ML and LiDAR to Map Aquatic Vegetation

Dr. Meisam Amani

Candace Macdonald

Dr. Sahel Mahdavi

Authors Amani, Macdonald, Mahdavi, Gullage, and So developed a machine learning algorithm to classify different aquatic vegetation.

Who should read this paper?

This paper will be of interest to those involved in marine science and remote sensing applications. The study investigates the application of LiDAR remote sensing data along with machine learning algorithms for aquatic vegetation mapping.

Why is it important?

The study utilizes remote sensing technology which is safer and cheaper for marine habitat mapping compared to traditional methods, such as field surveys. Moreover, the study employs advanced machine learning algorithms, which have more advantages compared to traditional methods, such as statistical models. The proposed method is an automatic algorithm which can be effectively applied to different locations. The ocean community can re-implement the proposed machine learning model to map aquatic vegetation in a cost-, time-, and computation-efficient approach.

About the authors

Dr. Meisam Amani is currently a senior remote sensing engineer and the key specialty leader of data analytics at Wood PLC, a global consulting and engineering company, where he manages and leads various industrial, governmental, and academic remote sensing projects worldwide. Over the past 11 years, he has worked on different applications of remote sensing, including but not limited to land cover/land use classification, soil moisture estimation, drought monitoring, water quality assessment, watershed management, power/transmission line monitoring, fog detection and nowcasting, and ocean wind estimation. To do these, Dr. Amani has utilized various remote sensing datasets (e.g., UAV, optical, LiDAR, SAR, scatterometer, radiometer, and altimeter) along with different machine learning and big data processing algorithms. A list of his research works, including over 50 peer-reviewed journal and conference papers, can be found at https://www.researchgate.net/profile/Meisam_Amani3.

Candace Macdonald, B.Sc., A. Dip., is a remote sensing and GIS specialist with a focus on topo-bathymetric LiDAR applications. Since 2010, she has contributed to various publications applying LiDAR analysis to storm surge and flood risk mapping and habitat classification, particularly intertidal macroalgae habitats. She has been with Wood since 2019 and is currently involved in a wide range of environmental assessment and remediation projects.

Dr. Sahel Mahdavi is currently affiliated with the Data Analytics team at Wood PLC. She has almost 11 years of academic and industrial background in remote sensing (RS) and is familiar with a wide array of topics relevant to RS/GIS and their applications in various environmental aspects. She has authored more than 40 peer-reviewed papers.

Mardi Gullage holds a bachelor of science and a postgraduate diploma in GISciences – both from Memorial University of Newfoundland. Previously, she has been employed as a research assistant at Memorial University and as a radar vision specialist at C-CORE in St. John's, using radar imagery to predict river ice jams, downstream flooding, and iceberg detection. Since 2009, she has been employed at Fisheries and Oceans Canada as a physical scientist and senior GIS analyst in the Ecosystems Management Branch where she supervises a GIS unit responsible for a variety of projects related to marine conservation, fish and fish habitat protection, species at risk, and aquaculture.

Mardi Gullage

Justin J. So, M.Sc., is trained in marine biology and science communications and has been working as a biologist with Wood since 2011. He has worked mainly in the fields of habitat monitoring, environmental effects monitoring, and environmental assessment. Mr. So has been involved in planning, implementation, and analysis for numerous coastal and offshore underwater visual surveys.

Justin J. So

AQUATIC VEGETATION MAPPING USING MACHINE LEARNING ALGORITHMS AND BATHYMETRIC LIDAR DATA: A CASE STUDY FROM NEWFOUNDLAND, CANADA

Meisam Amani[1], Candace Macdonald[2], Sahel Mahdavi[1], Mardi Gullage[3], and Justin So[4]
[1]*Wood Environment and Infrastructure Solutions, Ottawa, O.N., Canada*
[2]*Wood Environment and Infrastructure Solutions, Halifax, N.S., Canada*
[3]*Fisheries and Oceans Canada, St. John's, N.L., Canada*
[4]*Wood Environment and Infrastructure Solutions, St. John's, N.L., Canada*

ABSTRACT

Aquatic vegetation (AV) mapping is an important component of ecosystem-based management of coastal marine environments. A large portion of Newfoundland's coastal areas is covered by different AV types, which should be correctly mapped using recent advanced technologies. In this regard, remote sensing (RS) datasets, such as those acquired by a light detection and ranging (LiDAR) system, are great resources to accurately discriminate different AV. Therefore, in this study, bathymetric LiDAR data were used within an RS method to classify AV over Pistolet Bay and Bonne Bay on the west coast of Newfoundland. An unsupervised object-based classification algorithm was developed for this because reliable in-situ data (i.e., field survey GPS points), required for a supervised classification, were not available. Several post-processing steps were finally performed to reduce the errors and increase accuracy. It was visually observed that the accuracies of the produced AV maps were reasonable considering the limitations of the provided in-situ and LiDAR data. Moreover, the produced map of Pistolet Bay was compared with 48 reference samples of the eelgrass, rockweed, and non-vegetation classes, generated using Google Earth imagery. The results showed that 69% of the samples (33 samples) were correctly classified, indicating the high potential of the developed RS method for AV classification. Finally, several suggestions, including validating the results using in-situ data and utilizing better LiDAR data, were provided to improve the results in future studies.

KEYWORDS

Aquatic vegetation; Bathymetric LiDAR; Machine learning; Remote sensing; Classification

INTRODUCTION

Coastal marine environments are rich in biodiversity and form habitats that are populated by various types of flora and fauna. These environments are vulnerable to the effects of climate change and human-induced threats, such as the introduction of aquatic invasive species, fishing and aquaculture, as well as shipping and marine recreational activities [Klemas, 2016; Koch, 2001; Wedding et al., 2008]. Therefore, it is important to monitor these valuable natural resources. In particular, macroalgae and seagrasses are important components of coastal habitats as they enhance productivity, add habitat complexity, and function as nursery and rearing habitat for marine organisms [Teagle et al., 2017; DFO, 2009; 2013]. Changes to their community composition may also be reflective of influences of anthropogenic activities on the marine environment [Wells et al., 2007]. In this regard, one of the important steps is aquatic vegetation (AV) mapping using advanced techniques.

AV mapping can be generally conducted using field surveys and remote sensing (RS) methods. Although field survey is the most accurate approach for AV mapping, it has several limitations. For instance, field surveys can be time-consuming, costly, dangerous, and even sometimes impossible in inaccessible coastal environments. Additionally, they typically cover small areas and are discrete data points, requiring the user to interpolate the data gaps between locations. However, RS offers timely, cost-effective, and safe tools for AV mapping and monitoring purposes in almost any place in the world [Bostater, Jr. et al., 2004; Rowan and Kalacska, 2020].

Three types of RS datasets can be mainly used for AV mapping: optical, sound navigation and ranging (sonar), and dight detection and ranging (LiDAR). Each of these systems has its own advantages and disadvantages [Klemas, 2016]. For example, although optical satellite data are very cost-efficient and have few post-processing requirements, the accuracy highly depends on water quality and can only be used on non-cloudy days and over shallow water bodies [Ghirardi et al., 2019; Klemas, 2016]. Moreover, although sonar systems (e.g., multibeam echo sounder) have the potential to be used for accurate definition of seafloor habitat mapping, these systems usually lose efficiency as the water depth increases and the corresponding data require complex post-processing and interpretation steps. Sonar systems which are usually mounted on ships may not be able to easily access nearshore areas [Bio et al., 2020; Greene et al., 2018]. Finally, LiDAR systems provide relatively cost-effective datasets with a high spatial density. Various bathymetric LiDAR systems can reach wide ranges of depths (up to 70 m). However, the post-processing of LiDAR data can also be relatively time-consuming and the accuracy depends on the water condition [Collin et al., 2011; Wedding et al., 2008].

In western Newfoundland, many coastal environments can be challenging to map due to inaccessibility or lack of both high-quality sound and very high-resolution optical satellite data. However, bathymetric LiDAR data can be effectively used to characterize coastal habitats and can overcome some of these challenges. Therefore, through this study, we developed a machine learning algorithm to classify different AV in Bonne Bay and Pistolet Bay on the west coast of Newfoundland. To this end, we processed the bathymetric LiDAR data, and created several products to be used in the classification. The details of the methodology, analysis, and results along with several suggestions to improve the analysis in the future are provided in the following sections.

STUDY AREA AND DATA

Study Area

The study areas are Pistolet Bay and Bonne Bay, located on the island of Newfoundland (Figure 1). Pistolet Bay is located on the northern tip of the Great Northern Peninsula, with the approximate central geographical coordinates of 51° 33′ N, 55° 48′ W. Pistolet Bay can be categorized as fairly shallow throughout, with several rivers and estuaries along its coastline. It has a rugged coastline that ranges from exposed bedrock to multiple sandy beaches.

With the central geographical coordinates of 49° 33′ N, 57° 55′ W, Bonne Bay is located on the western side of Newfoundland and is adjacent to Gros Morne National Park. It is a naturally deep bay surrounded by steep mountains. The steep and rocky terrain extends into the water in many places and, thus, there are often steep drop-offs in the water adjacent to the shoreline.

AV Classes

There are different types of AV over the study areas. However, based on the discussions with several scientists, Kelp, Eelgrass, Irish Moss, and Rockweed are commonly found over these two study areas. Therefore, these four AV were initially used in this study. The characteristics of these four AV types are described below.

Figure 1: The locations of the study areas (Pistolet Bay and Bonne Bay).

It should be noted that these characteristics vary at different locations and in the reports by different research studies. Therefore, the characteristics which are potentially useful for the purpose of this study are summarized. These characteristics were obtained from discussions with several biologists who were familiar with the study areas as well as the following references: [DFO, 2009, 2013; Slater Museum of Natural History, n.d.; OCEANA, n.d.; Canadian Encyclopedia, n.d.; University of Maine, n.d.; Vandermeulen, 2005].

- **Kelp (e.g., *Laminaria sp.*):** (1) it usually grows at depths between 5-30 m; (2) its height is usually more than 1 m; (3) it can grow on rock, scallops, horse mussels, gravel, and mud bottoms; (4) it usually grows in subtidal zones in Newfoundland.
- **Eelgrass (e.g., *Zostera marina*):** (1) it usually grows at depths of 1-5 m, but can be sometimes observed at deeper areas as well; (2) its height is usually between 0.2-

1.5 m; (3) it usually grows on sand, mud, and pebble bottoms; (4) it usually grows in subtidal zones.

- **Irish Moss (e.g., *Chondrus crispus*):** (1) it usually grows at depths of 1-2 m; (2) its height is usually less than 0.2 m; (3) it usually grows on rocks near coastlines; (4) it usually grows in intertidal to shallow subtidal zones and it can move around.
- **Rockweed (e.g., *Ascophyllum nodosum*):** (1) it usually grows at depths of 0-3 m; (2) its height is usually between 0.2-0.5 m. However, rockweeds with the height of less than 0.2 m or between 0.5-1 m have been also observed in some regions and conditions; (3) it usually grows on rock and boulder bottoms near coastlines; (4) it usually grows in intertidal to shallow subtidal zones.

Considering the presence of other AV and non-vegetated areas over the study areas, as well as the challenges of correctly identifying the four main AV discussed above, the final AV classes which were used in this study were Eelgrass, Eelgrass/Kelp, Eelgrass/Other Vegetation, Eelgrass/Rockweed, Non-Vegetation, Non-Vegetation/Irish Moss, and Rockweed. It should be noted that having two different types of AV in a single class means that the area can be either of the two classes. For example, if an area is classified as Eelgrass/Kelp, it means that the area could be Eelgrass or Kelp.

Reference Data

Although some public reference data were available from the study areas, there were several problems associated with those data. The first issue was that only a small part of each study area was covered by the available field data. However, the larger issue was in the data themselves, which comprised large polygons representing AV types. Several polygons overlapped each other to a large extent, making them ineffective for identifying one type of AV from another. Field data to be used in a classification should represent one species and would ideally be discrete point data instead of polygonal areas. As a result, it was not possible to use these datasets for their intended purpose of training a supervised classification in this study.

Based on the above explanation, an unsupervised classification algorithm was developed for AV mapping, for which there is no need for reference/ground truth data. However, to perform accuracy assessment, we investigated the Pistolet Bay area using the visual interpretation of the multi-temporal high-resolution Google Earth images, and several areas of Eelgrass, Rockweed, and Non-Vegetation (e.g., mud, bedrock, and sand) were identified. Figure 2 demonstrates the distribution of these reference samples over Pistolet Bay. In total, seven, 23, and 18 samples of Eelgrass, Rockweed, and Non-Vegetation were generated, respectively. These reference samples were used in the statistical accuracy assessment of the produced maps.

LiDAR Data

LiDAR data were collected over both Pistolet Bay and Bonne Bay. The start and the end date of bathymetric data collections were August 15, 2017, and November 1, 2017, respectively. The positional and sounding accuracies of the data have been reported to be 0.1 m and 0.25 m, respectively, and the spatial resolution of the data was 5 m. Figure 3 illustrates the

Figure 2: The locations of the visually selected reference samples.

LiDAR data coverage over these two study areas. As is clear from the image, the coverage of water bodies over Pistolet Bay was higher than that of Bonne Bay, due to the shallower depths in Pistolet Bay.

METHODOLOGY

The flowchart of the method to produce the object-based AV maps for both study areas using the LiDAR data is illustrated in Figure 4. The detail of each step is also provided below.

1. The LiDAR point cloud data were provided in the LASer (LAS) format, which is an industry-standard binary format for storing airborne LiDAR data with a low size. In the first step, a LAS dataset was created from the LiDAR data for each of the study areas, and the correct datum (NAD 1983) and projection (UTM 21N) were assigned to the data. Then, the LAS point files were processed to generate several LiDAR raster products which contained useful information (e.g., AV height and digital elevation model (DEM) for AV mapping). These LiDAR products are illustrated in Figure 5 for the selected area indicated in Figure 3A and are briefly described below. Each of these products contains specific information, where combining them within a machine learning algorithm would improve the accuracy. It is, finally, worth noting that there were some flight line artifacts in the LiDAR data (see Figure 5E as an example) which caused a limitation in AV classification model. Since the data were provided as post-processed but unclassified LAS files lacking trajectory information, we were not able to redo any post-processing of the LiDAR points to remove these flight line artifacts.

o **DEM:** DEM shows the elevation of land surface or ocean floor for each resolution cell. Since each resolution cell in LiDAR data contains several returns,

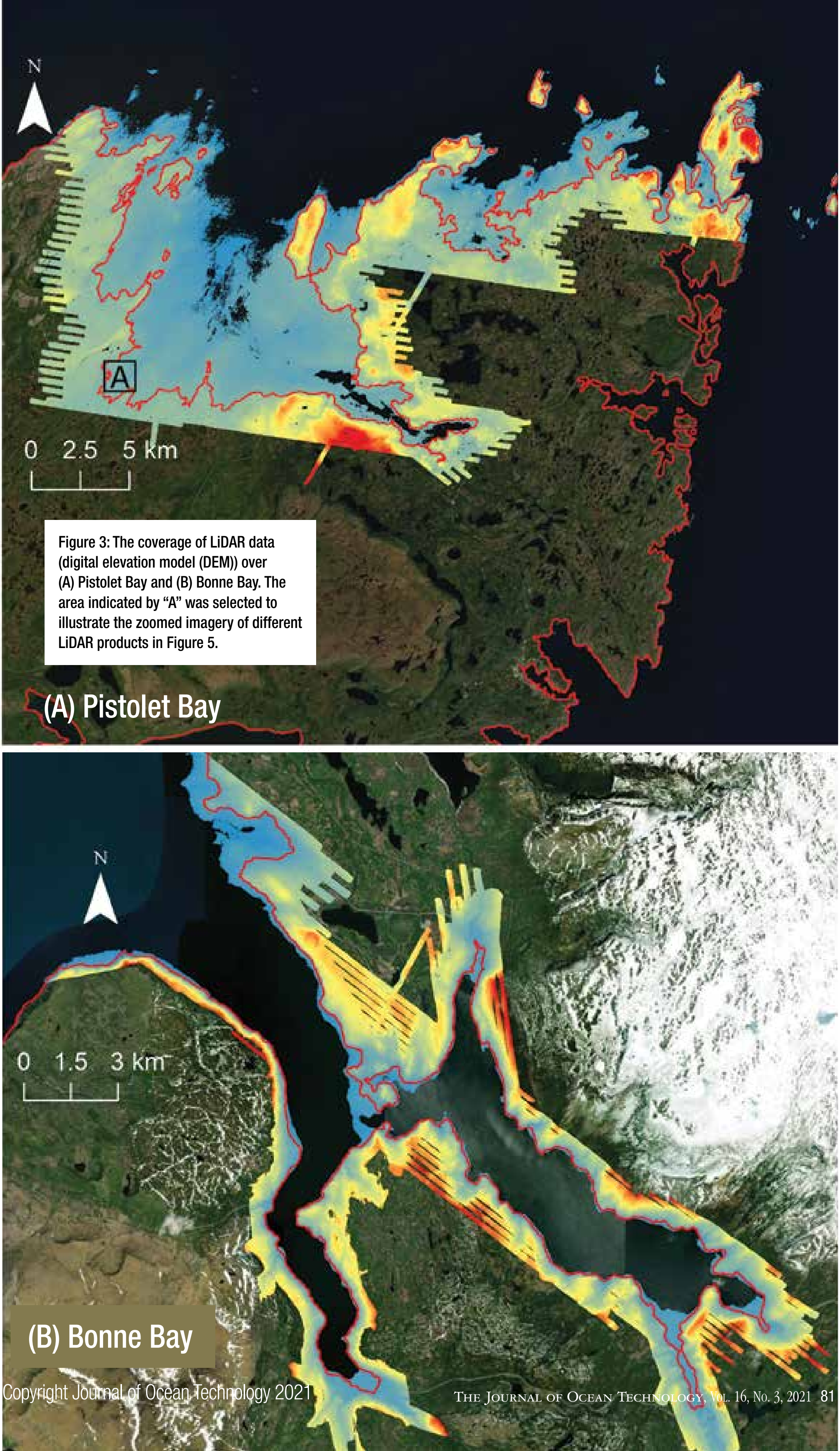


Figure 3: The coverage of LiDAR data (digital elevation model (DEM)) over (A) Pistolet Bay and (B) Bonne Bay. The area indicated by "A" was selected to illustrate the zoomed imagery of different LiDAR products in Figure 5.

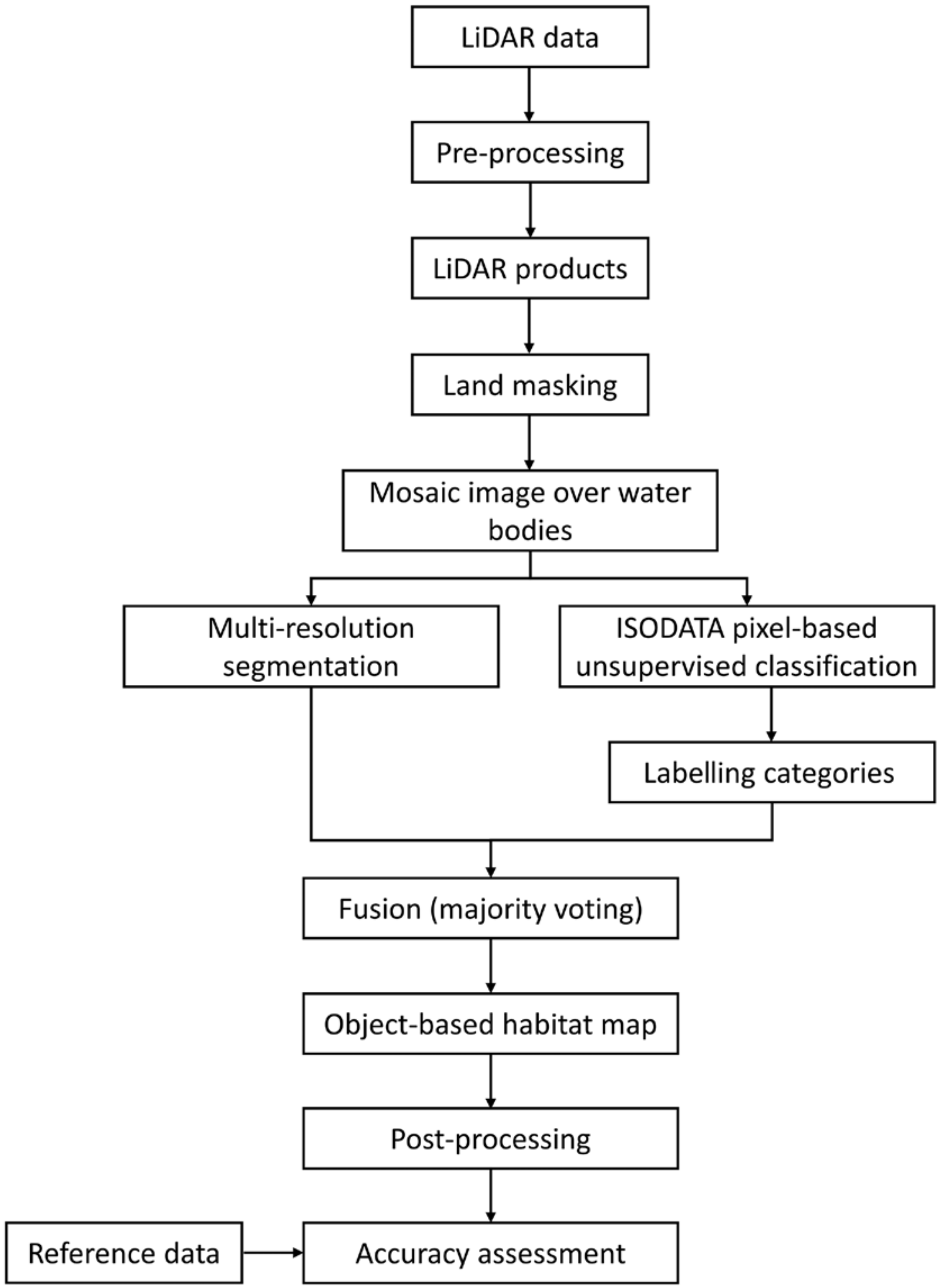


Figure 4: The flowchart of the classification model for aquatic vegetation (AV) mapping.

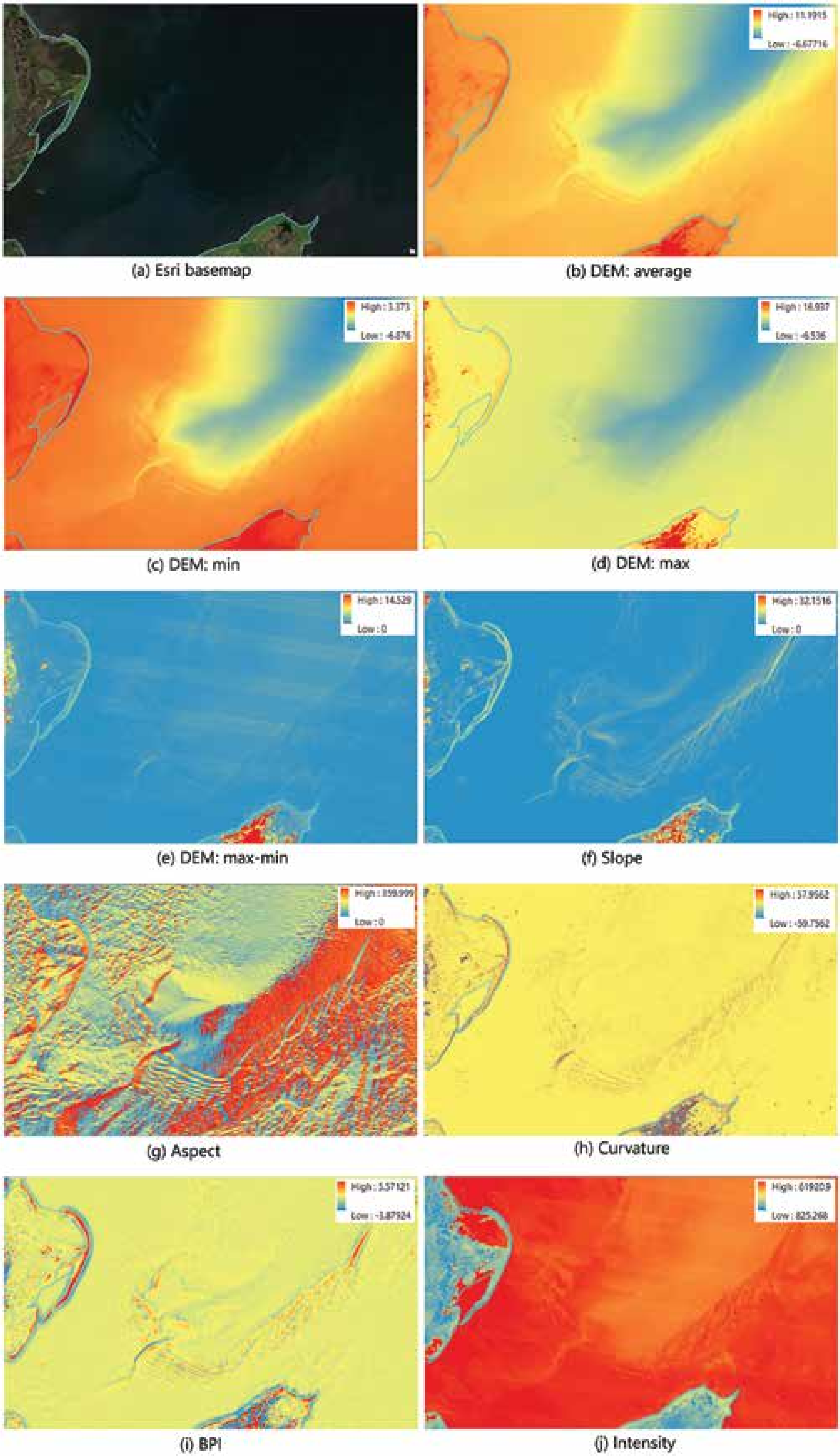


Figure 5: A part of the Pistolet Bay study area (Region A in Figure 3) along with the LiDAR parameters which were used in this study for aquatic vegetation (AV) classification.

DEM can be calculated from the average (Figure 5B), minimum (Figure 5C), and maximum (Figure 5D) returns. In this study, all three DEMs were created from the LiDAR data. The minimum DEM products can be used to estimate ocean depth, especially over non-vegetated areas. The difference between the DEM derived from the maximum and minimum returns was also generated (Figure 5E). This product is an indicator of the AV heights.

o **Slope and Aspect:** Slope and Aspect are created from the DEM. Slope shows the rate of the change of elevation and Aspect demonstrates the direction of slope. In this study, slope (Figure 5F) and aspect (Figure 5G) were generated from the DEM of the average return.

o **Curvature:** The curvature function displays the shape or curvature of the slope and shows its concavity or convexity. The curvature parameter was also derived from the DEM of the average return (Figure 5H).

o **Bathymetric Position Index (BPI):** This index is a second derivative of the DEM and is a measure of the elevation change of a set of cells in DEM relative to its surroundings (Figure 5I).

o **Intensity:** Intensity of LiDAR data was also used to improve the accuracy of classification (Figure 5J). Intensity indicates the strength of the return pulse. Since different AV types result in various strengths due to several factors, such as density and canopy structure, Intensity can be effectively used for AV discrimination. In fact, Intensity was the most useful layer for AV mapping within the developed machine learning algorithm in this study.

2. The land areas were masked and only LiDAR products over water bodies remained for the classification.
3. All the LiDAR products were overlayed and a single mosaic image over water bodies with nine layers was produced. These layers were the minimum, maximum, average, and difference between maximum and minimum returns of LiDAR data, as well as slope, aspect, curvature, BPI, and intensity. This mosaic image was the input for the classification model.
4. It has been widely reported that a supervised classification algorithm provides more accurate and more reliable results compared to unsupervised methods [Mahdavi et al., 2018]. This is true, but only when the field data used to train the supervised classification are highly accurate and are collected within a relatively short time before or after the collection of the LiDAR or other RS data used in the model. Consequently, considering the lack of in-situ data for conducting a supervised classification, an unsupervised classification method was implemented in this study. To this end, the mosaicked image was fed into an unsupervised classification algorithm, called the Iterative Self-Organizing Data Analysis Technique (ISODATA). ISODATA calculates the average of the classes evenly distributed in the data space, and classifies the pixels based on their distance to the class centres. Based on the pixels assigned to a class centre, new means are calculated. This process continues until a convergence is reached [Geoffrey and Hall, 1965].
5. Although an unsupervised classification algorithm divides the imagery into spectrally similar categories, it does not provide the label of classes and requires the modeller

to identify the ground or seabed feature in each class. Therefore, in this study, based on the characteristics of AV, as well as visual interpretation of Google Earth images, labels of different categories were assigned.

6. It has been extensively argued that object-based methods are superior to pixel-based algorithms for classification. One of the reasons is that the outputs of object-based classification methods are closer to real-world objects. Moreover, the classification does not have a noisy structure because single or a few pixels with very high or very low values were merged within larger segments. Therefore, in this study, an object-based classification method was implemented. In an object-based method, the first step is segmenting the study area. To this end, the multi-resolution segmentation algorithm, available in the eCognition software (https://geospatial.trimble.com/products-and-solutions/ecognition), was applied to segment the study areas into homogenous objects. This algorithm is a bottom-up region merging technique that groups neighbouring pixels based on the homogeneity criteria [Baatz and Schape, 2000].

7. After producing both classification and segmentation, they were automatically fused using the majority voting method. Figure 6 illustrates the procedure of the fusion through majority voting.

8. To improve the accuracy of the produced object-based map, several post-processing steps were performed. For example, the produced maps were compared to the multi-temporal high resolution Google Earth imagery and the label of the classes were revised where there was a misclassification.

9. Finally, the accuracy of the map was assessed both visually and statistically using Google Earth imagery and the reference data, respectively.

RESULTS

Figure 7A,B illustrates the initial unsupervised pixel-based classifications of Pistolet Bay and Bonne Bay using 20 and 15 categories. A zoomed area from Pistolet Bay is also demonstrated in Figure 7C to show more details in these maps. The maps are significantly noisy and contain no information about different AV classes over the study areas. Furthermore, the flight line artifacts are clear in these maps (see Figure 7C). These, in fact, show the importance of the subsequent process (i.e., labelling the classes, generating the object-based maps, and post-processing of the maps) to produce more accurate AV maps.

Figure 8 shows the final classified AV maps of both study areas obtained from the proposed classification model. A zoomed area of the map from Pistolet Bay, along with the corresponding very high-resolution Google Earth image, are also provided in Figure 9. The accuracies of these classified maps were visually assessed by comparing them with the high-resolution Google Earth images, and it was observed that the identified areas had a good match with the detectable AV types in the study areas. Moreover, it was observed that the produced final maps had more accuracy and are less noisy compared to the initial unsupervised pixel-based maps (compare Figure 7, Figure 8, and Figure 9).

The areas of different AV classes were also calculated from the final classified maps and

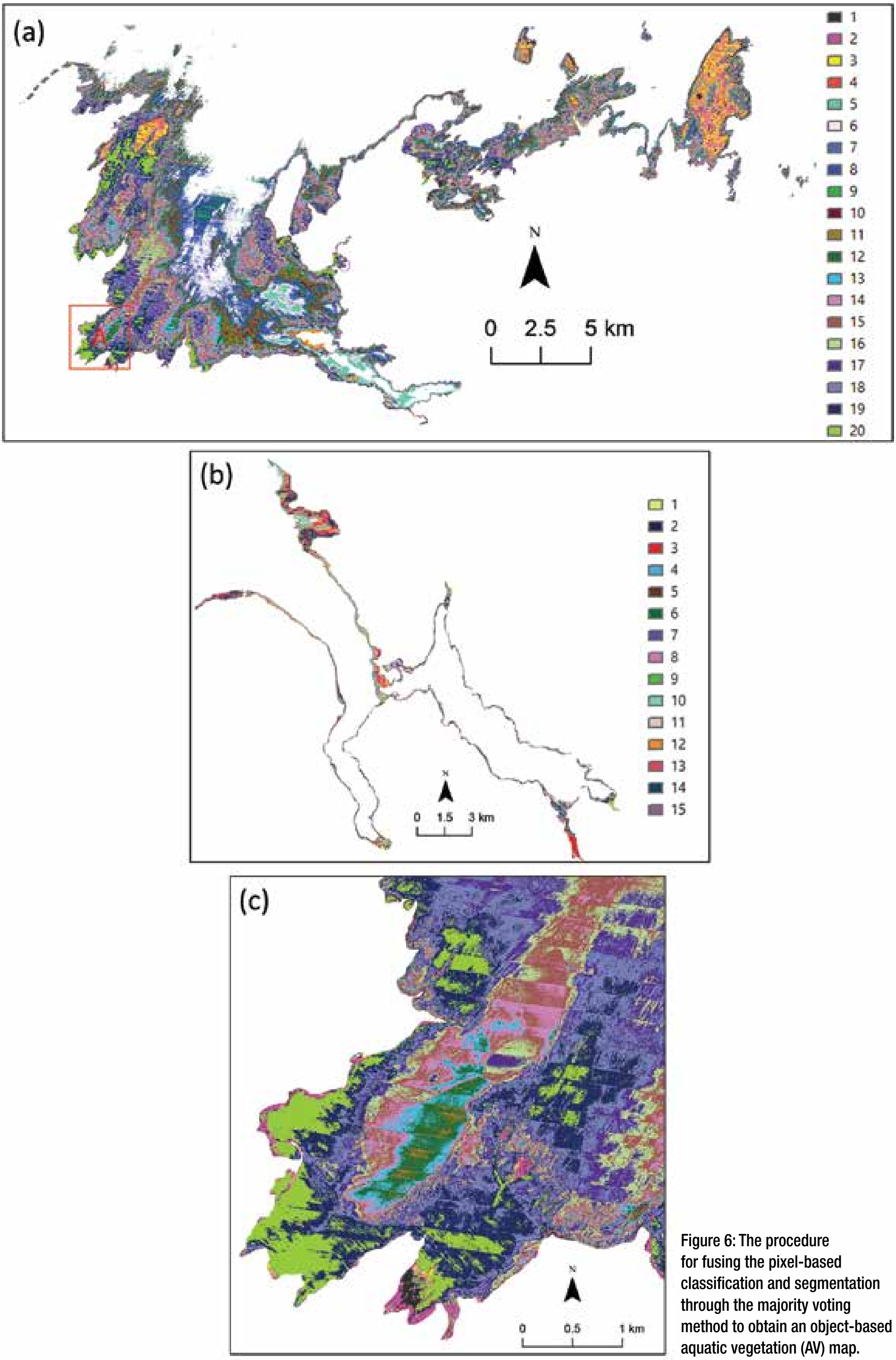


Figure 6: The procedure for fusing the pixel-based classification and segmentation through the majority voting method to obtain an object-based aquatic vegetation (AV) map.

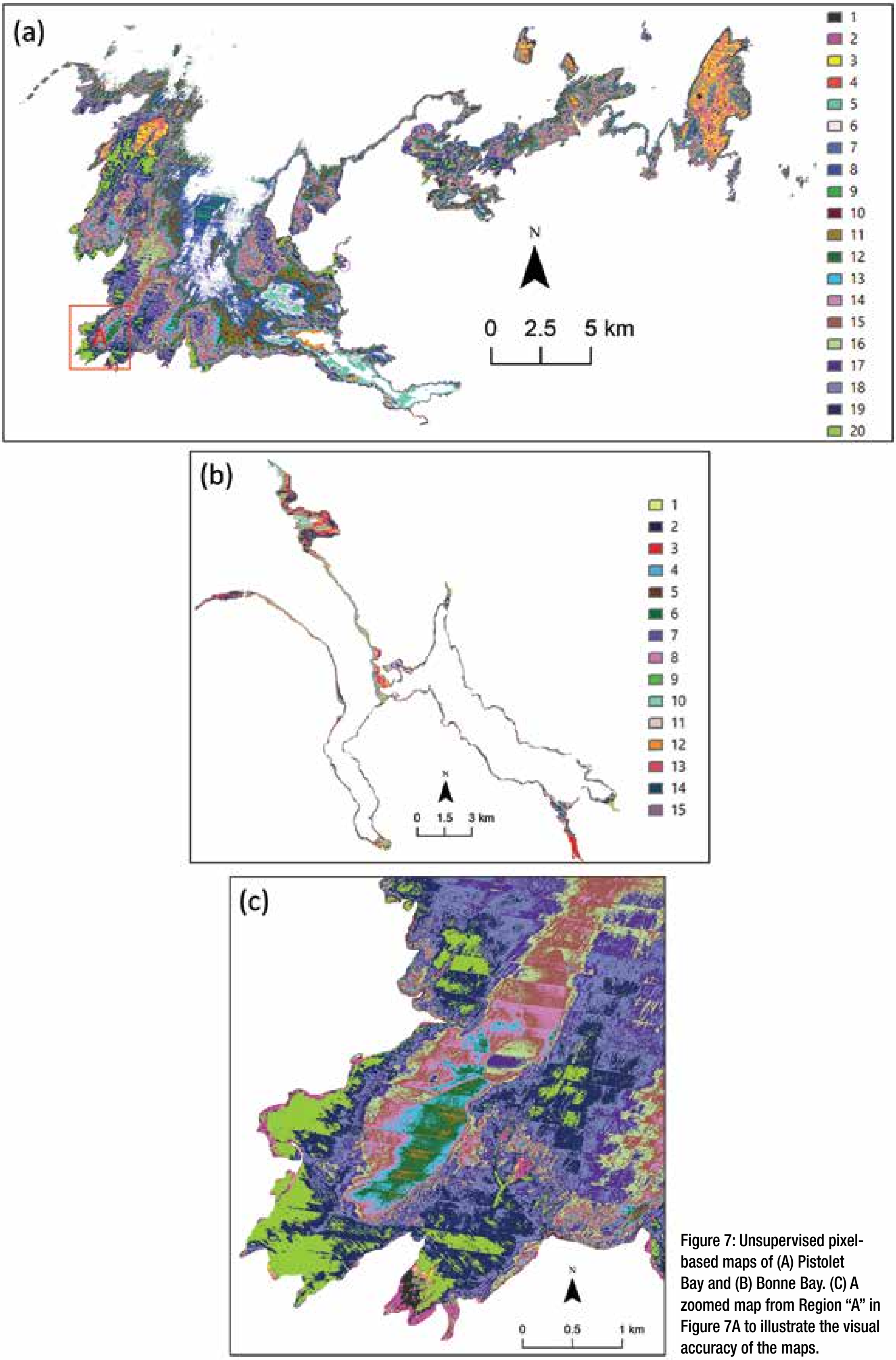
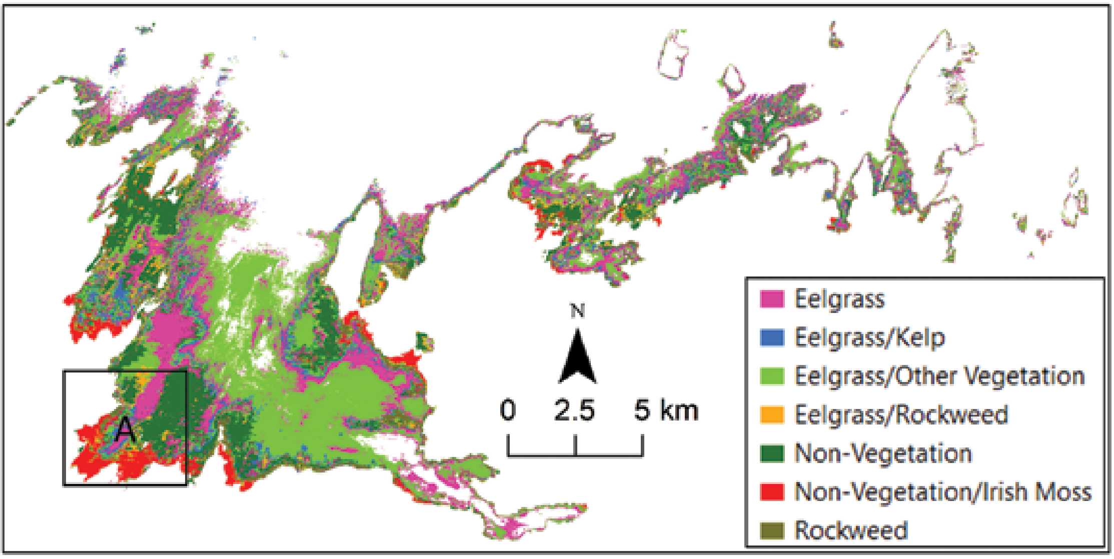
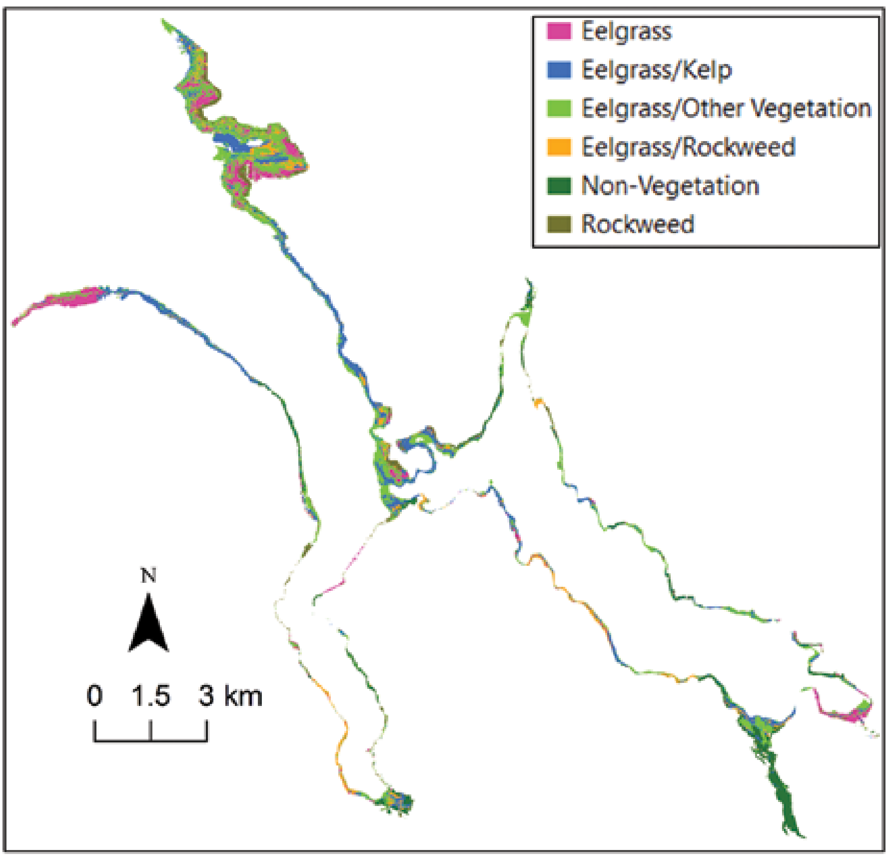


Figure 7: Unsupervised pixel-based maps of (A) Pistolet Bay and (B) Bonne Bay. (C) A zoomed map from Region "A" in Figure 7A to illustrate the visual accuracy of the maps.



(a)



(b)

Figure 8: Aquatic vegetation (AV) maps of (A) Pistolet Bay and (B) Bonne Bay.

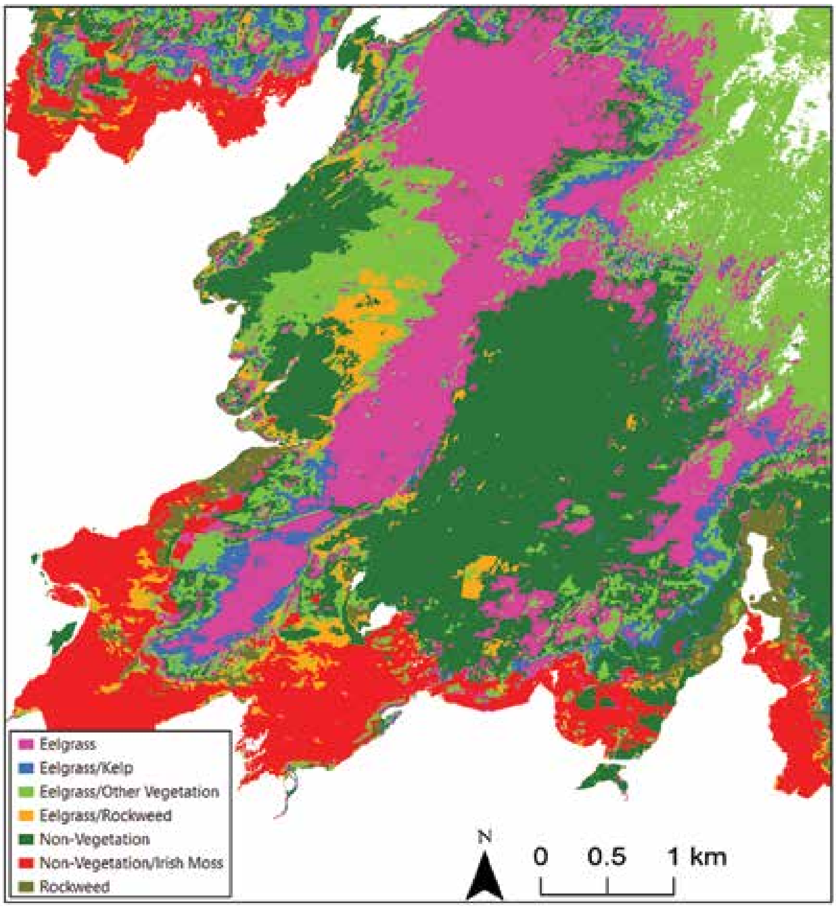


Figure 9: (A) A zoomed image from the aquatic vegetation (AV) map of Pistolet Bay (Region A in Figure 8) along with the (B) very high-resolution Google Earth imagery acquired in March 2020 to demonstrate the visual accuracy of the maps.

the results are reported in Table 1. In Pistolet Bay, most areas are dominated by Eelgrass/Other Vegetation (58.19 km^2), Eelgrass (37.30 km^2), and Non-Vegetation (30.27 km^2). In Bonne Bay, Eelgrass/Other Vegetation (5.41 km^2) and Eelgrass/Kelp (3.65 km^2) had the highest coverages.

to produce accurate AV maps. While the overall accuracy of the classification was reasonable, there were several limitations which prevented a better accuracy. In this section, these limitations along with several suggestions were provided to produce more accurate results in future studies.

Class	Area (km^2)	
	Pistolet Bay	Bonne Bay
Eelgrass	37.30	1.97
Eelgrass/Kelp	8.82	3.65
Eelgrass/Other Vegetation	58.19	5.41
Eelgrass/Rockweed	5.29	1.59
Non-Vegetation	30.27	2.00
Non-Vegetation/Irish Moss	7.55	0
Rockweed	5.38	0.97
Total	**152.8**	**15.59**

Table 1: Area of each aquatic vegetation (AV) class obtained from the produced maps for Pistolet Bay and Bonne Bay.

The accuracies of the maps were also statistically assessed using the generated reference samples and the results are provided in Table 2. To this end, the labels of the classes in the AV maps were compared with those of the reference samples. It was observed that five, 12, and 16 samples out of the seven, 23, and 18 samples of the Eelgrass, Rockweed, and Non-Vegetation were correctly identified, respectively. Therefore, the accuracy of the detection for the Eelgrass, Rockweed, and Non-Vegetation classes were, respectively, 71%, 52%, and 89%. Although the accuracy of the Rockweed class was relatively low, which was mainly due to misclassification with the Non-Vegetation class, these levels of accuracies were high considering the limitations of this study.

LIMITATIONS AND SUGGESTIONS

In this study, an RS method was implemented

One of the main limitations was related to the lack of in-situ data. As discussed, supervised classification algorithms require in-situ data (ideally point data) and provide higher classification accuracies compared to unsupervised methods. Additionally, a portion of available in-situ data are typically reserved to use as validation in post-classification accuracy assessments. Therefore, future studies should first aim to collect ample reliable in-situ samples representing the AV types of interest. The samples should represent locations of homogeneous AV classes (i.e., only eelgrass). High quality imagery from drone surveys would complement future AV characterizations and have advantages for work in more inaccessible environments. The availability of in-situ data will also facilitate a comprehensive accuracy assessment on the maps and will properly reveal the confusion between different classes and limitations of the maps.

Reference sample	AV map	Correctly (✓) and incorrectly (×) identified
Eelgrass	Non-Vegetation/Irish Moss	×
Eelgrass	Rockweed	×
Eelgrass	Eelgrass/Rockweed	✓
Eelgrass	Eelgrass/Other Vegetation	✓
Eelgrass	Eelgrass	✓
Eelgrass	Eelgrass	✓
Eelgrass	Eelgrass	✓
Rockweed	Rockweed	✓
Rockweed	Non-Vegetation	×
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Non-Vegetation	×
Rockweed	Non-Vegetation	×
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Rockweed	✓
Rockweed	Non-Vegetation/Irish Moss	×
Rockweed	Non-Vegetation/Irish Moss	×
Rockweed	Rockweed	✓
Rockweed	Non-Vegetation	×
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Eelgrass/Rockweed	✓
Rockweed	Non-Vegetation	×
Rockweed	Rockweed	✓
Rockweed	Eelgrass/Other Vegetation	×
Rockweed	Eelgrass/Other Vegetation	×
Rockweed	Rockweed	✓
Rockweed	Rockweed	✓
Rockweed	Non-Vegetation	×
Rockweed	Eelgrass	×
Non-Vegetation	Eelgrass/Other Vegetation	×
Non-Vegetation	Non-Vegetation/Irish Moss	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Eelgrass/Other Vegetation	×
Non-Vegetation	Non-Vegetation/Irish Moss	✓
Non-Vegetation	Non-Vegetation/Irish Moss	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation/Irish Moss	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓
Non-Vegetation	Non-Vegetation	✓

Table 2: The reference samples and their identified labels in the classified aquatic vegetation (AV) map of Pistolet Bay.

As discussed, there were some flight line artifacts in the raw LiDAR data (Figure 5). These artifacts are a result of the trajectory post-processing of the raw LiDAR points after the aerial data collection is complete. Although it was tried to improve this error with the classification model, these artifacts caused considerable misclassifications in the produced maps. Acquiring higher quality data in future studies will help to reduce these uncertainties. Another issue with the LiDAR data was the relatively lower vertical accuracy (i.e., 25 cm). It is well-known that there are many AV types, such as Irish Moss, with heights less than 25 cm. This was, in fact, the main reason that Irish Moss was not detected as a separate class in this study. Therefore, to identify short vegetation such as Irish Moss in future studies, other remote sensing products should be explored which could provide higher vertical accuracy. Finally, it is suggested to acquire LiDAR data between August and September, which is the peak growth of AV.

Although the four main AV classes were considered in this study, it is obvious that there are additional types of AV throughout these areas. Neglecting these classes in the classification model was also another limitation of this study. Therefore, in future studies, an investigation should first be performed to identify the dominant AV types in the areas and collect enough in-situ data from those classes. Additionally, the single non-vegetation class would benefit from splitting into different non-vegetation classes to improve the reliability of the final maps.

As discussed, there are inconsistencies in the definition of the characteristics of a single AV type and there is no comprehensive study about the ecological characteristics of various AV types in these areas. In this study, the labelling of the categories in the unsupervised classification were performed using visual interpretation, which has a limitation, and the AV characteristics, which were not consistent in the literature. Therefore, using these methods to identify the type of AV caused several misclassifications. If in-situ point data are available and a supervised classification is utilized in future studies, the corresponding uncertainties will be resolved.

Finally, another approach to improve the classification results is utilizing other types of RS data, including very high-resolution optical (e.g., multispectral satellite) and sonar data. Moreover, very high-resolution drone imagery can be also effectively used over relatively small areas. Each of the LiDAR, multispectral satellite, sonar, and drone datasets have their own advantages for AV classification and combining them will result in a higher classification accuracy. However, it should be noted that using all of these datasets will increase the cost and, thus, an optimum option should be selected based on the budget and level of accuracy.

CONCLUSION

Coastal marine environments are indicators of biodiversity and are habitats for several types of flora and fauna. These environments are negatively impacted by various natural and anthropogenic threats and, therefore, it is important to be monitored using advanced techniques. In this study, several AV were classified using a machine learning algorithm and bathymetric LiDAR data. It was observed that the results were reasonable considering several limitations that existed to obtain a high classification accuracy. It was also suggested to use reliable in-situ and LiDAR data, as well as very high-resolution satellite or drone imagery to obtain more accurate results in future studies.

ACKNOWLEDGMENTS

The authors thank Robert Gregory, Terry Wells, and Philip Sargent for their valuable comments on this paper.

REFERENCES

Baatz, M. and Schape, A. [2000]. *Baatz multiresolution segmentation: an optimization approach for high quality multiscale image segmentation*. Angew. Geogr. Informationsverarbeitung Vol. 12, pp. 12-23.

Bio, A.; Gonçalves, J.A.; Magalhães, A.; Pinheiro, J.; and Bastos, L. [2020]. *Combining low-cost sonar and high-precision global navigation satellite system for shallow water bathymetry*. Estuaries and Coasts. https://doi.org/10.1007/s12237-020-00703-6.

Bostater, Jr., C.R.; Ghir, T.; Bassetti, L.; Hall, C.; Reyeier, E.; Lowers, R.; Holloway-Adkins, K.; and Virnstein, R. [2004]. *Hyperspectral remote sensing protocol development for submerged aquatic vegetation in shallow waters*. Proceedings: SPIE 5233, Remote Sensing of the Ocean and Sea Ice, p. 199. https://doi.org/10.1117/12.541191.

Canadian Encyclopedia [n.d.]. *Irish moss*. Retrieved from: https://www.thecanadianencyclopedia.ca/en/article/irish-moss.

Collin, A.; Long, B.; and Archambault, P. [2011]. *Benthic classifications using bathymetric LiDAR waveforms and integration of local spatial statistics and textural features*. Journal of Coastal Research, Vol. 62, pp. 86-98. https://doi.org/10.2112/SI_62_9.

DFO [2009]. *Does eelgrass* (*Zostera marina*) *meet the criteria as an ecologically significant species*? Retrieved from: https://waves-vagues.dfo-mpo.gc.ca/Library/337549.pdf.

DFO [2013]. *Assessment of Information on Irish Moss, Rockweed, and Kelp Harvests in Nova Scotia*. Retrieved from: https://waves-vagues.dfo-mpo.gc.ca/Library/348493.pdf.

Geoffrey, B. and Hall, D. [1965]. *ISODATA, a novel method of data analysis and pattern classification*.

Ghirardi, N.; Bolpagni, R.; Bresciani, M.; Valerio, G.; Pilotti, M.; and Giardino, C. [2019]. *Spatiotemporal dynamics of submerged aquatic vegetation in a deep lake from Sentinel-2 data*. Water, Vol. 11, No. 563. https://doi.org/10.3390/w11030563.

Greene, A.; Rahman, A.F.; Kline, R.; and Rahman, M.S. [2018]. *Side scan sonar: a cost-efficient alternative method for measuring seagrass cover in shallow environments*. Estuarine Coastal and Shelf Science, Vol. 207, pp. 250-258. https://doi.org/10.1016/j.ecss.2018.04.017.

Klemas, V.V. [2016]. *Remote sensing of submerged aquatic vegetation*. Seafloor Mapping along Continental Shelves, pp. 125-140. https://doi.org/10.1007/978-3-319-25121-9_5.

Koch, E.W. [2001]. *Beyond light: physical, geological, and geochemical parameters as possible submersed aquatic vegetation habitat requirements*. Estuaries, Vol. 24, No. 1. https://doi.org/10.2307/1352808.

Mahdavi, S.; Salehi, B.; Granger, J.; Amani, M.; Brisco, B.; and Huang, W. [2018]. *Remote sensing for wetland classification: a comprehensive review*. GIScience and Remote Sensing, Vol. 55, pp. 623-658. https://doi.org/10.1080/15481603.2017.1419602.

OCEANA [n.d.] *Kelp forests*. Retrieved from: https://oceana.ca/en/marine-life/marine-ecosystems/kelp-forests.

Rowan, G. and Kalacska, M. [2020]. *Remote sensing of submerged aquatic vegetation: an introduction and best practices review*. https://doi.org/10.31219/osf.io/2gpxz.

Slater Museum of Natural History [n.d.] *Rockweed* (*Fucus distichus*). Retrieved from: https://www2.pugetsound.edu/academics/academic-resources/slater-museum/exhibits/marine-panel/rockweed/.

Teagle, H.; Hawkins, S.J.; Moore, P.J.; and Smale, D.A. [2017]. *The role of kelp species as biogenic habitat formers in coastal marine ecosystems*. Journal of Experimental Marine Biology and Ecology, Vol. 492, pp. 81-98.

University of Maine [n.d.]. *Rockweed fact sheet*. Retrieved from: https://extension.umaine.edu/signs-of-the-seasons/indicator-species/rockweed-fact-sheet/.

Vandermeulen, H. [2005]. *Assessing marine habitat sensitivity: a case study with eelgrass (Zostera marina L.) and kelps (Laminaria, Macrocystis)*. Ottawa. https://publications.gc.ca/collections/collection_2013/mpo-dfo/Fs70-5-2011-095-eng.pdf.

Wedding, L.M.; Friedlander, A.M.; McGranaghan, M.; Yost, R.S.; and Monaco, M.E. [2008]. *Using bathymetric lidar to define nearshore benthic habitat complexity: implications for management of reef fish assemblages in Hawaii*. Remote Sensings of Environment, Vol. 112, pp. 4159-4165. https://doi.org/10.1016/j.rse.2008.01.025.

Wells, E.; Wilkinson, M.; Wood, P.; and Scanlan, C. [2007]. *The use of macroalgal species richness and composition on intertidal rocky seashores in the assessment of ecological quality under the European Water Framework Directive*. Marine Pollution Bulletin, Vol. 55, pp. 151-161. https://doi.org/10.1016/j.marpolbul.2006.08.031.