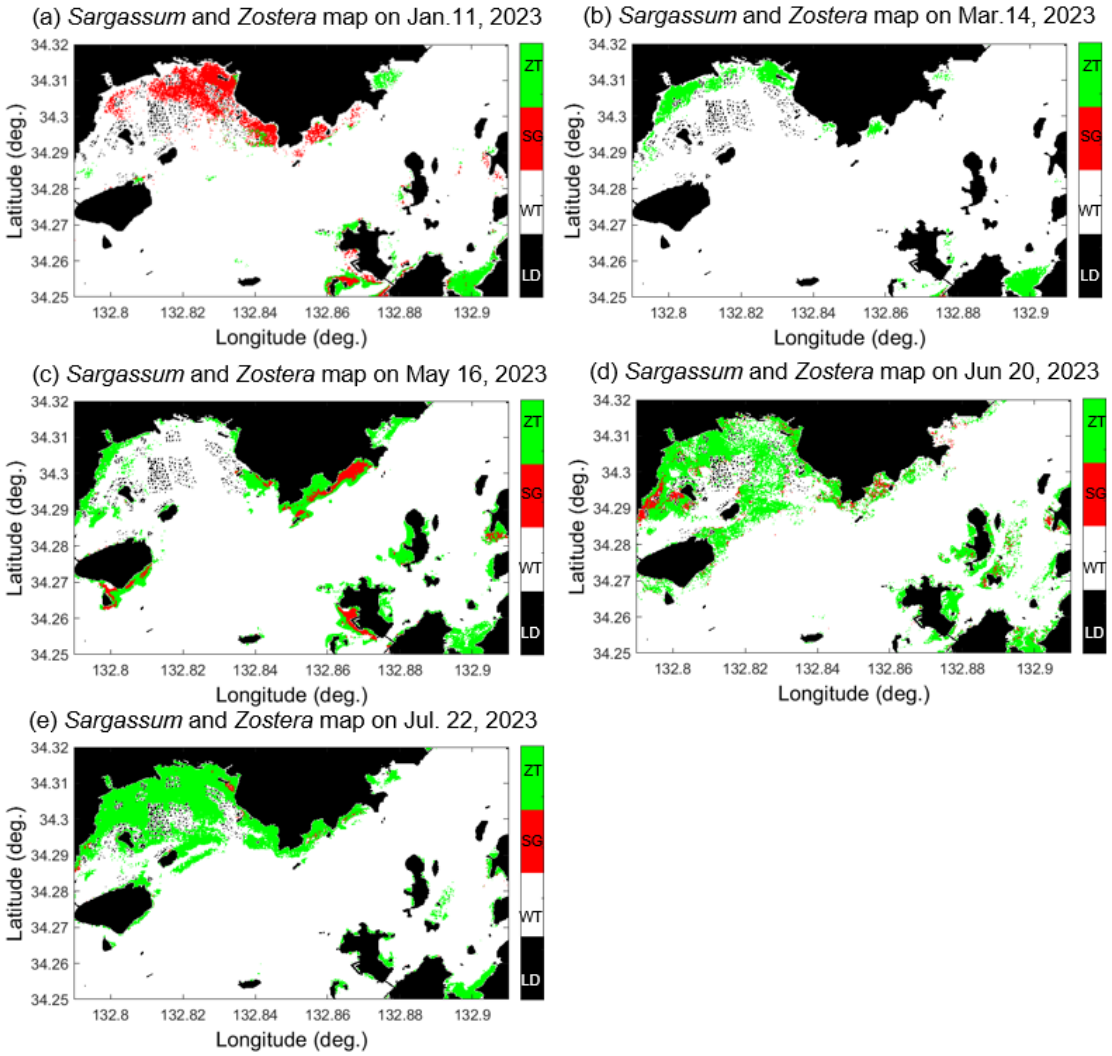


Figure 4.15 Distribution maps of *Sargassum* and *Zostera* in 2023. The green and red colors represent *Zostera* (EL) and *Sargassum* (SG), respectively. The white and black colors represent sea water (WT) and land (LD), respectively. (**a**–**e**) represents obtained *Sargassum* and *Zostera* maps on Jan. 11, Mar. 14, May 16, Jun. 20, and Jul. 22, 2023.

Table 4.6. Seagrass species observed at each location in 2023. "Other" indicates the SWSG beds other than *Sargassum* and *Zostera*. "Unknown" indicates a station wherein no survey was conducted on the day, and "None" indicates the absence of SWSG beds.

Date	Stations						
	St. 1	St. 2	St. 3	St. 4	St. 5	St. 6	St. 7
May. 25	Unknown	*Zostera*	None	*Zostera/Sargassum*	Unknown	*Zostera/Sargassum*	Unknown
Jun. 19	Unknown	None	*Zostera*	*Zostera*	Unknown	*Zostera*	Unknown
Jul. 19	Unknown	*Zostera/Sargassum*	None	Other	Unknown	*Zostera*	Unknown

4.5. Conclusions

This study proposed a method for separating *Sargassum* and *Zostera* in the Seto Inland Sea using Sentinel-2 data and examined the validity of the method. Consequently, the following points were clarified.

1. The BI method was an effective method, using a combination of Band 3 (green) and Band 4 (red) Sentinel-2 data for estimating the area of SWSG beds in the sea.
2. The spectral reflectance characteristics of *Sargassum* and *Zostera* measured on site were understood. In general, the characteristics exhibit low reflectance in the blue and red bands for both *Sargassum* and *Zostera*, but exhibit relative high reflectance in the green band for *Zostera*.
3. The SZDI, which used the height from the background of Sentinel-2 Band 3 (560 nm) as an index, was proposed as a method for separating *Sargassum* and *Zostera*. Its validity was confirmed through field surveys, the examination of past SWSG bed maps, and the consideration of biological characteristics.

In the future, we will improve methods for automatically determining BI thresholds, which are currently determined by trial and error. Furthermore, we intend to promote classification methods that use artificial intelligence and hyperspectral sensors, which have been rapidly developed in recent years, to classify kelp and green algae in addition to *Sargassum* and *Zostera* beds.

Chapter 5: Distinguishing Sargassum and Zostera Using Machine Learning

5.1. Introduction

5.1.1. Backgrounds

In Chapter 4, the proposed algorithm, SZDI, has demonstrated effectiveness in distinguishing *Sargassum* and *Zostera*. Nonetheless, challenges persist in the application of the SZDI algorithm. These included the subjective determination of the SZDI threshold and extended algorithm development time and program creation times. Consequently, the utilization of machine learning, knowing for its ability to conduct rapid and objective analyses, is anticipated in the remote sensing field.

5.1.2. Previous Studies

According to the study of Arthur Samuel, computers can be programmed to learn and play a better game of checkers more effectively than humans through machine learning technology. These learning schemes may also be applicable in many other fields [82,83]. With the development of machine learning and the availability of vast datasets, applications of machine learning in image preprocessing, classification, and target recognition in remote sensing have been realized. For instance, the 2001 National Land-cover Database (NLCD) land cover classification for the contiguous USA was produced by using decision trees (DTs) [84,85]. Furthermore, According to the nature of data processing, the machine learning method can be broadly categorized as either supervised or unsupervised, with the former requiring supervised classification before analysis and the latter performing direct classification without provided information

[86]. Due to the difficulty in obtaining prior knowledge owing to the varied characteristics and complex backgrounds, the application of unsupervised learning has been significant [87,88]. In remote sensing, three well-known unsupervised classification models exist. Lv et al. classified construction and vegetation in aerial remote sensing images of Shanghai districts using K-Means and analyzing CIELAB of pixels [89]. Irvin et al. employed ISODATA classification techniques to classify landform elements in southwestern Wisconsin, USA, and verified the effectiveness of this algorithm [90]. Wang and Cheng experimented with the Ant colony optimization algorithm (ACO) to divide TM images into 6 classes, finding it superior to the ISODATA algorithm and K-means algorithm. Anil K. Jain praises the K-Means clustering algorithm as a standard and simple algorithm, that can be widely applied in various fields and over 50 years [91](Jain, 2010). ABBAS and Wang have employed mean squared error (MSE) to evaluate the performance of image-processing models [87,92].

5.1.3. Objectives

The objective of this chapter is to utilize an unsupervised learning model, K-Means to distinguish *Sargassum* and *Zostera* in the same area as discussed in the previous chapter, using Setinel-2 images. Building upon the SZDI algorithm explored in the preceding chapter, the reflectances of band 2 (blue), band 3 (green), and band 4 (red) will be designated as the features for unlabeled data. Ultimately, the accuracy of this model will be evaluated using MSE.

5.2. Materials and methods

5.2.1. Study area and satellite data

To evaluate the accuracy of the chosen model, the binary SZDI images were assumed as the correct answers, and the same study area, satellite data, and dates are employed for the analysis in this chapter. Repeatedly, the study area was the Seto Inland Sea off the coast of Takahara City, located in the south-central part of Hiroshima Prefecture, as illustrated in Figure 4.1. The satellite data utilized consisted of atmospherically corrected reflectance (level 2) data from the Sentinel-2 multispectral instrument (MSI). The specific dates for the satellite data were March 19, May 23, and July 27, 2021.

5.2.2. Data collection and pre-processing

According to SZDI algorithm constitution outlined in equations (17 - 18) in subsection 4.3.3., the most significant differentiation between Sargassum and Zostera occurs in the reflectance values of band 2 (blue), band 3 (green), and band 4 (red). To achieve precise automated distinguishing, the reflectances of these three bands were selected as the features for K-means clustering. The training data were derived from a 770 × 1400 supervised-selected rectangle, which closely resembled the size of the study area depicted in Figure 4.1. The water dept effect corrected reflectance values for Band 2, Band 3, and Band 4 were utilized. After unrolling the matrix of 770 × 1400, a list containing 1,078,000 pixels was created. Lists of reflectance of Band 2, band 3, band 4, SZDI value and expected clusters were merged to form csv

files using Python in Jupyter Lab. The csv files for the three dates are presented in Table 5.1.

Table 5.1. (a), (b) and (c) depict the CSV data obtained after preprocessing all 1,078,000 pixels within the 770 × 1400 rectangle. The lists labeled "band 2", "band 3" and "band 4" contain the reflectance values of the corresponding bands. The list labeled "SZDI" contains the original SZDI values of these pixels, where "-1" represents land and "0" represents the sea deeper than -10m. The list labeled "Cluster" represents the clusters assigned after undergoing K-means clustering, with "-1" representing land, "1" representing *Sargassum*, and "2" representing *Zostera*.

(a)

	band2	band3	band4	SZDI	Cluster
1	NaN	NaN	NaN	-1	-1
2	NaN	NaN	NaN	-1	-1
…	…	…	…	…	…
739005	0.061228	0.064213	0.061228	0.019017	1
739006	0.060643	0.070313	0.060643	0.019481	2
739007	0.058927	0.071137	0.058927	0.02289	1
…	…	…	…	…	…
1077999	NaN	NaN	NaN	-1	-1
1078000	NaN	NaN	NaN	-1	-1

(b)

	band2	band3	band4	SZDI	Cluster
1	NaN	NaN	NaN	-1	-1
2	NaN	NaN	NaN	-1	-1
…	…	…	…	…	…
739005	0.060417	0.061763	0.060417	0.013986	2
739006	0.062294	0.063784	0.062294	0.014757	2
739007	0.060875	0.067472	0.060875	0.016509	2
…	…	…	…	…	…
1077999	NaN	NaN	NaN	-1	-1
1078000	NaN	NaN	NaN	-1	-1

(c)

	band2	band3	band4	SZDI	Cluster
1	NaN	NaN	NaN	-1	-1
2	NaN	NaN	NaN	-1	-1
…	…	…	…	…	…
738995	0.057685	0.051308	0.057685	0.010312	2
738996	0.058978	0.052721	0.058978	0.010842	2
738997	0.055735	0.051958	0.055735	0.010842	2
…	…	…	…	…	…
1077999	NaN	NaN	NaN	-1	-1
1078000	NaN	NaN	NaN	-1	-1

5.2.3. Application of unsupervised machine learning model

Let X= $\{X_i\}$, i =1,..., n be the set of d-dimensional points to be clustered into a set of K clusters, C=$\{C_k$; K=1,...,K$\}$.The K-means algorithm seeks a partition such that the squared error between the empirical mean of a cluster and the points in the cluster is minimized. Let μ_k be the mean of cluster C_k. The squared error between μ_k and the points in cluster C_k is defined as [91]:

$$J(C_k) = \sum_{X_i \in C_k} \|X_i - \mu_k\|^2 \qquad (26)$$

The objective of K-means is to minimize the sum of the squared error across all K clusters,

$$J(C\,) = \sum_{k-1}^{K} \sum_{X_i \in C_k} \|X_i - \mu_k\|^2 \qquad (27)$$

As the squared error cconsistently decreases with an increase in the number of clusters K (with J(C)=0 when K=n), it can only be minimized for a fixed number of clusters [91].

Given that this chapter aims to distinguish *Sargassum* and *Zostera* from the SWSG area, which has been extracted by the bottom index, the cluster K for this

chapter will be best to 2. Utilizing reflectances of band 2, band 3, and band 4 as features and employing the KMeans function imported from sklearn.cluster, the distribution of *Sargassum* and *Zostera* will be mapped.

5.2.4. Evaluation of the machine learning model

Two evaluation measures were utilized to assess K-means performance, one of the most common criteria for optimization is the mean square error (MSE) (equation 21). MSE is the sum of the squares of the differences between true values and predicted values for each data, divided by the total number of data (https://paddlepedia.readthedocs.io/en/latest/tutorials/deep_learning/loss_functions/MSE.html; Accessed on Dec. 5, 2023).

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \quad (28)$$

where m represents the number of samples (all pixels in the cropped study area), y_i is the true value (binarated SZDI results), and $\hat{y}_i$ is the predicted value (K-means clustering results). A smaller MSE value, closer to 0, indicates that the predicted value is closer to the true value. In this study, MSE value was calculated by employing the 'mse_squared_error' function from 'sklearn.metrics'.

The other measure is Cohen's Kappa coefficient. It is generally considered a robust statistic that is useful for evaluating the consistency between either interrater or intrarater agreements. The range of the kappa coefficient is from -1 to +1, with values closer to 1 indicating higher consistency. The kappa coefficient is defined as [93,94]:

$$k = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_0}{1 - p_e} \quad (29)$$

where p_o represents the actual consistency when measuring the same thing. And p_e represents the hypothetical probability of a random match [95]. In this study, the binarized SDZI value and the K-means clustering value are assumed to be two raters, and k value is considered as the assessment of consistency. Finally, the k value will be interpretated based on Table 5.1. In this study, k value is calculated by employing the 'cohen_kappa_score' function imported from 'sklearn.metrics'.

Table 5.1. Interpretation of Cohen's kappa.

Value of Kappa	Level of Agreement
0-0.20	None
0.21-0.39	Minimal
0.40-0.59	Weak
0.60-0.79	Moderate
0.80-0.90	Strong
Above 0.90	Almost Perfect

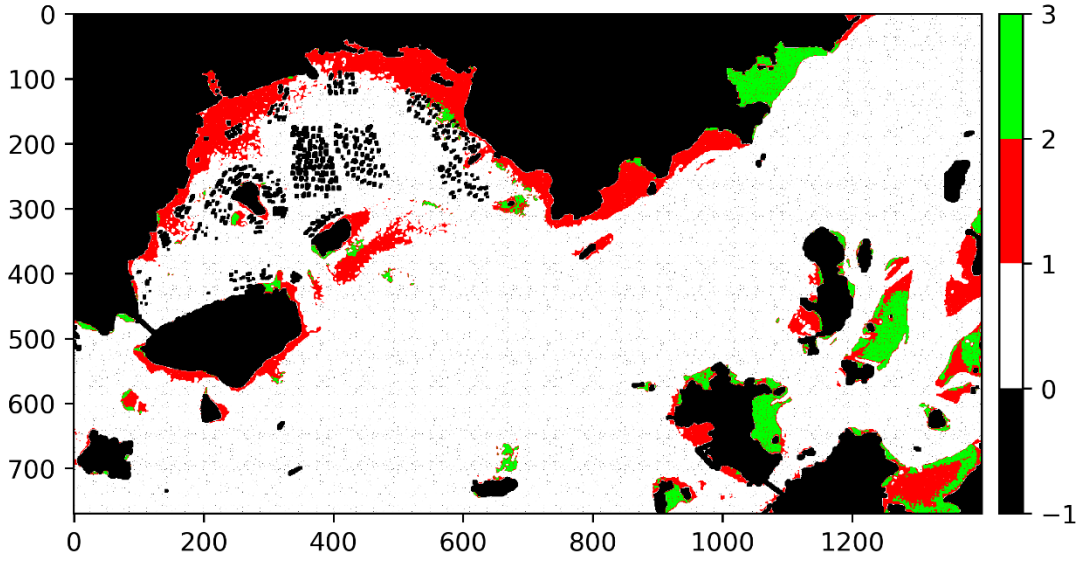
(https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/; Accessed on Dec. 5, 2023)

5.3. Results

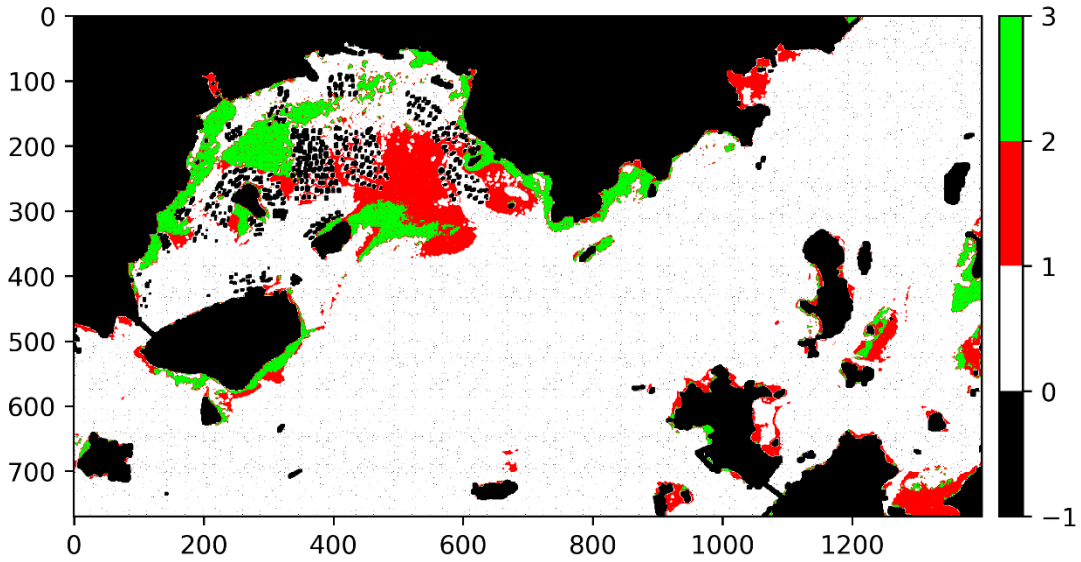
5.3.1. *Sargassum* and *Zostera* Distribution Obtained Using Machine Learning

As described in subsection 5.2, the K-means clustering was applied using the same study area and date. Then, two clusters were obtained, with the cluster representing Sargassum presented in red, and the cluster representing Zostera presented in green. The land is depicted in black, and the sea in white. The K-means images are shown in Figure 5.2.

(a)


0
100
200
300
400
500
600
700
0
200
400
600
800
1000
1200
3
2
1
0
−1

(b)


0
100
200
300
400
500
600
700
0
200
400
600
800
1000
1200
3
2
1
0
−1

(c)

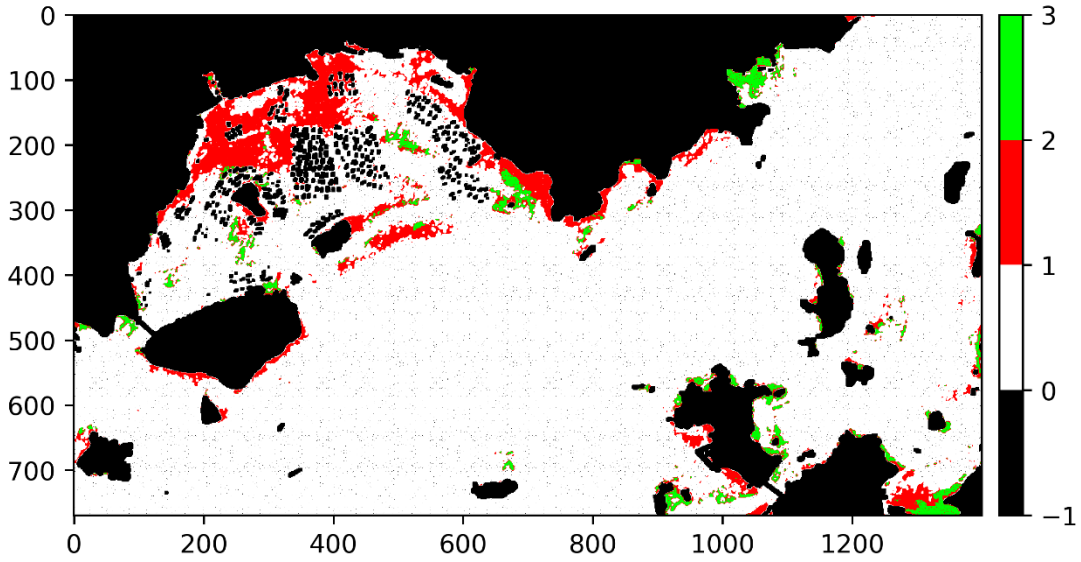


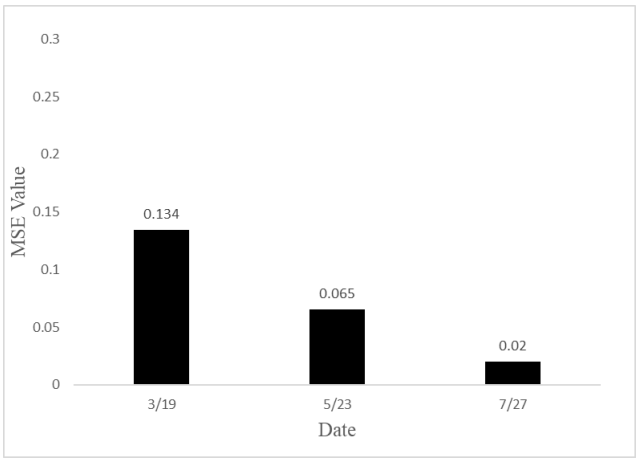
Figure 5.2. *Sargassum* and *Zostera* distribution maps using K-means clustering. (a), (b) and (c) represents the distribution map on Mar. 19, May 23 and Jul. 27, 2023. In these maps, *Sargassum* is represented by the red color, *Zostera* by the green color, land by the black color, and sea by the white color. The labels on the colorbars denote different cluster labels during calculations.

Figure 5.2 (a), (b), and (c) show high consistency with binarized SZDI images (depicted in Figure 4.6) by manual visual inspection. According to the figures, *Zostera* predominantly exists in the sandy mud areas shallower than the 2m depth, while *Sargassum* predominantly exists in the reef areas shallower than the 10m depth. The figures also suggest that the peak season for *Zostera* is around May, and the maturation period for the *Sargassum* extends from March to December. To precisely assess the accuracy of the K-means clustering model, MSE and Kappa coefficient methods were employed.

5.3.2. Evaluation of the Machine Learning Model

After comparing the K-means images and binarized SZDI images using the MSE and Kappa coefficient algorithm, the obtained MSE value and k value are as follows in Figure 5.3.

(a)



(b)

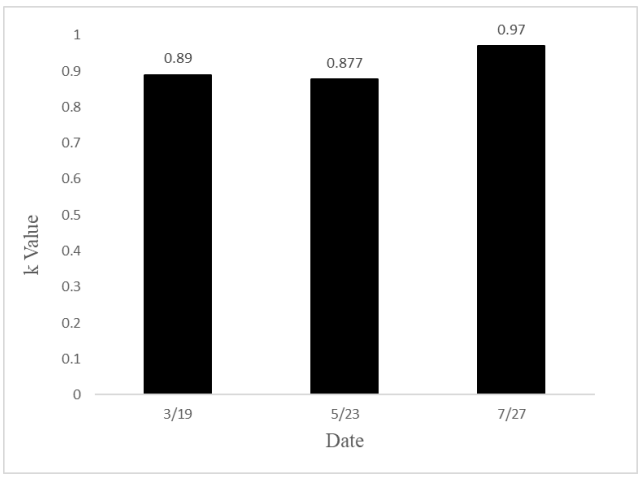


Figure 5.3. (a) illustrates the MSE value of K-means images compared to binarized SZDI images on Mar. 19, May 23, and Jul. 27, 2023, (b) illustrates the K value of K-means images compared to binarized SZDI images on the same dates.

The MSE values for each date are under 0.14, approaching 0, indicating that the K-means images closely resemble the binarized SZDI images. Furthermore, the k values for each date are over 0.87, signifying ‘strong’ and ‘almost perfect’ agreement between the K-means images and binarized SZDI images, according to Table 5.1. Thus, it is confirmed through quantitative evaluation that the K-means clustering model in this study is effective in certain aspects.

5.4. Discussion

5.4.1 Validity of the K-Means Model

In this study, when the verified binarized SZDI images were assumed to be the true dataset, the K-means images demonstrated high consistency with them. The consistency was confirmed through the MSE algorithm and Kappa coefficient. Subsequently, to explore wider applicability, band 1 to band 5 (after band 5, the reflectance is corrected to be zero, so it is not considered) were selected as five features of K-means clustering. The MSE value is 0.13, the K value is 0.89 on Mar. 19, 2023. The MSE value is 0.07, and the k value is 0.88 on May 23, 2023. The MSE value is 0.02, the k value is 0.97 on Jul. 27, 2023. These values are similar to those when features are selected as band 2 to band 4, indicating that the k-means clustering model can be validly used for classifying in this study area.

Moreover, it is noteworthy that, compared to the SZDI method, K-means clustering can demonstrate its effectiveness without subjective judgment, and the development of the algorithm and programming time can be substantially reduced. These advantages show high potential in the remote sensing field in the future.

5.4.2 Limitations of the K-Means Model

There still exist two major challenges with the practical application of the K-means model. The first challenge is that the results of K-means are difficult to accurately verified when the true dataset is absent. In this study, the true dataset was tentatively assumed as the binarized SZDI images obtained from Chapter 4. Although these binarized SZDI images have been verified, there may still be some errors compared with the true distribution. Moreover, in many cases, the true situation is not easy to grasp, making it more challenging to access the effectiveness of machine learning models. The second challenge is that the number of clusters needs to be defined before applying K-means. If the study area contains multiple seaweeds and seagrass species, determining the appropriate number of clusters will be challenging, resulting in inaccurate classification.

5.5 Conclusions

This chapter attempted an unsupervised machine learning model, K-means clustering to distinguish *Sargassum* and *Zostera* from the SGSW area obtained from Chapter 4. The results were then assessed by comparing them with the SZDI images using the MSE algorithm and Kappa coefficient. Consequently, the following points were clarified.

1. The K-means clustering is confirmed to be effective for classifying *Sargassum* and *Zostera* in this study area in some aspects.

2. Compared with the SZDI method proposed in Chapter 4, K-means clustering demonstrates prominent advantages, including an objective process and a shorter development time for algorithms and programming.

In the future, additional work is expected to further complete this study. This includes increasing field surveys to provide a true dataset for machine learning and developing an automated and reliable method to define the number of clusters. These enhancements are believed to broaden the scope of application of the methods.

Chapter 6: Conclusions

This study is primarily divided into three research components, including the detection of floating marine debris using Landsat-8 data, the classification of *Sargassum* and *Zostera* using Sentinal-2 data, and the classification of *Sargassum* and *Zostera* classification using machine learning.

Firstly, in the background of the yearly increasing rain disasters attributed to climate change, the damage caused by widespread and devastating floods, landslides, and mudflows has been more severe. These damages contain a significant amount of marine debris that flows into the sea, posing hazards to the environment, fishing industry, and navigation transportation. As a contribution to immediately removing these hazards and promoting post-disaster reconstruction, the first research component has clarified the following items:

(1) It is confirmed that the marine debris had a high peak reflectance in band 5 (central wavelength 865 nm), from the spectral reflectance data of Landsat-8 level-2

(2) It is confirmed that the corrected Floating Algae Index (cFAI) method performed better than the original Floating Algae Index (FAI) method in remove the background water signals from the Landsat-8 images.

(3) It is confirmed that the Otsu method, commonly used in image binarization, is effective in objectively determining images of cFAI. This makes it possible to automatically detect marine debris from complicated backgrounds.

(4) Based on the findings (1) - (3), the distribution and movement change of marine debris in the Seto Inland Sea off the coast of Hiroshima Prefecture after the heavy rain in July 2018 was reproduced and verified.

Secondly, considering the importance of conserving the blue carbon ecosystems, the distribution of the seaweed and seagrass beds has become crucial. Moreover, due to variations in the storage ability of organic carbon among species, there is a need for the development of a classification method. In response to this requirement, this study has clarified the following items:

(1) It is confirmed that the Bottom Index is effective for estimating the area of the seaweed and seagrass beds in the Seto Inland Sea, when using a combination of Band 3 (green) and Band 4 (red) of Sentinal-2 data.

(2) It is confirmed that the spectral reflectance characteristics of *Sargassum* and *Zostera* both exhibit low reflectance in the blue and red bands. But relatively high reflectance in the green band is observed for *Zostera*.

(3) Based on the findings (1) – (2), we proposed a classification method named Sargassum and Zostera Distinguishing Index (SZDI) and reproduced the distribution of the two species in the Seto Inland Sea. This method was verified by several field surveys and their specific biological characteristics.

Finally, considering the existing challenges of SZDI algorithm, including subjective determination of threshold and extended algorithm development time and program creation times, the machine learning methods, K-means clustering was employed for classification. Through the attempt with K-means clustering, this study has clarified the following items:

(1) It is confirmed that the 2-clustering K-means clustering algorithm is effective for distinguishing *Sargassum* and *Zostera* in this study in some aspects.

(2) It is confirmed that K-means clustering demonstrated prominent advantages, including an objective process and a shorter development time for algorithms and programming, compared with the SZDI method.

In the future, additional work is expected to further complete the study. This includes increasing the widening and frequency of field surveys, which are expected to be effective verification materials. Clarifying the biological characteristics is also important. Furthermore, the development of a more accurate water depth correction method is needed.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Associate Professor Yuji Sakuno of the Graduate School of Advanced Science and Engineering, Hiroshima University, Japan. Throughout my doctoral studies, he provided invaluable assistance in both my academic pursuits and personal life, creating an environment without the burden of undue pressure from a supervisor. Associate Professor Sakuno kindly showed care when I unfortunately suffered an injury in a field survey. He also encouraged me to participate in several domestic and international academic conferences to present my work. These experiences led to receiving the award of excellent contestant in the student award paper competition at the ISRS conference in 2023 in Korea. He also actively assisted me with scholarship applications. Even during my academic hiatus, he continued to express genuine concern for my well-being. Therefore, I would like to express my deepest heartfelt wishes for his future happiness and well-being.

My appreciation is extended to Professor Hidetsugu Iwashita in my lab and my sub-supervisor, Professor Hidemi Mutsuda, for their constructive comments and suggestions regarding my study and job hunting.

Then, I wish to express my deep appreciation to Associate Professor Taniguchi Naokazu, who not only provided extensive support and guidance for my field survey but also offered crucial insights and advice for my study. His passion for academic profoundly influenced my enthusiasm for study.

My appreciation is extended to the Administrative Office of the Graduate School of Advanced Science and Engineering and the Transportation and Environmental Systems Program, where the staff exhibited kindness during my Ph.D.

period. Additionally, I would like to thank my friends here for their unwavering support and companionship.

Finally, I would like to express my deepest gratitude to my parents for their unwavering financial support, enabling me to complete my academic pursuits. And my appreciation is also extended to my friends in China, Chen, Ye, and others, who consistently provided me with steadfast support, encouragement, and motivation.

References

1. Van Aalst, M.K. The Impacts of Climate Change on the Risk of Natural Disasters. *Disasters* **2006**, *30*, 5–18, doi:10.1111/j.1467-9523.2006.00303.x.
2. Trenberth, K. Uncertainty in Hurricanes and Global Warming. *Science* **2005**, *308*, 1753–1754, doi:10.1126/science.1112551.
3. Nicholls, R.J.; Mimura, N.; Topping, J.C. Climate Change in South and South-East Asia: Some Implications for Coastal Areas. *Journal of global environment engineering* **1995**, *1*, 137–154.
4. Ross, T.; Lott, N. A Climatology of 1980-2003 Extreme Weather and Climate Events.
5. Shultz, J.M.; Russell, J.; Espinel, Z. Epidemiology of Tropical Cyclones: The Dynamics of Disaster, Disease, and Development. *Epidemiologic Reviews* **2005**, *27*, 21–35, doi:10.1093/epirev/mxi011.
6. Japan Meteorological Agency. 2018. Strong Wind and Storm Surge Caused by Typhoon No. 21 in 2018, (in Japanese). Technical Report. Http://Www.Data.Jma.Go.Jp/Obd/Stats/ Data/Bosai/Report/2018/20180911/20180911.Html.
7. Mori, N.; Yasuda, T.; Arikawa, T.; Kataoka, T.; Nakajo, S.; Suzuki, K.; Yamanaka, Y.; Webb, A. 2018 Typhoon Jebi Post-Event Survey of Coastal Damage in the Kansai Region, Japan. *Coastal Engineering Journal* **2019**, *61*, 278–294, doi:10.1080/21664250.2019.1619253.
8. Cabinet Office, Government of Japan (2018a) Summary: The Damage Situations Caused by the Heavy Rain in July, 2018. Available Online, Http://Www.Bousai.Go.Jp/Updates/ H30typhoon7/Pdf/300905_1700_h30typhoon7.Pdf (in Japanese).
9. Hirota, K.; Konagai, K.; Sassa, K.; Dang, K.; Yoshinaga, Y.; Wakita, E.K. Landslides Triggered by the West Japan Heavy Rain of July 2018, and Geological and Geomorphological Features of Soaked Mountain Slopes. *Landslides* **2019**, *16*, 189–194, doi:10.1007/s10346-018-1100-3.
10. 裕司作野 現地調査と衛星データによる西日本豪雨前後の広島県沖における透明度分布特性. *土木学会論文集 B2(海岸工学)* **2019**, *75*, I_1045-I_1050, doi:10.2208/kaigan.75.I_1045.
11. Macreadie, P.I.; Anton, A.; Raven, J.A.; Beaumont, N.; Connolly, R.M.; Friess, D.A.; Kelleway, J.J.; Kennedy, H.; Kuwae, T.; Lavery, P.S.; et al. The Future of

Blue Carbon Science. *Nat Commun* **2019**, *10*, 3998, doi:10.1038/s41467-019-11693-w.

12. Hori, M.; Bayne, C.J.; Kuwae, T. Blue Carbon: Characteristics of the Ocean's Sequestration and Storage Ability of Carbon Dioxide. In *Blue Carbon in Shallow Coastal Ecosystems*; Kuwae, T., Hori, M., Eds.; Springer Singapore: Singapore, 2019; pp. 1–31 ISBN 9789811312946.
13. Takasugi, Y.; Hoshika, A.; Noguchi, H.; Tanimoto, T. The Role of Tidal Vortices in Material Transport around Straits. *Journal of Oceanography* **1994**, *50*, 65–80.
14. Takeoka, H. Progress in Seto Inland Sea Research. *Journal of oceanography* **2002**, *58*, 93–107.
15. Okaichi, T.; Yanagi, T. *Sustainable Development in the Seto Inland Sea, Japan: From the Viewpoint of Fisheries*; Terra Scientific Publishing Company, 1997; ISBN 978-4-88704-122-6.
16. Yoshida, G.; Hori, M.; Sakiyama, K.; Hamaguchi, M.; Kajita, A.; Nishimura, K.; Shoji, J. Distribution of Seaweed Bed and Tidal Flat and Their Correlations with Fisheries Catches in the Nine Sea Areas of the Seto Inland Sea. Fish Eng 2010, 47, 19-29.
17. Yoshida, T. Japanese Species of Sargassum Subgenus Bactrophycus (Phaeophyta, Fucales). *Journal of the Faculty of Science, Hokkaido University. Series 5, Botany* **1983**, *13*, 99–246.
18. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A Review of Vegetation Indices. *Remote Sensing Reviews* **1995**, *13*, 95–120, doi:10.1080/02757259509532298.
19. Rouse Jr, J.W.; Haas, R.H.; Schell, J.; Deering, D. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; 1973;
20. Rouse Jr, J.W.; Haas, R.H.; Deering, D.; Schell, J.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; 1974;
21. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote sensing of Environment* **1979**, *8*, 127–150.
22. Pearson, R.L.; Miller, L.D.; Program, U.S.I.B. *Remote Mapping of Standing Crop Biomass for Estimation of the Productivity of the Shortgrass Prairie, Pawnee National Grasslands, Colorado*; IBP Grassland Biome; Department of Watershed Sciences, College of Forestry and Natural Resources, Colorado State University, 1972;

23. Baret, F.; Guyot, G. Potentials and Limits of Vegetation Indices for LAI and APAR Assessment. *Remote Sensing of Environment* **1991**, *35*, 161–173, doi:10.1016/0034-4257(91)90009-U.

24. Elvidge, C.D.; Lyon, R.J.P. Influence of Rock-Soil Spectral Variation on the Assessment of Green Biomass. *Remote Sensing of Environment* **1985**, *17*, 265–279, doi:https://doi.org/10.1016/0034-4257(85)90099-9.

25. Huete, A.R.; Jackson, R.D.; Post, D.F. Spectral Response of a Plant Canopy with Different Soil Backgrounds. *Remote Sensing of Environment* **1985**, *17*, 37–53, doi:https://doi.org/10.1016/0034-4257(85)90111-7.

26. KAUFMAN, Y.J.; HOLBEN, B.N. Calibration of the AVHRR Visible and Near-IR Bands by Atmospheric Scattering, Ocean Glint and Desert Reflection. *International Journal of Remote Sensing* **1993**, *14*, 21–52, doi:10.1080/01431169308904320.

27. Hu, C. A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans. *Remote Sensing of Environment* **2009**, *113*, 2118–2129, doi:https://doi.org/10.1016/j.rse.2009.05.012.

28. Song, S.; Sakuno, Y.; Taniguchi, N.; Iwashita, H. Reproduction of the Marine Debris Distribution in the Seto Inland Sea Immediately after the July 2018 Heavy Rains in Western Japan Using Multidate Landsat-8 Data. *Remote Sensing* **2021**, *13*, 5048, doi:10.3390/rs13245048.

29. Lyzenga, D.R. Passive Remote Sensing Techniques for Mapping Water Depth and Bottom Features. *Appl. Opt.* **1978**, *17*, 379, doi:10.1364/AO.17.000379.

30. Lyzenga, D.R. Remote Sensing of Bottom Reflectance and Water Attenuation Parameters in Shallow Water Using Aircraft and Landsat Data. *International Journal of Remote Sensing* **1981**, *2*, 71–82, doi:10.1080/01431168108948342.

31. Sochea L.; Sakuno Y. EELGRASS BED MONITORING USING SATELLITE TERRA/ASTER DATA IN YOSHINA TIDAL FLAT. *PROCEEDINGS OF HYDRAULIC ENGINEERING* **2008**, *52*, 1381–1386, doi:10.2208/prohe.52.1381.

32. Matsunaga, T.; Hoyano, A.; Mizukami, Y. Monitoring of Coral Reefs on Ishigaki Island in Japan Using Multitemporal Remote Sensing Data.; Frouin, R.J., Kawamura, H., Kishino, M., Eds.; Sendai, Japan, January 5 2001; pp. 212–222.

33. Dierssen, H.M.; Chlus, A.; Russell, B. Hyperspectral Discrimination of Floating Mats of Seagrass Wrack and the Macroalgae Sargassum in Coastal Waters of Greater Florida Bay Using Airborne Remote Sensing. *Remote Sensing of*

Environment **2015**, *167*, 247–258, doi:https://doi.org/10.1016/j.rse.2015.01.027.

34. Johnsen, G.; Sakshaug, E. Biooptical Characteristics of PSII and PSI in 33 Species (13 Pigment Groups) of Marine Phytoplankton, and the Relevance for Pulse-amplitude-modulated and Fast-repetition-rate Fluorometry [1]. *Journal of Phycology* **2007**, *43*, 1236–1251, doi:10.1111/j.1529-8817.2007.00422.x.

35. Orzymski, J.; Johnsen, G.; Sakshaug, E. THE SIGNIFICANCE OF INTRACELLULAR SELF-SHADING ON THE BIOOPTICAL PROPERTIES OF BROWN, RED, AND GREEN MACROALGAE [1]. *Journal of Phycology* **1997**, *33*, 408–414, doi:10.1111/j.0022-3646.1997.00408.x.

36. Tsuguti, H.; Seino, N.; Kawase, H.; Imada, Y.; Nakaegawa, T.; Takayabu, I. Meteorological Overview and Mesoscale Characteristics of the Heavy Rain Event of July 2018 in Japan. *Landslides* **2019**, *16*, 363–371, doi:10.1007/s10346-018-1098-6.

37. Opfer, S.; Arthur, C.; Lippiatt, S. NOAA Marine Debris Shoreline Survey Field Guide. **2012**.

38. Moy, K.; Neilson, B.; Chung, A.; Meadows, A.; Castrence, M.; Ambagis, S.; Davidson, K. Mapping Coastal Marine Debris Using Aerial Imagery and Spatial Analysis. *Marine pollution bulletin* **2018**, *132*, 52–59.

39. Kubota, M. A Mechanism for the Accumulation of Floating Marine Debris North of Hawaii. *Journal of Physical Oceanography* **1994**, *24*, 1059–1064.

40. Isobe, A.; Iwasaki, S.; Uchida, K.; Tokai, T. Abundance of Non-Conservative Microplastics in the Upper Ocean from 1957 to 2066. *Nature communications* **2019**, *10*, 417.

41. Komatsu, T.; Sagawa, T.; Sawayama, S.; Tanoue, H.; Mohri, A.; Sakanishi, Y. Mapping Is a Key for Sustainable Development of Coastal Waters: Examples of Seagrass Beds and Aquaculture Facilities in Japan with Use of ALOS Images. *Sustainable development. Intech Publishing Co, Rijeka, Croatia* **2012**, 145–160.

42. Murataa, H.; Komatsub, T.; Yonezawac, C. Detection and Discrimination of Aquacultural Facilities in Matsushima Bay, Japan, for Integrated Coastal Zone Management and Marine Spatial Planning Using Full Polarimetric L-Band Airborne Synthetic Aperture Radar 5. **2019**.

43. Hu, C. A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans. *Remote Sensing of Environment* **2009**, *113*, 2118–2129, doi:10.1016/j.rse.2009.05.012.

44. Hu, L.; Zeng, K.; Hu, C.; He, M.-X. On the Remote Estimation of Ulva Prolifera Areal Coverage and Biomass. *Remote Sensing of Environment* **2019**, *223*, 194–207, doi:10.1016/j.rse.2019.01.014.

45. Sato, H.; Takeda, K.; Matsumoto, K.; Anai, H.; Yamakage, Y. Efforts for Disaster Prevention/Mitigation to Protect Society from Major Natural Disasters. *Fujitsu Sci. Tech. J* **2016**, *52*, 107–113.

46. Cheng, J.W.; Mitomo, H. Multi-Channel Information Dissemination for Disaster Evacuees–the Case of the 2016 Kumamoto Earthquake in Japan. **2018**.

47. Garcia, R.A.; Fearns, P.; Keesing, J.K.; Liu, D. Quantification of Floating Macroalgae Blooms Using the Scaled Algae Index. *JGR Oceans* **2013**, *118*, 26–42, doi:10.1029/2012JC008292.

48. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE transactions on systems, man, and cybernetics* **1979**, *9*, 62–66.

49. Oyama, Y.; Matsushita, B.; Fukushima, T. Distinguishing Surface Cyanobacterial Blooms and Aquatic Macrophytes Using Landsat/TM and ETM+ Shortwave Infrared Bands. *Remote Sensing of Environment* **2015**, *157*, 35–47.

50. Biermann, L.; Clewley, D.; Martinez-Vicente, V.; Topouzelis, K. Finding Plastic Patches in Coastal Waters Using Optical Satellite Data. *Scientific reports* **2020**, *10*, 5364.

51. Son, Y.B.; Choi, B.-J.; Kim, Y.H.; Park, Y.-G. Tracing Floating Green Algae Blooms in the Yellow Sea and the East China Sea Using GOCI Satellite Data and Lagrangian Transport Simulations. *Remote Sensing of Environment* **2015**, *156*, 21–33.

52. Qiu, Z.; Li, Z.; Bilal, M.; Wang, S.; Sun, D.; Chen, Y. Automatic Method to Monitor Floating Macroalgae Blooms Based on Multilayer Perceptron: Case Study of Yellow Sea Using GOCI Images. *Optics express* **2018**, *26*, 26810–26829.

53. Qi, L.; Hu, C.; Mikelsons, K.; Wang, M.; Lance, V.; Sun, S.; Barnes, B.B.; Zhao, J.; Van der Zande, D. In Search of Floating Algae and Other Organisms in Global Oceans and Lakes. *Remote Sensing of Environment* **2020**, *239*, 111659.

54. Smith, S.V. Marine Macrophytes as a Global Carbon Sink. *Science* **1981**, *211*, 838–840, doi:10.1126/science.211.4484.838.

55. Greiner, J.T.; McGlathery, K.J.; Gunnell, J.; McKee, B.A. Seagrass Restoration Enhances “Blue Carbon” Sequestration in Coastal Waters. *PLoS ONE* **2013**, *8*, e72469, doi:10.1371/journal.pone.0072469.

56. Lavery, P.S.; Mateo, M.-Á.; Serrano, O.; Rozaimi, M. Variability in the Carbon Storage of Seagrass Habitats and Its Implications for Global Estimates of Blue Carbon Ecosystem Service. *PLoS ONE* **2013**, *8*, e73748, doi:10.1371/journal.pone.0073748.

57. Tokoro, T.; Hosokawa, S.; Miyoshi, E.; Tada, K.; Watanabe, K.; Montani, S.; Kayanne, H.; Kuwae, T. Net Uptake of Atmospheric CO2 by Coastal Submerged Aquatic Vegetation. *Glob Change Biol* **2014**, *20*, 1873–1884, doi:10.1111/gcb.12543.

58. Ministry of the Environment. Available Online: Https://Www.Env.Go.Jp/Water/Heisa/Setonaikai_law_rev.Html (Accessed on 9 Nov., 2023).

59. Yamamoto, T. The Seto Inland Sea—Eutrophic or Oligotrophic? *Marine Pollution Bulletin* **2003**, *47*, 37–42, doi:10.1016/S0025-326X(02)00416-2.

60. Imai, I.; Yamaguchi, M.; Hori, Y. Eutrophication and Occurrences of Harmful Algal Blooms in the Seto Inland Sea, Japan. *Plankton Benthos Res* **2006**, *1*, 71–84, doi:10.3800/pbr.1.71.

61. Takeoka, H. Progress in Seto Inland Sea Research. J Oceanogr 2002, 58, 93-107.

62. Traganos, D.; Reinartz, P. Mapping Mediterranean Seagrasses with Sentinel-2 Imagery. *Marine Pollution Bulletin* **2018**, *134*, 197–209, doi:10.1016/j.marpolbul.2017.06.075.

63. Philpot, W.D. Bathymetric Mapping with Passive Multispectral Imagery. *Appl. Opt.* **1989**, *28*, 1569, doi:10.1364/AO.28.001569.

64. Stumpf, R.P.; Holderied, K.; Sinclair, M. Determination of Water Depth with High-Resolution Satellite Imagery over Variable Bottom Types. *Limnol. Oceanogr.* **2003**, *48*, 547–556, doi:10.4319/lo.2003.48.1_part_2.0547.

65. Biodiversity Center of Japan, [in Japanese]. Available Online: Https://Www.Biodic.Go.Jp/Moba/1_1.Html (Accessed on 20 Sep-Tember 2023).

66. Gower, J.; Hu, C.; Borstad, G.; King, S. Ocean Color Satellites Show Extensive Lines of Floating Sargassum in the Gulf of Mexico. *IEEE Transactions on Geoscience and Remote Sensing* **2006**, *44*, 3619–3625, doi:10.1109/TGRS.2006.882258.

67. Gower, J.F.R.; King, S.A. Distribution of Floating Sargassum in the Gulf of Mexico and the Atlantic Ocean Mapped Using MERIS. *International Journal of Remote Sensing* **2011**, *32*, 1917–1929, doi:10.1080/01431161003639660.

68. Hu, C.; Feng, L.; Hardy, R.F.; Hochberg, E.J. Spectral and Spatial Requirements of Remote Measurements of Pelagic Sargassum Macroalgae. *Remote Sensing of Environment* **2015**, *167*, 229–246, doi:https://doi.org/10.1016/j.rse.2015.05.022.

69. Yoshida, T. Japanese Species of Sargassum Subgenus Bactrophycus (Phaeophyta,Fucales).

70. Albert, A.; Mobley, C. An Analytical Model for Subsurface Irradiance and Remote Sensing Reflectance in Deep and Shallow Case-2 Waters. *Opt. Express* **2003**, *11*, 2873, doi:10.1364/OE.11.002873.

71. ODAGAWA, S.; TAKEDA, T.; YAMANO, H.; MATSUNAGA, T. A Proposed Method for Estimation of Cover Degree for Each Bottom-Type in Coral Reef Areas Using Hyperspectral Data. *Journal of The Remote Sensing Society of Japan* **2016**, *36*, 1–10, doi:10.11440/rssj.36.1.

72. Morel, A.; Prieur, L. Analysis of Variations in Ocean Color1. *Limnology & Oceanography* **1977**, *22*, 709–722, doi:10.4319/lo.1977.22.4.0709.

73. Matsunaga, T.; Hoyano, A.; Mizukami, Y. Spatial/Temporal Variation of Sea Water Extinction Coefficient Ratio and Temporal Change Detection in Coral Reef Using Bottom Index. In Proceedings of the Proceedings of the 26th Japanese Conference on Remote Sensing; 2000; pp. 281–282.

74. Qi, L.; Hu, C.; Wang, M.; Shang, S.; Wilson, C. Floating Algae Blooms in the East China Sea. *Geophysical Research Letters* **2017**, *44*, doi:10.1002/2017GL075525.

75. Ministry of the Environment; Biodiversity Center of Japan. Marine Biotic Survey (Coral Reef Survey), in the 4th National Survey on the Natural Environment [in Japanese]. 1991. Available Online: Http://Www.Biodic.Go.Jp/Reports2/4th/Coralreef/4_coralreef.Pdf. (Accessed on 20 September 2023).

76. Ministry of the Environment; Government of Japan. Mitsu Bay Area Healthy Plan [in Japanese]. 2014. Available Online: Http://Www.Biodic.Go.Jp/Reports2/4th/Coralreef/4_coralreef.Pdf. (Accessed on 20 September 2023).

77. Satomi Kakuta; Wataru Takeuchi; Anchana Prathep SEAWEED AND SEAGRASS MAPPING IN THAILAND MEASURED USING LANDSAT 8 OPTICAL AND TEXTURAL IMAGE PROPERTIES. *Journal of Marine Science and Technology* **2016**, *24*, doi:10.6119/JMST-016-1026-4.

78. Nellemann, C.; Corcoran, E. *Blue Carbon: The Role of Healthy Oceans in Binding Carbon: A Rapid Response Assessment*; UNEP/Earthprint, 2009;

79. Thorhaug, A.; Richardson, A.D.; Berlyn, G.P. Spectral Reflectance of *Thalassia Testudinum* (Hydrocharitaceae) Seagrass: Low Salinity Effects. *American J of Botany* **2006**, *93*, 110–117, doi:10.3732/ajb.93.1.110.

80. Shin, J.; Lee, J.-S.; Jang, L.-H.; Lim, J.; Khim, B.-K.; Jo, Y.-H. Sargassum Detection Using Machine Learning Models: A Case Study with the First 6 Months of GOCI-II Imagery. *Remote Sensing* **2021**, *13*, 4844, doi:10.3390/rs13234844.

81. Fujiwara, M. Research on Appropriate Creation Technology for Zostera Beds in and around the Coastal Shallow Areas, Seto-Inland Sea. 香水試研報 **2013**.

82. Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. **1959**.

83. Mahesh, B. Machine Learning Algorithms - A Review. **2018**, *9*.

84. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. *International Journal of Remote Sensing* **2018**, *39*, 2784–2817, doi:10.1080/01431161.2018.1433343.

85. Homer, C.; Huang, C.; Yang, L.; Wylie, B.; Coan, M. Development of a 2001 National Land-Cover Database for the United States. *photogramm eng remote sensing* **2004**, *70*, 829–840, doi:10.14358/PERS.70.7.829.

86. Turgay Celik Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering. *IEEE Geosci. Remote Sensing Lett.* **2009**, *6*, 772–776, doi:10.1109/LGRS.2009.2025059.

87. Abbas, A.W.; Minallh, N.; Ahmad, N.; Abid, S.A.R.; Khan, M.A.A. K-Means and ISODATA Clustering Algorithms for Landcover Classification Using Remote Sensing. *Sindh University Research Journal-SURJ (Science Series)* **2016**, *48*.

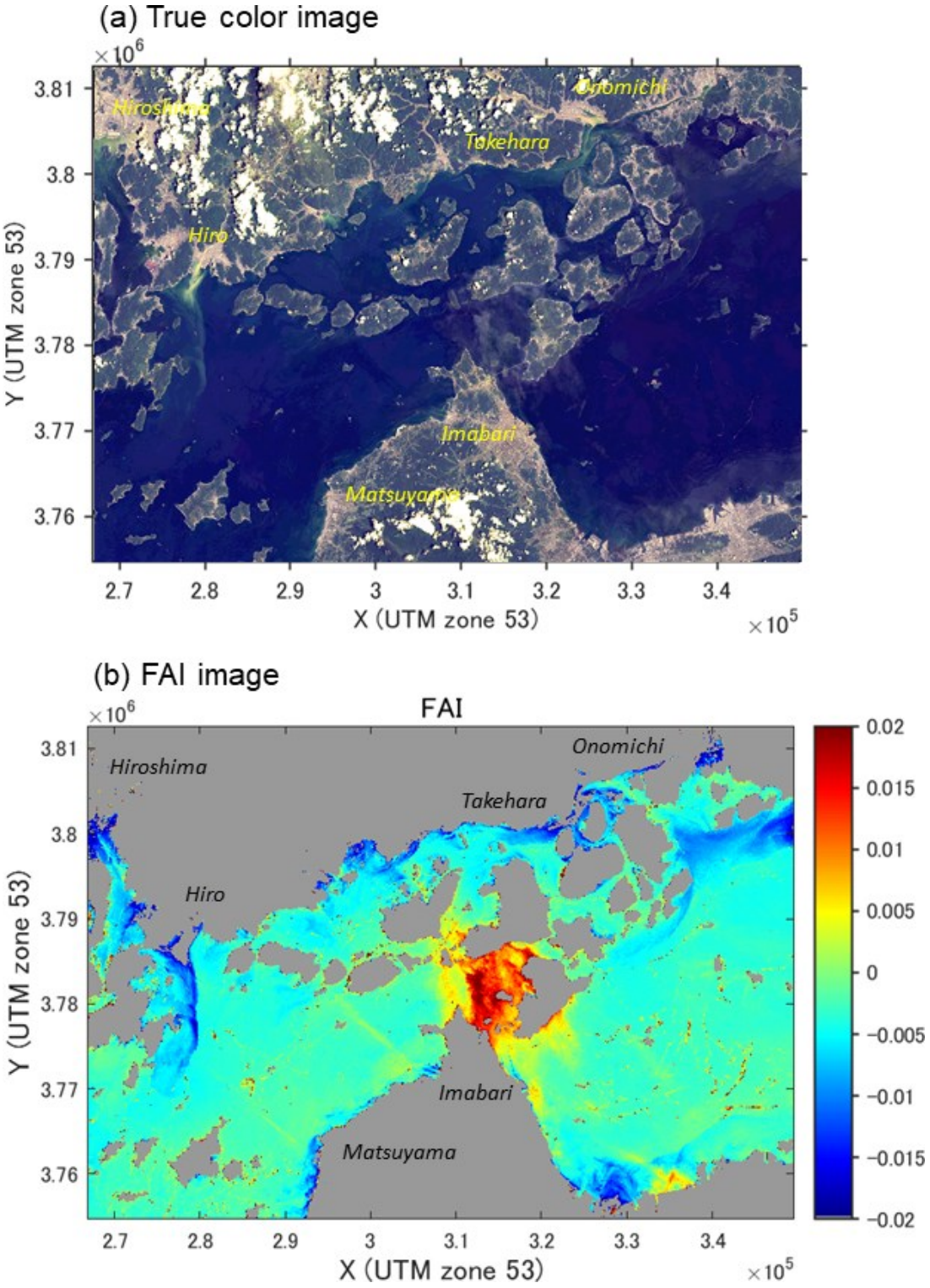
88. Wang, D.; Cheng, B. An Unsupervised Classification Method of Remote Sensing Images Based on Ant Colony Optimization Algorithm. In *Advanced Data Mining and Applications*; Cao, L., Feng, Y., Zhong, J., Eds.; Lecture Notes in Computer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; Vol. 6440, pp. 294–301 ISBN 978-3-642-17315-8.

89. Lv, Z.; Hu, Y.; Zhong, H.; Wu, J.; Li, B.; Zhao, H. Parallel K-Means Clustering of Remote Sensing Images Based on MapReduce. In *Web Information Systems and Mining*; Wang, F.L., Gong, Z., Luo, X., Lei, J., Eds.; Lecture Notes in

Computer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; Vol. 6318, pp. 162–170 ISBN 978-3-642-16514-6.

90. Irvin, B.J.; Ventura, S.J.; Slater, B.K. Fuzzy and Isodata Classification of Landform Elements from Digital Terrain Data in Pleasant Valley, Wisconsin. *Geoderma* **1997**, *77*, 137–154, doi:10.1016/S0016-7061(97)00019-0.
91. Jain, A.K. Data Clustering: 50 Years beyond K-Means. *Pattern Recognition Letters* **2010**, *31*, 651–666, doi:10.1016/j.patrec.2009.09.011.
92. Wang, H.; Zhang, X.; Wu, W.; Liu, H. Prediction of Soil Organic Carbon under Different Land Use Types Using Sentinel-1/-2 Data in a Small Watershed. *Remote Sensing* **2021**, *13*, 1229, doi:10.3390/rs13071229.
93. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **1960**, *20*, 37–46, doi:10.1177/001316446002000104.
94. McHugh ML. Interrater Reliability: The Kappa Statistic. Biochem Med (Zagreb). 2012;22(3):276-82. PMID: 23092060; PMCID: PMC3900052.Https://Www.Ncbi.Nlm.Nih.Gov/Pmc/Articles/PMC3900052/; Accessed on Dec. 12, 2023.
95. Cohen's Kappa, DATAtab, Https://Datatab.Net/Tutorial/Cohens-Kappa; Accessed on Dec. 5, 2023.

Appendix A


(a) True color image
×10^6
3.81
3.8
3.79
3.78
3.77
3.76
Y (UTM zone 53)
Hiroshima
Onomichi
Takehara
Hiro
Imabari
Matsuyama
2.7
2.8
2.9
3
3.1
3.2
3.3
3.4
X (UTM zone 53)
×10^5
(b) FAI image
FAI
×10^6
Hiroshima
Onomichi
Takehara
Hiro
Imabari
Matsuyama
0.02
0.015
0.01
0.005
0
−0.005
−0.01
−0.015
−0.02

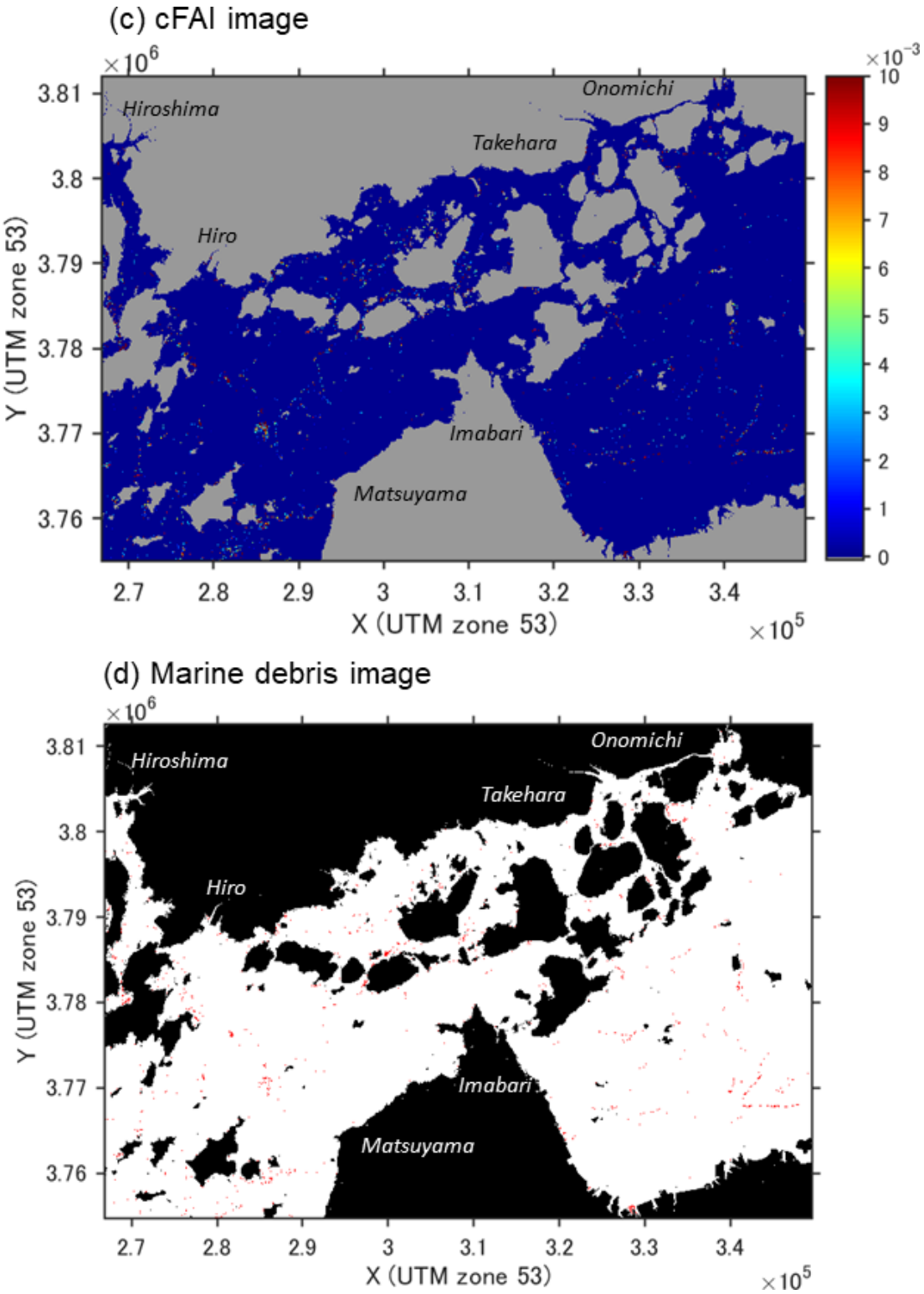


Figure A1. (**a**) True color, (**b**) FAI, (**c**) cFAI, and (**d**) marine debris images calculated from Landsat-8 data acquired on 9 July 2018. Results over a wider area than that in Figure 7 are shown.