

Use of multispectral UAV images for classification and biomass assessment of seaweed in intertidal rocky shores

Joelen Cruz da Silva

Mestrado em Deteção Remota

Departamento de Geociências, Ambiente e Ordenamento do Território

2021

Orientador

Doutor José Alberto Álvares Pereira Gonçalves, Professor Auxiliar, Faculdade de Ciências (FCUP)

Coorientadora

Doutora Carla Débora Gonçalves Borges, Investigadora de Pós-doutoramento do Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR)

Coorientadora

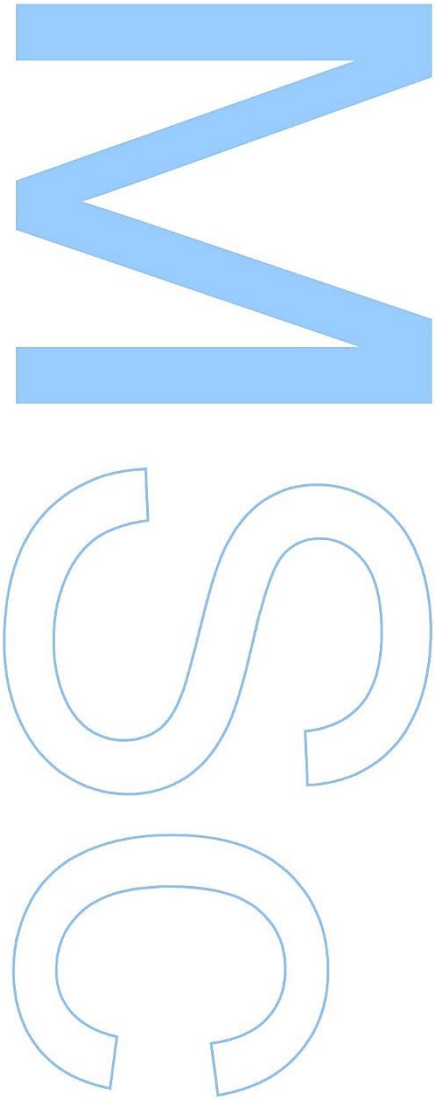
Doutora Lia Bárbara Cunha Barata Duarte, Professora Auxiliar, Faculdade de Ciências (FCUP)

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____



Dissertação submetida à Faculdade de Ciências da Universidade do Porto, para a obtenção do grau de Mestre em Deteção Remota, da responsabilidade do Departamento de Geociências, Ambiente e Ordenamento do Território.

A presente dissertação foi desenvolvida sob a orientação científica do Doutor José Alberto Álvares Pereira Gonçalves, Professor Auxiliar na Faculdade de Ciências (FCUP) com coorientações da Doutora Carla Débora Gonçalves Borges, Investigadora de Pós-doutoramento do CIIMAR e da Doutora Lia Bárbara Cunha Barata Duarte, Professora Auxiliar na Faculdade de Ciências (FCUP).

Acknowledgements

First of all, I would like to thank my supervisors Doctor José Alberto and Doctor Débora Borges for their valuable and essential teaching, collaboration and advice for this work, in addition to all the encouragement and motivation. Thank you for the opportunity and for the trust you have placed in me. I am grateful for being an effective part of the SWUAV project where I learned how to do science in Portugal. At SWUAV I feel part of a strong and friendly team.

To my colleagues from the master and from CIIMAR. I would like to name each one of you, but I was helped by so many people that I would not fit in these few lines here. Those encouraging words worked and I persevered.

To my beloved and dear family that strengthens me daily. My mother Joelma, my sister Joyce and my grandfather João. I left you with the other half of my heart across the Atlantic.

My old Chica, who received the call of God unexpectedly and left me here with the good memories of a happy childhood. I feel like a child when I remember you, my beloved grandmother, I'm sorry for not saying goodbye to you, I carry within me your strength and determination as a warrior woman who raised 9 children by herself.

I also dedicate this work to the person who was by my side, listening to me, supporting and strengthening me, Pétala Sousa. I could never have imagined that I would find someone so special in this adventure called master. I would like to take you to all the other adventures that will arise in this uncertain life I lead.

I cannot omit the technical contribution I received from my great friend Kênia Santos who, even far from me, was present in this study and significantly contributed to the production of Javascript codes, I am very happy to keep this friendship for so many years, I hope one day be able to repay the height.

The winds of education and science brought my sailboat to Portugal, I hope these same winds can take me to my beloved Marajó, where I opened my eyes for the first time and where I dream of closing them for the last time.

Unlike my graduation, where I was always surrounded by many friends and family, the master's was a long lonely journey. A trajectory filled with countless challenges that seem too big, sadness, uncertainty, joy and many setbacks, but despite the lonely process that any researcher is destined for, it brings together contributions from several people, indispensable to finding the best course at every moment of the journey. Traveling this long path was only possible with the support, energy and strength of several people, to whom I especially dedicate this project.

This work was carried out in the scope of the project “SWUAV-Intertidal zone mapping and seaweed biomass assessment using UAV images” (POCI-01-0145-FEDER-032233), supported by POCI (Programa Operacional Competitividade e Internacionalização) and FCT through OE. Within the scope of this project, I was awarded a research grant (SWUAV/BI/2019-021).

“Ao meu lugar encantado e a todas as pessoas, igualmente encantas, que lá habitam ...”
0° 43’23.85” S 48° 31’ 11.18” W
Made in Marajó.

Abstract

Monitoring the intertidal zone is traditionally undertaken by in situ survey methods which have assisted in understanding these complex habitats; however, often only a small area of large reefs is observed. Recent advancements in unmanned aerial vehicles (UAVs) give new options for monitoring large-scale coastal ecosystems by capturing centimeter-resolution images. This work aimed to assess the use of supervised classifications performed on RGB and multispectral images collected by UAVs, in order to map the predominant ecosystem elements of the intertidal zone of a rocky shore in northern Portugal. The supervised classifiers used were: Random Forest (RF), Support Vector Machine (SVM), and the Minimum Distance (MD) algorithm. The predominant ecosystem elements mapped were: seaweed (SW); rock, barnacles and limpets (RBL); sand (S); mussels and rock (MR) and submerged seaweed (SS). Samples of these predominant elements of the ecosystem (encompassing independent training and validation areas) were properly identified and labeled in the field. Data of three flights (July, September and November, 2020) were considered. From the multispectral bands several vegetation indices were calculated. To achieve the proposed objective, three methodologies were established: methodology 1 - classification of all stacked bands; methodology 2 - classification of the vegetation indices bands; and methodology 3 - classification of all stacked bands and vegetation indices. All methodologies were implemented JavaScript and run on the Google Earth Engine program. From the accuracy metrics obtained, it was concluded that methodology 3 presented better results for the Machine Learning classifiers. Additionally, during the study period, there was a decrease in the area covered by the seaweed classes, in line with the known seasonality patterns for those species. This research can serve as a basis for further studies and tests for the application of supervised classification algorithms in other coastal areas supporting future environmental monitoring programs. Such programs are essential for predicting and preventing the impact of climate change upon such sensitive areas and contribute to the conservation of the intertidal zone.

Keywords: Environmental monitoring; Google Earth Engine; High resolution images; Intertidal zone; Coastal mapping; Multispectral.

Resumo

A monitorização da zona intertidal é tradicionalmente realizada por métodos de investigação *in situ* que têm auxiliado na compreensão desses habitats complexos, sendo, contudo, apenas possível avaliar pequenas áreas de recifes extensos. Avanços recentes em Veículos Aéreos Não Tripulados (VANTs) oferecem novas opções para monitorizar ecossistemas costeiros a uma maior escala por meio da captura de imagens com resolução centimétrica. Este trabalho pretendeu avaliar o uso de classificações supervisionadas realizadas em imagens RGB e multiespectrais recolhidas por VANTs, de forma a mapear os elementos do ecossistema predominantes na zona intertidal de uma praia rochosa no norte de Portugal. Os classificadores supervisionados utilizados foram: *Random Forest* (RF), *Support Vector Machine* (SVM) e o *Minimum Distance* (MD). Os elementos do ecossistema mapeados predominantes foram: algas marinhas; rocha, cracas e lapas; areia; mexilhões e rocha; e algas submersas. Amostras desses elementos predominantes do ecossistema (abrangendo áreas de treino e validação independentes) foram devidamente identificadas e rotuladas em campo. Dados de três voos (Julho, Setembro e Novembro de 2020) foram considerados. A partir das bandas multiespectrais foram calculados vários índices de vegetação. Para atingir o objetivo proposto, foram estabelecidas três metodologias: metodologia 1 - classificação de todas as bandas empilhadas; metodologia 2 - classificação das bandas dos índices de vegetação; e metodologia 3 - classificação das bandas empilhadas e índices de vegetação. Todas as metodologias foram implementadas em JavaScript e executadas no programa *Google Earth Engine*. A partir das métricas de *accuracy* obtidas, concluiu-se que a metodologia 3 apresentou melhores resultados para os classificadores de *Machine Learning*. Adicionalmente, verificou-se que durante o período de estudo houve diminuição da área coberta pelas classes de algas de acordo com os padrões de sazonalidade conhecidos para as espécies abrangidas. Esta investigação poderá ser a base para novos estudos e testes de aplicação de algoritmos de classificação supervisionada em outras áreas costeiras, apoiando futuros programas de monitorização ambiental. Esses programas são essenciais para prever e prevenir o impacto das alterações climáticas nestas áreas sensíveis e contribuir para a conservação da zona intertidal.

Palavras-chave: Monitorização ambiental; Google Earth Engine; Imagens de alta resolução; Zona intertidal; Mapeamento costeiro, Multiespectral.

Index

1. Introduction 2

1.1 Objective 5
1.2 Dissertation structure 5

2. Literature Review 6

2.1 Intertidal zone and rocky shores 6
2.2 Unmanned Aerial Vehicle (UAV) 8
2.3 Image classification 10
2.3.1 Minimum distance 11
2.3.2 Random Forest 12
2.3.3 Support Vector Machine 13
2.4 Accuracy metrics 14
2.5 Vegetation Indices 16
2.6 Google Earth Engine 18

3. Methodology 19

3.1 Study area 19
3.2 Fieldwork 20
3.3 Images data 20
3.4 Preprocessing 22
3.5 Processing in Google Earth Engine 23

4. Results and Discussion 28

4.1 Exploratory analysis 28
4.2 Classifications accuracy and visual validation 30
4.3 Temporal variation of the covered area by classes 37

5. Conclusion 40

References 42

Annex I 48

Annex II 51

Annex III 54

Tables index

Table 1 - Confusion matrix. 14
Table 2 - Spectral bands of MicaSense RedEdge-3 sensor. 21
Table 3 - Name of the classes studied and number of training and validation areas by field collection....23
Table 4 - Computed vegetation indices found in the literature and their respective equations. 23
Table 5 - Overall Accuracy (O.A) and Kappa Index (K.I) results. 30
Table 6 - Assessing the accuracy of the classifiers for all months and all methodologies. 36
Table 7 - The parameters that were used for training (per classifier). 54

Figure index

Figure 1 - a) Exemple of a rocky shores (Praia da Congreira - Vila Chã); b) exemple of a sandy shores (Praia das Pastoras - Porto). 6
Figure 2 - Universal zoning scheme by Stephenson and Stephenson,1949 (three zone scheme) - adapted from Raffaelli and Hawkins, 1996. 7
Figure 3 - Search for articles published between the years 2012 and 2020 with the word UAV or DRONE in their titles. Source: Web of science (https://apps.webofknowledge.com/). 9
Figure 4 - Representation of the remote sensing data matrix with elements of the image (pixel) located in Rows (i), columns (j) and band (k). adapted from Jensen, 2007. 10
Figure 5 - Dataset representation and margin adapted from Alpaydin (2004). 13
Figure 6 - Location map of the study area. 19
Figure 7 - a) Ecosystem elements to be classified. b) Geographical coordinates collection for the central point of the quadrats with the GNSS receiver. 20
Figure 8 - a) MicaSense RedEdge-3 sensor. b) UAV DJI Matrice 200. 21
Figure 9 – a) Digital surface model (DSM) and b) photogrammetrically orthorectified image product mosaicked from an image collection (orthomosaic). 22
Figure 10 - a) The central point plotted in QGIS environment. b) Circle created with the buffer tool from the central point (0.25 m radius). 22
Figure 11 - Central code written in JavaScript, band stacking and cutting process. 24
Figure 12 - GEE code for methodology 1: insertion of training areas and supervised classification. 24
Figure 13 – GEE code for methodology 2: a) Calculation of VIs and stacking in a single variable. b) Insertion of training areas and supervised classification 25
Figure 14 - GEE code for methodology 3: creation of the variable containing the bands and VI. Supervised classification on this new variable. 25
Figure 15 – GEE code for insertion of validation samples. 26
Figure 16 - GEE code for calculating accuracy metrics. 26
Figure 17 - k-fold (5-fold) cross validation scheme. In each “iteration”, the examples in one of the data folds is left out for testing and the remaining examples in the k1 fold are used to train the classifier. he performance quality of the trained classifier is tested on the left-out fold (in each “experiment”), and the overall performance measure is then averaged over the k folds (k “experiments”) adapted from Reaelizadeh et al. (2009). 27

Figure 18 – a) Histogram of the multispectral bands for july; b) Histogram of the multispectral bands for september and c) Histogram of the multispectral bands for november.. 28
Figure 19 - a) Orthophoto for the month of July; b) Orthophoto for the month of September; c) Orthophoto for the month of November. In yellow, the region of interest (ROI).....29
Figure 20 - a) Best classifier for july and b) Worst classifier for july. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed. 30
Figure 21 - a) Orthophoto of the july; b) Best classification for july, Support Vector Machine methodology 3 (Overall accuracy: 0.991 and kappa index: 0.994) and c) Worst classification for the month of july, Minimum Distance methodology 2 (Overall accuracy: 0.480 and kappa index: 0.133).....31
Figure 22 - Best classifier for september and b) Worst classifier for september. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.....32
Figure 23 - a) Orthophoto of the September; b) Best classification for September, Random Forest methodology 3 (Overall accuracy: 0.942 and kappa index: 0.883) and c) Worst classification for September, Minimum Distance methodology 2 (Overall accuracy: 0.396 and kappa index: 0.232). 33
Figure 24 - Best classifier for the month of November and b) Worst classifier for the month of November. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed. 34
Figure 25 - Orthophoto of the November; b) Best classification for November, Random Forest methodology 3 (Overall accuracy: 0.937 and kappa index: 0.88) and c) Worst classification for November, Minimum Distance methodology 3 (Overall accuracy: 0.771 and kappa index: 0.543); 35
Figure 26 - Comparative thematic map between the supervised classifications of Random Forest (left) and Support Vector Machine (right) methodology 3 for july, september and november.....37
Figure 27 - Percentage of area covered by study class. RF: Random Forest, SVM: Support Vector Machine, JUL: July, SEP: September and NOV: November. 38

List of abbreviations

GEE - Google Earth Engine

GNDVI - Green Normalized Difference Vegetation Index

GRVI - Green Red Vegetation Index

KI - Kappa Index

MD - Minimum Distance

ML - Machine Learning

MP - Megapixel

MR - Mussels and rock class

NDRE - Normalized Difference Red Edge

NDVI - Normalized Difference Vegetation Index

NIR - Near Infra-Red

OA - Overall Accuracy

PA - Producer's accuracy

RBL - Rock, barnacles and limpets class

RDVI - Renormalized Difference Vegetation Index

RF - Random Forest

RGB - Red, Green and Blue

ROI - Region of interest

RS - Remote Sensing

S - Sand class

SR - Simple Ratio

SS - Submerged seaweed class

SVM - Support Vector Machine

SW - Seaweed class

SWUAV - Mapping the intertidal zone and assessing seaweed biomass using UAV images

UA - User accuracy

UAV - Unmanned Aerial Vehicle

1. Introduction

Remote Sensing is the science that allows analyzing Earth elements from electromagnetic sensors without being in direct contact with the object under study. Generally these sensors are present on orbital platforms or satellites, planes and at field level (Jensen 2007; Lillesand, Kiefer et al. 2015). However, satellites are considered to have too low spatial resolution to be effective in more specific zones, such as intertidal zones. Light aircraft are common platforms for intertidal mapping surveys as they offer higher spatial resolutions than satellites (Rossiter, Furey et al. 2020b). Over the past decade, image data acquired from Unmanned Aerial Vehicle (UAV), given their high spatial, spectral and temporal resolutions, has attracted increasing attention to discriminate biological objects (Raj, Kar et al. 2020). UAVs coupled with multispectral imagers offer a wide range of possibilities in environmental science: in the study of canopy and vegetation dynamics, precision agriculture, ecosystem processes, natural disaster management and spatial ecology (Anderson and Gaston 2013). In the marine and aquatic environment, UAV applications have focused on habitat characterization, coastal erosion (Chabot and Bird 2013), 2D/3D coastal environment characterization and seaweed monitoring (Murfitt, Allan et al. 2017b).

Sensors produce a lot of data that needs to be interpreted. In recent decades, a photointerpreter was responsible for transforming data into information, exploring the aspects of color, texture and geometry of images. In other words, the significance of the results achieved in the visual analysis of the photographic products was based on the experience of the photointerpreter, who used subjective criteria in the design of the thematic classes (Machado, Barbosa et al. 1994). Automatic classification methods were developed to shorten processing time and exclude subjective criteria. The classification via computer is based on statistical parameters, that is, the identification and discrimination of targets is done without the subjectivity of visual analysis.

Digital image processing aims to provide tools that facilitate the identification and extraction of the information contained in the images, for later interpretation. Thus, the basis of this process refers to a set of radiance classes defined in different bands of the electromagnetic spectrum (Jensen 2007; Kubat 2017). Dedicated systems are used for interactive analysis and manipulation of raw images. The result of this process is the production of other images, already containing specific information, extracted and highlighted from the raw images. There are several classification methods that seek accurately identify the information of each pixel of the image, classifying it into categories (Kubat 2017; Richards 1999).

In supervised classification the user or image analyst “supervises” the pixel classification process (Richards 1999). The user specifies the various pixels values or spectral signatures that should be associated with each class. This is done by selecting representative sample sites of a known cover type called training areas (Jensen 2007). The algorithm then uses the spectral signatures from these training areas to classify the whole image. Ideally, the classes should not overlap or should only minimally overlap with other classes in the image (Jensen 2007; Lillesand, Kiefer et al. 2015).

Another approach is through vegetation indexes where proportions of different spectral bands are used to qualitatively and quantitatively assess vegetative cover using spectral measurements (Bannari, Morin et al. 1995; Elvidge and Chen 1995; Richards 1999).

Monitoring of intertidal reefs is traditionally undertaken by on-ground survey methods which have assisted in understanding these complex habitats; however, often only a small spatial footprint of the reef is observed (Collin, Dubois et al. 2019). Recent developments in UAVs provide new opportunities for monitoring broad scale coastal ecosystems through the ability to capture centimeters resolution imagery and topographic data not possible with conventional approaches (Song and Park 2021).

Murfitt, Allan et al. (2017a) carried out studies with the aim of assessing the potential of UAVs to accurately monitor intertidal reefs located in Victoria Park, Australia. This was achieved by comparing dominant algal communities measured using on-ground quadrat surveys to those extracted from UAV captured photogrammetry. Results showed no significance difference between UAV and on-ground methods for macroalgal canopy cover. Quadrat surveys provide very thorough observations of targeted sections of the reef but fail to capture the complete interconnected reef due to often limited emersion times and size of some intertidal reefs. The results also reveal a difference in total time required to complete the survey of the intertidal reef, with UAV survey and analysis taking approximately half the time of on-ground quadrat observations. Authors conclude that although UAVs may not be able to replace on-ground monitoring techniques on intertidal reefs, they provide a complementary data source giving a more comprehensive understanding of intertidal reef assemblages.

In addition to algae, other organisms present in the intertidal are studied through images retrieved by UAV. Gomes, Peteiro et al. (2018) performed combinations of 3D photogrammetric models and spectral analysis derived from aerial images, with ground-based quadratic sampling to provide realistic measurements of mussel (*Mytilus galloprovincialis*) density, size and reproductive output. A UAV was used to conduct intertidal photographic surveys during low tides along the Portuguese central west coast. From the images, a mosaic was generated and they were georeferenced and a 3D

photogrammetric model was built. An analysis of the spectral signature for the different ground cover types was used to perform a maximum likelihood supervised classification with overall high classification accuracy (86.5% ± 4.3%, mean ± SD). Furthermore, the effect of environmental variables (substrate complexity and exposure to waves) on mussel density in the rocky substrate and on size was analyzed. The results indicate that maximum mussel density occurred at intermediate wave exposure values, while smaller mussels were found at high wave exposure values. Density, size and reproductive production maps were predicted for the study region. In this sense, the authors concluded that the information allows the validation of metapopulation models that address the persistence of spatially structured populations in fragmented habitats and, thus, allow the management of essential protected areas.

The use of auxiliary spectral bands has the potential to improve the mapping of intertidal zones that arise during low tide. Collin, Dubois et al. (2019) investigated the contribution of the drone-based red edge (RE), near-infrared (NIR), and digital surface models (DSM) into the classification accuracy of five main habitats of the largest intertidal biogenic reefs in Europe, built by the honeycomb worm *Sabellaria alveolata*. Based on photoquadrats and the maximum likelihood algorithm, overall, producer's and user's accuracies were distinctly augmented. When isolated, the DSM provided the highest gain percentage (3.42%), followed by the NIR (2.58%), and RE (2.02%). When joined, the combination of the DSM with both RE and NIR was the best contributor (4.98%), followed by the DSM with RE (4.80%), DSM with NIR (3.74%), and RE with NIR (3.22%). The authors concluded that on the class scale, all datasets increasingly advantaged sand, gravel, reef, mud and water. As short-term perspectives, the habitat classification will be very likely refined by the integration of spatially- and spectrally-derived products (such as runoff, topographic position index, rugosity index, normalized difference vegetation index, normalized difference water index) (Collin, Dubois et a 2019).

1.1 Objective

This work aims to evaluate the supervised classifications performed on RGB and multispectral images collected by UAVs, in order to map the predominant ecosystem elements in the rocky intertidal zones in northern Portugal.

The supervised classifiers used were: Random Forest (RF), Support Vector Machine (SVM), both Machine Learning algorithms and the Minimum Distance algorithm. The predominant ecosystem elements mapped were: 1) seaweed; 2) rock, barnacles and limpets; 3) sand; 4) mussels and rock and 5) submerged seaweed.

1.2 Dissertation structure

This dissertation is divided into five chapters. In this chapter, in addition to this section, there is a brief introduction and the objective of this project. Chapter 2 contains a literature review about the main fundamentals related to the intertidal zone and rocky shores, and also the recent technical developments in UAVs, and image classification algorithms, relevant to this dissertation.

Chapter 3 was dedicated to the methods and techniques used to achieve the objective of this dissertation. An introduction to the study area was presented as well as the description of the field work, image data, processing in Google Earth Engine were also presented.

Chapter 4 was dedicated to the presentation of the results and their discussion. For this purpose, this chapter is subdivided into: exploratory analysis, accuracy of classifications and visual validation and temporal variation of the covered area by classes. Finally, in chapter 5 a conclusion of the work was performed, in an attempt to withdraw some deductions.

2. Literature Review

This stage consisted of collecting and reading articles and books, in order to obtain a solid theoretical foundation of the various topics covered in this dissertation, of which we highlight: specific literature for the study area (Intertidal zone and rocky shores) and materials and techniques used (UAV's, image classification, accuracy metrics, vegetation indices and Google Earth Engine). Literature review was enhanced by researching topics of interest in academic consultation platforms, such as Google Scholar and Web of Science taking advantage of the availability of online databases.

2.1 Intertidal zone and rocky shores

The intertidal zone is the marine area that is most intensely studied when compared to any other marine habitat, because it is the most accessible place for humans (Helmuth, Mieszkowska et al. 2006; Nybakken 2001; Nybakken and Cummings 2019). It is the most temporally and spatially variable zone of all marine habitats. It ranges from plains of sand and mud to rocky reefs and allows the development of a wide variety of plants and animals (Raffaelli and Hawkins 1996; Wilson 2005). Along the coast there are several types of shores, divided into two large groups: rocky shores and sandy shores (Figure 1).

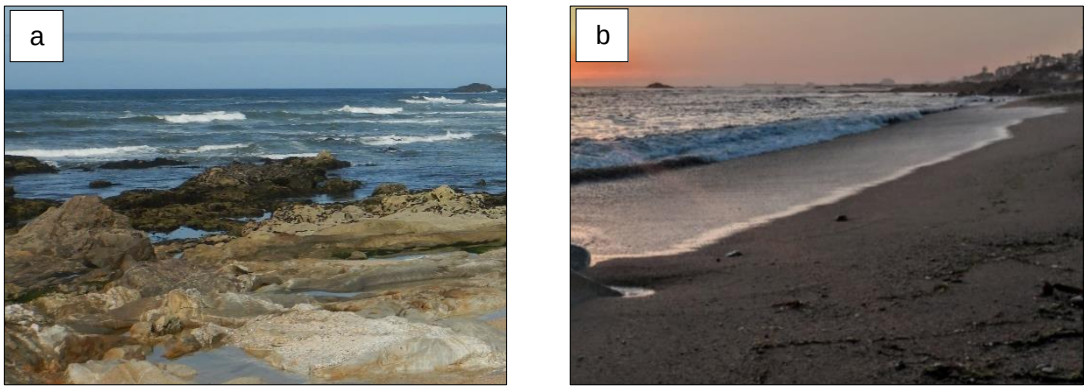


Figure 1 - a) Exemple of a rocky shores (Praia da Congreira - Vila Chã); b) exemple of a sandy shores (Praia das Pastoras - Porto).

A rocky shore is an intertidal area that consists of solid rocks. It is often a biologically rich environment and can include many different habitat types, such as steep rocky cliffs, platforms, rock pools and boulder fields (Raffaelli and Hawkins 1996; Tait and Dipper 1998). In Portugal, three main areas can be identified of rocky shores (north, center and south), separated by extensive areas of soft sediments (Carvalho, Oliveira et al. 1992; Santos, Lopes et al. 2014). In the northern region, from Moledo do Minho to Aguda, the rocky shores are typically granitic, except for Vila Chã, where the rocky platforms are made up of schists (Boaventura, Ré et al. 2002; Santos, Lopes et al. 2014). The central zone is composed of limestone and sandy limestone with smooth to irregular surfaces. Schists make up most of

the bedrock on the southwest coast and part of the south coast, being replaced by sandy limestone towards the east of the south coast (Boaventura, Ré et al. 2002).

The rocky shores appear due to the effect of the waves on the coast, that are mainly erosive, wearing out the softer materials, leaving the hardest rocks exposed. Most of the substrate is, therefore, stable and permanent, forming a secure surface on which a variety of organisms that require fixation can grow; for example, large algae, barnacles, mussels and limpets (Tait and Dipper 1998).

The intertidal zone of rocky shores (Figure 2) can be divided in three zones (Stephenson and Stephenson 1949). Supralittoral fringe or High-shore zone: a high coastal area flooded only during high tides, characterized by fouling lichens, cyanobacteria (bluegreen bacteria), barnacles and limpets. Midlittoral fringe or Midshore zone: this is a turbulent zone that is (un)covered twice a day. The zone extends from the lower limit of the barnacles band to the upper limit of large brown algae (e.g. Laminariales, Fucoidales). Common organisms are filtering organisms, bivalves and cirripeds like mussels and barnacles. Infralittoral fringe or Low-shore zone: this region is usually covered with water. It is only uncovered when the tide is extremely low. In contrast to the other zones, the organisms are not well adapted to long periods of dryness or to extreme temperatures. The common organisms in this region are brown seaweed and Echinoderms such as starfish and ophiuroids.

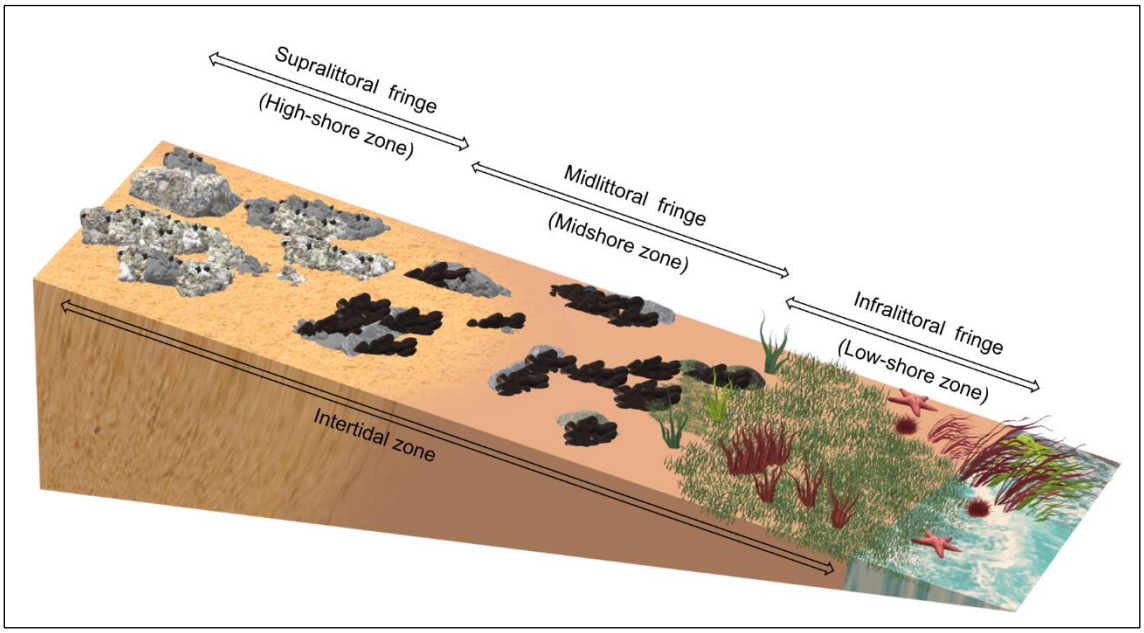


Figure 2 - Universal zoning scheme by Stephenson and Stephenson,1949 (three zone scheme) - adapted from Raffaelli and Hawkins, 1996.

These communities are subject to the many and frequent changes imposed by wave action and seasonal variations linked to tidal cycles (Littler and Murray 1975; Nybakken and Cummings 2019). Species that live in the rocky shores have to adapt to the sharp environmental gradient existent in this habitat. The vertical gradient, defined by the tides, produces periods of submersion and emersion, leading to contrasting conditions of temperature, salinity or desiccation. The horizontal gradient, caused by wave action, may have direct effects on species dislodgement or indirectly affect feeding or reproduction (Tait and Dipper 1998).

These intertidal zone conditions offer a series of challenges for mapping activities, offering a limited time window for survey (Rossiter, Furey et al. 2020a; Tait and Dipper 1998). Therefore, traditional field survey methods, whilst accurate, are time-consuming and limited in the area that can be covered during low tide. Remote Sensing (RS) permits large areas to be rapidly surveyed but the effectiveness of satellites and aircraft for mapping fine-scale ecosystem elements mapping is limited by their coarse spatial resolution and restricted operational flexibility (Jensen, 2007). The rapid development of Unmanned Aerial Vehicles (UAVs) and sensor technology can address these issues and provide a potential alternative to established remote sensing platforms (Thomas Rossiter, Thomas Furey, Tim McCarthy, & Dagmar B. Stengel, 2020b).

For the reasons explained above, monitoring marine biodiversity is considered a fundamental activity for the conservation of natural resources, the deepening of scientific and technical knowledge in a perspective of resource exploration and sustainable development (Stanners & Bourdeau, 1995). Therefore, there is a clear need to develop methods to monitor the extent and integrity of the rocky shores.

2.2 Unmanned Aerial Vehicle (UAV)

The term UAV is commonly used in the computer science, robotics and artificial intelligence, as well as the photogrammetry and RS communities. Additionally, synonyms like Remotely Piloted Vehicle (RPV), Remotely Operated Aircraft (ROA) or Remotely Piloted Aircraft (RPA) and Unmanned Aerial Systems (UVS) can also infrequently be found in the literature (Eisenbeiß 2009; Everaerts, Lewyckyj et al. 2004). These vehicles are remotely controlled, semi-autonomous, autonomous, or have a combination of these capabilities. Comparing UAVs to manned aircraft the main difference between the two systems is that on the UAVs no pilot is physically present in the aircraft (Everaerts 2008).

UAVs present themselves as a technology for the acquisition of aerophotogrammetric data (Eisenbeiß 2009), since it makes possible to obtain vertical aerial photographs, planimetric and altimetric components for building maps (Jensen 2007). Like all eligible RS systems, capturing photographs via UAVs, depends on energy that is

reflected and / or emitted from characteristics of the Earth's surface (Lillesand, Kiefer et al. 2015). In order to capture all that energy, UAVs are usually equipped with a photogrammetric evaluation system that can include video, thermal or infrared cameras, multispectral, hyperspectral and LiDAR sensors or a combination of these sensors (Simic Milas, Sousa et al. 2018).

The most well-known UAVs are associated with military uses. Typically, they are rather hefty, weaponized equipment with extremely costly sensors that are not intended for public usage. There are, however, numerous types of UAVs on the market that are light enough to be carried by a single person and have a high cost/benefit ratio. Most of these tiny UAVs are ideal for photogrammetric applications, particularly those equipped with a camera, GPS, inertial measurement unit (IMU), radio connection, and a small computer processor (Gonçalves and Henriques 2015).

Most of these UAVs are highly automated and easy to use. There is a wide range of potential applications for these devices, including surveying, mining, agriculture, environmental monitoring and management and conservation (Gonçalves and Henriques 2015). The UAV-images can be also used for the high resolution texture mapping on existing DSMs and 3D-models, as well as for image rectification (Everaerts 2008).

In the past 10 years, comparisons between traditional collection methods and RS methods have grown. This is due to the democratization of satellite images and the affordable price of UAVs (Murfitt, Allan et al. 2017a). A reflection of the increase in the use of UAVs by the civilian population can be seen in the gradual growth in the number of publications of articles that contain the word “UAV” or “DRONE” in their title (Figure 3).

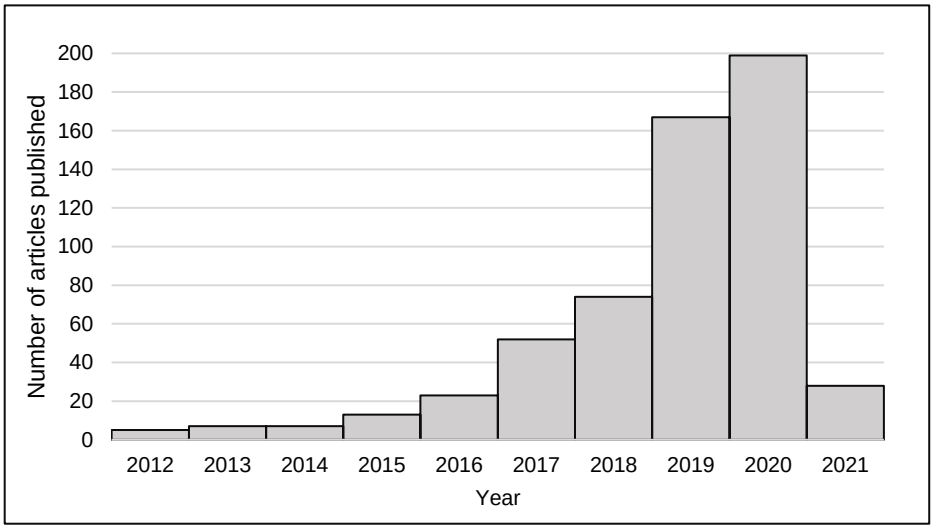


Figure 3 - Search for articles published between the years 2012 and 2020 with the word UAV or DRONE in their titles. Source: Web of science (https://apps.webofknowledge.com/).

Recent UAV developments provide new opportunities for monitoring coastal ecosystems on a large scale through the ability to capture centimeter-resolution images and

topographic data impossible with conventional approaches. Studies carried out on Parks Victoria's Intertidal Reefs in Australia demonstrate the benefits of low-cost UAVs for intertidal monitoring through rapid data collection, full coverage census, identification of the dominant canopy habitat and generation of geomorphic derivatives to explain biological variation (Murfitt, Allan et al. 2017a).

Starting in 2015 on the Portuguese northwest coast UAV flights were carried out with the aim of mapping and monitoring sand dunes and beaches with soil resolution better than 5 cm. Gonçalves et al. (2015) concluded that UAVs can replace many conventional flights, with considerable gains in the cost of data acquisition and without any loss in the quality of topographic data and aerial images.

2.3 Image classification

Digital remote sensor data are usually stored as a matrix (array) of numbers. Each digital value is located at a specific row (i) and column (j) in the matrix (Jensen 2007). A pixel is defined as a two-dimensional picture element that is the smallest no divisible element of a digital image (Richards 1999). Each pixel at row (i) and column (j) in the image has a spectral reflectance or brightness associated value (Figure 4).

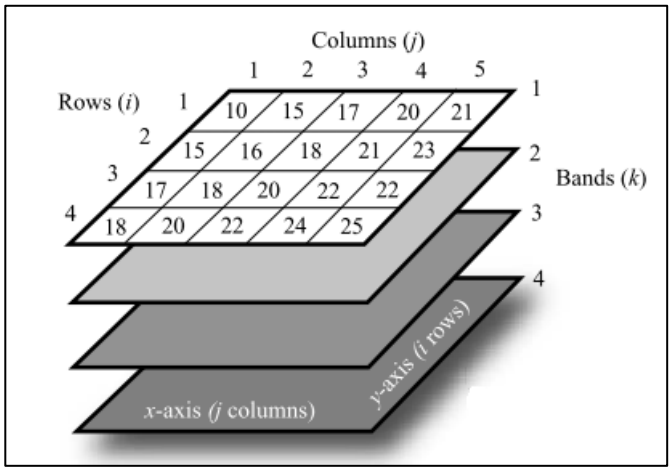


Figure 4 - Representation of the remote sensing data matrix with elements of the image (pixel) located in Rows (i), columns (j) and band (k). adapted from Jensen, 2007.

The overall objective of image classification procedures is to automatically categorize all pixels in an image into land cover classes or themes. Often this is done using spectral patterns; that is, pixels that share similar combinations of spectral reflectance or emissivity are grouped together in classes that are assumed to represent particular categories of surface features (Lillesand, Kiefer et al. 2015).

There is a wide variety of proposed algorithms in the literature for the classification of RS images. Different techniques and several research areas are being used as a basis

for the development of algorithms, among them: statistics (Dean and Smith 2003), neural networks (Gonçalves, de Andrade Netto et al. 2005), SVM (Pal and Mather 2005), and decision trees (Hansen, Dubayah et al. 1996).

In RS, traditionally, image classification techniques are grouped in three general categories: supervised, unsupervised and hybrid (Lillesand, Kiefer et al. 2015). In particular supervised classification can be defined informally as the process of using samples of known identity (i.e., pixels already assigned to informational classes) to classify pixels of unknown identity (i.e., to assign unclassified pixels to one of several informational classes) (Campbell and Wynne 2011). Samples of known identity are those pixels located within training areas, or training fields. The analyst defines training areas by identifying regions on the image that can be clearly matched to areas of known identity on the image. Such areas should typify spectral properties of the categories they represent and, of course, must be homogeneous in respect to the informational category to be classified. That is, training areas should not include unusual regions, nor should they straddle boundaries between categories. Size, shape, and position must favor convenient identification both on the image and on the ground. Pixels located within these areas form the training samples used to guide the classification algorithm to assign specific spectral values to appropriate informational classes. Clearly, the selection of these training data is a key step in supervised classification (Campbell and Wynne 2011).

2.3.1 Minimum distance

Minimum distance classifier is a supervised image classification technique, in which the pixels are classified based on their distance from the mean spectra of the predefined classes (Abburu and Golla 2015). In this method, first the mean vector for each class is calculated based on the training dataset. Next, using the Minimum distance algorithm, the Euclidean distance of every unclassified pixel from the mean vector is calculated. The pixel is then assigned to the class with the minimum distance. The distance (Dcx) of a particular pixel from different class mean vectors (x) is usually calculated using Euclidean distance: $Dcx = ||xi - ci||$.

This type of classifier is mathematically simple and therefore computationally less complex. It requires the least time for computation among all the other supervised classification techniques. The disadvantage of this technique is that it takes into account only the mean value, and so it is less efficient than the machine learning technique.

Supervised parametric approaches (like the Minimum Distance) or non-supervised approaches (like clustering) were the options available for image classification task in the past and are still widely used today. However, the rise of Machine Learning (ML) opened

many doors regarding the production and accuracy of classification maps. Since then, many papers bench-marking algorithm performance using Sentinel-2 images emerged, stating that the most commonly used algorithms nowadays are Random Forests (RFs), Support Vector Machines (SVMs) and Artificial Neural Network (ANN), usually rendering a very similar performance (Pirotti, Sunar et al. 2016; Talukdar, Singha et al. 2020).

Machine learning is about programming computers to optimize performance criteria by using data samples or past experiences (Alpaydin 2004). Machine learning domains focus on how machines learn the rules of the example. The goal of machine learning is to learn the pattern of the example and then to generalize it with a new example. Kubat (2017) divides supervised machine learning into 5 groups: logic based algorithms, perceptron based techniques, statistical learning algorithms, instance-based learning, and support vector machines.

2.3.2 Random Forest

The Random Forest (RF) classifier consists of a combination of decision trees classifiers where each classifier is generated using a random vector sampled independently from the input vector, and each tree casts a unit vote for the most popular class to classify an input vector (Breiman 2001; Breiman, Friedman et al. 1984; Kubat 2017).

Two steps involving random selection are used to create each decision tree. First, a bootstrap strategy is utilized to randomly select only 2/3 of the training samples with replacement to build each decision tree. The remaining 1/3 of the training samples are defined as out-of-bag (OOB) data, which are used for inner cross-validation to evaluate the performance of Random Forest. Second, only a subset of the predictor variables is randomly selected to split each tree using the Gini index (Breiman 2001).

Nowadays, the Gini index is the most frequently used attribute selection measure for splitting each decision tree, which measures the impurity of a given attribute with respect to the rest of the classes. Both of the two random selection steps result in less correlation between the decision trees and thus a higher generalization capability of the Random Forest classifier (Breiman 2001).

Since the major problem of the decision trees classifier is overfitting[1], RF overcomes it by constructing an ensemble of decision trees (Breiman, Friedman et al. 1984). A vast number of trees of this method provide better accuracy in the field of image classification and land-use modelling (Feng, Gong et al. 2015). The RF has become popular in remote

[1] An overfitting scenario occurs when, in the training data, the model has an excellent performance, however when using the test data, the result is bad. In this case, the model learned the relationships existing in the training so well that it ended up just memorizing what should be done, and when receiving the information from the predictor variables in the test data, the model tries to apply the decorated rules, but with different data rule has no validity, and performance is affected. In this scenario, the trained model has no generalization capacity (Breiman et al. 1984).

sensing as a non-linear and non-parametric alternative, with promising predictive capabilities for high-dimensional data sets (Pal 2005).

2.3.3 Support Vector Machine

Support Vector Machines (SVM) are based on statistical learning theory and have the aim of determining the location of decision boundaries that produce the optimal separation of classes (Vapnik 1999). SVM became popular in the last decade for solving problems in classification and regression. The SVM approaches classification problem through the concept of the margin, which is defined to be the smallest distance between the decision boundary and any of the samples. The decision boundary is chosen to be the one for which the margin is maximized. The margin is defined as the perpendicular distance between the decision boundary and the closest of the data points. Maximizing the margin leads to a particular choice of decision boundary (Alpaydin 2004). The location of this boundary is determined by a subset of the data points, known as support vectors (Figure 5).

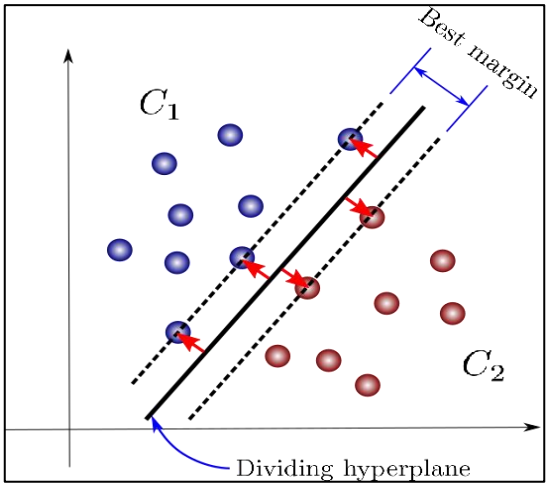


Figure 5 - Dataset representation and margin adapted from Alpaydin (2004).

In figure 5, we can clearly distinguish two categories, each identified by C1 and C2. The linear function in between (hyperplane) ensures the best possible separation between the two categories.

The SVMs have been applied to several pattern recognition tasks, obtaining results superior to those of other learning techniques in various applications (Lorena and de Carvalho 2007). SVM is a classifier that usually generates good results for noisy and complex data. In general, SVM-oriented algorithms extract the parameters that allow generalization. In addition, it has great perception capacity for data with high entropy, defined as the amount of associated uncertainty (randomness), that is, the higher the

entropy, the more atypical (out of the pattern) the scene will be, events in which other classifiers encounter problems (Alpaydin 2004).

2.4 Accuracy metrics

The evaluation of the accuracy of the classified images helps to assess the quality of the obtained classification (Stehman and Czaplewski 1998). Accuracy is the degree of correspondence between the classification and the visual references of the image, and is expressed through error matrices. The error matrices compare the relationship between the known reference data (validation area) and the corresponding results of the classification procedure(Congalton 1991).

Normally, the reference data set is represented along the columns, which is compared with the classification product data, and these are represented along the lines. Table 1 exemplifies a representation of a confusion matrix. The elements of the main diagonal (in bold) indicate the level of success (or agreement) between the two sets of data and the total number of data collected (Chuvieco 1991).

Table 1 - Confusion matrix.

classification	Reference data			Total in the lines X_{i+}
	1	2	C	
1	$\mathbf{X_{11}}$	X_{12}	X_C	X_{1+}
2	X_{21}	$\mathbf{X_{22}}$	X_C	X_{2+}
C	X_{31}	X_{32}	$\mathbf{X_C}$	X_{C+}
Total in the columns X_{+i}	X_{+1}	X_{+2}	X_{+C}	**n**

The results of the confusion matrix allow to calculate of the following indexes:

i. Overall Accuracy (OA);

ii. Producer's Accuracy (PA) of the classes;

iii. User Accuracy (UA) of the classes;

iv. Kappa index (KI),

The Overall accuracy allows to calculate the global viability of the map, relating the diagonal elements with the total of collected samples, given by the following formula (Watzlawick, Kirchner et al. 2003):

$$OA = \frac{\sum Xi}{n} \times 100$$

Where:

Xi = diagonal elements

n = total samples

The distribution of accuracy across individual categories is not shown in overall accuracy. However, the accuracy of an individual category is obtained from the producer's accuracy in the class and the user accuracy in the class. Producer's accuracy indicates the probability of the sample being correctly classified, being really a measure of omission error, and user accuracy indicates the probability that a classified pixel in the image has to represent that category in the field (Congalton and Green 2019), calculated by the following formulas:

$$PA = \frac{X_{jj}}{n_{+j}}$$

$$UA = \frac{X_{jj}}{n_{j+}}$$

Where:

X jj = number of observations of the validation sample classified in class j of the map and corresponding to class j in the reference information;

n_{j+} = total observations of the validation sample classified in the map with class j;

n_{+j} = total observations of the validation sample belonging to class j.

An analysis of the Kappa index (KI) is a discrete multivariate technique used to assess thematic accuracy (Cohen 1960). The Kappa coefficient is a measure of actual agreement (indicated by the diagonal elements of the confusion matrix), or that is, it is a measure of how well the rating conforms to the baseline data.

$$KI = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} (X_{i+} \; X_{+i})}{N^2 - \sum_{i=1}^{C} (X_{i+} \; X_{+i})}$$

Where:

N = total number of observations;

r = number of lines in the confusion matrix;

X_{ii} = observations in row i and column i;

X_{i+} = total marginal of line i;

X_{+i} = total marginal of column i.

2.5 Vegetation Indices

Vegetation indices (VI) are quantitative measurements indicating the vigor of vegetation (Campbell and Wynne 2011). The applications of VIs in the monitoring and discrimination of land uses on the Earth's surface have several advantages. The following have been reported by Liu (2015): a) the compact data of VIs are calculated with the reflectance data of the various spectral bands of electromagnetic radiation energy; b) the relative values compensate for the variation in radiation intensity; c) partially correct the effects of local changes in atmospheric conditions and variations in the viewing angle of sensors and solar angles, considering whether these effects in the variation of reflectance received by the sensors in the bands that compose the indices are of the same magnitude. In addition, such VIs can be used to determine a range of biophysical parameters and vegetation characteristics such as the leaf area index, biomass, percentage of soil cover, photosynthetic activity, water deficiencies and even productivity estimates (Motohka, Nasahara et al. 2010).

Over forty VIs have been developed during the past 20 years to try to explore the applications of spectral reflectance in monitoring vegetation (Bannari, Morin et al. 1995). The different vegetative covers can be distinguished according to their unique spectral behavior in relation to overall ground elements (Tucker 1979). Visible radiation in the red (630-690 nm) is absorbed by chlorophyll while radiation in the near infrared (760-900 nm) is strongly reflected by leaf cellular structures. When examining the general reflectance curve of vegetation, the deviation observed between the red and near infrared constitutes a variable sensitive to the presence of green vegetation (Bannari, Morin et al. 1995). The spectral response of vegetation in the red is strongly correlated with chlorophyll concentration while the spectral response in the near infrared is controlled by the leaf area index and green vegetation density (Gitelson, Kaufman et al. 1996). Normalized Difference Vegetation Index (NDVI) is the combination of these two spectral domains that allows differentiating vegetation from soils and determining photosynthetically active biomass through the density of vegetation cover (Tucker 1979).

The Green Normalized Difference Vegetation Index (GNDVI) is a VIs studied initially by Gitelson (1996) that combines the bands of green and NIR. This index was considered

much more sensitive to the concentration of chlorophyll over a wide range of chlorophyll variations than NDVI, and allowed an accurate estimate of the pigment concentration. The green channel is above the green edge in the reflectance spectrum, around 520 nm, on the long wavelength side of the blue Chlorophyll-a absorption band (Horler, Dockray et al. 1983). The spectral behavior of this edge was found to be very similar to that of the red edge. The difference between them is that the green edge is primarily determined by chlorophyll-a, chlorophyll-b and CAR absorption, while the red edge is governed by chlorophyll-a and chlorophyll-b. At longer wavelengths this spectral band is located before the large range of absorption by chlorophylls. Thus, in the range 530-570 nm, two strong absorption processes reach their minimum, producing the monotonous relationship of chlorophyll-a versus green reflectance with a high sensitivity to chlorophyll-a concentration (Gitelson, Kaufman et al. 1996).

Normalized Difference Red Edge (NDRE) can be a better indicator of vegetation vigor than NDVI or other indices. In more advanced states of maturity, the leaves accumulate high levels of chlorophyl, and while red radiation is susceptible to high levels of absorption, the red-edge band can be more reflected and therefore better serve the purpose of indicator of vigor vegetative (Barnes, Clarke et al. 2000).

Green Red Vegetation Index (GRVI) is a useful indicator for vegetation phenology, especially for leaf color (Tucker 1979). In other words, the GRVI is an index that is sensitive enough to detect the color change of the green leaf to yellow or red, since the green reflectance is used in the calculation of the GRVI (Motohka, Nasahara et al. 2010).

Simple Ratio (SR) is probably the first VI used and is based on the principle that leaves absorb relatively more electromagnetic radiation in the red region (visible) than in the infrared region. Therefore, the greater the amount of leaves in the canopy, the greater the value of the ratio (Birth and McVey 1968).

A peculiarity attributed to NDVI is its rapid saturation, which makes it insensitive to the increase in plant biomass after a certain stage of development, that is, the index stabilized in a level has the same value, although there is an increase in canopy density (Baret and Guyot 1991; Carneiro, Furlani et al. 2019). An attempt to circumvent the NDVI saturation problem suggests the renormalized difference vegetation index (RDVI). As a non-linear index, RDVI is not only less sensitive to variations in the geometric and optical properties of the unknown foliage, but also less affected by solar and visual geometry. It aims to linearize the relationship between the index and biophysical parameters (Baret and Guyot 1991).

2.6 Google Earth Engine

Google Earth Engine combines data catalogs of satellite imagery with computing power optimized for parallel processing of geospatial data, which is usually very heavy. Besides having Application Programming Interfaces (APIs) for both Python and JavaScript, an online Integrated Development Environment using JavaScript, called GEE Code, allows the user to perform spatial analysis, visualization tasks and download the necessary imagery.

Users can sign up for access at the Earth Engine homepage (https://earthengine.google.com) (Google 2015). While prior experience with GIS, RS and scripting make is easier to get started, they are not strictly required, and the user's guide is oriented towards domain novices. Accounts come with a quota for uploading personal data and saving intermediate products, and any inputs or results can be downloaded for offline use. The GEE also provides a set of the state-of-the-art machine learning classifiers for pixel-based classification that can be used for high resolution images (Farda 2017).

The GEE is in use across a wide range of disciplines, ranging from global forest change (Hansen, Potapov et al. 2013), global wetland surface change (Pekel, Cottam et al. 2016), crop yield estimation (Lobell, Thau et al. 2015), rice mapping (Dong, Xiao et al. 2016), urban mapping (Zhang, Li et al. 2015), flood mapping (Coltin, McMichael et al. 2016), restoration fire (Soulard, Albano et al. 2016), malaria risk mapping (Sturrock, Cohen et al. 2014), invasive species detection (Elkind, Sankey et al. 2019).

All the studies mentioned above used satellite images processed in a GEE environment. The use of images collected by UAVs (RGB, multispectral and hyperspectral) in GEE environment is relatively innovative. Bennett, Younes et al. (2020) presented a widely accessible and reproducible semi-automatic workflow that enables rapid processing of drone imagery for image analysis using GEE. Under this workflow, were processed and classified 230 drone images of Heron Reef, Australia and produced a substrate classification with an overall accuracy of 86%, which mapped coral cover with 92% accuracy. This study provided cheap and efficient methods to process and assess drone imagery that will benefit monitoring.

3. Methodology

3.1 Study area

Vila Chã is a coastal parish in the municipality of Vila do Conde, inserted in the Metropolitan Area of Porto, in the northwest coast of Portugal. The study area was located at Congreira beach, Vila Chã rocky shores (41º17'45''N - 41º17'43''N and 08º44'12''W - 08º44'19''W). It was characterized by a rocky outcrop colonized with diverse seaweeds, mussels, barnacles, limpets, with sandy patches, following the zonation patterns of temperate rocky shores (Figure 6).

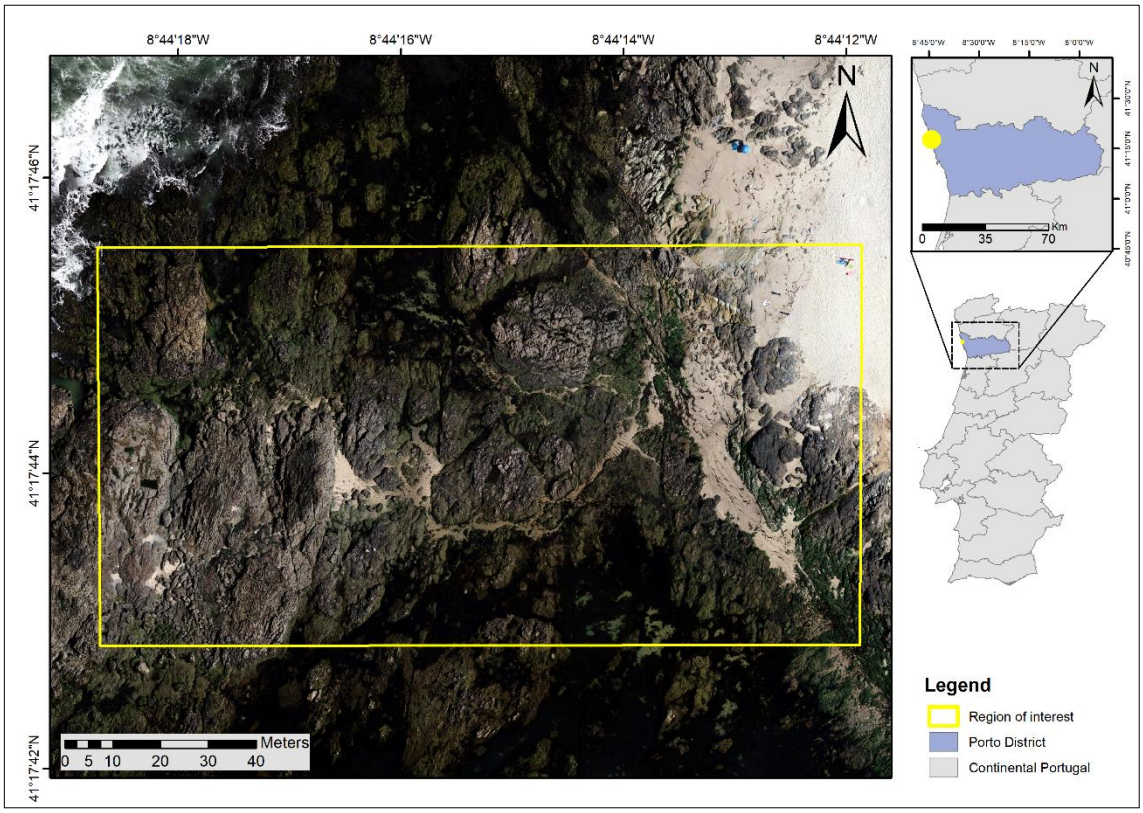


Figure 6 - Location map of the study area.

According to the Köppen climate classification (Monteiro, Ferreira et al. 2005; Peel, Finlayson et al. 2007), the study area is described by a mild climate, with dry and warm summers (Csb- Warm-summer Mediterranean climate). The average annual temperature is 15.2 °C (minimum value in January: 9.5 °C; maximum value in July and August: 20.8 °C) and the annual thermal amplitude is low because of the oceanic influence. The average annual rainfall is 1236 mm/year, ranging from 181 mm in the wettest month (December), to 20.4 mm in the driest month (July). The area has a water deficit from June to September, particularly in July and August (Zeferino, Carvalho et al. 2021).

3.2 Fieldwork

Before each UAV flight, information on the intertidal ecosytem was collected during the low tide. Thus, 50 x 50 cm quadrat frames were placed upon representative areas of the following elements: 1) seaweed; 2) rock, barnacles and limpets; 3) sand and 4) mussels and rock (Figure 7 - a). With the aid of a precision GNSS receiver, the geographical coordinates of the central points of the quadrats were collected, with an accuracy of 2 cm (Figure 7 - b).

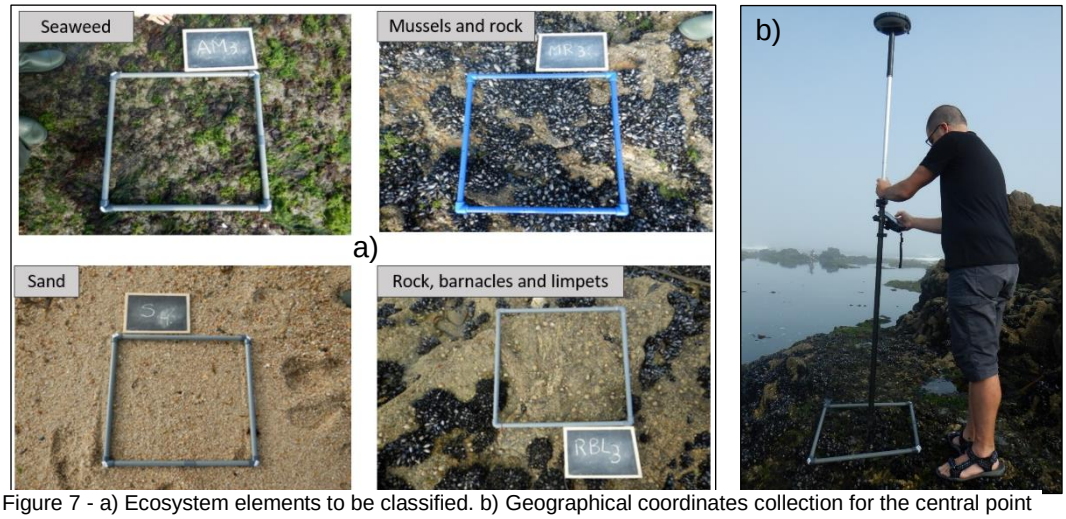


Figure 7 - a) Ecosystem elements to be classified. b) Geographical coordinates collection for the central point of the quadrats with the GNSS receiver.

The quadrats (encompassing training and validation areas) were properly identified and label in the field, being removed from the shore after GPS reading. At the end of the field work a total of 40 GPS reading were obtained (10 for class/element). The UAV performed the flight to collect the RGB and Multispectral images after collecting all the GPS readings. The UAV flights were performed in July, September and November 2020.

3.3 Images data

The Matrice 200, a DJI quadcopter, was used to survey the study area. It was equipped with a Zenmuse X4S sensor that captures information in the visible spectrum (RGB) with a resolution of 20 MP. Along with the RGB sensor, a MicaSense RedEdge-3 multispectral sensor was also coupled to the UAV (Figure 8).



Figure 8 - a) MicaSense RedEdge-3 sensor. b) UAV DJI Matrice 200.

This sensor enables to acquire data in the Blue, Green, Red, Red-edge, and Near-Infrared parts of the electromagnetic spectrum, with 1.2 MP resolution (Table 2).

Table 2 - Spectral bands of MicaSense RedEdge-3 sensor.

Band	**Central wavelength (nm)**	**Bandwidth (nm)**
Blue	475	20
Green	560	20
Red	668	10
Red-edge	717	10
Near-InfraRed	840	40

All image bands were georeferenced and orthorectified. From the RGB images, an orthophoto mosaic was generated and from the multispectral images, the reflectance of each individual band was extracted. For future comparisons, the resolutions of all bands captured with the multispectral sensor were standardized at 2.8 cm/pixel.

To obtain good georeferencing results, ground control points (GCPs) were marked on the ground throughout the study area and collected with the aid of a precision GNSS receiver. Although the UAV has a navigation GPS receiver attached, with an accuracy of a few meters, enough for the UAV navigation, it is not good enough for the precision georeferencing of the images.

Images were processed in Agisoft Photoscan, a photogrammetric software that has a large degree of automation. First, common points between overlaping images were obtained automatically by image processing algorithms, and then the relative orientation between all the image block were obtained. This operation is designated as image alignment. Then all the GCPs were identified on all the images, and an absolute orientation of the images was obtained. Then a dense point cloud was generated and a DSM, which is a regular grid. This DSM is used to orthorectify the images and build an orthomosaic. Figure 9 shows, for the case of the RGB images, the DSM and the orthomosaic.

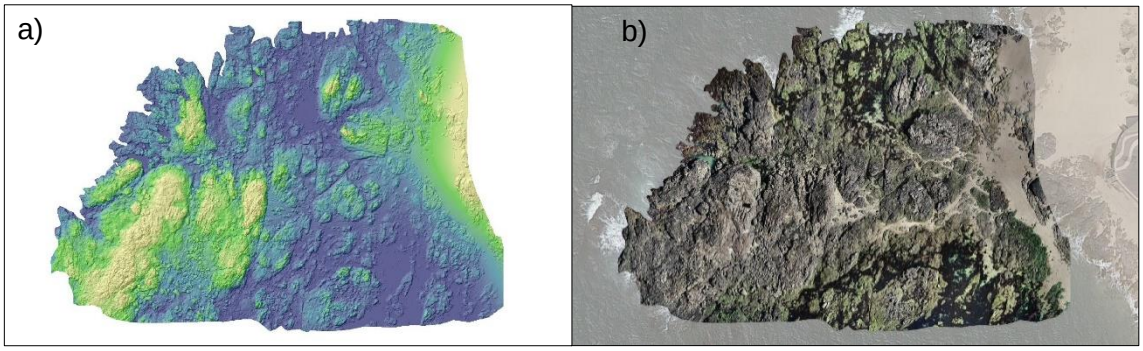


Figure 9 – a) Digital surface model (DSM) and b) photogrammetrically orthorectified image product mosaicked from an image collection (orthomosaic).

The process is common to RGB and multispectral images. The work of image classification is done over the orthomosaics.

3.4 Preprocessing

In a computational environment through the QGIS 3.4 program, the central points of the quadrats were plotted on the orthophoto produced from the RGB bands collected in the field. With the buffer tool, points were transformed into circles with a radius of 25 cm (Figure 10 - a). The buffer tool was used to expand the point (GPS precision coordinate), so that the circle occupies almost the entire reference area highlighted by the quadrat (Figure 10 - b). In this way, shapefiles for the training areas were defined and later used in the classification algorithms.

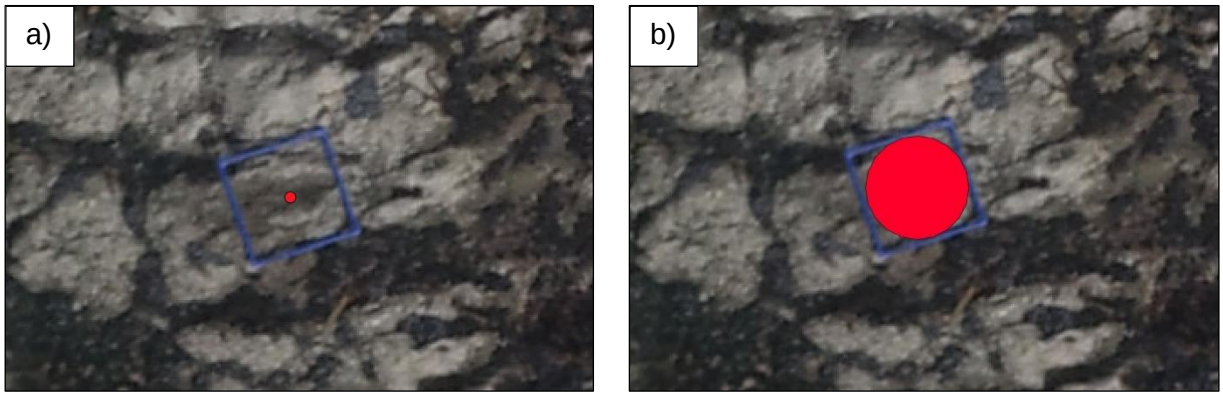


Figure 10 - a) The central point plotted in QGIS environment. b) Circle created with the buffer tool from the central point (0.25 m radius).

Also, in the QGIS environment, an additional “submerged seaweed” class was defined, through visual interpretation of orthophoto. To ensure a "fair" assessment of the generalization of a classifier, the data used to train the classifier should be separated from the data that is used to assess its accuracy. Therefore, the 10 field samples per class were separated into training areas and validation sets, assuming that 70% of the samples were

use for training the classifiers and 30% of the samples were use evaluating the accuracy of the classification results (Table 3).

Table 3 - Name of the classes studied and number of training and validation areas by field collection.

Abbreviation	Class	Training areas	Validation areas
SW	1) seaweed	7	3
RBL	2) rock, barnacles and limpets	7	3
S	3) sand	7	3
MR	4) mussels and rock	7	3
SS	5) submerged seaweed	7	3
	total	35	15

3.5 Processing in Google Earth Engine

For each month of research (July, September and November 2020) and for each classifier algorithm (MD, RF and SVM), three methodologies were tested for multispectral data; methodology 1: classification of the all the bands; methodology 2: classification of VI and methodology 3: classification of the bands and vegetation indexes stacked in a single variable. The most relevant VI suggested in the literature were selected (Albetis et al. 2017). Table 3 shows the indices with their respective formulas.

Table 4 - Computed vegetation indices found in the literature and their respective equations.

Name	Equation	References	Related
Normalized Difference Vegetation Index	$\mathrm{NDVI} = \frac{\mathrm{N}-\mathrm{R}}{\mathrm{N}+\mathrm{R}}$	(Rouse et al. 1974)	Biomass
Green Normalized Difference Vegetation Index	$\mathrm{GNDVI} = \frac{\mathrm{N}-\mathrm{G}}{\mathrm{N}+\mathrm{G}}$	(Gitelson et al. 1996)	Chlorophyll concentration
Green Red Vegetation Index	$\mathrm{GRVI} = \frac{\mathrm{G}-\mathrm{R}}{\mathrm{G}+\mathrm{R}}$	(Tucker 1979)	Vegetation phenology
Normalized Difference Red Edge	$\mathrm{NDRE} = \frac{\mathrm{N}-\mathrm{RE}}{\mathrm{N}+\mathrm{RE}}$	(Barnes et al. 2000)	Vegetative vigor
Simple Ratio	$\mathrm{SR} = \frac{\mathrm{N}}{\mathrm{R}}$	(Birth and McVey 1968)	Leaves in the canopy
Renormalized Difference Vegetation Index	$\mathrm{RDVI} = \frac{\mathrm{N}-\mathrm{R}}{\sqrt{\mathrm{N}+\mathrm{R}}}$	(Roujean and Breon 1995)	Non-saturation in increased biomass

Legend: R: red band; N: near infrared band; G: green band; RE: red-edge band

A central code written in JavaScript was developed to achieve the proposed methodologies. Therefore, the necessary data for the classifications were loaded into the GEE cloud: bands (Red, Green, Blue, Red-edge and NIR), shapefile with ROI, shapefile with training area and shapefile with validation area. In the sequence, the bands were

stacked in a single variable and that variable was cut from the ROI (Figure 11). This step was common to all methodologies. From this stage, the code was suited to each tested methodology.




```
//save the image in a variable
var image = ee.Image('users/joelencruz/bandas_julho_2020/BGRREN_Vila_Cha_06_07_2020')
print (image)


//specify the band names
var image = image.select(['b1', 'b2', 'b3', 'b4', 'b5'], ['BLUE', 'GREEN',  'RED','RedEdge', 'NIR'])


//load cutting shape
var corte = ee.FeatureCollection('users/joelencruz/bandas_julho_2020/area_padrao_vila_cha_2020')


// cut bands
var corte_image = image.clip(corte)
```

Figure 11 - Central code written in JavaScript, band stacking and cutting process.

When methodology 1 was applied, right after the insertion of the training areas, the variable containing the stacked bands was classified (Figure 12).

```
//Insert training samples
var samples = ee.FeatureCollection('users/joelencruz/dissertacao/00_shape_treino_julho_2020_com_algas_sub')
print(samples.aggregate_array("classe").distinct())


var am = samples.filter(ee.Filter.stringStartsWith('classe', 'am')).map(function(feature){return feature.set('class', 1)})
var rbl = samples.filter(ee.Filter.stringStartsWith('classe', 'rbl')).map(function(feature){return feature.set('class', 2)})
var s = samples.filter(ee.Filter.stringStartsWith('classe', 's')).map(function(feature){return feature.set('class', 3)})
var mr = samples.filter(ee.Filter.stringStartsWith('classe', 'mr')).map(function(feature){return feature.set('class', 4)})
var as = samples.filter(ee.Filter.stringStartsWith('classe', 'alga sub')).map(function(feature){return feature.set('class', 5)})

var samples = am.merge(rbl).merge(s).merge(mr).merge(as)

print("samples",samples)


var samples = corte_image
  .sampleRegions({
  collection: samples,
  scale:0.05 ,
  tileScale: 6
  })


//classification
var classifier = ee.Classifier.minimumDistance({
  })
  .train(samples, 'class',corte_image.bandNames())

var result = corte_image
  .classify(classifier)
```

Figure 12 - GEE code for methodology 1: insertion of training areas and supervised classification.

For methodology 2, before the insertion of the training areas, the VIs were properly calculated (Table 4), a variable was created from the stacking of all VIs (Figure 13 - a). Then, the training samples were inserted and the classification performed on the variable containing the stacking of the VIs (Figure 13 - b).

a)

```
////////////////////////////////////////////////////////////////////
//Calculate vegetation indices
var ndvi = corte_image
  .normalizedDifference(['NIR', 'RED'])
  .rename("NDVI")//.abs()

var gndvi = corte_image
  .normalizedDifference(['NIR', 'GREEN'])
  .rename("GNDVI")//.abs()


var sr = corte_image.expression('NIR / RED',{
    'NIR': image.select('NIR'),
    'RED': image.select('RED')
  }).rename("SR")//.abs()

var grvi = corte_image
  .normalizedDifference(['GREEN', 'RED'])
  .rename("GRVI")//.abs()

var ndre = corte_image
  .normalizedDifference(['NIR', 'RedEdge'])
  .rename("NDRE")//.abs()

var rdvi = corte_image.expression('(NIR - RED) / ((NIR + RED)**(1/2))',{
    'NIR': image.select('NIR'),
    'RED': image.select('RED')
  }).rename("RDVI")//.abs()


  // joining VIs into a single variable
var imageWithNewBands = ndvi.addBands(gndvi)
var imageWithNewBands = imageWithNewBands.addBands(sr)
var imageWithNewBands = imageWithNewBands.addBands(grvi)
var imageWithNewBands = imageWithNewBands.addBands(ndre)
var imageWithNewBands = imageWithNewBands.addBands(rdvi)


print(imageWithNewBands)
```

b)

```
//Insert training samples
var samples = ee.FeatureCollection('users/joelencruz/dissertacao/00_shape_treino_julho_2020_com_algas_sub')
print(samples.aggregate_array("classe").distinct())


var am = samples.filter(ee.Filter.stringStartsWith('classe', 'am')).map(function(feature){return feature.set('class', 1)})
var rbl = samples.filter(ee.Filter.stringStartsWith('classe', 'rbl')).map(function(feature){return feature.set('class', 2)})
var s = samples.filter(ee.Filter.stringStartsWith('classe', 's')).map(function(feature){return feature.set('class', 3)})
var mr = samples.filter(ee.Filter.stringStartsWith('classe', 'mr')).map(function(feature){return feature.set('class', 4)})
var as = samples.filter(ee.Filter.stringStartsWith('classe', 'alga sub')).map(function(feature){return feature.set('class', 5)})

var samples = am.merge(rbl).merge(s).merge(mr).merge(as)

print("samples",samples)


var samples = imageWithNewBands
  .sampleRegions({
  collection: samples,
  scale:0.05 ,
  tileScale: 6
  })


//classificação MD
var classifier = ee.Classifier.minimumDistance({
  })
  .train(samples, 'class',imageWithNewBands.bandNames())

var result = imageWithNewBands
  .classify(classifier)
```

Figure 13 – GEE code for methodology 2: a) Calculation of VIs and stacking in a single variable. b) Insertion of training areas and supervised classification

And for methodology 3, after the creation of the variable containing the stacking of the VIs and the association of the bands in the same variable, the classification was performed under this new variable (Figure 14).

```
  // add the bands with the VI in a single variable
var imageWithNewBands = corte_image.addBands(ndvi)
var imageWithNewBands = imageWithNewBands.addBands(gndvi)
var imageWithNewBands = imageWithNewBands.addBands(sr)
var imageWithNewBands = imageWithNewBands.addBands(grvi)
var imageWithNewBands = imageWithNewBands.addBands(ndre)
var imageWithNewBands = imageWithNewBands.addBands(rdvi)

print ("Bandas finais",imageWithNewBands.bandNames())


//insertion of training areas
var samples = ee.FeatureCollection('users/joelencruz/dissertacao/00_shape_treino_julho_2020_com_algas_sub')
print(samples.aggregate_array("classe").distinct())


var am = samples.filter(ee.Filter.stringStartsWith('classe', 'am')).map(function(feature){return feature.set('class', 1)})
var rbl = samples.filter(ee.Filter.stringStartsWith('classe', 'rbl')).map(function(feature){return feature.set('class', 2)})
var s = samples.filter(ee.Filter.stringStartsWith('classe', 's')).map(function(feature){return feature.set('class', 3)})
var mr = samples.filter(ee.Filter.stringStartsWith('classe', 'mr')).map(function(feature){return feature.set('class', 4)})
var as = samples.filter(ee.Filter.stringStartsWith('classe', 'alga sub')).map(function(feature){return feature.set('class', 5)})

var samples = am.merge(rbl).merge(s).merge(mr).merge(as)

var samples = imageWithNewBands
  .sampleRegions({
  collection: samples,
  scale:0.05 ,
  tileScale: 6
  })

// supervised classification
var classifier = ee.Classifier.minimumDistance({
  })
  .train(samples, 'class',imageWithNewBands.bandNames())

var result = imageWithNewBands
  .classify(classifier)
```

Figure 14 - GEE code for methodology 3: creation of the variable containing the bands and VI. Supervised classification on this new variable.

After the classifications, the script becomes the same as for all methodologies. At this point, the accuracy of the classifications was tested. The next step in the central script was to insert the validation samples and perform the accuracy calculations (Figure 15).

```
//insert validation areas
var valicao = ee.FeatureCollection('users/joelencruz/dissertacao/00_shape_validacao_julho_2020_com_algas_sub')
print(valicao.aggregate_array("classe").distinct())


var am = valicao.filter(ee.Filter.stringStartsWith('classe', 'am')).map(function(feature){return feature.set('class', 1)})
var rbl = valicao.filter(ee.Filter.stringStartsWith('classe', 'rbl')).map(function(feature){return feature.set('class', 2)})
var s = valicao.filter(ee.Filter.stringStartsWith('classe', 's')).map(function(feature){return feature.set('class', 3)})
var mr = valicao.filter(ee.Filter.stringStartsWith('classe', 'mr')).map(function(feature){return feature.set('class', 4)})
var as = valicao.filter(ee.Filter.stringStartsWith('classe', 'alga sub')).map(function(feature){return feature.set('class', 5)})

var validacao = am.merge(rbl).merge(s).merge(mr).merge(as)


print("validação", validacao)

var validation = result.sampleRegions({
  collection: validacao,
  properties: ['class'],
  scale: 0.05,
  geometries: true})
```

Figure 15 – GEE code for insertion of validation samples.

The code written to perform the calculations of the accuracy metrics is shown in Figure 16.

```
//Error Matrix
var testAccuracy = validation.errorMatrix('class', 'classification');

print('Validation error matrix: ', testAccuracy);
print('Validation overall accuracy: ', testAccuracy.accuracy());
print('Kappa',testAccuracy.kappa())
print('Exatidão Global',testAccuracy.accuracy())
print('User ACC - Comisison',testAccuracy.consumersAccuracy())
print('Producers ACC - Omission',testAccuracy.producersAccuracy())
```

Figure 16 - GEE code for calculating accuracy metrics.

The metrics calculated were: overall accuracy, kappa index, producer's accuaracy and user accuracy. The producer's accuracy is the accuracy of the map from the point of view of the creator of the map (the producer). The producer's accuracy is a complement of the error of omission, that is, the *producer's accuracy = 1 - omission error*. User accuracy is the accuracy from a map user's point of view, not the map creator. User accuracy is the complement of the commission error, *user accuracy = 1 - commission error*.

Finally, the classification results were exported in tif format, as thematic maps using the "Export.image.toDrive" function. For each tested methodology, classifying algorithm and month of study, a script was replicated. In this way, 3 months of study, 3 methodologies and 3 classifiers, that is, 27 scripts were run to the methodologies.

Other evaluation method also applied in this dissertation is cross-validation. The cross-validation, also known as k-fold cross-validation, in which the data are divided into k subsets (optimally 5 or 10, to allow a less biased estimation) with k "experiments" (Rodriguez, Perez et al. 2009). The cross-validation procedure ensures that each example is included exactly once in the test fold and that each example in the test fold is not used to train the classifier (Figure 17).

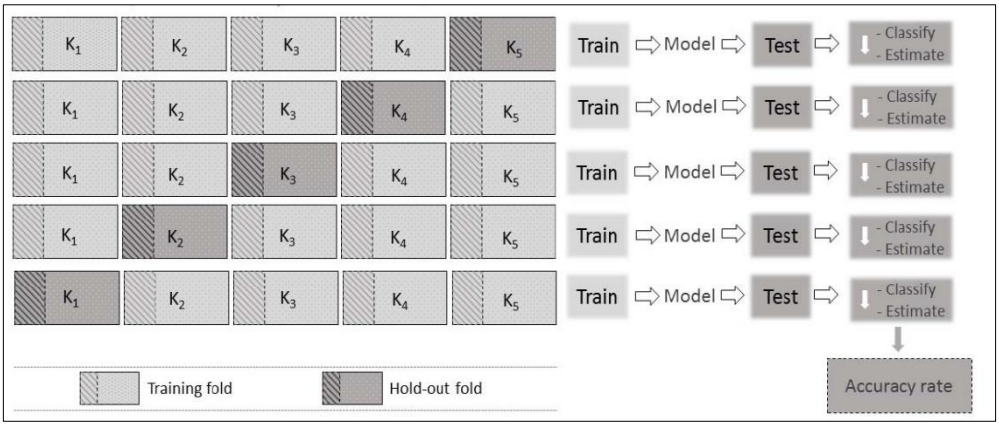


Figure 17 - k-fold (5-fold) cross validation scheme. In each “iteration”, the examples in one of the data folds is left out for testing and the remaining examples in the k1 fold are used to train the classifier. he performance quality of the trained classifier is tested on the left-out fold (in each “experiment”), and the overall performance measure is then averaged over the k folds (k “experiments”) adapted from Reaelizadeh et al. (2009).

For example, given 5 subsets K1, K2 ... K5 the first step of K-Fold is using K1 for testing and from K2 to K5 for training. In the second step, K2 is used for testing and everything else for training, including K1 that was used for testing in the first step, in the third step to tenth the same logic will be applied successively. The final result of the K-Fold validation is the classifier's average performance in the K tests. The objective of repeating the tests several times is in order to increase the reliability of the estimation of the classifier's accuracy (Arlot and Celisse 2010; Refaeilzadeh, Tang et al. 2009).

4. Results and Discussion

4.1 Exploratory analysis

The figure below shows pixel frequency number histograms for each band of the multispectral sensor for each month of study. On the x-axis the pixel count values and on the y-axis the reflectance between 0 and 1.

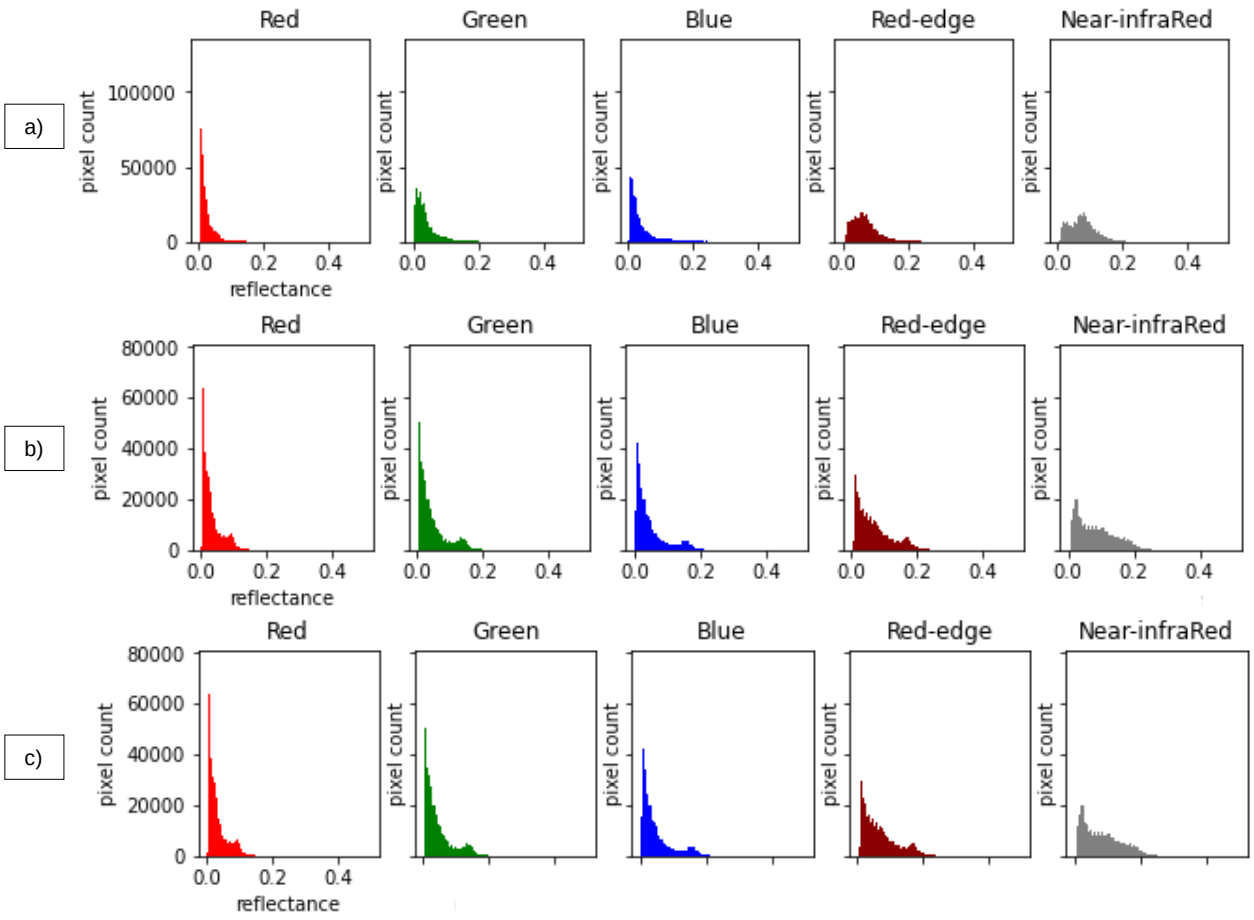


Figure 18 – a) Histogram of the multispectral bands for July; b) Histogram of the multispectral bands for September and c) Histogram of the multispectral bands for November.

Exploratory analysis of spectral datasets in the visible and infrared region is a necessary stage in the image processing procedures. The main objective is to validate if the collection of bands was carried out and if they are well positioned in the electromagnetic spectrum. The visible bands (Red, Green and Blue) have a higher frequency of pixels in areas with low reflectance, very close to zero, which makes these bands darker. The Red-edge and NIR bands have more homogeneous pixel frequency values on the reflectance axis (Figure 18).

Figure 19 shows the orthophoto produced from the bands collected by the RGB sensor. ROI with area of 13287 m^2, is highlighted. Orthophotos have a spatial resolution of 1.1 cm/pixel. In the Northeast of Figure 19, a patch of sand was observed. Submerged seaweed were observed in the north and southwest. Rocks, barnacles and limpets have been checked to the southwest and in higher zones located to the northwest and northeast.

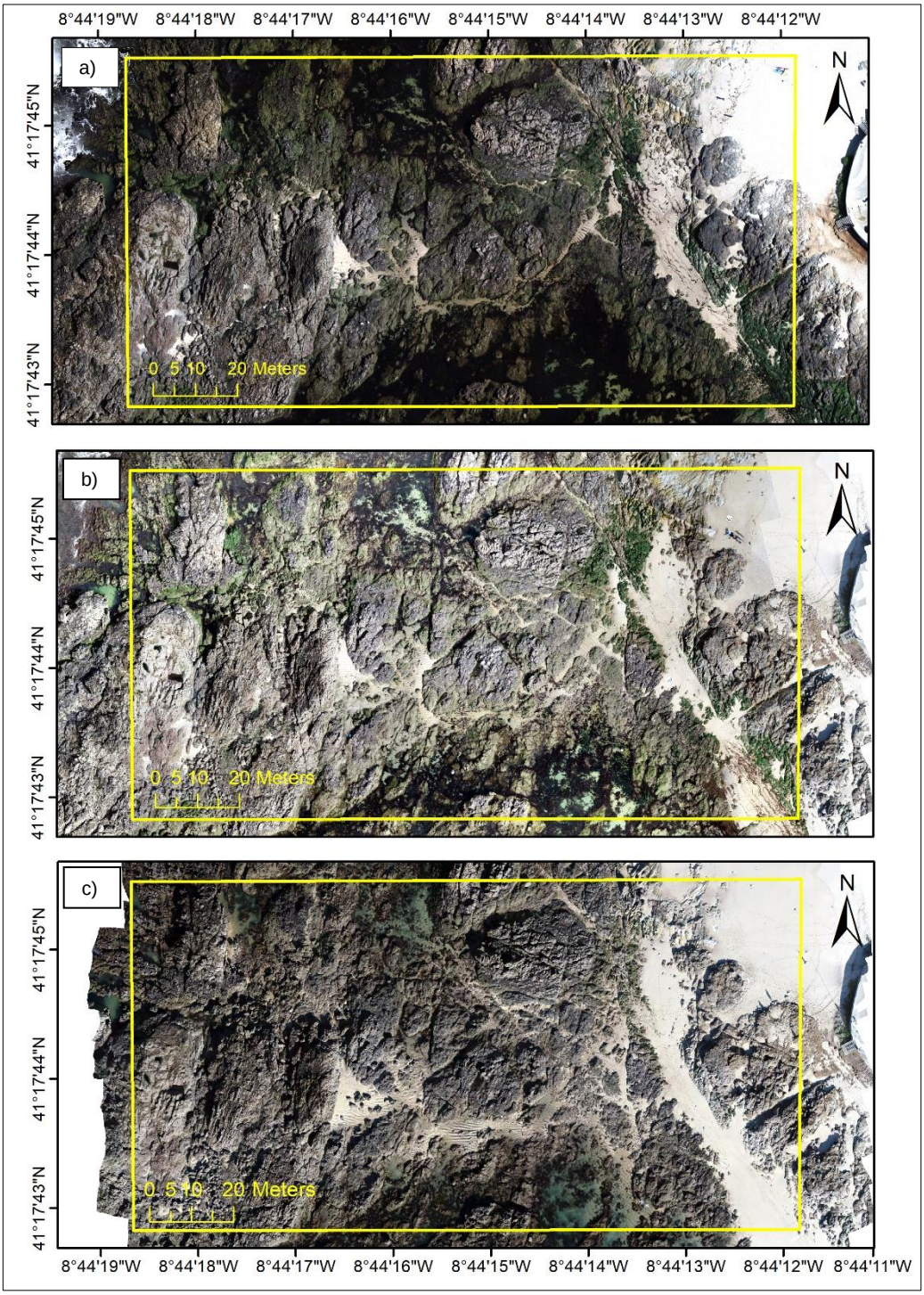


Figure 19 - a) Orthophoto for the month of July; b) Orthophoto for the month of September; c) Orthophoto for the month of November. In yellow, the region of interest (ROI).

4.2 Classifications accuracy and visual validation

The performance results, in terms of Overall Accuracy (O.A.) and Kappa Index (K.I), of the evaluated classifiers in all methodologies and all the months of study are presented in Table 5. Among all methodologies and all months of study, the MD classifier presented the worst results in O.A. and K.I. The Machine Learning classifiers achieved higher values but similar between them. In some cases, the results of O.A. were equal and the K.I. has a difference of 0.003, as observed between the RF and SVM classifiers methodology 3 for July (O.A.: 0.991 / K.I: 0.945 and O.A.: 0.991 / K.I: 0.948 respectively).

Table 5 - Overall Accuracy (O.A) and Kappa Index (K.I) results.

		Minimum Distance		Random Forest		Support Vector Machine	
Methodology	Months	O.A	K.I	O.A	K.I	O.A	K.I
Methodology 1	July	0.963	0.798	0.987	0.922	0.978	0.870
	September	0.821	0.647	0.920	0.838	0.861	0.692
	November	0.850	0.710	0.936	0.881	0.907	0.823
Methodology 2	July	0.480	0.133	0.991	0.943	0.990	0.942
	September	0.396	0.232	0.923	0.843	0.922	0.846
	November	0.850	0.710	0.933	0.874	0.932	0.872
Methodology 3	July	0.489	0.137	0.991	0.945	0.991	0.948
	September	0.398	0.234	0.942	0.883	0.932	0.862
	November	0.711	0.543	0.937	0.881	0.932	0.872

Based on the accuracy metrics, for each month of research (Annex II), the best and the worst classifiers among all the methodologies were selected. Figure 19 shows the results, for July, of the best and worst classifier. The best classifier for July was the Support Vector Machine in methodology 3 (O.A.: 0.991 and K.I.: 0.948). For this classifier the values of user accuracy per class were: SW: 0.93; RBL: 0.78; S: 1.00; MR: 0.92 and SS: 1.00. The results of producer's accuracy per class were: SW: 0.94; RBL: 0.94; S: 1.00; MR: 0.73 and SS: 1.00 (Figure 20 - a). The worst classifier for the month of July was the Minimum Distance methodology 2 (O.A.: 0.480 and K.I.: 0.133). The user accuracy values per class were: SW: 0.05; RBL: 0.31; S: 0.95; MR: 0.04 and SS: 0.98. The producer's accuracy per class were: SW: 0.55; RBL: 0.96; S: 1.00; MR: 0.45 and SS: 0.46 (Figure 20 - b).

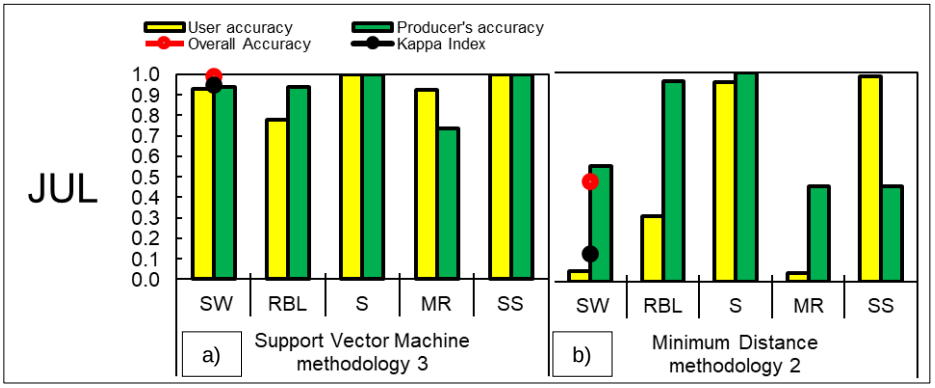


Figure 20 - a) Best classifier for july and b) Worst classifier for july. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

A comparative layout was created between: a) the orthophoto of the month; b) thematic map of the best classification; and c) thematic map of the worst classification (Figure 21).

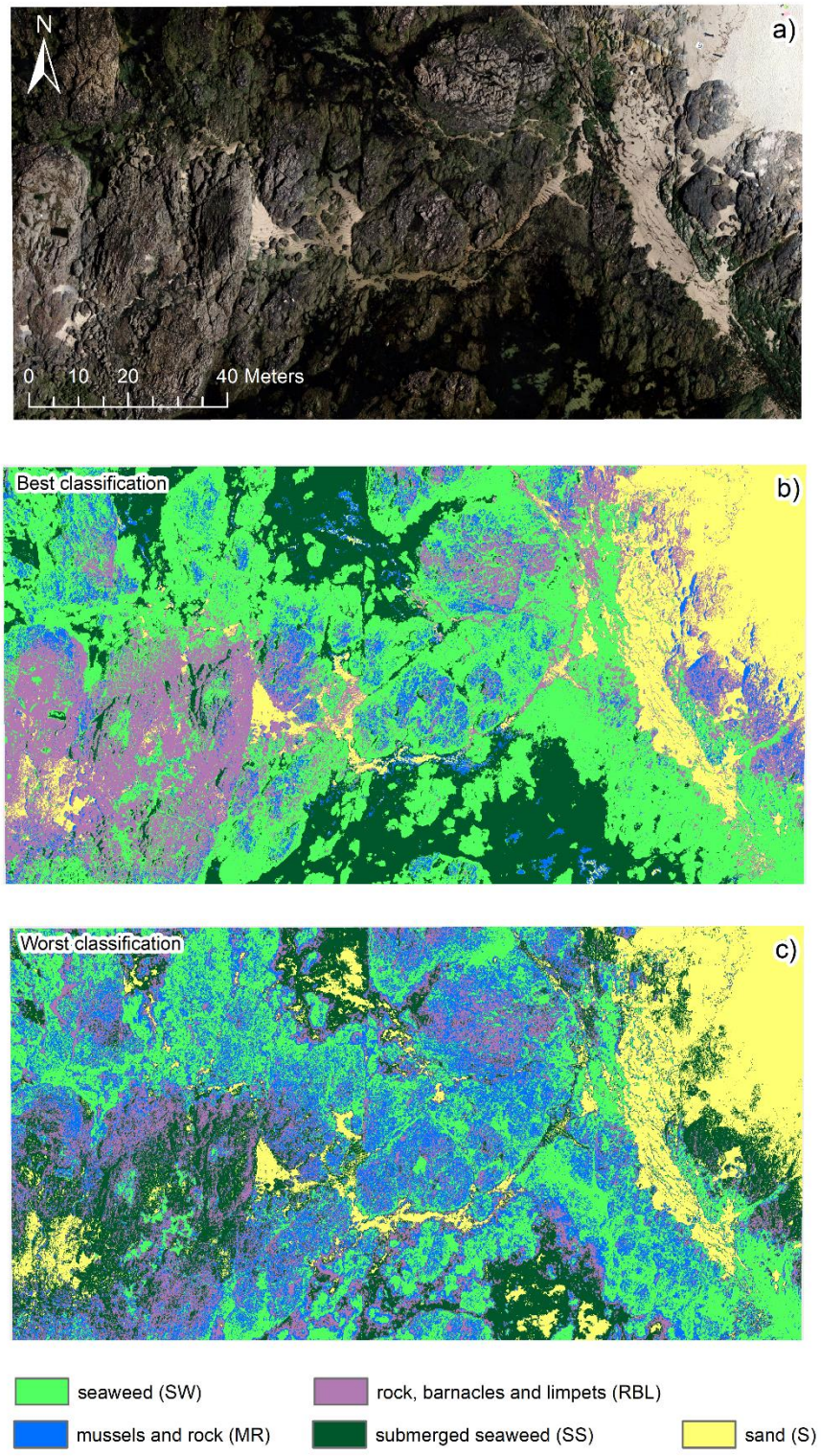


Figure 21 - a) Orthophoto of the July; b) Best classification for july, Support Vector Machine methodology 3 (Overall accuracy: 0.991 and kappa index: 0.994) and c) Worst classification for the month of July, Minimum Distance methodology 2 (Overall accuracy: 0.480 and kappa index: 0.133).

Figure 21 - b) shows the thematic map of the best classifier for July (Support Vector Machine, methodology 3). Class S was well allocated to the northeast, although the values of user accuracy and producer's accuracy were both calculated as 1.00, it is noted that to the southeast, in a submerged zone, part of the sand observed in the orthophoto was classified erroneously as MR, the user accuracy and producer's accuracy for the MR class were 0.92 and 0.73, respectively. In the north and southeast of Figure 21 - a), the zone of submerged seaweed was verified in the orthophoto. This zone is well allocated in the SS class with user accuracy and producer's accuracy of 1.00. The zones of rock, barnacles and limpets were verified in the orthophoto in the north and southwest of Figure 21 - a). These zones were allocated to the RBL class with user accuracy of 0.78 and producer's accuracy of 0.94. However, there was a considered commission error, where the class is overestimated with false positives, which can be seen in the east of Figure 21 - b).

Figure 21 - c) shows the thematic map for the worst classifier in July (Minimum Distance, methodology 2). The sand zone was well classified to the northeast with user accuracy and producer's accuracy of 0.95 and 1.00 respectively. To the north and southeast, the submerged seaweed areas were underestimated, with omission error, with a producer's accuracy of 0.44. There was also the commission error, observed in the east zone where the SS class was classified in a dry zone, the user accuracy was of 0.98. The SW and MR classes showed large commission and omission errors, with user accuracy of 0.05 and 0.04 and producer's accuracy of 0.55 and 0.45 for SW and MR respectively.

For the month of September, the best classifier was the Random Forest methodology 3 (O.A.: 0.942 and K.I.: 0.883). For this classifier the values of user accuracy per class were: SW: 0.67; RBL: 0.84; S: 0.98; MR: 0.87 and SS: 1.00. The results of producer's accuracy per class were: SW: 0.80; RBL: 0.98; S: 0.98; MR: 0.89 and SS: 0.95 (Figure 22 - a). The worst classifier for September was the Minimum Distance methodology 2 (O.A: 0.396 and K.I: 0.232), the results of user accuracy per class were: SW: 0.60; RBL: 0.79; S: 0.90; MR: 0.93 and SS: 1.00 and producer's accuracy by class: SW: 0.77; RBL: 0.96; S: 1.00; MR: 0.87 and SS: 0.93 (Figure 22 - b).

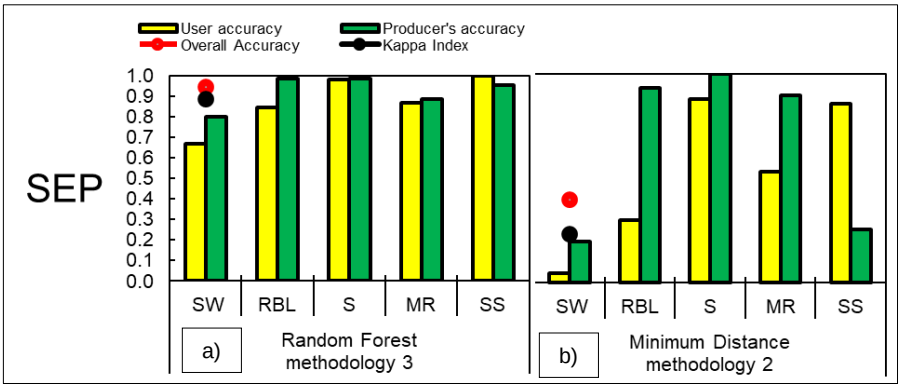


Figure 22 - Best classifier for September and b) Worst classifier for September. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

Figure 23 shows the comparative map between the orthophoto, the best classifier and the worst classifier for September.

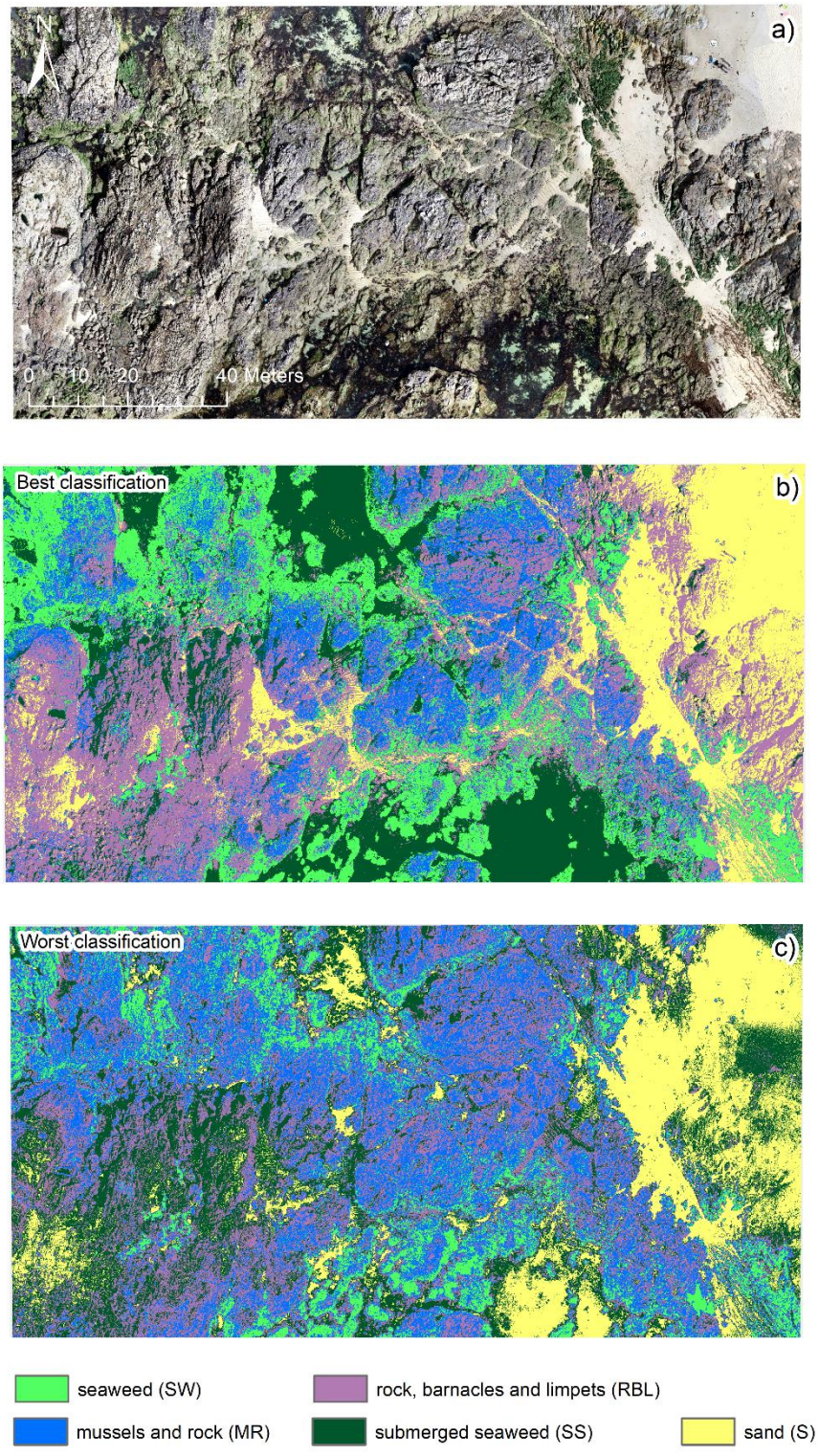


Figure 23 - a) Orthophoto of the September; b) Best classification for September, Random Forest methodology 3 (Overall accuracy: 0.942 and kappa index: 0.883) and c) Worst classification for September, Minimum Distance methodology 2 (Overall accuracy: 0.396 and kappa index: 0.232).

Figure 23 - b) shows the thematic map of the best classifier for September (Random Forest, methodology 3), class S was well allocated to the Northeast, with user accuracy and producer accuracy of 0.98 for both accuracy metrics. It is observed that to the Southeast, in a submerged zone, part of the sand verified in the orthophoto (Figure 23 - a) was erroneously classified as SS. To the north and southeast of Figure 23 - b) the SS class was allocated with results of user accuracy and producer's accuracy of 1.00 and 0.95 respectively. To the north and southwest the RBL class was well allocated, however the user accuracy and producer's accuracy metrics were 0.84 and 0.94, showing commission and omission errors. The SW and MR classes showed commission and omission errors, with user accuracy of 0.67 and 0.87 and producer's accuracy of 0.80 and 0.89 respectively.

Figure 23 - c) shows the thematic map for the worst classifier for September (Minimum Distance, methodology 2). The sand zone was well classified to the northeast with user accuracy and producer's accuracy of 0.88 and 1.00 respectively. However, false-positive errors were observed in the submerged zone, to the southeast, there was an overestimation of the real submerged sand zone (Figure 23 - c). To the north and southeast, the SS class areas were not well allocated, with user accuracy and producer's accuracy of 0.86 and 0.26. Visual verification corroborates the producer's low accuracy value, so the SS class was underestimated by the classifier, there is also an overestimation of the SS class in the eastern and southeastern zones. Also, in the southeast zone, the RBL class was not well allocated, the class can be verified in almost all the zones of the figure with the exception of the northeast zone, the precision metrics form 0.30 and 0.93 for user accuracy and producer's accuracy, respectively.

For the month of November the best classifier was the Random Forest methodology 3 (O.A.: 0.937 and K.I.: 0.88). For this classifier the values of user accuracy per class were: SW: 0.66; RBL: 0.83; S: 0.94; MR: 0.86 and SS: 1.00. The results of producer's accuracy per class were: SW: 0.83; RBL: 0.64; S: 0.99; MR: 0.87 and SS: 0.95 (Figure 24 - a). The worst classifier for the month of November was the Minimum Distance methodology 3 (O.A: 0.711 and K.I: 0.54), the results of user accuracy per class were: SW: 0.65; RBL: 0.28; S: 0.79; MR: 0.45 and SS: 1.00 and producer's accuracy by class: SW: 0.62; RBL: 0.71; S: 0.99; MR: 0.44 and SS: 0.73 (Figure 24 - b).

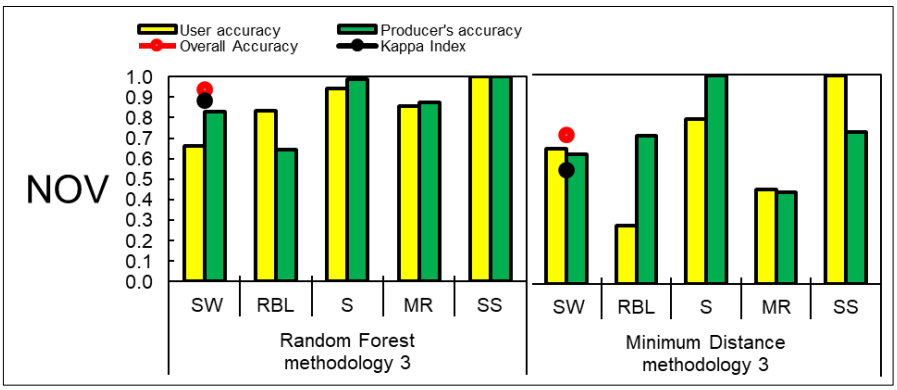


Figure 24 - Best classifier for the month of November and b) Worst classifier for the month of November. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

Figure 25 shows the comparative map between the orthophoto, the best classifier and the worst classifier for the month of November.

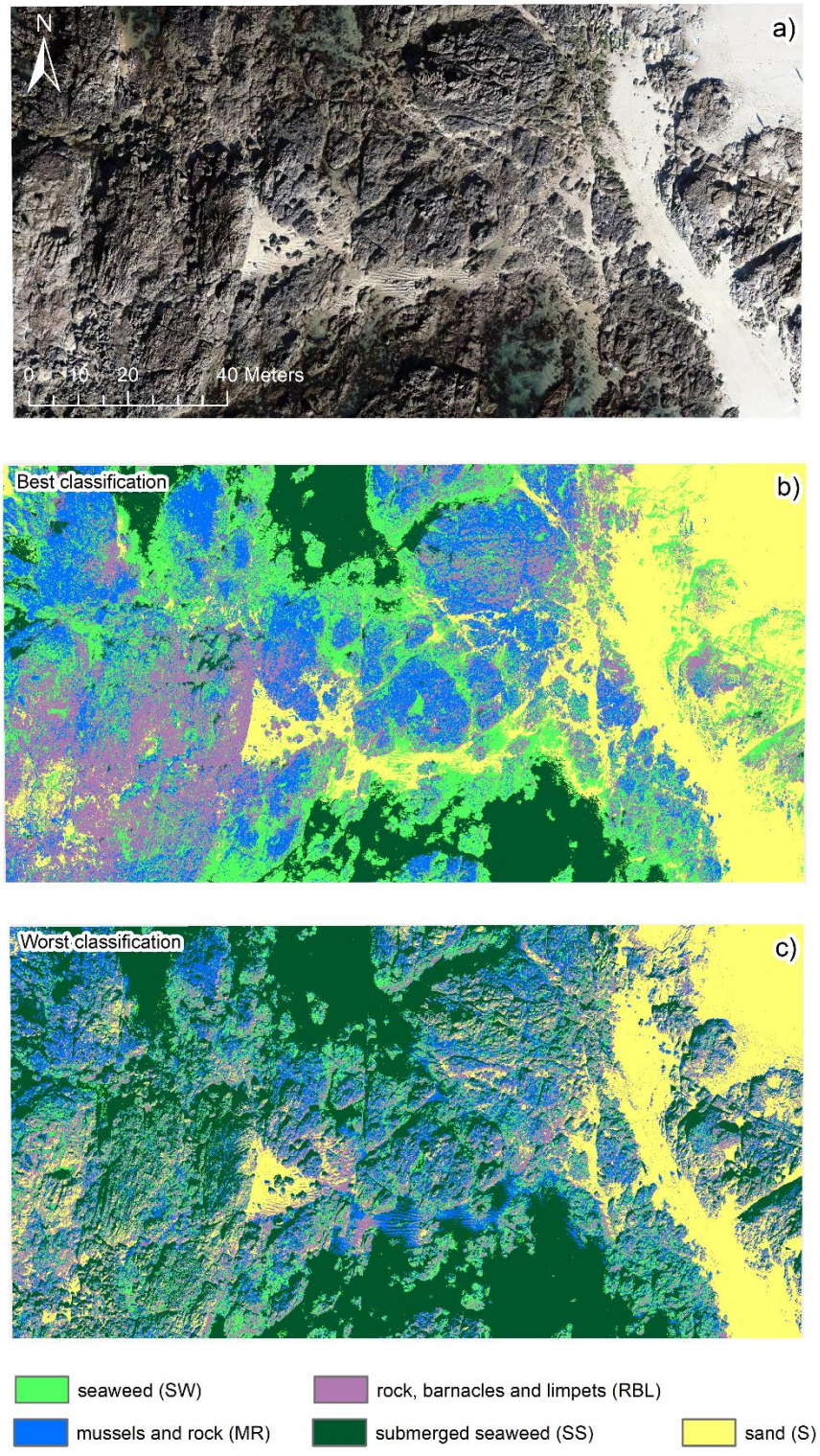


Figure 25 - Orthophoto of the November; b) Best classification for November, Random Forest methodology 3 (Overall accuracy: 0.937 and kappa index: 0.88) and c) Worst classification for November, Minimum Distance methodology 3 (Overall accuracy: 0.771 and kappa index: 0.543);

Figure 25 - b) shows the thematic map for the best classifier for November (Random Forest, methodology 3), class S was well allocated, to the northeast, with values of 0.94 and 0.99 for user accuracy and producer's accuracy. However, to the north and southeast, in the submerged zones, a submerged sand zone was erroneously classified as the SS class, this class had accuracy metrics of 1.00 and 0.95 for user accuracy and producer's accuracy respectively. For the North and Southwest, the RBL class was not well allocated, showing precision metrics of 0.83 and 0.64 for user precision and producer precision. The SW and MR classes showed commission and omission errors, with user accuracy of 0.66 and 0.86 and producer's accuracy of 0.83 and 0.87 for SW and MR respectively.

Figure 25 - c) shows the thematic map for the worst classifier for November (Minimum Distance, methodology 3). To the northeast the sand zone was rated with user accuracy and producer's accuracy of 0.79 and 0.99, respectively. Omission errors would be present in the submerged zones located to the north and southwest, where a sand zone observable in the orthophoto was erroneously allocated to class SS. Comission errors were wrongly allocated to the west of Figure 25 c), for S class, this zone is identifiable in the orthophoto as a zone of rocks, barnacles and limpets. The SW and MR classes showed commission and omission errors, with user accuracy of 0.64 and 0.45 and producer's accuracy of 0.62 and 0.44 for SW and MR respectively.

Table 6 shows the result for cross-validation method applied to the data.

Table 6 - Assessing the accuracy of the classifiers for all months and all methodologies.

Methodology	Months	MD	RF	SVM
Methodology 1	July			
	September			
	November			
Methodology 2	July			
	September			
	November			
Methodology 3	July			
	September			
	November			
k - fold (10) Cross-validadtion in %	<80	80 - 90	90 - 95	> 95

The Machine Learning algorithms presented values above 95% of overall accuracy for all months of study and for all methodologies, except for the month of November in all methodologies. The Minimum distance algorithm obtained the worst performance with an accuracy of less than 95% for all months of study in all methodologies, except for the month of September methodology 1.

4.3 Temporal variation of the covered area by classes

From the accuracy metrics, it was noticed that methodology 3 presented lightly better results for the Machine Learning classifiers. Thus, Figure 26 shows the thematic maps resulting from the supervised classification of the Random Forest and Support Vector Machine algorithms, both produced by methodology 3 for July, September and November.

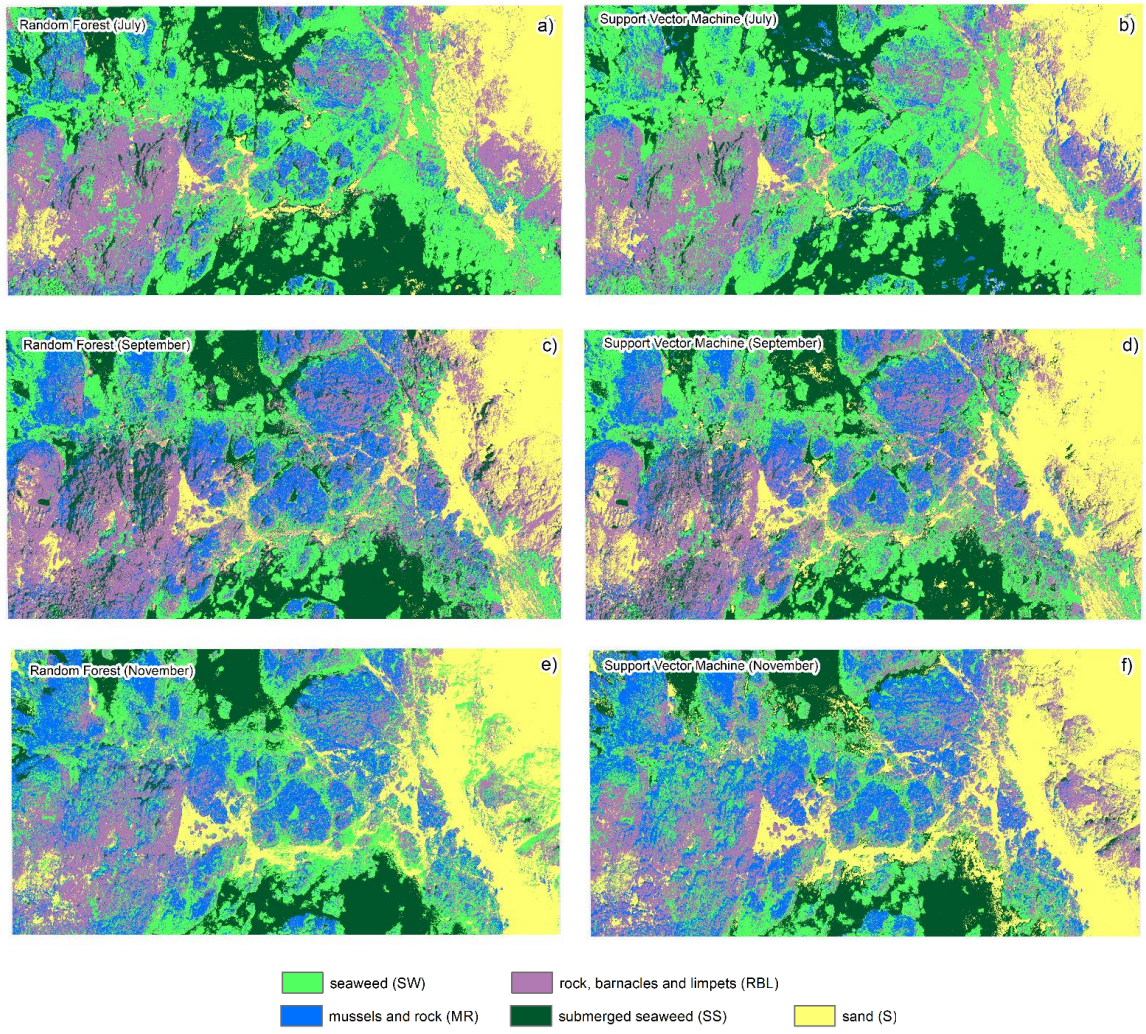


Figure 26 - Comparative thematic map between the supervised classifications of Random Forest (left) and Support Vector Machine (right) methodology 3 for july, september and november.

From the visual interpretation of the thematic maps it stands out the reduction of the area coved by the seaweed (SW) and submerged seaweed (SS) classes from July to November for both classifiers (Figure 26). To better characterize the temporal variation of the covered area by classes, Figure 27 was assembled presenting the months of study on

the x-axis and the percentages of covered area per class on the y-axis. Between July and November, for both classifiers, there was a decrease in the area covered by the SS class.

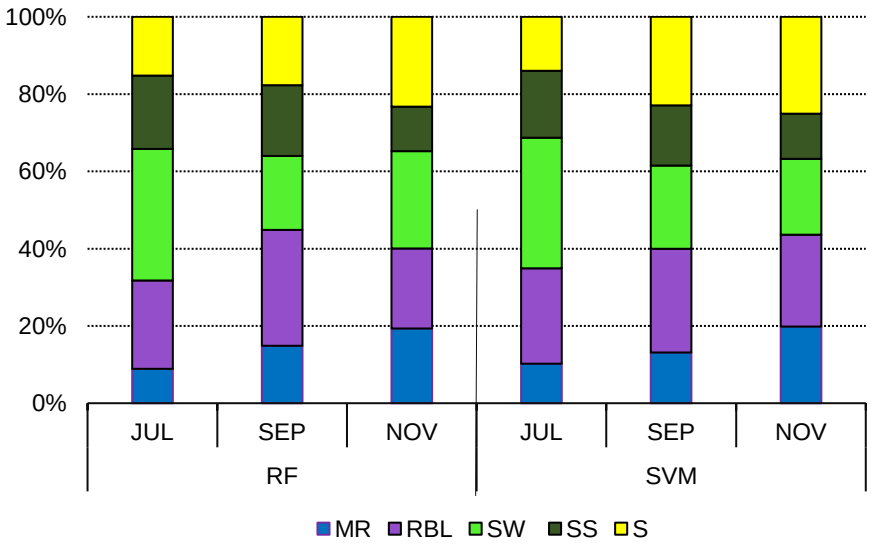


Figure 27 - Percentage of area covered by study class. RF: Random Forest, SVM: Support Vector Machine, JUL: July, SEP: September and NOV: November.

This class is mainly composed by *Laminaria* spp. and *Saccorhiza polyschides* so variations in area covered by this class could be associated with the annual variation of the *Saccorhiza polyschides*. This species has a seasonal cycle that starts in mid-March, with the germination of young sporophytes and lose their stems between October and November (Tuya, Larsen et al. 2011). Therefore, the decrease in the area covered by the SS class may be related to the decrease in the biomass of *Saccorhiza polyschides*. This reduction in the covered area from July to November is also noticeable for the SW class, as intertidal seaweeds tend to become less abundant with the approaching of the winter months (Raffaelli and Hawkins 1996). As result of the reduction in the area covered by these two classes, when grouping them together it is evident the decrease in the corresponding covered area which is in line with their seasonal patterns (Figure 28).

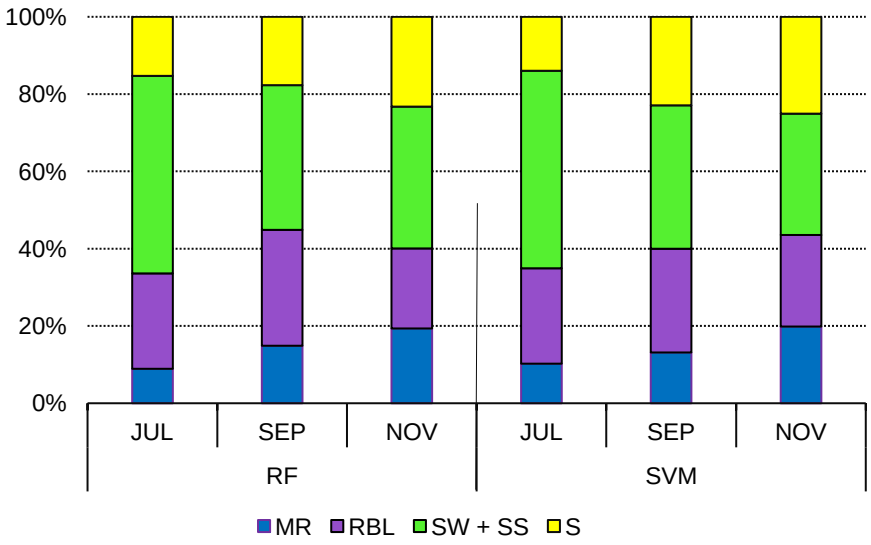


Figure 28 - Percentage of area covered by study group (with associated SW and SS classes). RF: Random Forest, SVM: Support Vector Machine, JUL: July, SEP: September and NOV: November

Class S, composed of sand, increased in area between July and November for both classifiers. This fact may be related the hydrodynamics of the rocky shore, as well as the to speed of the waves, which can be potentially important agents driving the change in the reef's geomorphology, moving sand around, despite its relatively low energy (Harris, Vila-Concejo et al. 2014; Kench 1998).

The MR class is formed by mussels and rocks. The mussels present in the ROI are of the species: *Mytilus galloprovincialis* and *Mytilus edulis*. These mussel species are an important ecological component of temperate rocky coastal communities (Hilbish, Mullinax et al. 2000). The variation of the area covered by MR class increased from July to November, probably related to the recruitment of new individuals that occurred during summer and their respective growth (Dodet, Bertin et al. 2010; Gomes, Peteiro et al. 2018).

The RBL class is composed of rocks, barnacles and limpets. The barnacle species found in the ROI are *Chthamalus stellatus* and *Chthamalus montagui*. The limpet species, present in the ROI, belong to the genus *Patella.* The variation of the area covered by these species from July to November is not so coherent for both classifiers demanding further studies.

5. Conclusion

In this work, multispectral images collected by UAV in an intertidal environment were classified. Three supervised classification algorithms were tested using three different methodologies. The results suggest that the classification methods used were highly capable of producing excellent results, with highly accurate classification for the different parameters, due to strict validation conditions (i.e., using polygon pixels not used for model training). Such results were obtained following careful definition of the training and validation areas, which were selected to include only relatively homogeneous patches.

Of the three tested methodologies, methodology 3 (with creation of one variable containing the stacking of the VIs and the association of the multispectral bands before the classification was performed) presented the best results. For such methodology both machine learning classifiers (Random Forest and Support Vector Machine) presented similar results in terms of precision metrics. On the other hand, the supervised minimum distance classifications (for the three tested methodologies) presented the worst results in terms of visual validation and precision metrics.

In the ROI, can be found several structuring species that define the zonation patterns of temperate rocky shores, such as *Laminaria* spp., *Saccorhiza polyschides*, *Mytilus* spp., *Chthamalus* spp. and *Patella* spp. These species are considered structuring species as they contribute the most to zonation (Helmuth, Mieszkowska et al. 2006). The fact that these species are sessile, enable the formation of assemblages with other species, giving them shelter and protecting them from physical factors, such as the action of the tides and sun exposure (Boaventura, Ré et al. 2002). Through the analysis of the covered area variation along the months of study it was possible to corroborate the known seasonal patterns particularly for the seaweed classes following the application of methodology 3 with the Random Forest and Support Vector Machine classifiers.

Therefore, it can be said that for the study area of Congreira beach at Vila Chã, rocky shore, methodology 3 with Random Forest and Support Vector Machine classifiers was able to accurately map and describe coverage seasonal patterns along the months of study particularly for seaweed classes.

The processing time of this work was considerably reduced due to the use of cloud programming made available free of charge by Google Earth Engine (GEE), a technological platform for the analysis of environmental data on a planetary scale that provides satellite images produced in the last 40 years. Such images are updated daily, and GEE provides the necessary tools and massive computational power to detect changes and trends in the earth's surface, oceans and atmosphere. All the computational power of GEE cloud programming was used for this research and were applied to the UAV collected images.

Therefore, using such tools, data processing becomes independent of the physical computational capacity and as a consequence, the processing time drops drastically.

This research can serve as a basis for further studies and tests for the application of supervised classification algorithms in other coastal areas supporting future environmental monitoring programs.

References

Abburu, S., & Golla, S.B. (2015). Satellite image classification methods and techniques: A review. *International journal of computer applications, 119*

Alpaydin, E. (2004). *Introduction to machine learning*. London, England: MIT press

Anderson, K., & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment, 11*, 138-146

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys, 4*, 40-79

Bannari, A., Morin, D., Bonn, F., & Huete, A.R. (1995). A review of vegetation indices. *Remote sensing reviews, 13*, 95-120

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment, 35*, 161-173

Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R.J., Li, H., & Moran, M.S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In, *5th International Conference on Precision Agriculture and other resource management*. Bloomington

Bennett, M.K., Younes, N., & Joyce, K. (2020). Automating drone image processing to map coral reef substrates using google earth engine. *Drones, 4*, 50

Birth, G.S., & McVey, G.R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer 1. *J Agronomy Journal, 60*, 640-643

Boaventura, D., Ré, P., Cancela da Fonseca, L., & Hawkins, S.J. (2002). Intertidal rocky shore communities of the continental Portuguese coast: analysis of distribution patterns. *Marine Ecology, 23*, 69-90

Breiman, L. (2001). Random Forests *Mach Learn*, 5–32

Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. (1984). *Classification and regression trees*. CRC press

Campbell, J.B., & Wynne, R.H. (2011). *Introduction to remote sensing*. Guilford Press

Carneiro, F.M., Furlani, C.E., Zerbato, C., Menezes, P.C.d., & Gírio, L.A.d.S. (2019). Correlations among vegetation indices and peanut traits during different crop development stages. *Engenharia Agrícola, 39*, 33-40

Carvalho, D.d., Oliveira, J., Pereira, E., Ramalho, M., Antunes, M., & Monteiro, J. (1992). Carta Geológica de Portugal na escala 1: 500.000. *Folha Sul Serv Geol de Portugal*

Chabot, D., & Bird, D.M. (2013). Small unmanned aircraft: precise and convenient new tools for surveying wetlands. *Journal of Unmanned Vehicle Systems, 1*, 15-24

Chuvieco, E. (1991). Fundamentos de teledetección espacial. *Estudios Geográficos, 52*, 371

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational psychological measurement, 20*, 37-46

Collin, A., Dubois, S., James, D., & Houet, T. (2019). Improving intertidal reef mapping using UAV surface, red edge, and near-infrared data. *Drones, 3*, 67

Coltin, B., McMichael, S., Smith, T., & Fong, T. (2016). Automatic boosted flood mapping from satellite data. *International Journal of Remote Sensing, 37*, 993-1015

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35-46

Congalton, R.G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: principles and practices*. CRC press

Dahlhoff, E.P., & Menge, B.A. (1996). Influence of phytoplankton concentration and wave exposure on the ecophysiology of Mytilus californianus. *Marine Ecology Progress Series, 144*, 97-107

Dean, A., & Smith, G. (2003). An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing, 24*, 2905-2920

Dodet, G., Bertin, X., & Taborda, R. (2010). Wave climate variability in the North-East Atlantic Ocean over the last six decades. *Ocean Modelling, 31*, 120-131

Dong, J., Xiao, X., Menarguez, M.A., Zhang, G., Qin, Y., Thau, D., Biradar, C., & Moore III, B. (2016). Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sensing of Environment, 185*, 142-154

Eisenbeiß, H. (2009). UAV photogrammetry. In, *University of Technology Dresden* (p. 236): ETH Zurich

Elkind, K., Sankey, T.T., Munson, S.M., Aslan, C.E., & Conservation (2019). Invasive buffelgrass detection using high-resolution satellite and UAV imagery on Google Earth Engine. *Remote Sensing in Ecology, 5*, 318-331

Elvidge, C.D., & Chen, Z. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment, 54*, 38-48

Everaerts, J. (2008). The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping. *The International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences, 37*, 1187-1192

Everaerts, J., Lewyckyj, N., & Fransaer, D. (2004). Pegasus: design of a stratospheric long endurance UAV system for remote sensing. *The International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences, 35*

Farda, N. (2017). Multi-temporal land use mapping of coastal wetlands area using machine learning in Google earth engine. In, *IOP Conference Series: Earth and Environmental Science* (p. 012042): IOP Publishing

Feng, Q., Gong, J., Liu, J., & Li, Y. (2015). Flood mapping based on multiple endmember spectral mixture analysis and random forest classifier—The case of Yuyao, China. *Remote Sensing, 7*, 12539-12562

Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment, 58*, 289-298

Gomes, I., Peteiro, L., Bueno-Pardo, J., Albuquerque, R., Pérez-Jorge, S., Oliveira, E.R., Alves, F.L., & Queiroga, H. (2018). What's a picture really worth? On the use of drone aerial imagery to estimate intertidal rocky shore mussel demographic parameters. *Estuarine, Coastal and Shelf Science, 213*, 185-198

Gonçalves, J.A., & Henriques, R. (2015). UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS Journal of Photogrammetry Remote Sensing, 104*, 101-111

Gonçalves, M.L., de Andrade Netto, M.L., Costa, J.A.F., & Zullo, J. (2005). Automatic remotely sensed data clustering by tree-structured self-organizing maps. In, *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS'05.* (p. 4 pp.): IEEE

Google (2015). Accessed in: 20/12/2020. Access link:https://earthengine.google.com/

Hansen, M., Dubayah, R., & DeFries, R. (1996). Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing, 17*, 1075-1081

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S., Goetz, S.J., & Loveland, T.R. (2013). High-resolution global maps of 21st-century forest cover change. *Science, 342*, 850-853

Harris, D.L., Vila-Concejo, A., & Webster, J.M. (2014). Geomorphology and sediment transport on a submerged back-reef sand apron: One Tree Reef, Great Barrier Reef. *Geomorphology, 222*, 132-142

Helmuth, B., Mieszkowska, N., Moore, P., & Hawkins, S.J. (2006). Living on the Edge of Two Changing Worlds: Forecasting the Responses of Rocky Intertidal Ecosystems to Climate Change. *Annual Review of Ecology, Evolution, and Systematics, 37*, 373-404

Hilbish, T.J., Mullinax, A., Dolven, S.I., Meyer, A., Koehn, R.K., & Rawson, P.D. (2000). Origin of the antitropical distribution pattern in marine mussels (Mytilus spp.): routes and timing of transequatorial migration. *Marine Biology Research, 136*, 69-77

Horler, D., Dockray, M., & Barber, J. (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing, 4*, 273-288

Jensen, J.R. (2007). *Remote Sensing of the Environment-An Earth Resource Perspective Prentice Hall*. New Jersey: Upper Saddle River

Kench, P. (1998). A currents of removal approach for interpreting carbonate sedimentary processes. *Marine Geology, 145*, 197-223

Kubat, M. (2017). *An introduction to machine learning*. Springer

Lillesand, T.M., Kiefer, R.W., & Chipman, J.W. (2015). *Remote Sensing and Image Interpretation.* University of Wisconsin, Madison

Littler, M.M., & Murray, S.N. (1975). Impact of sewage on the distribution, abundance and community structure of rocky intertidal macro-organisms. *Marine Biology, 30*, 277-291

Lobell, D.B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment, 164*, 324-333

Lorena, A.C., & de Carvalho, A.C. (2007). Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada, 14*, 43-67

Machado, R., Barbosa, V., Liporace, F., Santos, J., & Venturieri, A. (1994). Deforestation monitoring of the Amazon region using neural networksa comparison between different photointerpreters and networks. *Resource environmental monotoring, 30*, 49-55

Monteiro, A., Ferreira, C., & Madureira, H. (2005). Atlas Agroclimatológico do Entre Douro e Minho: Relatório do Projecto POCTI/GEO/14260/1998

Motohka, T., Nasahara, K.N., Oguma, H., & Tsuchida, S. (2010). Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing, 2*, 2369-2387

Murfitt, S.L., Allan, B.M., Bellgrove, A., Rattray, A., Young, M.A., & Ierodiaconou, D. (2017a). Applications of unmanned aerial vehicles in intertidal reef monitoring. *Scientific reports, 7*, 1-11

Murfitt, S.L., Allan, B.M., Bellgrove, A., Rattray, A., Young, M.A., & Ierodiaconou, D.J.S.r. (2017b). Applications of unmanned aerial vehicles in intertidal reef monitoring*, 7*, 1-11

Nybakken, J.W. (2001). *Marine biology: an ecological approach*. Benjamin Cummings San Francisco

Nybakken, J.W., & Cummings, B. (2019). *Marine Biology: An Ecological Approach*. (6ª edition ed.). Scientific e-Resources

Pal, M., & Mather, P.J.I.j.o.r.s. (2005). Support vector machines for classification in remote sensing*, 26*, 1007-1011

Pal, M.J.I.j.o.r.s. (2005). Random forest classifier for remote sensing classification*, 26*, 217-222

Peel, M.C., Finlayson, B.L., McMahon, T.A.J.H., & sciences, e.s. (2007). Updated world map of the Köppen-Geiger climate classification*, 11*, 1633-1644

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A.S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature, 540*, 418-422

Pirotti, F., Sunar, F., Piragnolo, M.J.I.A.o.t.P., Remote Sensing, & Sciences, S.I. (2016). BENCHMARK OF MACHINE LEARNING METHODS FOR CLASSIFICATION OF A SENTINEL-2 IMAGE*, 41*

Raffaelli, D., & Hawkins, S.J. (1996). *Intertidal ecology*. Springer Science & Business Media

Raj, R., Kar, S., Nandan, R., & Jagarlapudi, A. (2020). Precision Agriculture and Unmanned Aerial Vehicles (UAVs). *Unmanned Aerial Vehicle: Applications in Agriculture and Environment* (pp. 7-23): Springer

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. *Encyclopedia of database systems* (pp. 532-538)

Richards, J.A. (1999). *Remote sensing digital image analysis*. Springer

Rodriguez, J.D., Perez, A., & Lozano, J.A. (2009). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis machine intelligence, 32*, 569-575

Rossiter, T., Furey, T., McCarthy, T., & Stengel, D.B. (2020a). Application of multiplatform, multispectral remote sensors for mapping intertidal macroalgae: A comparative approach. *Aquatic Conservation-Marine and Freshwater Ecosystems, 30*, 1595-1612

Rossiter, T., Furey, T., McCarthy, T., & Stengel, D.B. (2020b). UAV-mounted hyperspectral mapping of intertidal macroalgae. *Estuarine, Coastal and Shelf Science*, 106789

Santos, F., Lopes, A., Moniz, G., Ramos, L., & Taborda, R.J.R.d.G.d.t.d.L. (2014). Gestão da Zona Costeira–O Desafio da Mudança. In

Simic Milas, A., Sousa, J.J., Warner, T.A., Teodoro, A.C., Peres, E., Gonçalves, J.A., Delgado Garcia, J., Bento, R., Phinn, S., & Woodget, A. (2018). Unmanned Aerial Systems (UAS) for environmental applications special issue preface. In: Taylor & Francis

Song, B., & Park, K. (2021). Comparison of Outdoor Compost Pile Detection Using Unmanned Aerial Vehicle Images and Various Machine Learning Techniques. *J Drones, 5*, 31

Soulard, C.E., Albano, C.M., Villarreal, M.L., & Walker, J.J. (2016). Continuous 1985–2012 Landsat monitoring to assess fire effects on meadows in Yosemite National Park, California. *Remote Sensing, 8*, 371

Steffani, C.N., & Branch, G.M. (2003). Growth rate, condition, and shell shape of Mytilus galloprovincialis: responses to wave exposure. *Marine Ecology Progress Series, 246*, 197-209

Stehman, S.V., & Czaplewski, R.L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment, 64*, 331-344

Sturrock, H.J., Cohen, J.M., Keil, P., Tatem, A.J., Le Menach, A., Ntshalintshali, N.E., Hsiang, M.S., & Gosling, R.D. (2014). Fine-scale malaria risk mapping from routine aggregated case data. *Malaria journal, 13*, 1-9

Tait, R., & Dipper, F. (1998). *Elements of Marine Ecology*. Oxford, Butterworth Heinemann

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A., & Rahman, A.J.R.S. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—A review*, 12*, 1135

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*, 127-150

Tuya, F., Larsen, K., & Platt, V.J.H. (2011). Patterns of abundance and assemblage structure of epifauna inhabiting two morphologically different kelp holdfasts*, 658*, 373-382

van Erkom Schurink, C., & Griffiths, C. (1993). Factors affecting relative rates of growth in four South African mussel species. *Aquaculture Asia, 109*, 257-273

Vapnik, V.N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks, 10*, 988-999

Watzlawick, L.F., Kirchner, F.F., & de Azambuja Madruga, P.R. (2003). Avaliação de classificação digital em imagens de vídeo multiespectral utilizando matriz confusão. *RECEN-Revista Ciências Exatas e Naturais, 5*, 47-57

Wilson, J.G. (2005). *The intertidal ecosystem: the value of Ireland's shores*. Royal Irish

Zeferino, J., Carvalho, M.R., Lopes, A.R., Jesus, R., Afonso, M.J., Freitas, L., Carvalho, J.M., & Chaminé, H.I. (2021). Assessment of future trends on groundwater quality in a nitrate vulnerable zone (Esposende–Vila do Conde sector, NW Portugal): towards a combined conceptual and mass transport modelling. *Hydrogeology Journal*, 1-17

Zhang, Q., Li, B., Thau, D., & Moore, R. (2015). Building a better urban picture: Combining day and night remote sensing imagery. *Remote Sensing, 7*, 11887-11913

Annex I

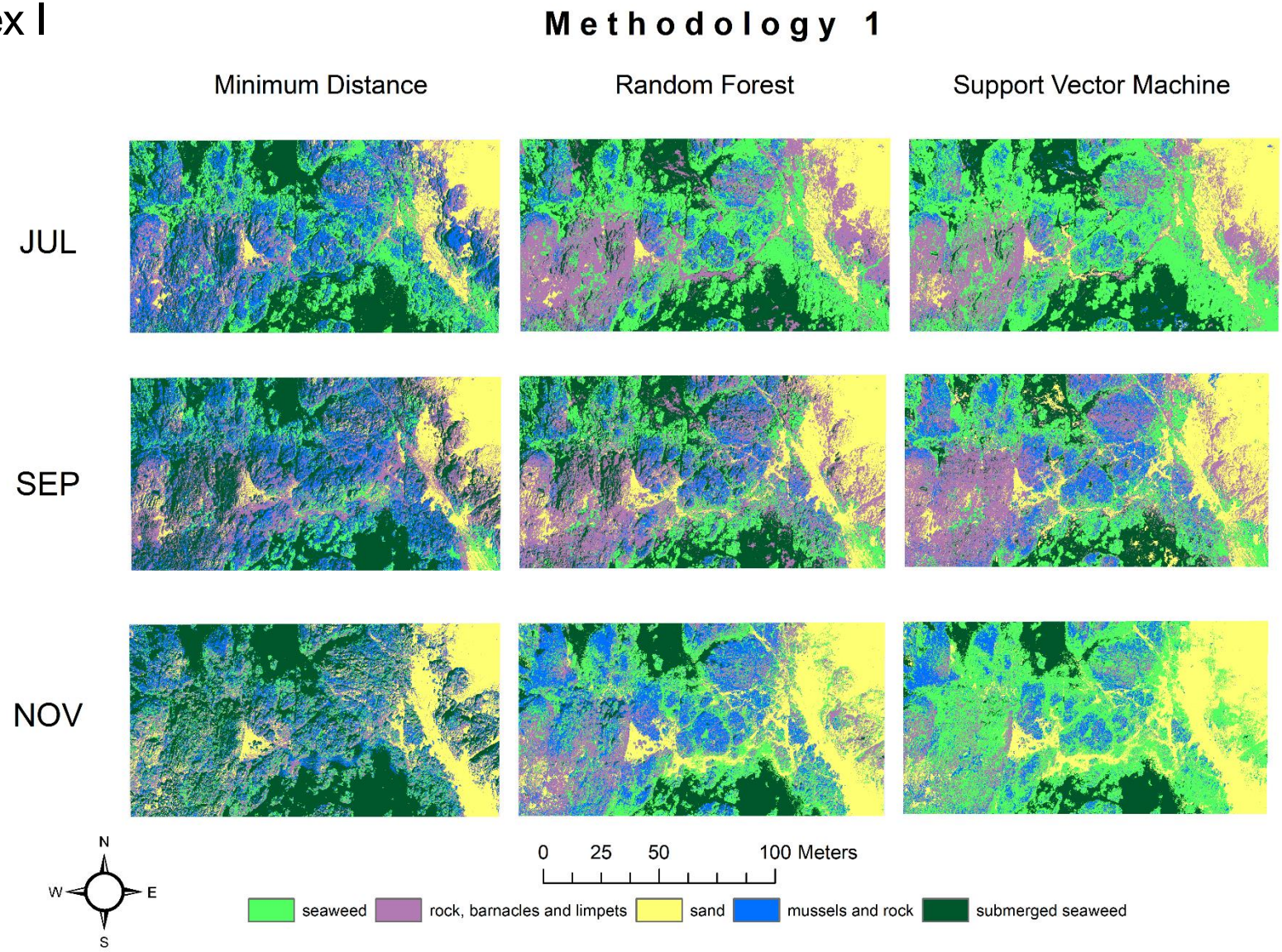


Figure 29 - Thematic maps of the classifications generated with methodology 1.

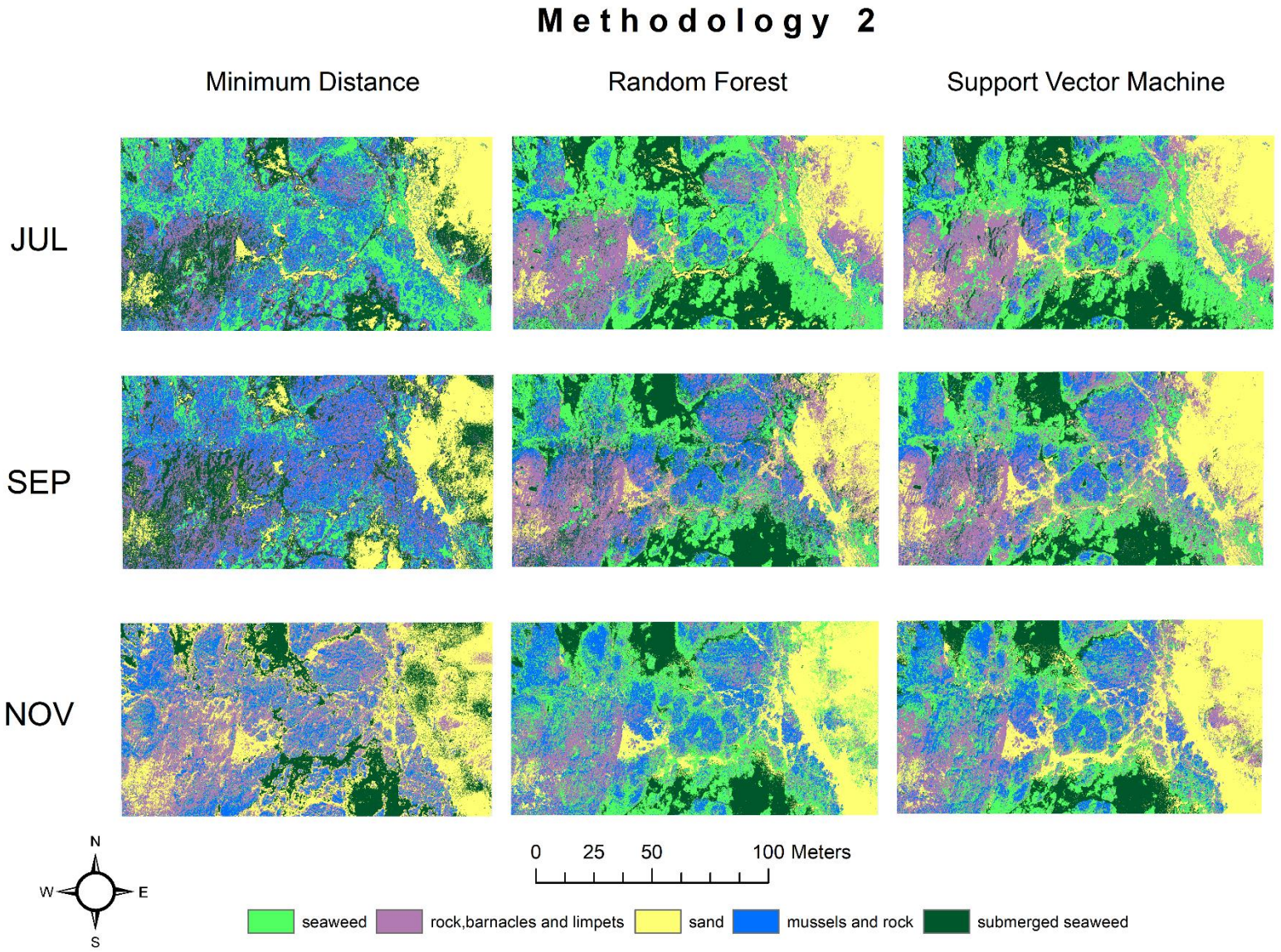


Figure 30 - Thematic maps of the classifications generated with methodology 2.

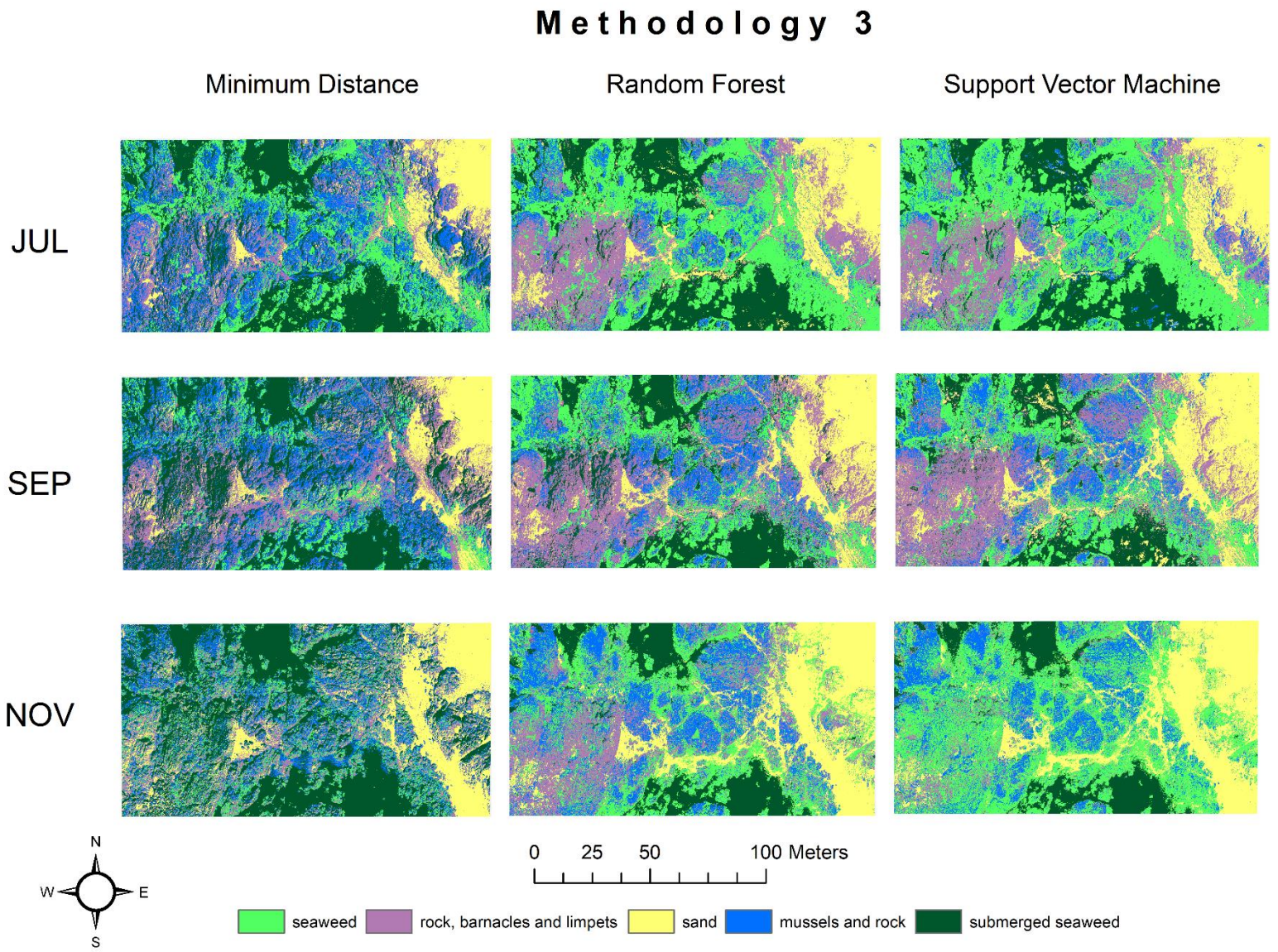


Figure 31 - Thematic maps of the classifications generated with methodology 3.

Annex II

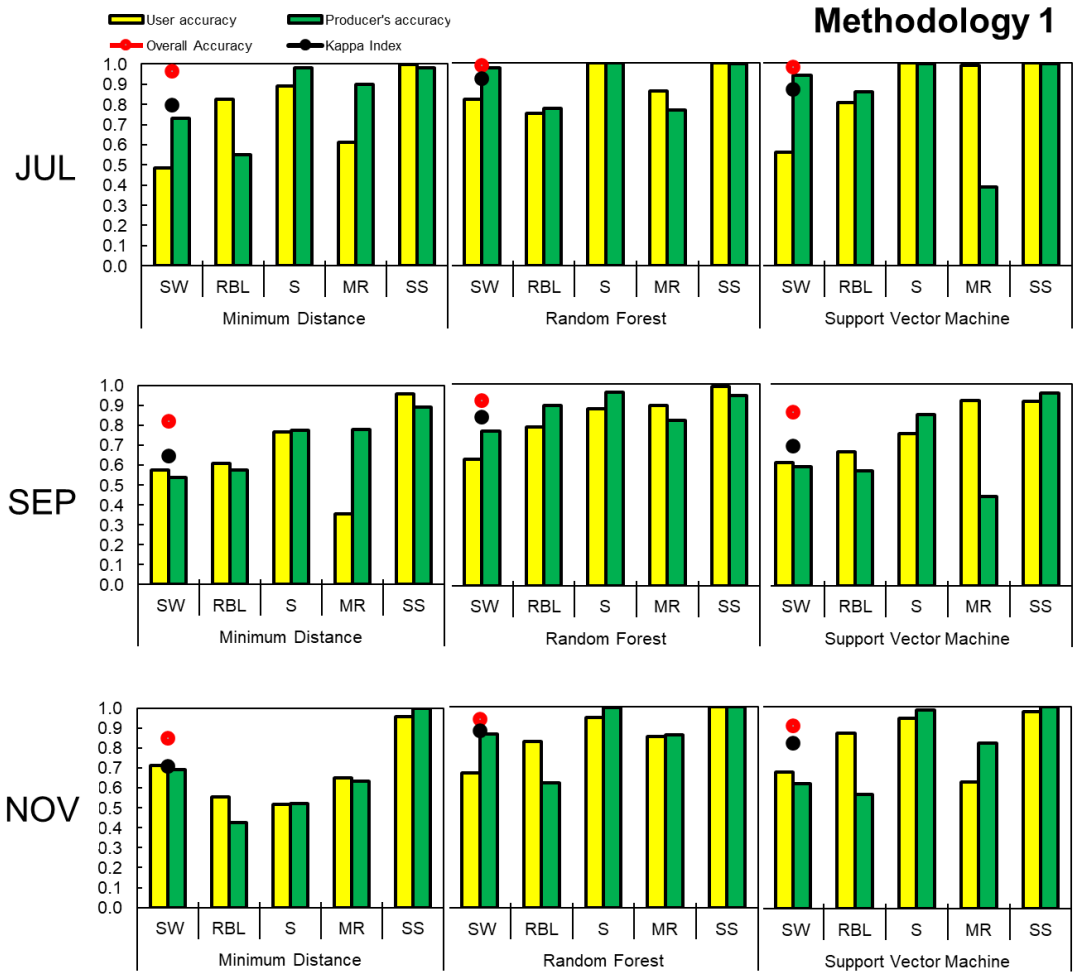


Figure 32 - Accuracy metrics for all months and all classifier algorithms using methodology 1. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

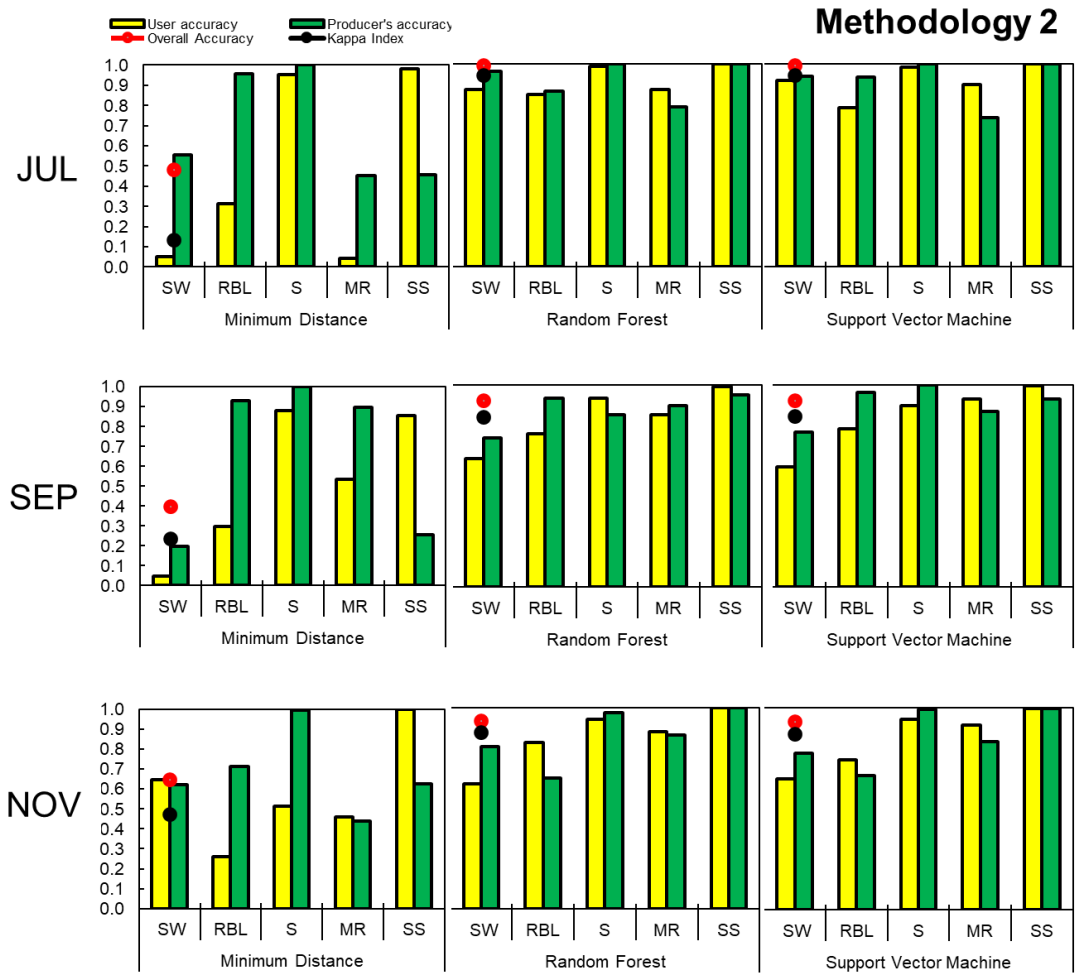


Figure 33 - Accuracy metrics for all months and all classifier algorithms using methodology 2. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

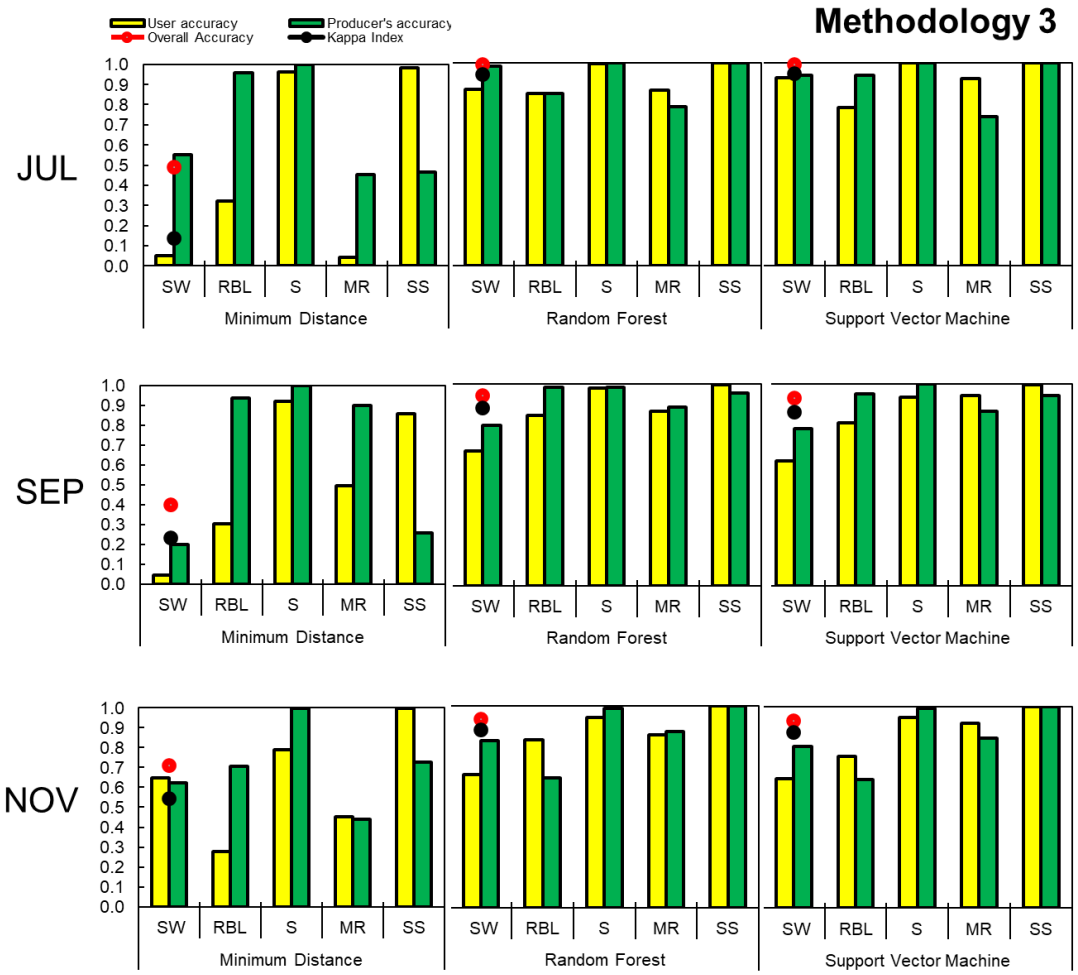


Figure 34 - Accuracy metrics for all months and all classifier algorithms using methodology 3. SW: seaweed; RBL: rock, barnacles and limpets; S: sand; MR: mussels and rock, SS: submerged seaweed.

Annex III

Table 7 - The parameters that were used for training (per classifier).

Minimum Distance	
The distance metric	Euclidean
K Nearest	1
Random Forest	
The number of decision trees to create	100
The number of variables per split. If set to 0 (default), defaults to the square root of the number of variables	0
The minimum size of a terminal node	no limit
The fraction of input to bag per tree	0.5
Support Vector Machine	
Decision Procedure	Voting
sum Type	C_SVC
Kernel Type	Linear
Degree	null
The gamma value in the kernel function	null
The coef0 value in the kernel function	null
The cost (C) parameter	1
The nu parameter	0.5