Development of seagrass monitoring techniques using remote sensing data

DISSERTATION

Zur Erlangung
des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
des Fachbereichs Mathematik/Informatik
der Universität Osnabrück

Vorgelegt von

Dimosthenis Traganos

Prüfer der Dissertation:
Prof. Dr. Peter Reinartz, Universität Osnabrück
Prof. Dr. Kostas Topouzelis, University of the Aegean, Griechenland

Tag der mündlichen Prüfung: 09. July, 2020

Abstract

Our planet is traversing the age of human-induced climate change and biodiversity loss. Projected global warming of 1.5 °C above pre-industrial levels and related greenhouse gas emission pathways will bring about detrimental and irreversible impacts on the interconnected natural and human ecosystem. A global warming of 2 °C could further exacerbate the risks across the sectors of biodiversity, energy, food, and water. Time- and cost-effective solutions and strategies are required for strengthening humanity's response to the present environmental and societal challenges.

Coastal seascape ecosystems including seagrasses, corals, mangrove forests, tidal flats, and salt marshes have been more recently heralded as nature-based solutions for mitigating and adapting to the climate-related impacts. This is due to their ability to absorb and store large quantities of carbon from the atmosphere. Focusing on seagrass habitats, although occupying only 0.2% of the world's oceans, they can sequestrate up to 10% of the total oceanic carbon pool, all the while providing important food security, biodiversity, and coastal protection.

But seagrass ecosystems, as all of their blue carbon seascape neighbors, are losing 1.5% of their extent per year due to anthropogenic activities. This has adverse implications for global carbon stocks, coastal protection, and marine biodiversity. Seagrass and seascape recession necessitates their science and policy-based management, protection, conservation which will ensure that our planet will remain within its sustainable boundaries in the age of climate change.

The present PhD Thesis and research aim is to develop algorithms for seagrass mapping and monitoring leveraging the recent emergences in remote sensing technology—new satellite image archives, machine learning frameworks, and cloud computing—with field data from multiple sources.

The main PhD findings are the demonstration of the suitability of Sentinel-2, RapidEye, and PlanetScope satellite imagery for regional to large-scale seagrass mapping; the introduction and incorporation of machine learning frameworks in the context of seagrass remote sensing and data analytics; the development of a semi-analytical model to invert the bottom reflectance of seagrasses; the design and implementation of multi-temporal satellite image approaches in coastal aquatic remote sensing; and

the introduction, design and application of a scalable cloud-based tool to scale up seagrass mapping across large spatial and temporal dimensions.

The approaches of the present PhD cover the gaps of the existing scientific literature of seagrass mapping in terms of the lack of spatial and temporal scalability and adaptability; the infancy in seagrass and seascape-related artificial intelligence endeavours; the restrictions of local server and mono-temporal approaches; and the absence of new methodological developments and applications using new (mainly open) satellite image archives.

I anticipate and envisage that the near-future steps after the completion of my PhD will address the scalability of the designed cloud-native, data-driven mapping tool to standardise, automate, commercialise and democratise mapping and monitoring of seagrass and seascape ecosystems globally. The synergy of the developed momentum around the global seascape with the technological potential of Earth Observation can contribute to humanity's race to adapt to and mitigate the climate change impacts and avoid cross tipping points in climate patterns, and biodiversity and ecosystem functions.

Zusammenfassung

Unser Planet durchläuft das Zeitalter des vom Menschen verursachten Klimawandels und des Verlusts der biologischen Vielfalt. Die prognostizierte globale Erwärmung von 1,5 °C über dem vorindustriellen Niveau und den damit verbundenen Treibhausgasemissionswegen wird nachteilige und irreversible Auswirkungen auf das vernetzte natürliche und menschliche Ökosystem haben. Eine globale Erwärmung um 2 °C könnte die Risiken in den Bereichen Biodiversität, Energie, Lebensmittel und Wasser weiter verschärfen. Zeit- und kosteneffiziente Lösungen und Strategien sind erforderlich, um die Reaktion der Menschheit auf die gegenwärtigen ökologischen und gesellschaftlichen Herausforderungen zu stärken.

Küstenökosysteme wie Seegras, Korallen, Mangrovenwälder, Wattenmeer und Salzwiesen wurden in jüngster Zeit als naturnahe Lösungen zur Minderung und Anpassung an die klimabedingten Auswirkungen angekündigt. Dies liegt an ihrer Fähigkeit, große Mengen an Kohlenstoff aus der Atmosphäre aufzunehmen und zu speichern. Sie konzentrieren sich auf Lebensräume für Seegras, obwohl sie nur 0,2% der Weltmeere einnehmen, können aber bis zu 10% des gesamten ozeanischen Kohlenstoffpools beschlagnahmen, während sie gleichzeitig eine wichtige Rolle bei der Ernährungssicherheit, der biologischen Vielfalt und dem Küstenschutz spielen.

Aber Seegras-Ökosysteme, wie alle ihre Nachbarn in der Blaukohle-Seelandschaft, verlieren aufgrund anthropogener Aktivitäten 1,5% ihres Umfangs pro Jahr. Dies hat nachteilige Auswirkungen auf die weltweiten Kohlenstoffbestände, den Küstenschutz und die biologische Vielfalt der Meere. Die Rezession von Seegras und Seelandschaft erfordert ihre wissenschaftlich und politisch fundierte Verwaltung, ihren Schutz und ihre Erhaltung, die sicherstellen wird, dass unser Planet im Zeitalter des Klimawandels innerhalb seiner nachhaltigen Grenzen bleibt.

Das Ziel der vorliegenden Dissertation und Forschung ist es, Algorithmen für die Seegraskartierung und -überwachung zu entwickeln, die die jüngsten Entwicklungen in der Fernerkundungstechnologie - neue Satellitenbildarchive, maschinelle Lernframeworks und Cloud Computing - mit Felddaten aus verschiedenen Quellen nutzen.

Die wichtigsten Ergebnisse der Doktorarbeit sind der Nachweis der Eignung von Sentinel-2, RapidEye und PlanetScope Satellitenbildern für regionale bis groß angelegte Seegraskartierungen; die Einführung und Integration von maschinellen

Lernrahmen im Kontext der Seegrasfernerkundung und Datenanalyse; die Entwicklung eines semi-analytischen Modells zur Umkehrung der Bodenreflexion von Seegras; das Design und die Implementierung von multitemporalen Satellitenbildansätzen in der aquatischen Küstenfernerkundung; und die Einführung, das Design und die Anwendung eines skalierbaren Cloud-basierten Tools zur Skalierung der Seegraskartierung über große räumliche und zeitliche Dimensionen hinweg.

Die Ansätze der vorliegenden Dissertation decken die Lücken der bestehenden wissenschaftlichen Literatur zur Seegraskartierung im Hinblick auf den Mangel an räumlicher und zeitlicher Skalierbarkeit und Anpassungsfähigkeit, die Anfänge der Bemühungen um künstliche Intelligenz bei Seegras und Seelandschaft, die Einschränkungen der lokalen Server- und monozeitlichen Ansätze sowie das Fehlen neuer methodischer Entwicklungen und Anwendungen unter Verwendung neuer (hauptsächlich offener) Satellitenbildarchive.

Ich erwarte und stelle mir vor, dass die nahen zukünftigen Schritte nach Abschluss meiner Doktorarbeit die Skalierbarkeit des entwickelten Cloud-nativen, datengesteuerten Kartierungsinstruments zur Standardisierung, Automatisierung, Kommerzialisierung und Demokratisierung der Kartierung und Überwachung von Seegras- und Meeresökosystemen weltweit betreffen werden. Die Synergie der entwickelten Dynamik um die globale Meereslandschaft herum mit dem technologischen Potenzial der Erdbeobachtung kann dazu beitragen, dass die Menschheit sich an die Auswirkungen des Klimawandels anpasst und mildert und Kreuzungspunkte in den Klimamustern sowie in den Funktionen der biologischen Vielfalt und des Ökosystems vermeidet.

Acknowledgments

"Keep in mind how fast things pass by and are gone—those that are now, and those to come. Existence flows past us like a river: the "what" is in constant flux, the "why" has a thousand variations. Nothing is stable, not even what's right here."

-Marcus Aurelius – Meditations

First, I would like to acknowledge the German Aerospace Center (DLR) and the German Academic Exchange Service (DAAD) for funding my four-year PhD studies and related developments. I am particularly indebted to my supervisor Prof. Dr. Peter Reinartz for always being present in a calm and didactic way. Without his initial acceptance, guidance and granted research freedom, I wouldn't be able to reach my current level of knowledge and skills.

Second, I would like to thank the people and colleagues with whom I shared my space and time at the Optical Remote Sensing of Water Team in DLR Berlin these four years. I am also grateful to two students that I had the pleasure to supervise for their BSc and MSc studies, Yannik Esser and Bharat Aggarwal.

Third, I am very happy and thankful that I had the opportunity to visit the Biophysical Remote Sensing Group at the School of Earth Sciences of the University of Queensland to work with and learn from Dr. Chris Roelfsema and the rest of the team. It was a thoroughly enjoyable and full of images (literally and metaphorically) one-month trip to Australia which I will always consider as a gift that keeps on giving.

Fourth, I am also thankful to Dr. Dimitris Poursanidis, Assistant Prof. Konstantinos Topouzelis, Aurélie Shapiro, and Dr. Daniele Cerra who shared their time, knowledge, insights, and ideas with me. Their experience and enthusiasm about science and life have motivated me during various times of my PhD.

Fifth, I would like to express my deepest gratitude to my entire family: my parents, Christina and Christos, who have provided me with the right genes, bacteria, inspiration and support during my entire life and have fostered my relationship with seawater before I could even speak; my parents-in-law, Eleni and Thodoros, who have additionally supported me with their unconditional love and understanding through calm and rough times; my sister and sister-in-law, Maria and Joan, who have provided me with significant support and balance through their intelligence and humor.

And last but not least, there will never be that many words to thank my wife and partner in life, in bliss and in crime, Miro. Thank you for being there and loving me the way I am, even when I wasn't; for guiding me not to waste time on nonsense and *noise*; for your industriousness which has extensively inspired me to become my current self; and for your acuteness which has energetically aided me to *see* and fulfill my potential, foremost as a human and furthermore as a scientist. For the above, I wholeheartedly dedicate this PhD to you.

Dimosthenis Traganos
Berlin, September 2020

Contents

1 Introduction

1.1 Motivation

In the age of Climate Change, global biodiversity loss and the Fourth Industrial Revolution that we are traversing, the planetary seascape ecosystem—including habitats like seagrasses, corals, mangroves, tidal flats—can offer an impressive and highly-valued panoply of ecosystem services for our protection. Seagrasses, underwater vegetation forming kilometer-wide meadows in the global seascape, once on land around the time that dinosaurs were roaming on Earth and then in the water, due to the toxic atmosphere, have been the center of interest of Aristotle when he was laying the foundations of the modern marine biology. Featuring the oldest living organism of nature at the moment (a *Posidonia oceanica* seagrass species in the Western Mediterranean of a 200,000-yr lifespan), seagrasses, despite their relatively small ocean coverage, have been more lately explored by scientists and policy-makers for their large carbon sinking abilities and services. In addition, seagrass habitats can protect the planetary coastline from erosion by trapping sediment in their roots; can clean the water column from pathogens and bacteria; offer food and nursery to thousands of marine organisms worldwide; all of which highlight their significant natural, cultural and financial value.

Seagrasses are known to cover a seabed area more or less equal to the size of Germany, in all continents but Antarctica, in exposed to the sunlight regions to 80 m of depth, in very clear to very turbid waters, in single patches to some kilometers of homogeneous beds, and in occurrence of single species to mixed grounds of more than a dozen species. But like all of their seascape neighbors, due to a plethora of mainly anthropogenic impacts, seagrasses have been dramatically degraded and decreased in area, limiting the magnitude and scale of their support to natural and human ecosystems, and to the multitude of global problems that relate to the coastal marine environment. This problem is further exacerbated by the currently limited information on the extent and diversity of seagrass beds; we cannot answer the why without knowing the where, the what and the how much. And without the three latter, nor can we realise effective protection, management and conservation of these vital organisms.

The need to find out, understand the where, what, how much and why of the seagrass ecosystems and add one small piece, hopefully, to the puzzle of the required holistic solution provided extrinsically the motivation for this PhD research for the last four years. The timing for the sparked motivation was exquisite due to the technological developments of the Fourth Industrial Revolution in artificial intelligence, cloud computing, and satellite sensor and data availability and capacity. All these offered the necessary symbiotic environment to develop more focused and state-of-the-art tools to explore and study seagrasses. Since the onset of my PhD, I have been lucky enough to experience the practical expansion of artificial intelligence, the realisation of the power of cloud computing, the exponential growth in government-funded and commercial satellite sensors and the resulting Big-Data paradigms of science and management. Without the aforementioned technological advances, my whole PhD would be rich in theories, but pure in tools, solutions and impacts.

As such, the designed and developed go-to approach of this PhD combines multi-scale satellite data (in single scenes to multi-temporal mosaics), powerful machine learning classifiers, and thousands of reference data (collected in the field or annotated on very high-resolution satellite images) to feed the artificial intelligence models within a cloud computing environment. This mainly top-down approach can be adapted in various time and spatial scales (from seasonal to inter-decadal and from single coasts to the planetary scale), and remote sensing data input to provide large-scale maps of seagrasses and their seascape roommates (e.g., corals, mangroves, tidal flats, kelps, macroalgae) and to better understand and protect the coastal marine benthos.

Optimistically, in the increasingly-threatened interconnected natural and human ecosystem, my proposed approach and solution could strengthen the understanding and the momentum around seagrass science, mapping, and management; and motivate other people, from students to managers and ministers, to mind seagrasses the same way I have minded them since my infancy.

1.2 Objectives

Since the very start of this PhD, the main objective has been to develop algorithms for seagrass mapping and monitoring using satellite data from multiple sources: government-funded, open and freely available satellite image archives from the Sentinel and Landsat series, and private-funded commercial from e.g., Planet (PlanetScope and

RapidEye), WorldView, etc. The algorithmic design and development have taken place from local servers and specialized off-the-shelf softwares to open cloud-based online platforms. To achieve my main objective, I had to set and meet a plethora of smaller sub-objectives:

a) To demonstrate the suitability of the aforementioned sensors for regional to large-scale mapping of seagrasses at various spatial scales (0.5 to 30 m).
b) To introduce and incorporate machine learning frameworks in the context of the remote sensing (RS) of seagrasses and more generally coastal aquatic habitats.
c) To select and annotate suitable and independent large-scale training and validation (reference) datasets to feed and test the machine learning algorithms in use.
d) To introduce, design and implement adaptive web-based serverless tools in the coastal aquatic context to effectively scale up seagrass mapping in the spatial and temporal dimensions.
e) To develop a semi-analytical model to invert the seabed (i.e. bottom or benthic) reflectance of seagrasses and similar nearby habitats.

All of the above five sub-objectives are interconnected and essential components of the cloud-based mapping and monitoring tool whose successful near-future scalability and implementation will resolve past and existing trends in seagrass ecosystems, unravel data issues, and aid their management, protection and conservation worldwide.

1.3 Contributions

There have been six publications that compose the sub-chapters of Chapter 4 of this PhD Thesis, reflecting its main objectives and contributions. More specifically, from the oldest to the newest, the most prominent contributions of my PhD research are the following:

a) The demonstration of the suitability of Sentinel-2, RapidEye, and PlanetScope satellite sensors for multi-scale baseline and time-series mapping of seagrasses (4.1 – 4.6).
b) The incorporation and democratization of Support Vector Machines and Random Forests for seagrass RS (4.1 – 4.6).
c) The identification of the degradation of *Posidonia oceanica* seagrass habitats with a trend of ~-0.1 km^2/yr between 2011 and 2016 in the largest gulf of the Aegean Sea

(Thermaikos Gulf, Greece) through a time-series of RapidEye satellite images (4.5).

d) The remotely sensed identification of seagrass beds up to 33 m of depth in the south of Crete (South Aegean Sea, eastern Mediterranean) on Sentinel-2 imagery (4.3).
e) The implementation and designation of the metrics of probability and uncertainty in the context of coastal aquatic RS (4.3).
f) The design of a semi-analytical model following a machine learning-based approach to retrieve water column and seabed characteristics of Mediterranean seagrasses from Sentinel-2 data (4.4).
g) The introduction and designation of multi-temporal approaches (in contrast to single-image approaches) in the coastal aquatic RS context aided by and hosted in the cloud environment (4.6).
h) The establishment of an end-to-end cloud-based methodological workflow for large-scale, high spatio-temporal baseline mapping and monitoring of seagrasses, combining the use of the cloud, machine learning models, public satellite imagery from Sentinel-2, and low cost in situ data. The first demonstration of the adaptability, efficiency and accuracy of the workflow is reflected in the mapping of 2,510 km^2 of seagrasses throughout 40,950 km^2 of the country-scale seabed extent of Greece using a multi-temporal composite of 1,045 Sentinel-2 images (4.6).

1.4 Structure of the Thesis

The structure of the present PhD Thesis is as follows: Chapter 2 describes important background information regarding both the Organism of interest here and the Technology in use to study and understand the Organism.

Chapter 3 introduces you to the State-of-the-Art mainly concerning the last twenty years of research in the remote sensing of seagrass ecosystems worldwide.

Chapter 4 presents the research contributions of my PhD: six publications related to the remote sensing of seagrass ecosystems and neighboring seascape, structured as sub-chapters that follow an increasing scalability axis.

Chapter 5 discusses the connection of Chapter 4 contributions and publications, and highlights the most vital findings of this PhD within a seascape connectivity, climate change and biodiversity loss context.

Finally, Chapter 6 puts forth the conclusions and takeaways of the present PhD Thesis, and finalizes it with my outlook for both the short and long-term future of my research and field.

2 Background

2.1 The Organism

Seagrasses (and not seaweed or algae) are marine flowering plants that can form extensive seafloor, monospecific meadows limited to the coastal continental shelf (0–80 m), where there is enough light for their growth via photosynthesis (Fig. 2.1). There are 72 species of these marine flowering plants which cover 344,958 km^2 of seabed in all continents but Antarctica (UNEP-WCMC, Short, 2018). Fig. 2.2 and Fig. 2.3 present the best current information on the global seagrass extent and species distribution. These 72 seagrass species extend their area in six global seagrass bioregions: The Temperate North Atlantic, Tropical Atlantic, Tropical Indo-Pacific, Temperate Southern Oceans, Temperate North Pacific, and Mediterranean (Green and Short et al. 2003; Fig. 2.3).

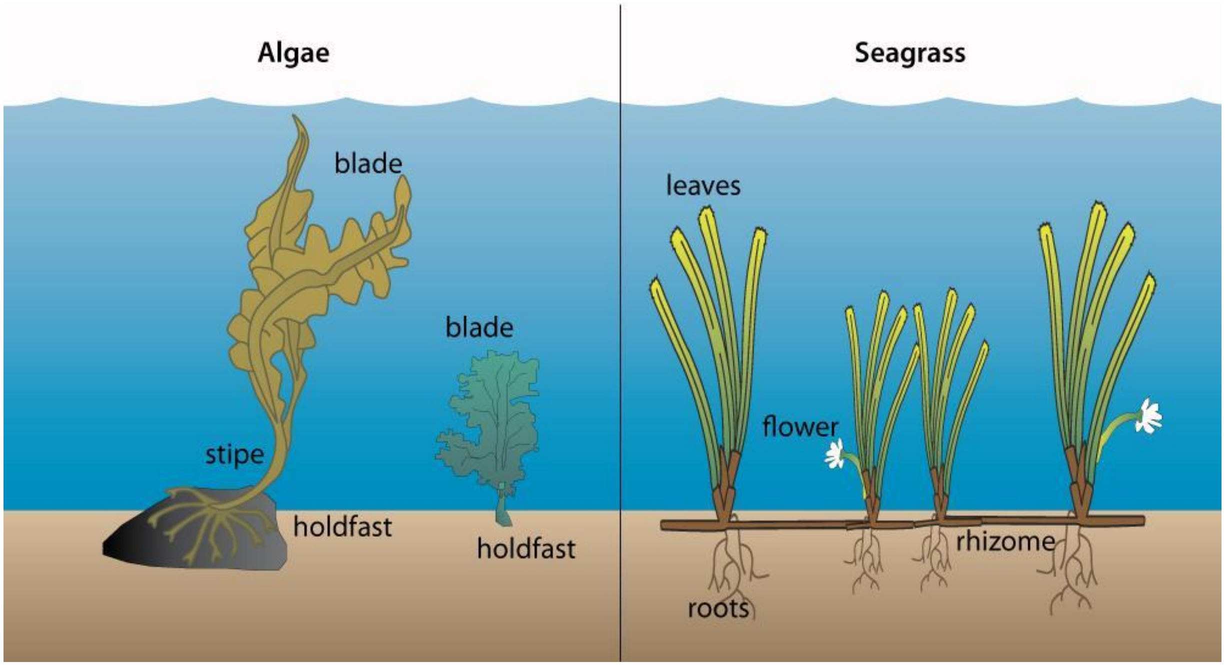


Fig. 2.1. Physiology of Seagrass vs Algae or “Seaweed” (Courtesy of the Integration and Application Network (ian.umces.edu), University of Maryland Center for Environmental Science).

The global species diversity of seagrasses is low, but species can extend over thousands of kilometers of coastline. They form a critical marine ecosystem for carbon storage, fisheries production, sediment accumulation, and stabilization (Green and Short, 2003). They contribute to the function of ocean ecosystems by providing an important nursery area for many species that support offshore fisheries and for adjacent habitats such as salt marshes, shellfish beds, coral reefs, and mangrove forests. Seagrass ecosystems are critical for threatened species (i.e. sirenians (dugong and manatee), sea turtles, and seahorses) which have high cultural, aesthetic and intrinsic values. The ecosystem functions of the seagrass meadows include maintenance of genetic variability, the resilience of the coastal environment through protection from erosion, and carbon sequestration by removing carbon dioxide from the atmosphere and binding it as organic matter (Nordlund et al., 2016). Their high productivity attributes them a disproportionate influence on oceanwide primary productivity, typically producing considerably more organic carbon than the seagrass ecosystem requires (Githaiga et al., 2017). Carbon storage by seagrasses is essentially an effective removal of carbon dioxide from the ocean–atmosphere system which plays a significant role in the amelioration of climate change impacts (Fourqurean et al., 2012).

Zooming in the seagrass bio-region of the Mediterranean, the basis for the majority of the herein presented research, the latter supports a great biodiversity of up to 18% of all known marine species (~17,000) (Coll et al., 2010). Seven of the total 72 seagrass species are found in the Mediterranean Sea (Short et al., 2007). *Posidonia oceanica* is the dominant Mediterranean seagrass species and although it has a slow horizontal growth of ~1 cm/yr, it is considered as one of the longest-living organisms with meadows reaching a thousand years of age (Arnaud-Haond et al., 2012).

Various studies have highlighted the numerous ecosystem services of the large *P. oceanica* meadows and the Mediterranean seagrass species in general including nursery grounds, buffering from coastline erosion, wave protection, carbon sequestration (the so-called 'blue carbon') and bio-indicator amongst others (Campagne et al., 2014; Fourqurean et al., 2012; Mtwana Nordlund et al., 2016). The largest stocks of organic carbon among all seagrasses have been documented in the extensive rhizomes of *P.* oceanica; three times greater mean living biomass is present in the Mediterranean seagrasses than the global mean of six seagrass bioregions (Fourqurean et al., 2012). Furthermore, the economic value of *P. oceanica* meadows arising from their protection from coastal erosion and sediment retention services has been estimated between 51.5 thousand €/km^2/yr and 1.72 million €/km^2/yr (Campagne et al., 2014; Vassallo et al., 2013).

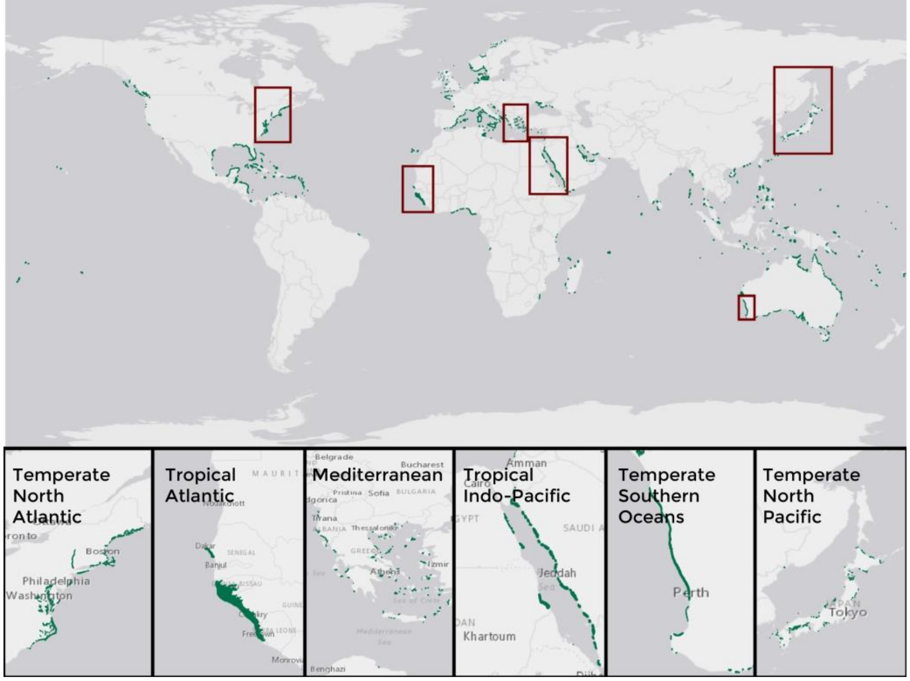


Fig. 2.2. Global Extent of Seagrasses (modified from UNEP-WCMC, Short, 2019) and examples from the six global seagrass bioregions.

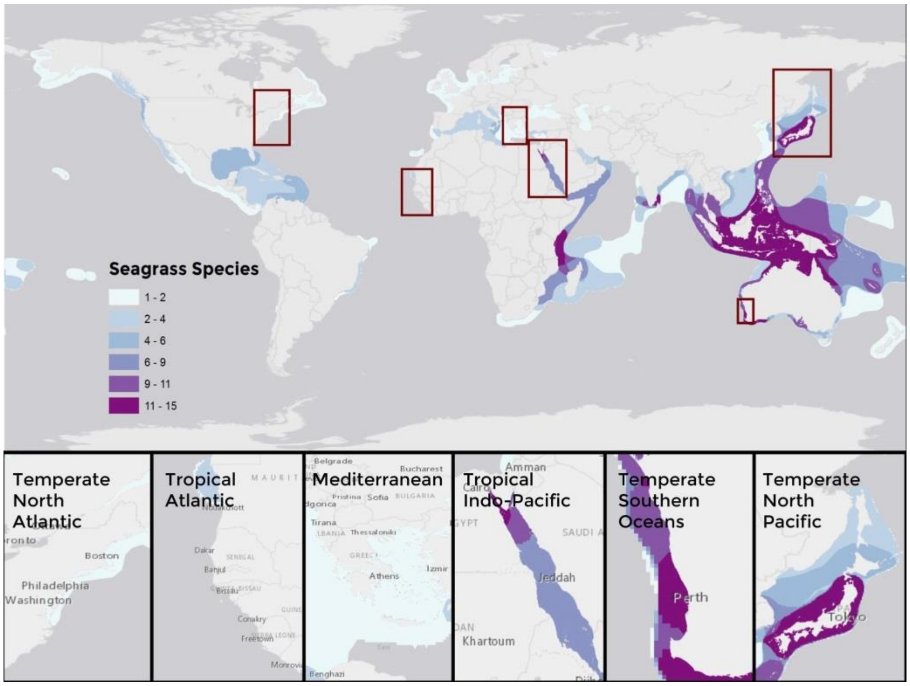


Fig. 2.3. Global Seagrass Species Distribution (modified from Green and Short, 2003) and examples from the six global seagrass bioregions.

P. oceanica species is protected by legislation under the EU Habitat Directive (92/43/CEE), the Bern (Annex II, Strictly Protected Flora Species) and the Barcelona (dedicated Action Plan under the "Protocol concerning Specially Protected Areas and Biological Diversity in the Mediterranean") Conventions as well as other legislations at a national level. Furthermore, according to the Marine Strategy Framework Directive (MFSD; 2008/56/EC), *P. oceanica* is selected as a representative species of the angiosperm quality elements for the Mediterranean marine environment and a 'Good Environmental Status' shall be achieved by all Member States regarding the angiosperm habitats. More recently, the EU regulated a ban on destructive fishing activities such as trawling within the first 50 m of depth in the Mediterranean, a measure which is bound to protect seagrass species indirectly (EC Council Regulation No. 1967, 21/12/2016) (Pergent et al., 2010).

Despite being a conservation priority of EU legislation, Mediterranean seagrass habitats are experiencing a decreasing trend in their coverage. Telesca et al. (2015) estimated an average area loss of 10.1% of *P. oceanica* meadows during the last 50 years throughout the Mediterranean seabed, but excluding areas without historical information on seagrass beds (mainly in the eastern Mediterranean), this number increases to 33.6%. The latter increase is close to the global report of a 29% decline of seagrass meadows since 1879. Declines of populations of Mediterranean seagrass ecosystems have been attributed mainly to anthropogenic activities including trawling, anchoring, coastal development, eutrophication and climate change (Bonacorsi et al., 2013; Waycott et al., 2009).

The major problem with rectifying this decreasing trend is the sparsity or absence of data on the extent of Mediterranean seagrass habitats, especially in the eastern and southern parts of the Mediterranean basin (Telesca et al., 2015). Present data in this part of the Mediterranean Sea are based on experts' knowledge and are principally point-based information on seagrasses' presence/absence (Giannoulaki et al., 2013). Lack of these valuable data hinders any effort for effective management and conservation of Mediterranean seagrass habitats.

Zooming out to the planetary-scale seagrass extent and condition, the existence of seagrass datasets at local to global scale can support resource management, strengthen decision-making and facilitate tracking of progress towards global conservation targets set by multilateral environmental agreements, such as the Aichi Biodiversity Targets of the United Nations' (UN's) Strategic Plan for Biodiversity 2011-2020, the Ramsar Convention, and the Sustainable Development Goals (SDG) of the UN's 2030 Agenda for Sustainable Development—particularly Goal 14 "Conserve and sustainably use the oceans, seas and marine resources for sustainable development" of the UN SDG 2030 (UN General Assembly, 2015).

2.2 The Technology

Modern problems require modern solutions. Technology could solve the problems associated with seagrasses today and strengthen our capacity to effectively protect and manage them globally. I present here the three satellite-related technologies that I have mainly exploited during my PhD and whose symbiosis will allow large-scale seagrass mapping and monitoring in the future in multiple spatio-temporal scales and in near real-time: a) the Satellite Sensors, b) the Cloud Computing Platforms, and c) the Artificial Intelligence.

2.2.1 Satellite Sensors

Optical satellite technology has been around since the early 70s with the first of the Landsat Series, the Landsat 1. Since then, and especially in the last 20 years, exponential advances in the so-called *Earth Observation* (i.e. the remote-sensing-based acquisition of information about the biophysiochemical properties of Earth's surface) have yielded hundreds of satellite sensors in single and constellation missions. The instruments that these sensors carry measure the sunlight which has travelled through the atmosphere, has penetrated the water surface and the water column and has been reflected by the seabed on its trip back through the water column and atmospheric media to the satellite instrument (Fig. 2.4). These sensor instruments measure the sunlight in different spatial (level of image detail), spectral (wavelengths in which light is measured), temporal (revisit time of measurement) and radiometric (colour depth) resolutions. Different parameters of seagrasses (e.g., extent, distribution, leaf area index, and bathymetry) can be quantified depending on the different satellite resolutions (Fig. 2.5). The associated resolutions to seagrass detection, mapping and monitoring are between 0.3 to 30 m of spatial resolution, revisit time between 1 and 17 days, and spectral wavelengths mainly in the visible range—400 and 700 nm (Fig. 2.6).

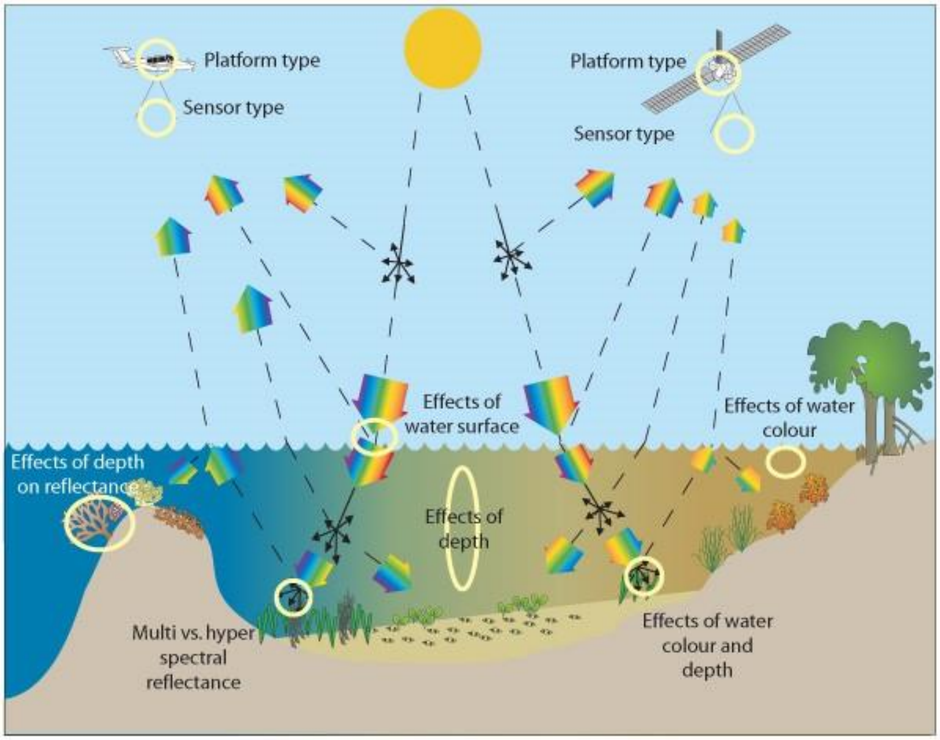


Fig. 2.4. Interactions of atmosphere, water surface, water column and seabed with sunlight over shallow coastal waters (from Roelfsema et al., 2017).

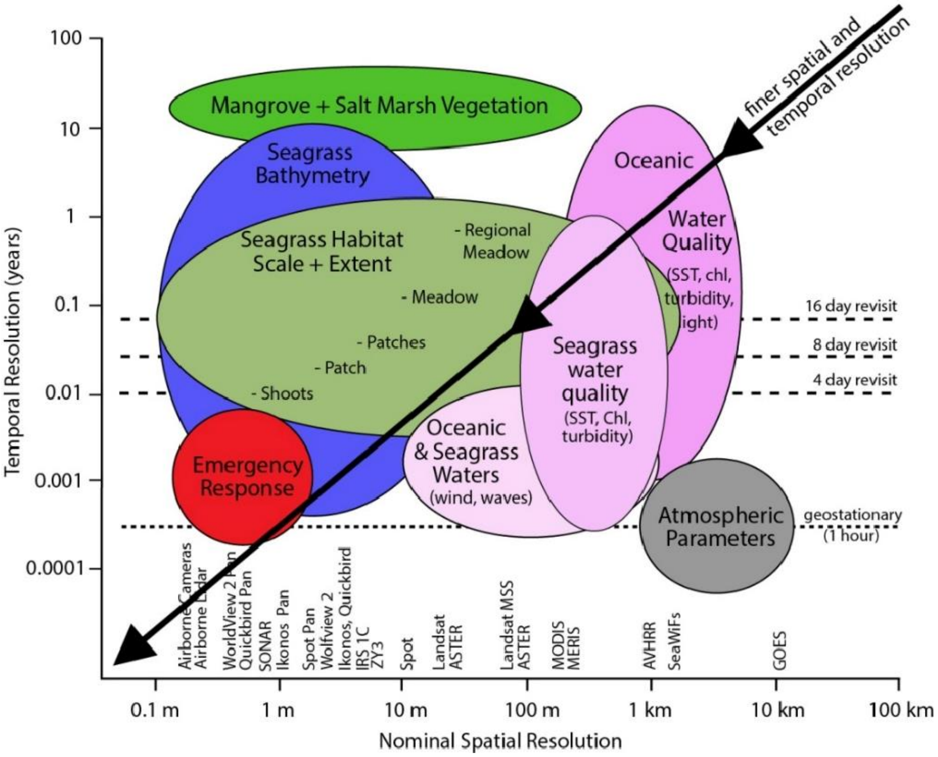


Fig. 2.5. Spatial and temporal resolution of satellites relevant to mapping of specific seagrass parameters (from Hedley et al., 2016).

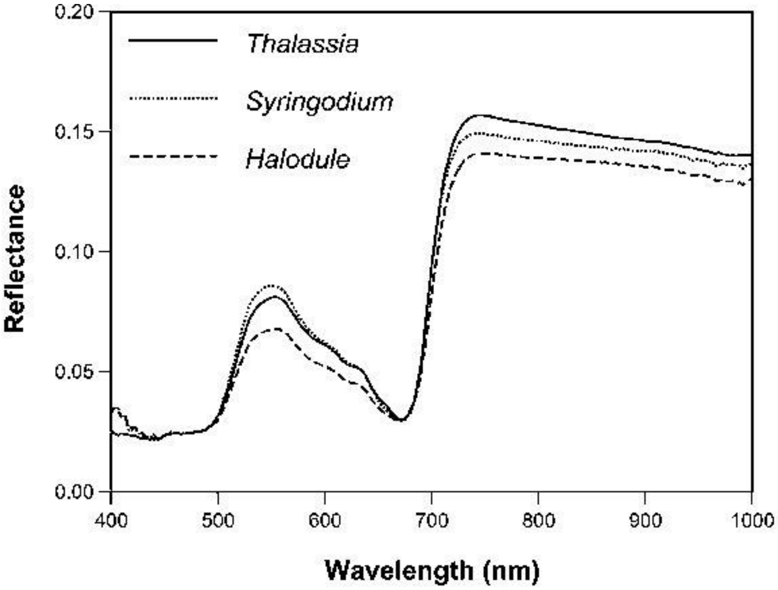


Fig. 2.6. Spectral profile of single blades of three seagrass species (*Thalassia testudinum, Syringodium filiforme,* and *Halodule wrightii;* from Thornhaug et al., 2007).

Today's satellites which satisfy the aforementioned characteristics and have been used to map seagrasses in a plethora of geographic areas in the last decade are either commercial or privately funded: IKONOS[1], Worldview series[2], Quickbird-2[3], Planet's PlanetScope and RapidEye[4]; and government-funded with open, public and free image archives: European Union's Copernicus Sentinel-2[5] and NASA Landsat Series[6]. Fig. 2.7 shows the commonly used for seagrass mapping spectral wavelength ranges of the aforementioned sensors. As regards specifically to Planet's PlanetScope, they are part and product of the increasing trend in the development of spaceborne sensors, Cubesats (National Academies of Sciences, Engineering and Medicine, 2016). CubeSats are small satellites whose size is expressed in increments of 10 cm cubes, such that 1 cube is 1U or "unit", 2 cubes are 2U (two 10 cm cubes together), etc. Earth observation scientists have barely assessed the potential of these shoebox-sized satellites in terms of lower-cost, but faster science return (National Academies of Sciences, Engineering and Medicine,

---

[1] https://www.satimagingcorp.com/satellite-sensors/ikonos/
[2] https://www.digitalglobe.com/company/about-us
[3] https://www.euspaceimaging.com/about/satellites/quickbird/
[4] https://assets.planet.com/docs/combined-imagery-product-spec-final-august-2019.pdf
[5] https://sentinel.esa.int/web/sentinel/missions/sentinel-2
[6] https://www.usgs.gov/land-resources/nli/landsat

2016). Having launched 88 CubeSats on 14.02.2017 (the so-called ‘Doves’), Planet has reached an unprecedented milestone of operating approximately 120 satellites in parallel in low Earth orbit. Planet’s satellite constellation of CubeSats 3U will allow daily imaging revisit of the entire Earth once it enters normal imaging operation.

The commercial satellite image datasets feature higher spatial and temporal resolution to their public counterparts, but at a higher cost as well. Fig. 2.8 showcases the effect of varying spatial scale of different sensors (including a drone) on detecting seagrasses. Another drawback with using expensive commercial satellite images is their limited sensed area; for example, a single CubeSat-derived PlanetScope image contains a slice of the Earth of 24.6x6.4 km in size, while in contrast, an individual Sentinel-2-derived tile contains a 63-times larger slice—100x100 km in size. The characteristics of the satellite sensors, the area of the seagrass of interest, and the available funds will ultimately dictate the selected instrument in use for seagrass mapping and monitoring. More information about the exploited satellite sensor technology in each research contribution can be found in the Materials and Methods sections of Chapter 4.

2.2.2 Cloud Computing Platforms

In the last five years, various cloud computing platforms have been realised and grown. These platforms can support Big-Data paradigms of science and management with an emphasis on data-driven and data-intensive analyses, time and cost-efficient data access, large computational resources, and high-end visualization (Goodchild et al., 2012). In the seagrass and seascape context, the paradigm shift of cloud computing environments could allow global-scale estimations, information, insights and management. Cloud serverless platforms provide five main advantages in comparison to the traditionally used local servers and softwares:

a) Petabyte-scale image archives of mostly open, free and public satellite, geospatial and auxiliary non-geospatial datasets.
b) High-performance computation, processing and analysis through massive parallelization—across thousands of computers in contrast to a single one.
c) Multi-temporal approaches following the implementation of all available images within a selected time period (e.g., season, year) instead of the traditional single image/scene approaches; multi-temporal composites theoretically contain fewer interferences (e.g., clouds, sunglint, waves) than single scenes.

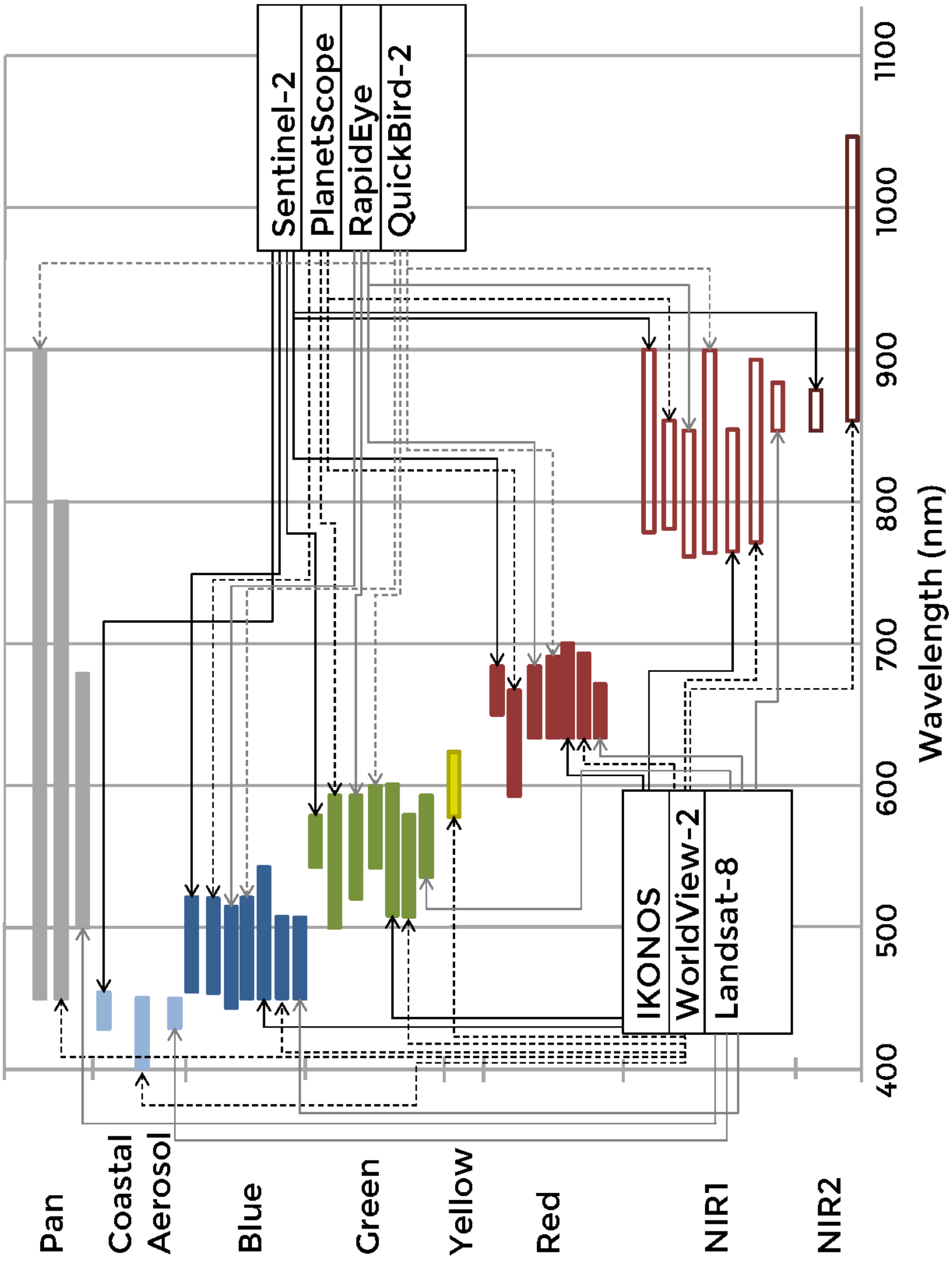


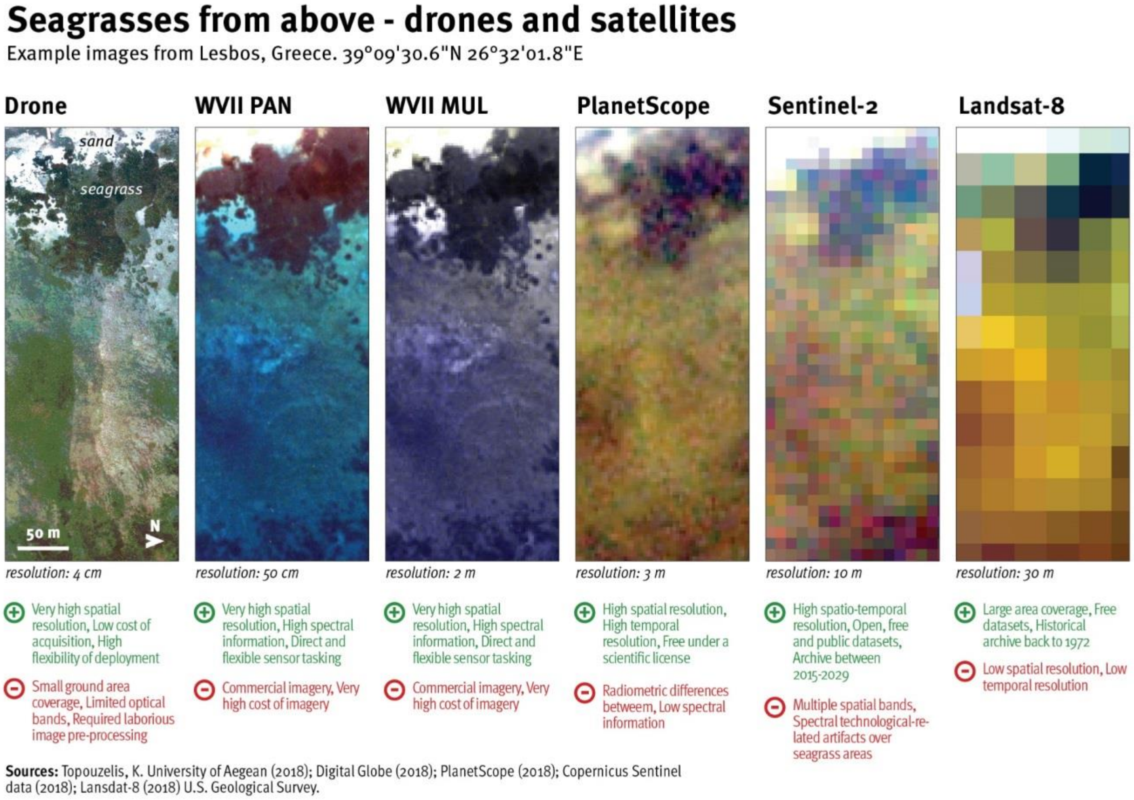
Fig. 2.7. Spectral wavelengths of the main multispectral satellites which have been commonly used in seagrass mapping.

a) Wide set of available artificial intelligence frameworks within the cloud environment.
b) Online visualization capability which allows an initially fast and easy qualitative assessment of cloud-based products.

Until the first semester of 2019, four platforms have been developed and are currently offering their cloud environment for storing, processing, analyzing, classifying, and visualizing Big Satellite Data: Google Earth Engine (Gorelick et al., 2017), Amazon Web Services (Amazon AWS, 2019), Microsoft Azure (Microsoft Azure, 2019), and the European Commission's Copernicus DIAS (Data and Information Access Services; Copernicus DIAS, 2019).

Scientists have harnessed the serverless platform of the Google Earth Engine to estimate the planetary extent and trajectory of forests (Hansen et al., 2013), surface water (Pekel et al. 2016), and tidal flats (coastal neighbors to seagrasses; Murray et al., 2019). The above and similar global research efforts have given birth to important baseline mapping and monitoring exercises and services: the Global Forest Watch, the Global Mangrove Watch, and the Allen Coral Atlas; these state-of-the-art planetary-scale efforts should provide inspiration, motivation, and technical capacity towards similar planetary-scale estimations in the seagrass and the holistic seascape extent.

Sub-chapter 4.6 presents how the symbiotic synergy between the cloud environment of the Google Earth Engine, artificial intelligence algorithms, and the open satellite data of Sentinel-2 allowed the first paradigm of large-scale baseline mapping at relatively high accuracies (Traganos et al. 2018a). The end-to-end cloud-based workflow, once cached, required less than 10 seconds to ingest more than 1,000 satellite images at 10-m spatial resolution, mosaic, pre-process for corrections and classify the multi-temporal composite and calculate the seagrass distribution in more than 40,000 km^2. I envisage that such a paradigm can and will pave the way for similar near-future endeavors in mapping of various seagrass and seascape biophysical parameters in multiple spatio-temporal scales, and with improved accuracies and automation.



Seagrasses from above - drones and satellites

Example images from Lesbos, Greece. 39°09'30.6"N 26°32'01.8"E

Sources: Topouzelis, K. University of Aegean (2018); Digital Globe (2018); PlanetScope (2018); Copernicus Sentinel data (2018); Lansdat-8 (2018) U.S. Geological Survey.

Fig. 2.8. Effect of varying spatial resolution between different satellite and drone sensors above a seagrass bed. WVII PAN and MUL are the pan-sharpened and multispectral imagery of WorldView-2. (taken from the Global Seagrass Report (in press).

2.2.3 Artificial Intelligence and Machine Learning

The foresaid large-scale endeavors would certainly be unable to be scaled up and achieved had it not been for the artificial intelligence. Artificial intelligence (AI) had theoretically emerged in the middle 1950s, but its first practical applications, usability and value were highlighted only fifty years later, around the end of the 2000s and the start of the 2010s. AI has grown into the catch-all term to include Machine and Deep Learning, and Artificial General Intelligence (potential future generalized version of AI systems) algorithms and frameworks. Here I focus more on machine learning, as I have mainly explored such pathways in classifying seagrasses in the present PhD. Machine learning algorithms are split into Supervised and Unsupervised Learning ones which vary in the inclusion or exclusion of human input. The Supervised Learning Machine Learning is further split into classification (which output classes e.g., seagrass extent) and regression problems (which predict a continuous variable e.g., seagrass leaf area index).

In the context of seagrass, scientists can solve the common classification problem of seagrass extent (the output class) with the use of satellite imagery by feeding supervised machine learning frameworks with training data—a human-designed labelled set which the framework exploits to learn potentially relevant patterns in the data and estimate the seagrass extent. The machine learning algorithm will then learn characteristics indicative of both seagrass and non-seagrass presence and will interpolate its learnings by classifying all the pixels of the implemented satellite imagery given the human input.

The most commonly employed machine learning algorithms in the remote sensing of seagrass have been Linear Regression, Random Forests (Breiman, 2001) and Support Vector Machines (Vapnik, 1995), mainly powered by their availability in specialized geospatial softwares and their straightforward parameterization and implementation. These algorithms are characterized by several pros and cons e.g., sensitivity to noise of the human-designed label dataset, computational speed and complexity, required number of input features, scalability. Also, they are characterized of varying degrees of bias (i.e. noise); bias can be introduced into such models and their products, for example, through poor spatial representation of an area which has sparse training labels for several reasons. To mitigate noise and their impacts in the context of seagrass classification, the human analyst should select and design suitable sets of training data by ensuring adequate representation in terms of the habitat scale (i.e. seagrass beds of varying density) and depth gradient (i.e. shallow to deep seagrasses).

In this PhD, I have mainly exploited the machine learning frameworks of Random Forests and Support Vector Machines. More information about the algorithms, their herein parameterization, their accuracies and their results can be found in the respective sections of Materials and Methods, and Results of Chapter 4. Albeit a yet non-existent deep learning framework and application for seagrass detection and mapping, in Chapter 5, I identify and articulate three main areas in which deep learning could bring about breakthrough innovations in data-driven approaches to seagrass mapping; these are primarily related to the foreseen improvements in automation, accuracies and streamlining of seagrass and more broadly seascape intelligence.

3 State of the Art

This section is not an extended aide-memoire about the history of the remote sensing of seagrasses. This section is about the technological advances, the scientific approaches and efforts, and the resulting publications and projects of the last 20 years; and how these have increased our capacity around and have aided sophisticated developments in the RS of seagrasses. The efforts and publications by Dekker et al. (2006), Hossain et al., (2014), and Purkis and Roelfsema (2015) serve a very good hors d'oeuvre to the most recent history of seagrass RS (using both optical and acoustic methods) and this State-of-the-art chapter. I hope that my take on this two-decade history succeeds into motivating the necessity and novelty of the present PhD thesis and its related publications.

The state-of-the-art in seagrass RS here refers to the available sensors (mainly satellites), and image processing, analysis and classification techniques to develop understanding and retrieve meaningful information about seagrasses. Section 2.2 has already provided an introduction to the technology and related advances, including satellite sensors, artificial intelligence, and cloud computing, regarding the domain of seagrass remote sensing. Building on this section for the same domain, there have been various developments and applications concerning the various image analysis, processing, and correction and classification methods to apply to the aforementioned satellite data.

3.1 Remote Sensing Approaches for Seagrass Mapping

Remote sensing technology can sense seagrass meadows which occupy the optically shallow seabed where part of the surface reflectance contains a bottom signal. Optically shallow areas comprise one of the most challenging environments for remote sensing due to the numerous interferences of parameters with the sunlight before this is detected and measured by a spaceborne sensor. To invert bottom and water column properties, and then to obtain quantitative information on optically shallow beds, one has to correct the interferences of the atmosphere, air-water interface and water column. A number of researchers have developed a range of empirical, semi-analytical and analytical solutions to retrieve properties of interest and ultimately untangle the initially weak bottom signal (Brando et al. 2009; Hedley, Roelfsema and Phinn, 2009; Klonowski, Fearns and Lynch, 2007; Lee et al. 1998, 1999; Lyzenga, 1978, 1981; Maritorena et al. 1994; Mobley et al. 2005). All these solutions have been the basis for understanding, studying, and assessing the coastal aquatic seabed using remote sensing from space and in situ in the last 40 years. A good review of the majority of the water column correction algorithms sits in Zoffoli et al. (2014) and a thorough intercomparison between empirical, analytical and hybrid approaches is described in Dekker et al. (2011). More elaboration on the Lyzenga (1978, 1981), Maritorena et al. (1994), and Lee et al. (1998, 1999) solutions can be found in the Materials and Methods sections of the present Thesis' chapter 4. Additional context and information on the practical implementation of these algorithms (especially the Lyzenga's approach) can be read in Green et al. (2010).

Various researchers have exploited remote sensing to map seagrasses in different spatio-temporal scales. I have selected to refer to only the last twenty years of optical remote sensing of seagrasses due to the launch of the first commercial high-resolution (4 m) multispectral satellite in 1999, IKONOS. IKONOS was a game-changer mainly in scientific endeavors in the Earth Observation domain and enabled respective high-resolution mapping approaches in the seagrass domain, as well. Mumby and Edwards (2002) used Lyzenga (1981) solution to map 13 shallow (<20 m) water environments (including the seagrasses of *Thalassia tesudinum* and *Syringodium filiforme*) with very high accuracies in the seagrass class (89% user accuracy) in ~60 km^2 of the Caribbean Sea. This paper implemented the Maximum Likelihood Classification (MLC) supervised classification framework to classify the 13 classes. Utilizing ten IKONOS images (Fig. 3.1), Andréfouët et al. (2003) also utilized Lyzenga (1981) empirical solutions and a combination of unsupervised and MLC classifier to map various seagrass species (e.g., *Thalassodendron* sp., *Thalassia* sp., *Halophila* sp., and *Halodule sp.*) up to 30 m of depth in ten coral-dominated tropical sites. The overall accuracies of their varying multi-class assessments using IKONOS were 15-20% improved in comparison to Landsat-7-derived ones.

3.2 Remote Sensing of the Mediterranean Seagrass Bioregion

Focusing on the Mediterranean, Fornes et al. (2006) mapped the extent of the therein dominant *P. oceanica* seagrass with the use of IKONOS and the MLC framework up to depths of 15 m in Mallorca Island, Balearic Islands, Spain. They achieved a producer and user accuracy of 92.9% and 91%, respectively, but in a small area of 1.31 km^2. Other remote sensing studies in the Mediterranean seagrass bioregion feature, chronologically from the oldest to the most recent one, those of: Pasqualini et al. (2005), which mapped *P. oceanica* beds up to 20 m in the Marine National Park of Zakynthos, Greece, with overall accuracies between 73 and 96%, implementing two SPOT 5 satellite images (2.5 and 10-m spatial resolution); Borfecchia et al. (2013), which charted the same seagrass species up to 16 m of depth, using high-resolution, multispectral Quickbird data in the coastal Lazio region of the middle Tyrrhenian Sea, Italy with an overall accuracy of 84%; and Matta et al. (2014), whose seagrass mapping was between a depth range of 0–10 m in the Gulf of Oristano, W Sardegna, Italy, with an overall accuracy of 84%, using RapidEye imagery at 5-m spatial resolution. All aforementioned mapping efforts addressed small-scale seagrass beds, both spatially and temporally (i.e. single-date mapping approaches).

3.3 Change Detection of Seagrasses

Analysis of multi-date spaceborne remote sensing data allows retrospective quantitative assessment of seagrass meadows (Dekker et al., 2006). Applied to the coastal environment, spaceborne image archives allow multi-temporal analysis and change detection of submerged ecosystems which could in turn permit the identification of possible degradation rates and boost conservation efforts of these problematic areas (Purkis and Roelfsema, 2015). Researchers have previously employed spaceborne time series to map seasonal to decadal change detection of seagrasses (Fig. 3.2) (Dekker et al., 2005; Knudby et al., 2010; Lyons et al., 2012; Pu et al., 2014; Roelfsema et al., 2014). The time-series analysis of seagrass communities is as accurate as the classification algorithms in use (Palandro et al., 2003). In the Mediterranean, the spatio-temporal dynamics and sources of observed variations of seagrass habitats, namely the intertidal *Zostera noltii* and *C. nodosa* species, have been assessed before through spaceborne and airborne time series (Barillé et al., 2010; Garrido et al., 2013).

3.4 Machine Learning-based Classification of Seagrasses

All the aforesaid methodological frameworks and applications implemented simple classification algorithms e.g., Maximum Likelihood, Nearest Neighbours. In the last four years (until 2019), machine learning algorithms (e.g., Random Forests, Support Vector Machines, k-nearest neighbors) have overruled the simple old and traditional classification algorithms in the remote sensing literature (Gislason et al., 2006;

Mountrakis et al., 2011). However, machine learning has been sparsely implemented in the quantitative assessment of coastal environments (Fig. 3.3) (Zhang, 2015).

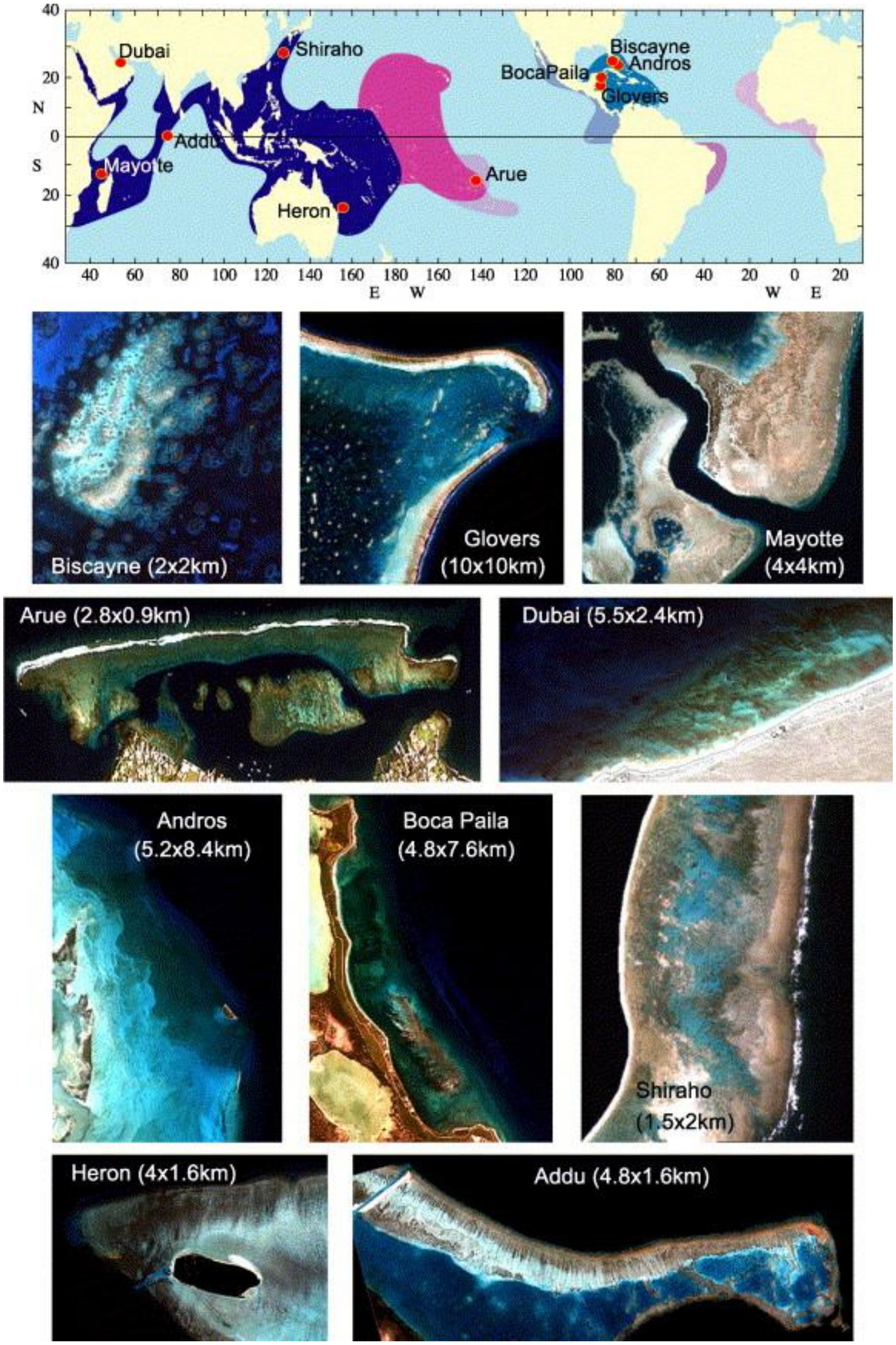


Fig. 3.1. Examples of IKONOS RGB high-resolution satellite imagery over seagrass and coral- dominated seabed areas in the world's tropics (from Andréfouët et al. (2003)).

3.5 Remote Sensing of Seagrass' Biophysical Parameters

There is an additional plethora of worth-mentioning research efforts and papers in the last two decades which increased the momentum and capacity in seagrass mapping (of various biophysical parameters) using remote sensing data. Dierssen et al. (2003) leveraged high-resolution airborne hyperspectral imagery from PHILLS and field measurements to quantitatively estimate leaf area index ((LAI: ratio of leaf area to substrate area) of *Thalassia testudinum* seagrass beds in Lee Stocking Island, Bahamas (Fig. 3.4). They achieved r^2 between 0.88 – 0.98. Knudby and Nordlund (2011) utilized IKONOS satellite imagery to map seagrass distribution and biomass in a 4.1-km^2 inter and sub-tidal multi-seagrass (seven species) habitat environment in Zanzibar, Tanzania (Fig. 3.5). Despite the optically complex environment, they achieved a high overall accuracy of 77.7% in habitat mapping, and r^2 of 0.88 and 0.91, respectively, for subtidal and intertidal seagrass species. In an exemplary endeavour for sophisticated seagrass mapping, Hedley et al. (2016) developed a physics-based inversion method to map LAI of *Thalassia testudinum* meadows (Fig. 3.6). Sensitivity analysis and application of the method to hyperspectral airborne imagery of the Portable Remote Imaging Spectrometer (PRISM) in Florida Bay, USA demonstrated good agreement with field data. Additionally, the method lacked empirical calibration data, while the paper provided important per-pixel uncertainty estimates which are an often missing component in the confidence and validation of seagrass and neighbouring habitats.

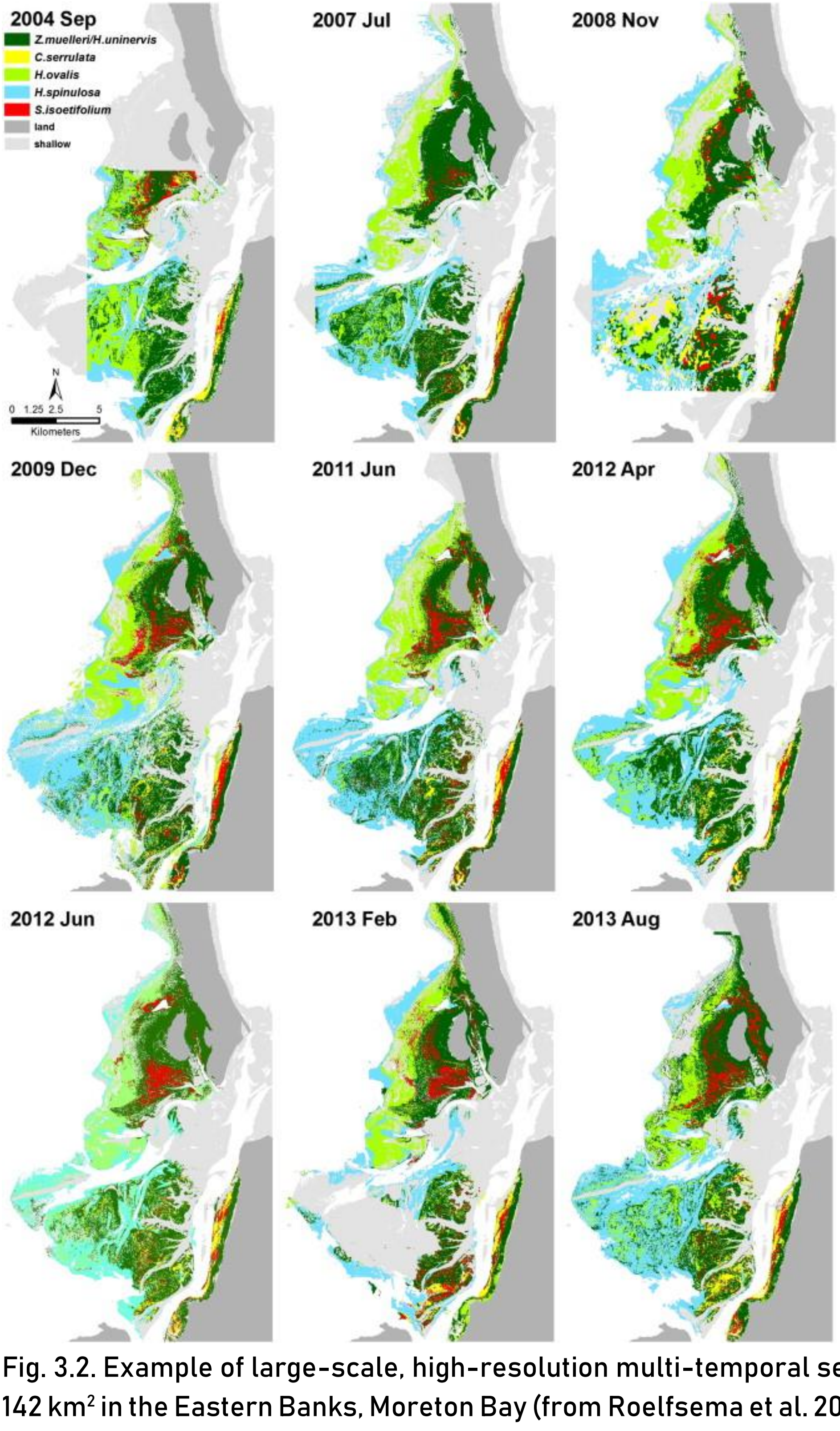


Fig. 3.2. Example of large-scale, high-resolution multi-temporal seagrass mapping over 142 km^2 in the Eastern Banks, Moreton Bay (from Roelfsema et al. 2014).

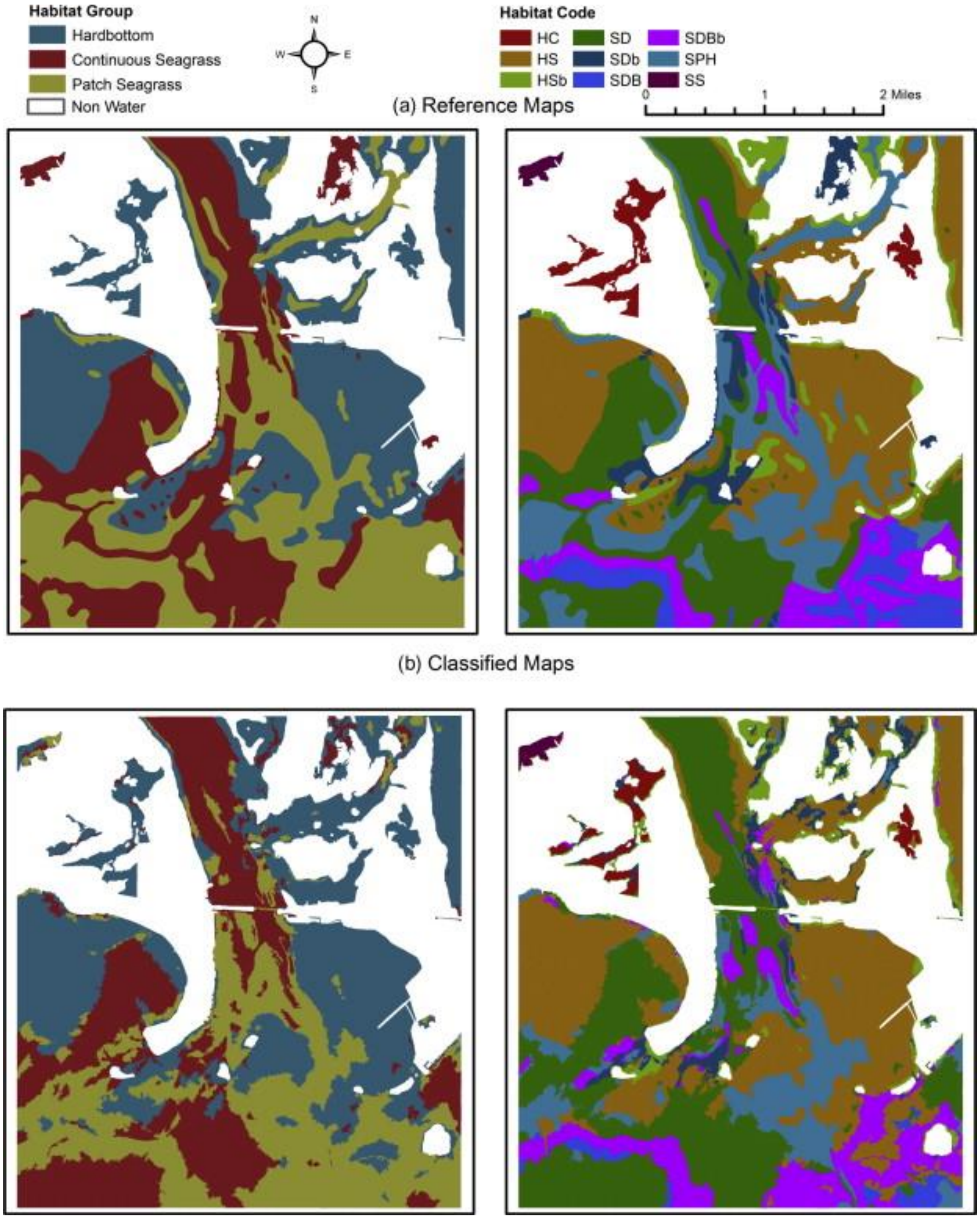


Fig. 3.3. Reference and classified habitat maps including seagrass habitats from the fusion of hyperspectral, aerial and bathymetry data and ensemble analysis of Random Forests, Support Vector Machines and k-nearest neighbors ((from Zhang, 2015).

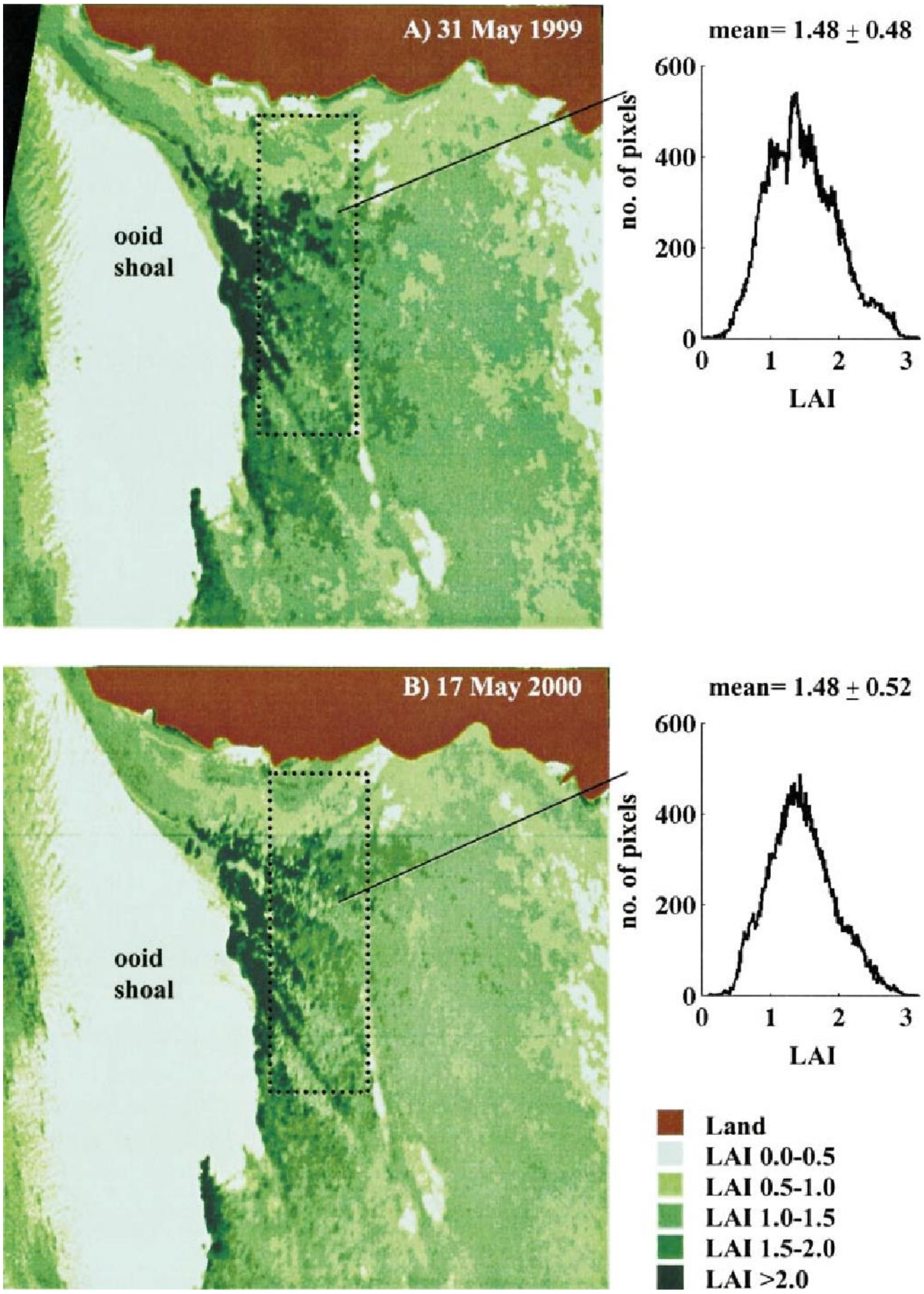


Fig. 3.4. Example of leaf area index (LAI: ratio of leaf area to substrate area) mapping using two high-resolution PHILLS imagery at Lee Stocking Island, Bahamas. Histograms feature the distribution of LAI frequency within the outlined box of the respective left panels (from Dierssen et al. 2003).

3.6 Mind the Gaps

In addition to the novelties at the time of their publication, the aforementioned state-of-the-art seagrass mapping approaches and papers feature several commonalities and gaps for today's state of advances and requirements in remote sensing:

a) The restricted seagrass mapped areas i.e lack of large-scale seagrass mapping efforts
b) The lack of methodological developments and applications using new spaceborne remote sensing datasets (e.g., Sentinel-2, Planet)
c) The lack of scalability and adaptability
d) The infancy of seagrass-related artificial intelligence and machine learning-based mapping
e) The lack of democratization in seagrass mapping and algorithms
f) The relatively low availability of open seagrass data

As I highlighted in the Introduction of this Thesis (section 1.1), today, the existing regression of and threats to seagrass habitats requires global-scale, data-driven measurements in a repeatable fashion for their science, management, protection and conservation. The exponential emergences in satellite remote sensing allow the development of Big Data-related methodologies and algorithms which could enable, in turn, scalable and adaptable mapping and monitoring of seagrasses worldwide. Scientifically, the present PhD thesis has been inspired by the pre-existing literature of remote sensing of seagrasses, but also built on their five main, aforementioned gaps.

As such, the following state-of-the-art contributions of this Thesis to seagrass mapping present:

a) Exploitation and applications of the new satellite remote sensing datasets of Sentinel-2, and Planet's PlanetScope and RapidEye
b) A highly adaptable and scalable cloud-based framework for multi-temporal seagrass mapping and monitoring
c) Introduction and application of machine learning frameworks
d) The advance in the democratization of seagrass mapping

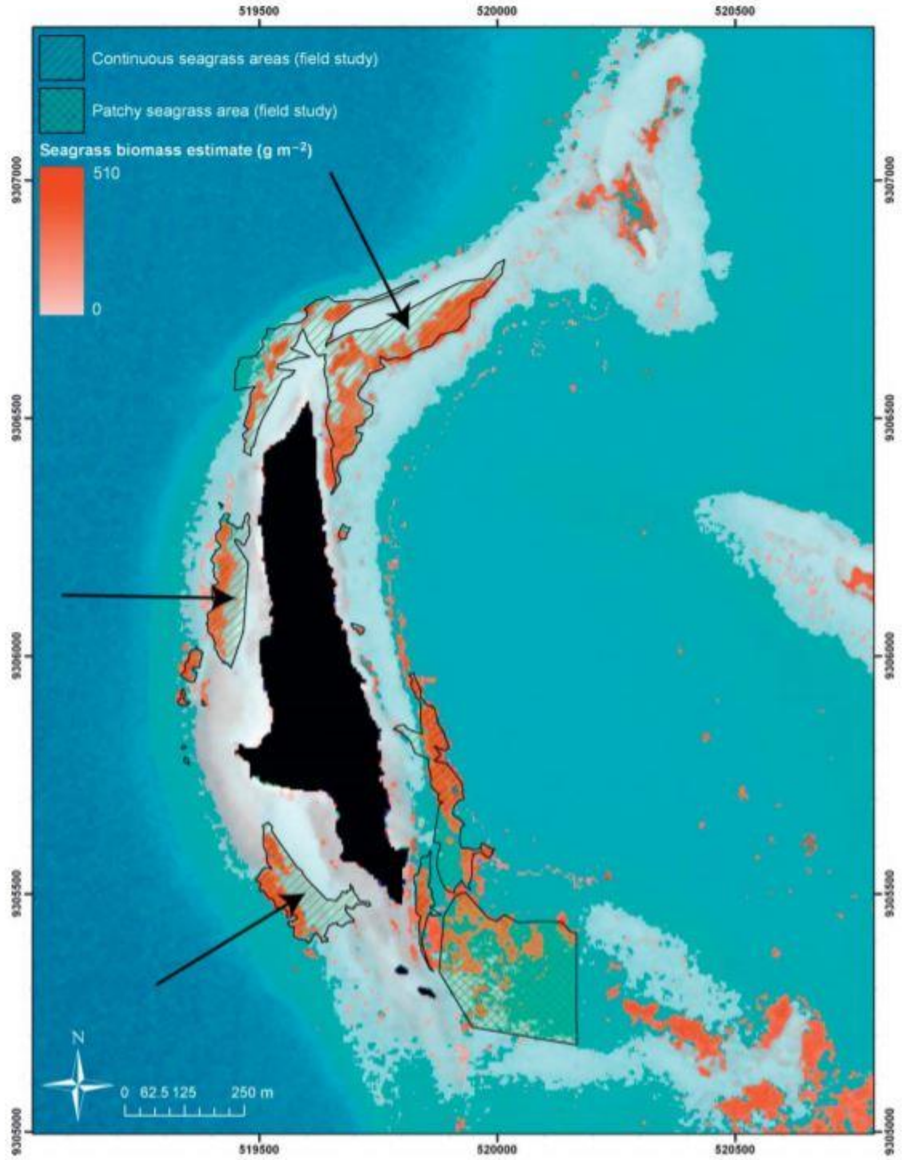


Fig. 3.5. Example of seagrass biomass mapping using IKONOS satellite imagery and field data in Zanzibar, Tanzania (from Knudby and Nordlund, 2011).

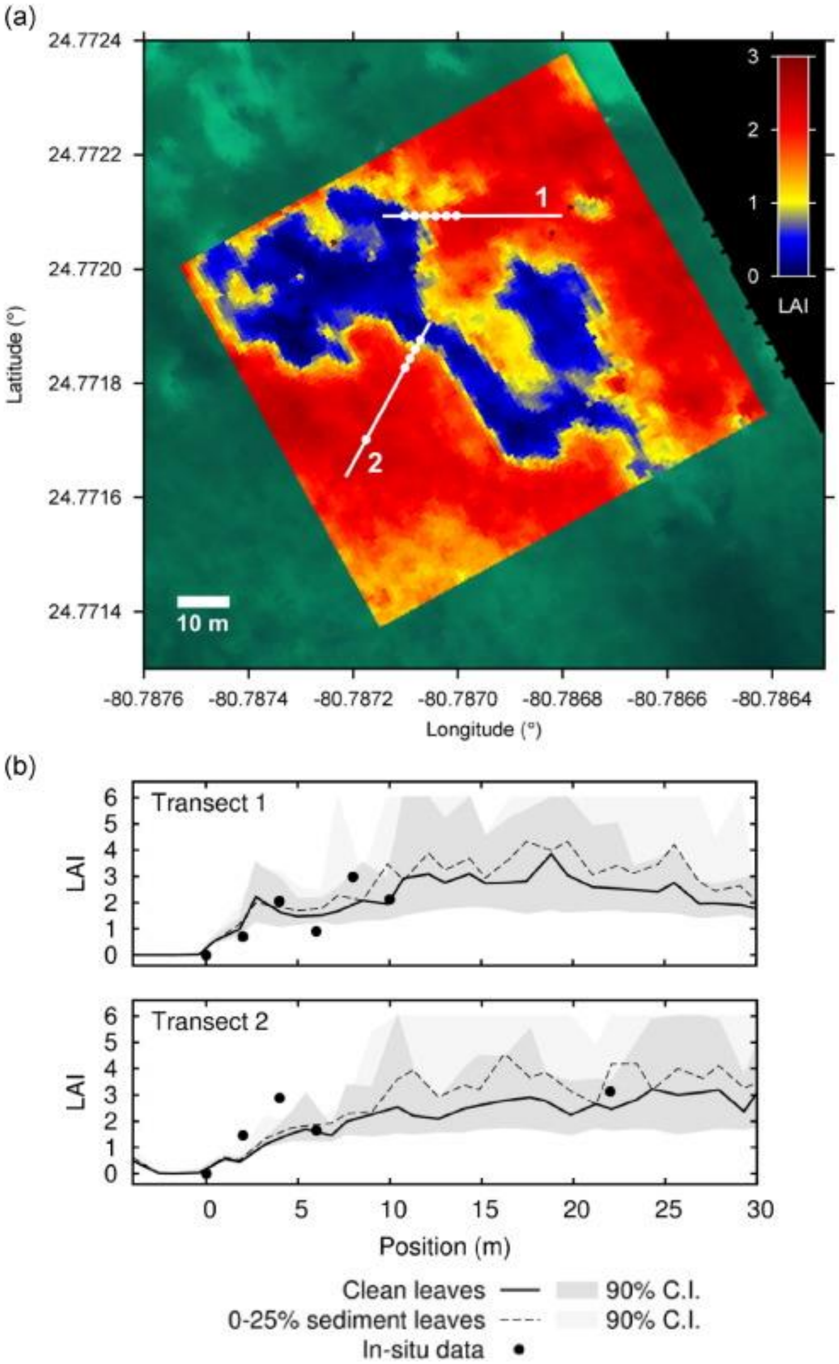


Fig. 3.6. Example of leaf area index (LAI: ratio of leaf area to substrate area) mapping using a physics-based inversion model, PRISM hyperspectral imagery and field data in Florida Bay, USA (from Hedley et al. 2016). The transects of (b) represent the white lines of (a) while the black dots are quadrat-based in-situ data, lines are model inversion results for only clean leaves and 0-25% coverage by sediment. Grey areas represent the 90% confidence intervals based on uncertainty propagation.

4 Contributions

I herewith provide the six main published contributions of my PhD research as sub-chapters of Chapter 4. These sub-chapters relate to the remote sensing of seagrass ecosystems, exploiting different spatio-temporal scales, satellite data input, machine learning classifiers in local servers and serverless environments. They are namely: **4.1** - Mapping Mediterranean Seagrasses with Sentinel-2 Imagery (Traganos and Reinartz, 2018a); **4.2** - Cubesat-derived Detection of Seagrasses using Planet Imagery following Unmixing-based Denoising: Is Small the Next Big? (Traganos et al. 2017); **4.3** - On the use of Sentinel-2 for Coastal Habitat Mapping and Satellite-derived Bathymetry Estimation using Downscaled Coastal Aerosol Band (Poursanidis et al. 2019); **4.4** -Machine Learning-based Retrieval of Benthic Reflectance and *Posidonia oceanica* Seagrass Extent using a Semi-Analytical Inversion of Sentinel-2 Satellite Data (Traganos and Reinartz, 2018b); **4.5** - Interannual Change Detection of Mediterranean Seagrasses using RapidEye Image Time Series (Traganos and Reinartz, 2018c); and **4.6** - Towards Global-Scale Seagrass Mapping and Monitoring using Sentinel-2 on Google Earth Engine: The case study of the Aegean and Ionian Seas (Traganos et al. 2018a).

4.1 Mapping Mediterranean Seagrasses with Sentinel-2 Imagery

4.1.1 Abstract

Mediterranean seagrasses have been hailed for their numerous ecosystem services, yet they are undergoing a decline in their coverage. The major complication with resolving this tendency is the sparsity of data on their overall distribution. This study addresses the suitability of the recently launched Sentinel-2 satellite for mapping the distribution of Mediterranean seagrass meadows. A comprehensive methodology is presented which applies atmospheric and analytical water column corrections and compares the performance of three different supervised classifiers. Remote sensing of the Thermaikos Gulf, northwestern Aegean Sea (Greece, eastern Mediterranean Sea) reveals that the utilization of Support Vector Machines on water column corrected reflectances yields best accuracies. Two Mediterranean seagrasses, *Posidonia oceanica* and *Cymodocea nodosa*, cover a total submerged area of 1.48 km^2 between depths of 1.4–16.5 m. With its 10-m spatial resolution and 5-day revisit frequency, Sentinel-2 imagery can mitigate the Mediterranean seagrass distribution data gap and allow better management and conservation in the future in a retrospective, time- and cost-effective fashion

4.1.2 Materials and Methods

Study site

The study site is a ~3.3 km^2 area located within the SE part of the Thermaikos Gulf, NW Aegean Sea, Greece (Fig. 4.1.1; eastern Mediterranean Sea). With a coastline that exceeds 350 km in length and 52,300 km^2 in area, Thermaikos is the largest gulf of the Aegean Sea and its coastal system is set within the humid mesothermal climatic zone and a principally tideless marine environment. The climate of the Thermaikos Gulf coastal system is defined as a semi-arid Mediterranean type with cold winters, mean annual air temperatures between 9 °C and 17.5 °C, an annual precipitation between 400 mm and 1300 mm and mean annual wind speeds between 2.5 and 15 m/s (Poulos et al., 2000). Within the inner shelf of the Thermaikos Gulf, Thermaikos Bay, surface water temperatures range between winter water values of 9 °C and summer values of 25 °C, while surface salinities vary between 28 psu (winter) and 35 psu (summer). The Thermaikos Gulf is considered as a microtidal marine environment where the major semi-diurnal constituents M2 and S2 seldom exceed 10 cm and 7 cm, respectively, while the major diurnal constituents M1 and

S1 display extremely small amplitudes of 2.7 cm and 1.7 cm, respectively. In addition, current meter observations exhibited speeds in the order of 5–20 cm/s near the water surface and a maximum of 9 cm/s near the bottom (Robles et al., 1983). In general, an anti-clockwise, thermohaline water circulation, the prevailing northerly winter winds in addition to the dominant south/southwesterly summer winds, and the mixing of different water masses regulate the hydrodynamic regime of the Thermaikos Gulf, restricting water constituents concentrations to a minimum in the eastern continental shelf of the

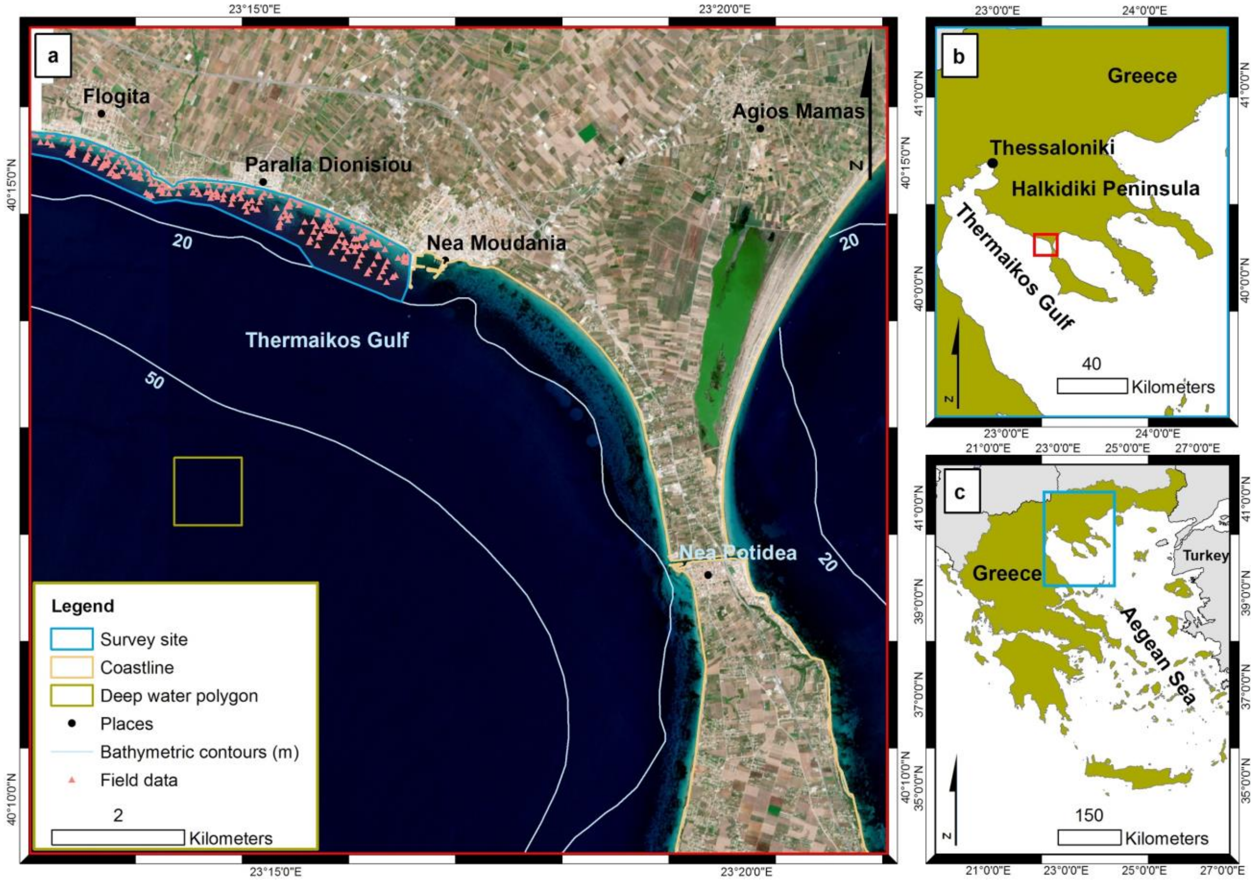


Fig. 4.1.1. Location of survey site within (a) Thermaikos Gulf, (b) Halkidiki Peninsula, (c) Aegean Sea, Greece. The depicted Sentinel-2A satellite image in (a) is an atmospherically corrected, true color (band 2 as blue, band 3 as green, band 4 as red) composite in UTM (zone 34) system/WGS84 projection. The outline colors of the left and upper right frame indicate their exact location. The deep water polygon is a selected 115 × 115 pixel window used in the water column correction step as it has very little water leaving radiance values in all visible bands.

gulf. In fact, this part of the shelf shows particulate matter concentrations (PMC) of < 1 mg/l, surface to bottom (Zervakis et al., 2005). The inner continental shelf which includes our survey site is also shallower than 40 m and features slopes of between 2% and 5%, while the outer continental shelf is deeper than 80 m. Finally, the prevailing sediments in the inner continental shelf are sand of > 50% content, silt of < 40% content and clay content of < 20% (Poulos et al., 2000; Robles et al., 1983; Tsimplis et al., 1995).

The area under investigation is situated in the wider vicinity of the Municipality of Nea Propontida, where several socioeconomic activities including agriculture, aquaculture, industry, tourism, fishing, and trade directly affect its coastal system. Nea Propontida has a population of 36,500 (National Statistics Service of Greece; census 2011) and the region of Central Macedonia of which it is part, exhibits the second largest quantity of fishing catch in the whole Greek fishing area with 11,869 t, hence 18.5% of the total Greek fishing catch of 64,236 t (National Statistics Service of Greece; census 2015). In total, 19.3% of the population of this neighboring region to our study area is employed in the primary sector (National Statistics Service of Greece; census 2011), while there are 91 hotels and 66 factories operating in the entire coastal region.

Data and methodology

Satellite data

The satellite imagery used in this study was acquired by Sentinel2A, the first satellite of the twin polar-orbiting Sentinel-2 satellites. Sentinel-2A was launched on 23 June 2015 while Sentinel-2B was launched on 07 March 2017. Both are part of a European fleet of satellites aiming to deliver core data to the European Commission's Copernicus program, a program whose services address six thematic areas: land, marine, atmosphere, climate change, emergency management and security. The Sentinel-2 satellites carry a single optical instrument payload, the MultiSpectral Imager (MSI). In a sun-synchronous orbit and at a mean altitude of 786 km above the Earth's surface, the MSI samples 13 spectral bands in the visible-near infrared (VNIR) and short wave infrared (SWIR) spectral range at 3 different spatial resolutions (10, 20, 60 m) and allows for a 290-km swath width with a high revisit frequency of 10 days (the combined constellation revisit frequency will be 5 days). Sun-synchronous orbits ensure that satellites overfly any given point on Earth at the same local solar time (ESA, 2015).

In this study, we processed and analyzed a Sentinel-2A image sensed on 10 June 2016 at 09.09 am UTC (Table 4.1.1). The imagery was downloaded from the Sentinels Scientific Data Hub (https://scihub.copernicus.eu/) which provides unlimited open access to the

Sentinel-2 Level-1C (L1C) user products. L1C products, also called granules or tiles, consist of 100 km^2 orthorectified and spatially registered images in Universal Transverse Mercator (UTM)/World Geodetic System 1984 (WGS84) projection. We chose the tile identified as S2A_OPER_MSI_L1C_TL_MTI__20160614- T160614_A005052_T34TFK_N02.02 (Table 4.1.1) for further analysis as it contains our study site (Fig. 4.1.1).

Field data

Generally, satellite remote sensing requires the presence of accurate field data in order to improve any imagery interpretation and feature extraction, aid supervised classifications and determine a subsequent accuracy assessment. Coastal areas are challenging environments for remote sensing assessment due to various parameters which can create spectral confusion i.e. depth gradient, numerous habitats, water quality. To diminish this effect, we adapted a stratified random sampling that ensures adequate representation of all physical habitats in their different depth range (Congalton, 1991).

Table 4.1.1
Sentinel-2A image characteristics

Tile ID	S2A_OPER_MSI_L1C_TL_MTI__20160614T160614_A005052_T34TFK_N02.02
Acquisition date	10 June 2016
Acquisition time	09:09 am UTC
Solar zenith angle	22.4°
Solar azimuth angle	134.5°
Sensor incidence angle	5°
Sensor view azimuth angle	51°
Sensor view azimuth angle	786 km

Field data were collected in our study site during a boat-based survey between 10 and 13 July 2016. The temporal difference of one month between the acquisition of the satellite image and the collection of the field data is acceptable because Mediterranean seagrass meadows undergo changes in larger time scales and in addition tidal changes between the two dates are negligible. Snorkeling and free diving (when needed in larger depths, where the bottom was not visible from the water surface) were chosen as the optimum sampling method. In total, 228 points (Fig. 4.1.1) were visited, photographed from directly above the seabed (Fig. 4.1.2) and related to specific habitat and coordinates, while 53 points were associated with specific depths measured by the Garmin Fishfinder 160C mounted on the bottom of the keel of the boat. A negative number was pre-entered to offset the distance between the bottom of the keel and the water surface. The camera used for the photographs was a waterproof GoPro HERO4 Session that captures up to 8-megapixel images. Coordinates were taken from the Android application GPS Waypoint Finder 1.0 operating on a Moto G (3rd Generation) Smartphone. The coordinates were noted by a surveyor as near as possible to the water surface, above the seabed point observed by the snorkeler, to guarantee greater accuracy, which was determined 4.2 m at all measurements. All the measurements were exported as GPX files and imported into the ArcGIS 10.3.1 platform using the GPX To Features (Conversion) tool. Four different habitats were identified in the survey site: *Posidonia oceanica* seagrass, *Cymodocea nodosa* seagrass, photophilous algae on rocks and sand (Fig. 4.1.2).

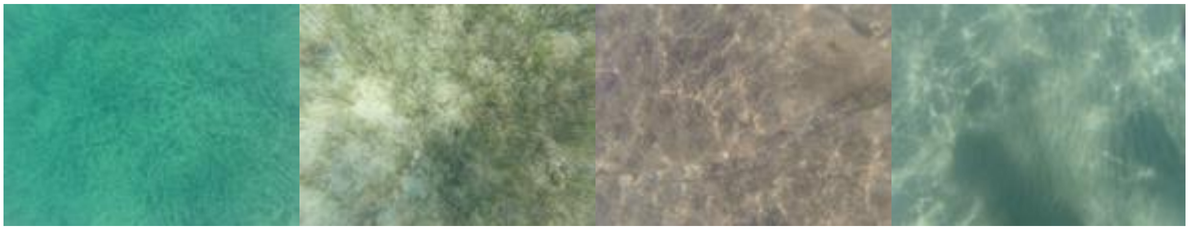


Fig. 4.1.2. Habitats exhibited in the survey site: (from left to right) Dense *Posidonia oceanica* meadows; Shallow *Cymodocea nodosa* habitat; Photophilous algae on rocks; Sand.

Auxiliary data

Bathymetry data were also used as an auxiliary dataset to aid bathymetry estimation and subsequent validation. These data are offered by the Hellenic Navy Hydrographic Service as a Digital Terrain Model (DTM) of the whole extent of the Greek Seas at a resolution. We used ArcGIS Spatial Analyst to create a contour dataset from the aforementioned DTM

Methodology

A schematic representation of the adapted methodology herein is shown in Fig. 4.1.3. To derive quantitative information about submerged habitats from remotely sensed images, one has to address the interference of the atmosphere and water column on the satellite image. A passive remote sensing instrument like Sentinel-2A MSI measures sunlight that has passed previously through the atmosphere, the water column, has reached the seabed and then has been reflected back to cross again the water column and atmosphere and be received by the satellite sensor. Therefore, the first steps of our methodology deal with the corrections of the interferences that determine the fate of light in the atmosphere and water before deriving any quantitative information on the aquatic habitats with focus on seagrasses. We used only the 10-m Sentinel-2A bands 2, 3, and 4 for the analysis (with the exception of the 10-m NIR band 8 to mask out land), as they are all within the visible spectrum (approximately 390 to 700 nm) and most importantly penetrate the water column deeper and provide sensible quantitative data on bottom reflectances.

Atmospheric & BRDF correction – radiometric calibration

Sentinel-2 L1C data are top-of-the-atmosphere (TOA) products. ATCOR 9.1.0 software (Richter, 2007) was implemented to correct the influence of the atmosphere, viewing and solar illumination geometry on the L1C data. The influence of the latter geometries on the imagery is described by the bidirectional reflectance distribution function (BRDF). ATCOR incorporates a sensor-specific atmospheric database of look-up tables (LUTs), which contain the results of radiative transfer calculations. ACTOR2 for flat terrain was used to process Sentinel-2A data, because it accounts for the angular variation of view and solar geometry over the image. During the atmospheric correction step, visibility was set to 45.6 km by ATCOR to avoid negative reflectance pixels. Aerosol type was chosen to the rural setting because experiments with a maritime type of aerosol resulted to negative reflectance pixels. An ozone column of 368 DU (Dobson units) was further applied. Finally, the range of adjacency effect was set to 1 km to avoid any interference in the coastline boundary between the brighter reflectances of land and darker reflectances of adjacent waters. The resulting bottom-of-the-atmosphere (BOA) water surface reflectances (hereafter denoted as Rw) are expressed in Integers, thus we divided them

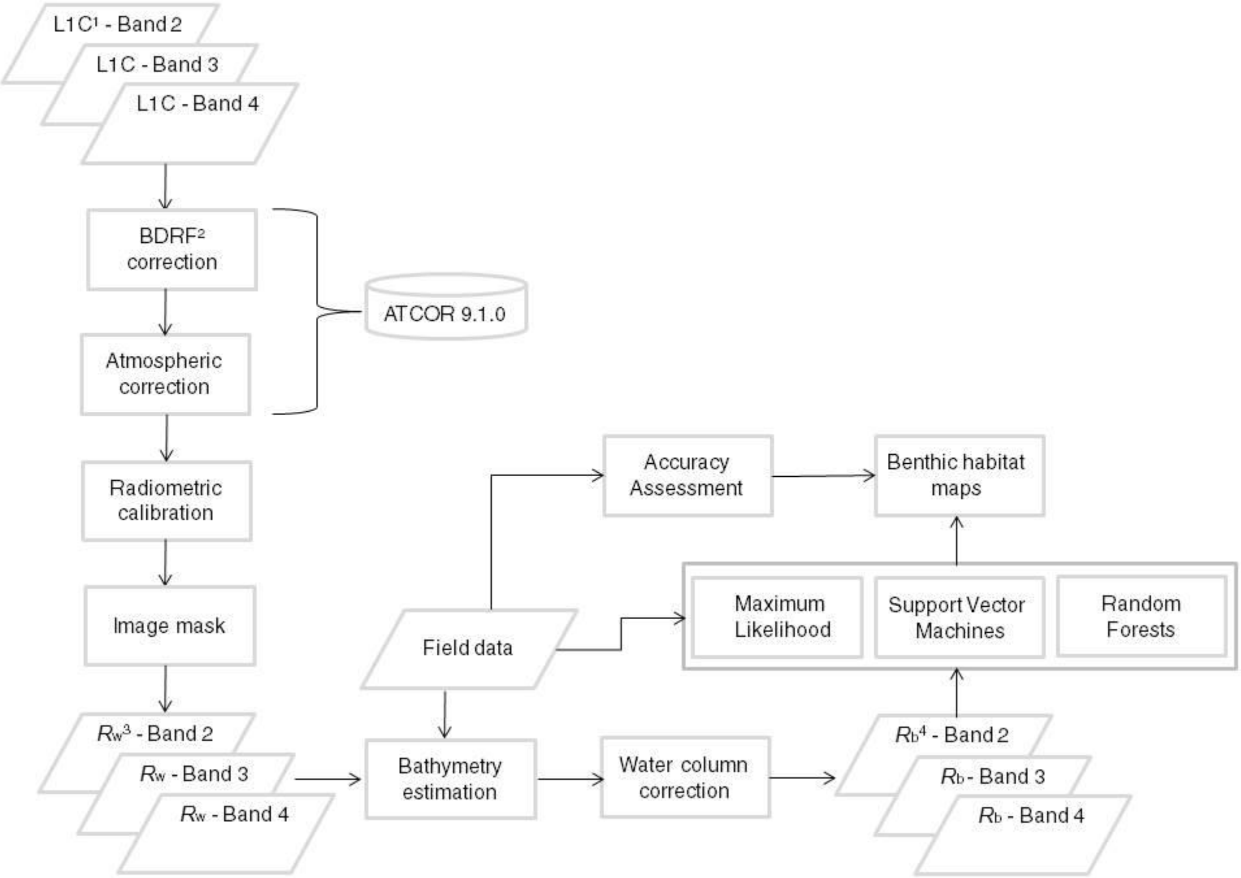


Fig. 4.1.3. Schematic presentation of the methodology. L1C[1] products are the initial orthorectified and spatially registered Sentinel-2A granules in UTM/WGS1984 projection, BDRF[2] is the bidirectional reflectance distribution function that defines how light is reflected at an opaque surface, R_w[3] represents water surface reflectances and R_b[4] are bottom reflectances.

by a special quantification value included in the Sentinel-2A metadata (10000) to derive Rw values in the range 0–1.

Image mask

Image mask is a crucial step in coastal remote sensing due to the enhancement of coastal water features by masking all terrestrial features. The atmospherically-corrected near-infrared (NIR) band 8 and the bathymetric contour of 20 m (Fig. 4.1.1a) were used to clip the initial granule to the coastal study area, masking out land and deeper waters than 20 m where it is difficult to derive reliable quantitative estimations of the seafloor. NIR Band 8 is implemented due to the great water absorption in this wavelength. We applied an Iso Cluster Unsupervised Classification to classify the Sentinel-2A granule into two classes containing land and water. Subsequently, we extracted our masked study area using the water class.

Water column correction

Water column imposes an additional interference on the remotely sensed measurement of submerged habitats and thus its effect requires correction. We utilized the analytical model of Maritorena et al. (1994) for optically shallow waters (where part of the surface reflectance is composed of a bottom signal) which is based on:

$$R_{\mathrm{w}} = R_{\infty} + (R_{\mathrm{b}} - R_{\infty})e^{[-2K_{\mathrm{d}}Z]} \tag{1}$$

Assuming a vertical and horizontal homogeneity in addition to high water transparency, which is the case in the Thermaikos survey site, the model equals the atmospherically corrected surface reflectance, R_w, to the reflectance of an infinitely deep water column, R_∞, plus substrate contrast (bottom reflectance, $R_b - R_\infty$) after correction of the water depth effect (the term $e^{[-2K_d Z]}$), where Kd is the light attenuation coefficient in the water and Z is the water depth. Maritorena et al.'s algorithm, therefore, can be inverted to derive bottom reflectances, R_b from measurements of R_w, R_∞, Z in each pixel of the satellite image and K_d. The following chapters explain in detail the approaches that we followed to estimate bathymetry, the infinitely deep water column reflectance, and the attenuation coefficient.

Bathymetry estimation and validation

Knowledge of bathymetry of a coastal area can be used to eliminate reflectance changes attributed to variable depth and water column attenuation. An additional problem that needs to be tackled by coastal remote sensing researchers is the reflectance differences between substrates. Seagrasses like *P. oceanica* tend to form dense meadows that appear very dark on satellite images, exhibiting smaller reflectance than deep water. In contrast, a sandy substrate appears very bright because of its higher light reflectance. As a result, dense *P. oceanica* meadows will appear deeper than sand at the same true depth. To overcome these difficulties and to measure bathymetry, researchers have proposed the implementation of band ratios based on the assumption that the ratio of the reflectances in these bands will remain constant regardless of the substrate in a single scene (Dierssen et al., 2003; Lyzenga, 1978; Stumpf et al., 2003). As such, we developed a site-specific algorithm, modifying the band ratio of Stumpf et al. (2003), to map bathymetry using Sentinel-2A band 2 and band 3 (at 490 and 560 nm respectively). These two bands can penetrate deeper into the water column than band 4 at 665 nm, and were identified as constant for all bottom types (Fig. 4.1.4). The reflectance ratio was furthermore filtered with a median 5 × 5 cell to reduce the resulting noise from the ratio calculation. Median filtering ranks neighboring pixels (in this case a 5 × 5 neighborhood) according to their brightness (intensity) and the median value becomes the new value from the central pixel. Median filtering is preferred over mean filtering because it is less sensitive to extreme values (outliers) and therefore it removes them more efficiently. The log-transformed and passed with a median filter reflectance ratio of blue and green bands was plotted against in situ measured bathymetry (*Z*) to develop a second-order polynomial

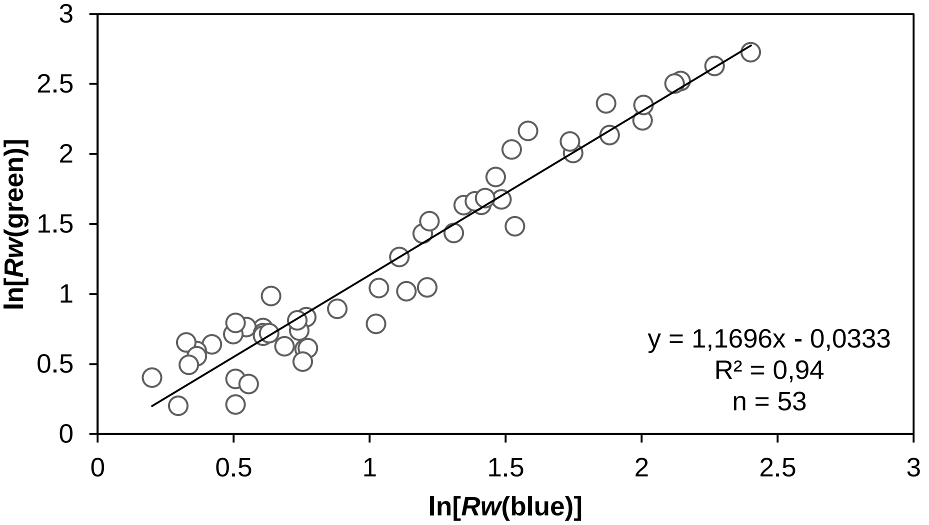


Fig. 4.1.4. Regression between log-transformed water-surface Sentinel-2A blue (band 2) versus Sentinel-2A green (band 3) reflectances over various habitats. An offset is applied to both logtransformed reflectances to facilitate presentation.

$$Z = 178.22x^2 - 428.78x + 259.17 \tag{2}$$

$$x = \frac{\ln[R_W(490)]}{\ln[R_W(560)]} \tag{3}$$

which explained > 96% of the variation ($p < 0.001$) in measured bathymetry in 15 points (Fig. 4.1.5) that covered the entire depth range and presence of habitats (seagrasses *Posidonia* *oceanica* and *Cymodocea nodosa*, sand and photophilous algae on rocks). It is worth mentioning that the deepest point on Fig. 4.1.5 (0.9012, 20.0) was obtained by taking the median value (0.9012) of 41 of the 107 pixels of the scene that fell onto the 20-m contour (Fig. 4.1.1; the remaining 66 pixels were used in the validation of the depth measurements). The initial in situ depth measurements spanned a depth range between 0 and 12 m and thus the 20- m image pixel value ensured that the calibration of the reflectance ratio with the surveyed depth measurement will result in sensible values over the whole observed depth range. Subsequently, an Eq. (2) was used to create a bathymetry map for every pixel of the atmospherically corrected scene.

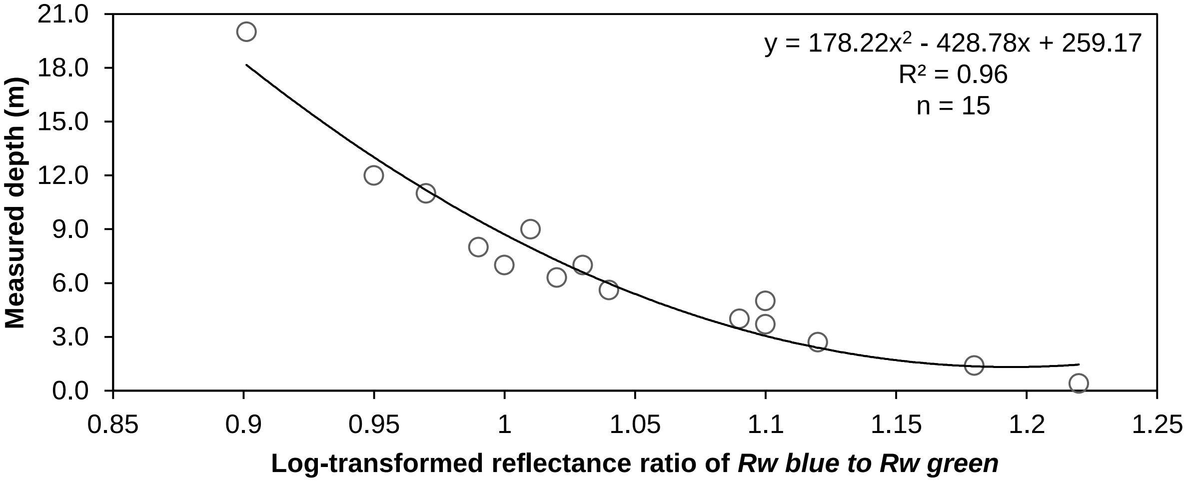


Fig. 4.1.5. Polynomial regression between log-transformed and passed with a 5 × 5 median filter water surface blue to green reflectance ratio and in situ depth measurements from the Thermaikos survey site. The displayed polynomial equation was used to measure bathymetry in the study site.

Infinitely deep water reflectance estimation

Water column correction following Maritorena et al. (1994) requires estimation of the parameter R∞, namely the reflectance measured over an infinitely deep water column. Here, we followed an image-based calculation of R_∞ from water surface reflectances, R_w by extracting the values of an 115 × 115 pixel polygon (indicated with a gold frame in Fig. 4.1.1) in waters deeper than 50 m where the reflectances are not affected by bottom signal (optically deep area). To avoid outliers, we calculated the median value of all 13,287 R_w pixel values for all 3 bands for subsequent use in the water column correction Eq. (1).

Diffuse attenuation coefficient estimation

The attenuation coefficient K_d represents the diffuse attenuation coefficient for both downwelling and upwelling light in the water column. Water attenuation coefficient is spectral dependent and increases with longer wavelengths, immersing in sensor noise in deep waters (with decreasing surface reflectance R_w). This fact further commends the use of Sentinel-2A bands 2, 3 and 4 which achieve reasonable water penetration. Based on Bierwirth et al. (1993), we calculated water attenuation coefficients for Sentinel-2A bands 2, 3 and 4 implementing the spatially co-registered modeled bathymetry and satellite image in a region of the survey site where the substrate and water quality remain constant, as the water correction model of Maritorena et al. (1994) dictates, yet depth is variable. These conditions were identified in an area on the SE of our survey site, where sand extents from 1 to 5 m of depth.

Supervised classifications

An important step in the majority of remote sensing studies of terrestrial or aquatic environments is to identify and categorize distinctive cover types of the studied environment on a satellite image into sensible classes, namely the classification procedure. Conventionally, image classifications are split into supervised and unsupervised, where the supervision factor is provided by field data associated with all the different classes apparent within the extent of the satellite image and taken during fieldwork synchronized to the overpass of the satellite sensor. In coastal remote sensing research, supervised classification is performed ideally on bottom reflectances which have already been corrected for the double pass of the signal through two mediums (atmosphere and water) before it reaches the sensor. In the coastal aquatic environment, characteristic substrate types are submerged aquatic vegetation, corals, algae and other cover classes. Generally, the outcome of an image classification is to derive some sort of quantitative measurement of a certain ecological value, which could be applied in other areas besides our study area.

In this study, we applied three different supervised classifiers, Maximum Likelihood, Support Vector Machines and Random Forests in three methodological considerations: a) before applying any atmospheric correction (L1C reflectances; Fig. 4.1.6A), b) after applying atmospheric correction (Rw, water surface reflectances; Fig. 4.1.6B), and c) after applying atmospheric and water column corrections (Rb, bottom reflectances; Fig. 4.1.6C). As a result, 9 different habitat maps were created. The aim was to quantitatively indicate the need for the correction steps prior to the application of any classification on the remotely sensed coastal image. Furthermore, we compared the three supervised classifiers, which are based on different principles and assumptions in order to be more accurate in our methodology than using a single classifier. The outcome of the image classifications here is a four-class habitat map (*Posidonia oceanica* and *Cymodocea nodosa* seagrasses, sand and photophilous algae on rocks) that delineates the distribution of each of these four classes. In the following sections, we provide a short introduction on all classifiers for optimum clarification.

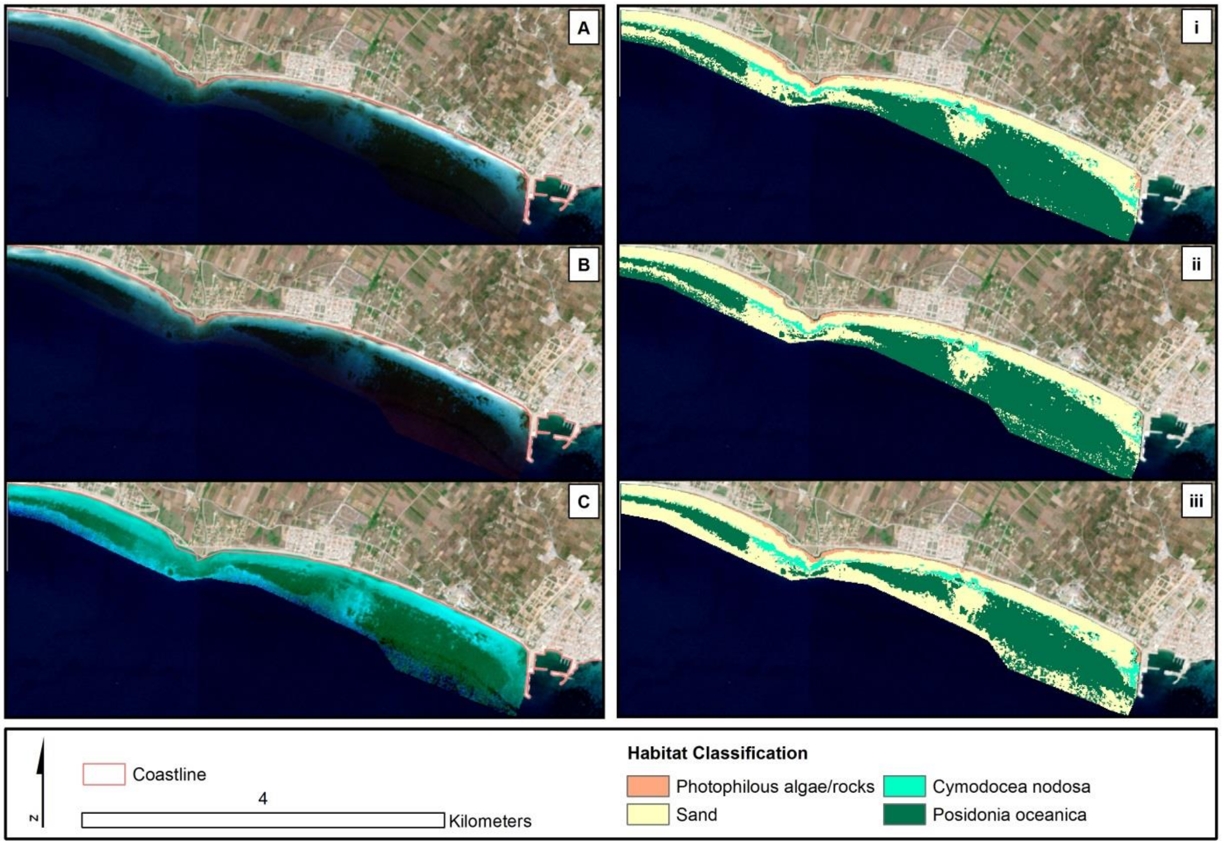


Fig. 4.1.6. Sentinel-2A RGB (10-m bands 2, 3, 4) composites of the various methodological steps and corresponding best performing classifiers in the Thermaikos Gulf study site. A. Masked L1C-composite (no correction applied), B. R_w, water surface reflectance composite (with atmospherically and geometric correction applied), C. R_b, bottom reflectance composite (with water column correction applied), i. Support Vector Machine (SVM) classification of A-composite, ii. Random Forest classification of B-composite, iii. SVM classification of C-composite. All 6 images are draped over atmospherically corrected Sentinel-2A RGB imagery in UTM (zone 34) system/WGS84 projection.

Maximum likelihood

The maximum likelihood classifier (MLC) was chosen for this study due to its simple adaptation and wide use in remote sensing studies of shallow aquatic habitats. MLC assumes a normal distribution of pixels in each class sample and follows the Bayesian decision rule. This decision rule takes into account both covariance/variance of data and the probability of each pixel of a scene belonging to each habitat class and outputs the distance between a candidate pixel and the mean of all habitat classes. This probability is considered to be equal for the selected classes a priori, an assumption that we also followed in our application of MLC for our 4 different habitat classes. Of the initial 228 visited sites in situ (Fig. 4.1.1), 85 sites were used to train MLC and the other two supervised classifiers altogether, while the other 143 sites were used to assess the accuracy of the supervised classification results. The 0.6 ratio of training to accuracy data follows the recommendations of Congalton (1991) and Green et al. (2000).

Support vector machines

The Support Vector Machines (SVM) represent a group of theoretically superior, supervised classification algorithms based on statistical learning theory (Vapnik, 1995). SVM are essentially machine learning algorithms, whose objective is to differentiate two classes by fitting an optimal separating hyperplane (decision boundary) to the training samples of the related classes. The samples closest to this hyperplane consist of the so-called support vectors. SVM have been applied efficiently to deal with both linearly and, more importantly, with nonlinearly separable classes applying a kernel function into a higher dimensional space, whose new data distribution allows better fitting of a linear hyperplane (Mountrakis et al., 2011). In this study, we applied SVM on our training data using the universal Gaussian radial basis function kernel (RBF)

$$K(x, x_i) = \exp(-g|x - x_i|^2 \qquad (4)$$

where parameter *g* defines the width of the RBF. Moreover, a regularization parameter, *C*, which regulates the degree of acceptable misclassification was set. A two-dimensional grid search in a predefined range was performed to find an adequate pair values for *g* and *C* for our image classification. The SVM classification experiments were run using the open-source LIBSVM library (Chang and Lin, 2013) in EnMAP-Box software (van der Linden et al., 2015).

Random forests

Random Forests (RF) is an ensemble supervised classification algorithm that combines multiple selflearning decision trees to solve classification problems. RF relies on the assumption that different independent tree predictors provide incorrect predictions in different areas. By combining the predictions in a "forest", RF can improve the efficiency of the model (Breiman, 2001). The RF has been vastly applied into the classification of remote sensing data due to their insensitivity to both noisy data and overtraining as well as their good performance with small datasets (Gislason et al., 2006). The user of RF needs to define two parameters; the number of decision trees and the number of randomly selected features (n_r) from all features (n_a). The latter were related in our application of RF with the function

$$n_r = \sqrt{n_a} \tag{5}$$

while 100 decision trees were found to provide adequate results in all 3 experiments (described in the second paragraph of section 2.3.4.). Moreover, the Gini Index was adopted for the measurement of the best split selection (Breiman, 2001). RF parameterization and run were also performed using the EnMAP-Box software.

Accuracy assessment

In remote sensing, the accuracy assessment step is crucial, because it tests the accuracy of image classification (i.e. the resulting habitat maps) and builds the link between image and reality. We created error matrices for the image classifications. As discussed in section 2.3.4.1., 143 out of the total 228 training sites were processed for the accuracy assessment. The error matrix is the foundation of the accuracy assessment. It contains a square array of rows and columns, where each row and column is one habitat class in the classification. Each cell in the error matrix represents the number of classified training samples, while the rows represent classified training data and the columns indicate reference data for the assessment of the classified data. In whole, overall, producer and user accuracy are the outputs of the error matrix (Congalton, 1991). The overall accuracy expresses the ratio of the number of correctly classified validation samples (diagonal values in the error matrix) to the entire number of validation samples regardless the class. Producer accuracy corresponds to the ratio of the number of validation samples in a given class that are correctly classified to the actual number of validation samples in this class. On the other hand, user accuracy comprises the number of the correctly identified validation samples in a given class, divided by the total number of validation samples claiming to be in this class. While producer accuracy is a great statistical metric for the remote sensing scientist creating the habitat map, user accuracy is more significant in a management context of a given region, because it reports a quantitative probability for the actual presence of a studied habitat in this region i.e. *Posidonia oceanica* meadows.

4.1.3 Results

Pre-classification steps

Differences between the various methodological approaches that we followed in this study are displayed in succession in Fig. 4.1.6. The masked, non-atmospherically corrected (L1C), the atmospherically-corrected (R_w) and the water-column corrected (R_b) reflectances in the Thermaikos Gulf survey site are depicted as A, B, C in Fig. 4.1.6. It is apparent that the L1C and R_b composites differ a lot in sharpness following an initial visual inspection. Seabed habitats are largely enhanced in Fig. 4.1.6C as a result of the atmospheric, BRDF and water column correction all of which increase the spectral variability.

Water column correction was performed using Eq. (1) which was inverted and solved for Rb for the blue, green and red bands of Sentinel-2A. The equation takes R_w, Z (bathymetry), R∞ (reflectance of an infinitely deep water column), in each pixel of the scene and K_d (diffuse attenuation coefficient). Using a site-specific polynomial algorithm (2) on the previously atmospherically corrected R_w, we created a bathymetry map of the study site (Fig. 4.1.7). The mean depth of the whole site is 8.3 m with a mean slope of 3.2°. The validation of the modeled water depth was performed using 39 in situ depth points (of the initial 53 points), plus the median value of 66 of the 107 pixels that fell onto the 20-m contour (Fig. 4.1.1), for a sum of 40 depth points (Fig. 4.1.8). It is worth noting here that 25 of these 40 depth points were measured over *P. oceanica* beds. These depth points allowed the estimation of an rsquared value of 0.92 with an RMSE (root-mean-square-error) of 1.3 m, where an even distribution of calibrated data pointed is exhibited (Fig. 4.1.8). In addition, residuals between modeled depth by the site-specific algorithm and surveyed depth were calculated but did not show any particular pattern of over/underestimation (Fig. 4.1.9). Furthermore, the optically deep water column reflectances were estimated using the median values of 13,287 R_w pixels of the ~115 × 115 gold-framed polygon in Fig. 4.1.1. These values were determined as 0.0172, 0.0122 and 0.0092 for R_w490,R_w560 and R_w665, respectively. The diffuse attenuation coefficient for the three Sentinel2A bands was calculated plotting the modeled depths versus the logarithms of surface reflectances (Fig. 4.1.10). The elongated data clouds for all 3 bands indicate that the identified submerged sandy region of our survey site is adequate for the calculation of attenuation coefficients. The latter were estimated using the slope of the three regression lines of Fig. 4.1.10 which represent the parameter $-2K_d(\lambda)$ (λ denominates each of the three used bands). Fig. 4.1.11 displays the spectral dependence and increasing pattern of attenuation coefficient K_d with R_w490, R_w560 and R_w665. Values of attenuation coefficient are in meters because they are depth-specific, as the reflectances are also unitless.

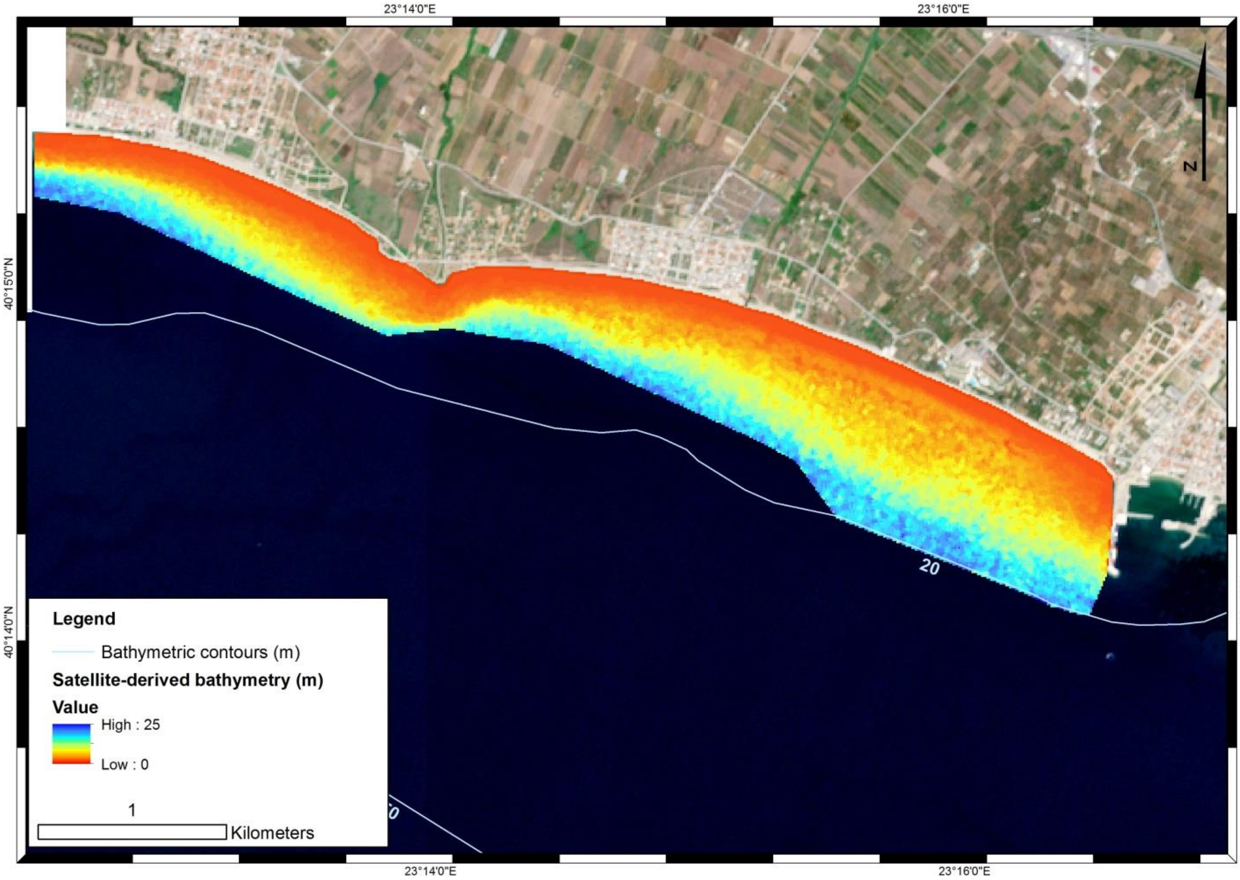


Fig. 4.1.7. Bathymetry map of the survey site draped over an atmospherically corrected S2A RGB composite. Projection is UTM (zone 34) system/WGS84.

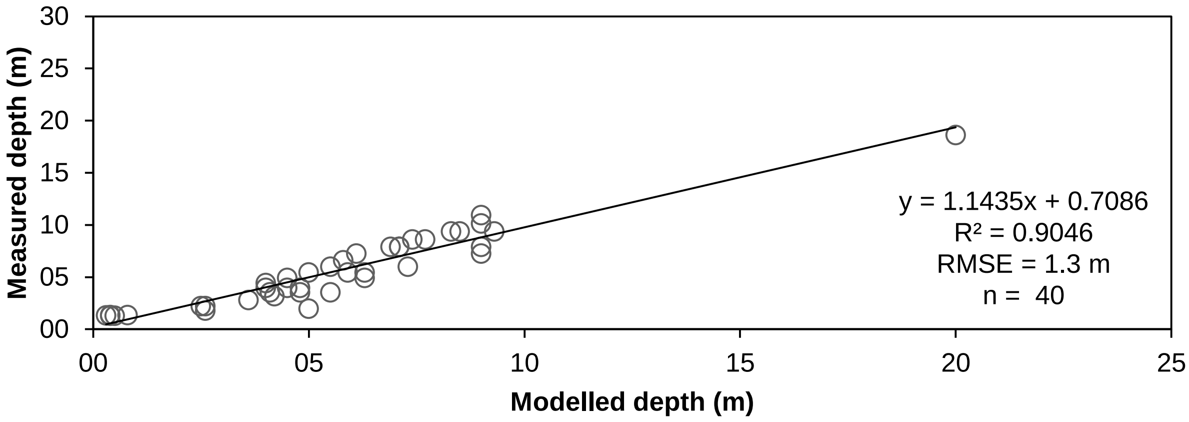


Fig. 4.1.8. Plot of modeled depth versus in situ measured depth for the validation of the bathymetry map of the survey site. Depth model was derived from the relationship of Fig. 4.1.5.

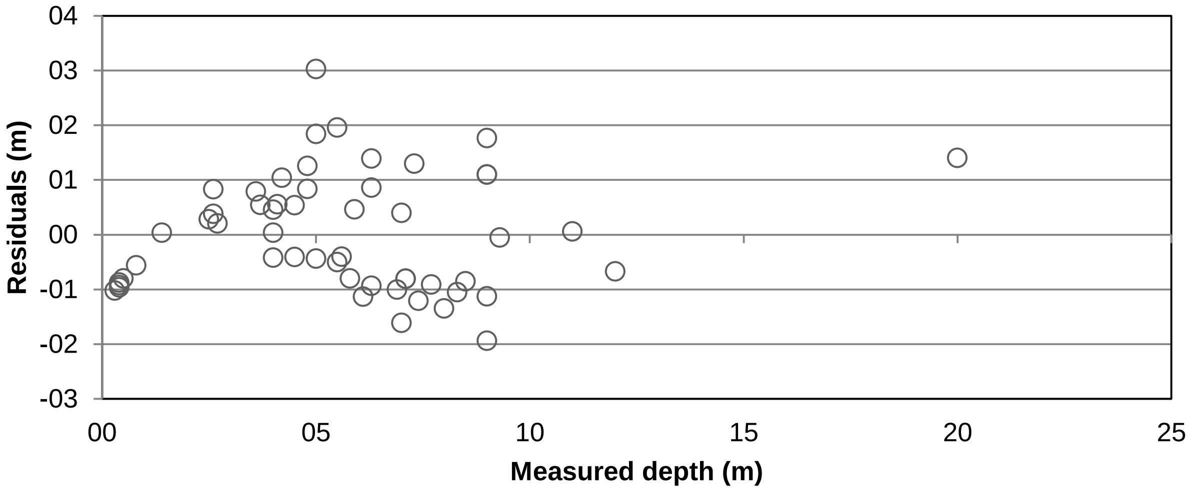


Fig. 4.1.9. Plot of the depth residuals from the regression of Fig. 4.1.7, between modeled depth and in situ measured depth. Note that RMSE is 1.3 m.

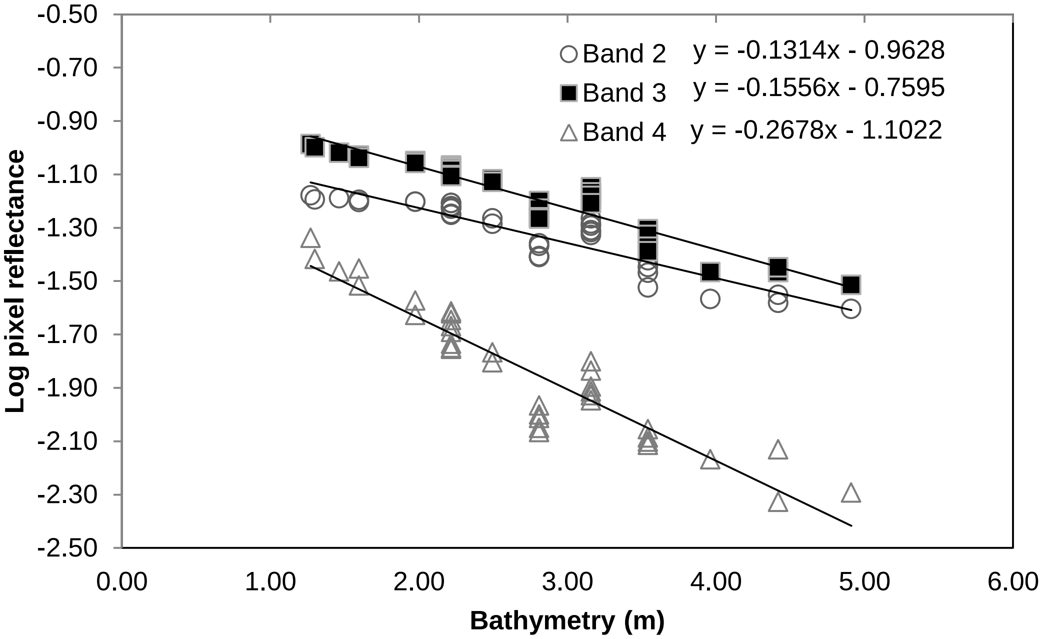


Fig. 4.1.10. Modeled bathymetry versus log-transformed water surface reflectances R_w490, R_w560 and R_w665. The surface reflectance pixels were extracted from a sub-area on the satellite image that exhibited a constant sandy substrate but a depth variance from 1 to 5 m. Attenuation coefficients were estimated using the slope of each regression line which represents the quantity $-2K_i$ (Bierwirth et al., 1993).

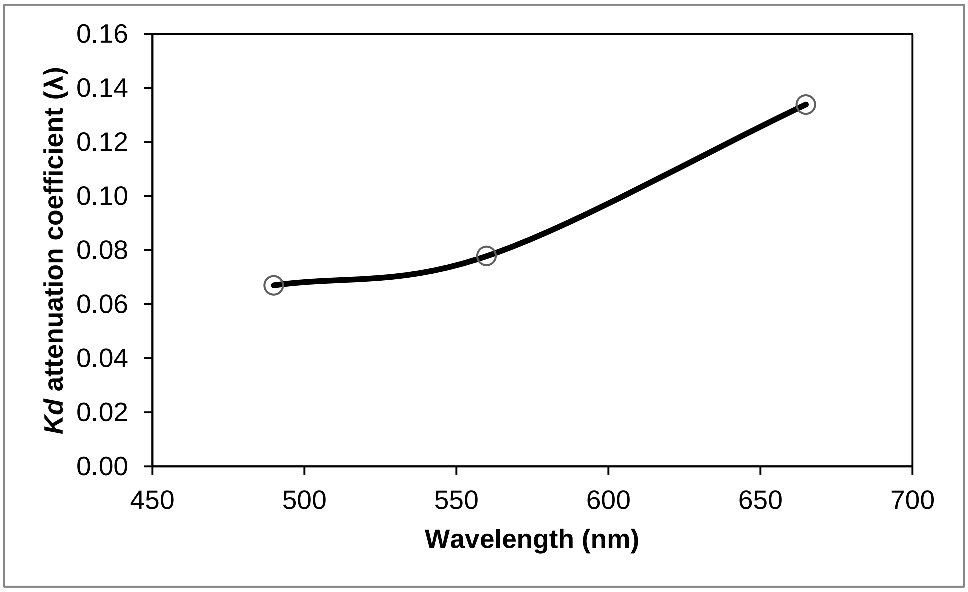


Fig. 4.1.11. Spectral dependence of diffuse attenuation coefficient for water surface reflectances R_w490, R_w560 and R_w665. Diffuse attenuation coefficients estimations were derived from the regressions of Fig. 4.1.9.

Supervised classifications

Three different supervised classifiers, Maximum Likelihood, Support Vector Machine and Random Forests, were applied to three different rasters for a total of 9 different, four-class habitat maps of the SE Thermaikos Gulf study site. The first classified raster was the S2A L1C composite (Fig. 4.1.6A), the second was the R_w composite (Fig. 4.1.6B) and the third was the Rb composite (Fig. 4.1.6C), the result of the water column correction. Therefore, quantitative comparisons were performed both in terms of classifier's performance and of the application of image corrections (i.e. atmospheric and/or water column). The accuracy assessment of the 9 habitat maps was conducted based on 143 in situ points and is expressed with the error matrix. Table 4.1.2 shows the error matrices for the best 3 classifiers for the three imagery composites, while the habitat maps resulting from the best classifier's performance for each classified imagery can be seen in the right column images of Fig. 4.1.6 (i, ii, iii). SVM classifier performed better than MLC and RF in classifying the non-atmospherically corrected data (Fig. 4.1.6A - i), yielding an overall accuracy of 86%. As seen from the error matrix, the two seagrass species under consideration here, *P. oceanica* and *C. nodosa* exhibited a mean producer accuracy and user accuracy of 75.8% and 85.6%, respectively. Errors in producer and user accuracies of both seagrass classes were solely attributed to confusion with the class of sand (Table 4.1.2). RF produced the best results classifier-wise in the classification of the atmospherically corrected reflectances (Table 4.1.2). Using 100 trees, RF increased the overall accuracy to 94.4%. Following atmospheric correction of the satellite imagery, mean producer accuracy of seagrass species increased to 87.5%, while user accuracy of their classes grew to 98.2% (Fig. 4.1.6B – ii). Seagrass habitats were again confused for sandy habitat, however atmospheric correction was more beneficial for the mapping of *C. nodosa* given the growth of 21.5% in its producer accuracy (Table 4.1.2; Fig. 4.1.6B - ii). After applying the water column correction, SVM classifier yielded an identical overall accuracy to the RF classification of the atmospherically-corrected image (94.4%) (Fig. 4.1.6C – iii). While the SVM classification of bottom reflectances performed better than the RF classifier on water surface reflectances as regards to producer accuracies of classified seagrass meadows (mean producer accuracy of 92.8% to 87.5%), it surprisingly exhibited smaller values as regards to mean user accuracies (87.4% to 98.2%). Nevertheless, on visually examining Fig. 4.1.6 and comparing classifications i and ii with iii, one can observe that in both i and ii classifications (with and without atmospheric correction), the deeper limit of *P. oceanica* lacks clear delineation in contrast with iii (with applied water column correction) where *P. oceanica* displays clearer boundaries in its seaward extent. On examining the classified habitat maps, both quantitatively

Table 4.1.2

Best error matrixes and statistical tests for accuracy assessment using different combination of datasets and classifiers.

Classes	**Rocky algae**	**Sand**	***Cymodocea nodosa***	***Posidonia oceanica***	**Total**	**User accuracy (%)**
Rocky algae	18	1	0	0	19	94.7
Sand	3	45	6	3	57	79
Cymodocea nodosa	0	2	8	0	10	80
Posidonia oceanica	0	5	0	52	57	91.2
Total	21	53	14	55	143	
Producer accuracy	85.7	84.9	57.1	94.6		
Kappa coefficient	0.94	0.78	0.78	0.86		
Overall accuracy: 86%, Overall kappa coefficient: 0.79						

Notes: Non-atmosperically corrected image (L1C) - Support Vector Machine classifier

Classes	**Rocky algae**	**Sand**	***Cymodocea nodosa***	***Posidonia oceanica***	**Total**	**User accuracy (%)**
Rocky algae	20	0	0	0	20	100
Sand	1	51	3	2	57	89.5
Cymodocea nodosa	0	0	11	0	11	100
Posidonia oceanica	0	2	0	53	55	96.4
Total	21	53	14	55	143	
Producer accuracy (%)	95.2	96.2	78.6	96.4		
Kappa coefficient	1.0	0.83	1.0	0.94		
Overall accuracy: 94.4%, Overall kappa coefficient: 0.92						

Notes: Atmosperically-corrrected water surface image (R_w) - Random Forest classifier (100 trees)

Classes	**Rocky algae**	**Sand**	***Cymodocea nodosa***	***Posidonia oceanica***	**Total**	**User accuracy (%)**
Rocky algae	20	0	0	0	20	100
Sand	0	48	2	0	50	96

Cymodocea nodosa	0	3	12	0	15	80
Posidonia oceanica	0	2	0	55	58	94.8
Total	21	53	14	55	143	
Producer accuracy (%)	95.2	90.6	85.7	100		
Kappa coefficient	1.0	0.94	0.78	0.92		
Overall accuracy: 94.4%, Overall kappa coefficient: 0.92						

Notes: Water-column corrected bottom image (R_b) - Support Vector Machine classifier

and qualitatively, we considered the SVM-derived habitat map of the bottom reflectances (Fig. 4.1.6C – iii) to be the one closest to reality from the total 9 habitat maps of the Thermaikos Gulf survey site.

Satellite-derived mapping of the SE Thermaikos Gulf reveals that *C. nodosa* covers an area of 0.17 km^2 at a depth range between 1.8 and 9.3 m with a mean depth presence of 3.8 m. Its beds lack a clear homogenous shape and usually precede *P. oceanica* meadows in terms of depth. On the other hand, *P. oceanica* meadows cover ~8 times larger area than *C. nodosa* (1.31 km^2) and span a 15-m bathymetry range (1.4–16.5 m) with a mean depth of 8.3 m. The species forms 3 clearly delineated, dense meadows with the largest in the southeasternmost part of the Thermaikos Gulf, extending to a length and width of 1.96 and 0.53 km, respectively (Fig. 4.1.6 - iii). Generally, the two seagrass species were distributed in the 45% of the total surveyed area of 3.3 km^2.

4.1.4 Discussion

Sentinel-2 application to Mediterranean seagrass meadows

The main aim of this study was to test the efficiency of the multispectral Sentinel-2A mission in quantitatively assessing seagrass meadows and, more broadly, coastal habitats. Here, two of the most dominant seagrass species in the Mediterranean Sea, *P. oceanica* and *C. nodosa*, were detected and mapped with an accuracy of > 90% up to depths of 16.5 m in highly transparent waters, following atmospheric and water column correction on Sentinel-2A imagery. These corrections are mandatory as > 80% of the total signal received by a satellite sensor over shallow aquatic habitats are attributed to atmospheric effects (Mishra et al., 2005). Additionally, the exhibited overall accuracy of submerged habitat maps on the corrected Sentinel-2A image was 94%.

The Sentinel-2 mission was conceived to support various thematic areas including water monitoring with open data access policy. These data will provide continuity to the archive data of Landsat and SPOT heritage missions. With its twin-satellite capability in full function, Sentinel-2 will allow high spatial and temporal resolution coverage over all coastal waters up to 20 km away from the shore, all islands > 100 km^2, all EU islands, the whole Mediterranean Sea and all inland waters between latitudes 56° S and 83° N (ESA, 2015). In regard to seagrass habitats in the Mediterranean Sea, Sentinel-2 mission enables fine-scale mapping and monitoring of seasonal, annual, inter-annual and decadal changes of these compelling ecosystems from a small coast to the whole extent of the Mediterranean Sea. Moreover, its 12-bit radiometric resolution allows for a potential number of 4096 brightness levels which is beneficial for the fine differentiation between the seaward limit of the dominant *P. oceanica* and an optically deep column. Generally, knowledge of the retrospective distribution of Mediterranean seagrass is crucial towards understanding the thresholds that govern seagrass distribution and drive these changes. Ultimately, this will allow more efficient restoration, conservation and management of seagrasses in the vastly biodiverse Mediterranean waters.

Supervised classifications

Three different supervised classifiers were compared here in terms of their classification performance of Mediterranean underwater habitats and specifically seagrasses. In addition, classifier's performance was assessed through non-atmospherically corrected, atmospherically corrected and water-column corrected images. The Support Vector Machine classifier exhibited better accuracies than Random Forests and Maximum Likelihood classifiers with the exception of the atmospherically-corrected image, where RF was superior. MLC displayed the lowest accuracies in all 9 classifications, which is attributed to its assumption of normal distribution of classes, a rare occurrence, in reality, for sample classes. The machine-learning RF and SVM methods exhibited same overall accuracies on classifying the atmospherically corrected image and the water-column corrected image, respectively. However, we observe that the RF classifier overestimates *P. oceanica* seaward limit due to spectral confusion with optically deeper waters prior to the application of water column correction. In 8 out of 9 classification results, *C. nodosa* beds displayed lower accuracies than *P. oceanica* beds. This is possibly attributed to the lower number of field points of the former (23) than the latter (88). Additionally, *C. nodosa* species were studied in sparse beds mixed with sand. This is further exhibited in the error matrices where *C. nodosa* class was mainly confused with sand class (Table 4.1.2). Therefore, the 10-m pixel basis (spatial resolution of Sentinel-2A bands 2, 3 and 4), on which all classifications were performed, is expected to contain both classes, resulting in the socalled 'mixed' pixels. Object-based classifications, linear unmixing models or higher resolution imagery can potentially provide better classifications of *C. nodosa* beds. In contrast, the 10-m pixel basis is adequate for highly accurate mapping of the distribution of dense canopy *P. oceanica* meadows. Generally, all three classifiers yielded satisfying overall accuracies of > 83% in all 9 classifications.

Error matrices of Table 4.1.2 demonstrate the spectral discrimination between *P. oceanica* and *C. nodosa* beds in the optically clear waters of the Thermaikos Gulf. This is due to their large reflectance difference. *P. oceanica* exhibits often much lower reflectance than an optically deep column and naturally lower than the sparse *C. nodosa* beds. Following accurate pre-processing, processing and analysis of its data, we expect Sentinel-2 to map seagrass distribution in more challenging and multispecies submerged environments than the one unveiled in the Thermaikos Gulf. Nevertheless, we suggest that the spatial organization of seagrass species comprises the greatest challenge in classifying them. In more complex submerged areas, in order to overcome the intrinsic spatial heterogeneity of seagrasses, multispectral-based or ideally hyperspectral-based approaches should focus on distinct pigment-related spectral features of each seagrass and other submerged aquatic vegetation within the wavelengths that concern water column corrections (500–680 nm) (Fyfe and Dekker, 2001; Dekker et al., 2006). Fyfe and Dekker (2001) observed the clear separation between 3 southeastern Australian seagrass species (Posidonia australis, Halophila ovalis, Zostera capricorni) following measurements of pigment-related spectral features with a field spectrometer. Laboratory measurements of the spectral reflectances of three Western Atlantic subtropical/tropical seagrasses (Thalassia testudinum, Halodule wrightii and Syringodium filiforme) revealed similar curve shapes but different peak heights between the three seagrasses (Thorhaug et al., 2015). In satellite-derived approaches, however, the challenges added to the spatial heterogeneity of seagrasses in accurately classifying multi-species environments are numerous including the interactions at the air-water interface, optical components of the overlying water column, variable depth, noise ratio, radiometric and spectral resolution. A characteristic example of the complexities that a multi-species environment imposes to remote sensing approaches is shown in Chumbe Island, Zanzibar, Tanzania where seven different seagrass species form mixed-species meadows of varying densities (Knudby and Nordlund, 2011).

Pre-classification steps

Bathymetry estimation

Bathymetry measurements coincident in spatial scale to the processed satellite image are integral to correct water column attenuation and derive bottom reflectances. A simple method to estimate bathymetry over an optically shallow region is the band ratio (Lyzenga, 1978). Mapping depth over shallow waters that contain seagrasses can be problematic, as these beds have usually lower reflectances than neighboring optically deep waters probably due to the shading which occurs within the canopy (Dekker et al., 2006). Stumpf et al. (2003) developed a ratio algorithm for relatively clear waters to deal with the reflectance issue that requires small tuning with available in situ depth data. In this study, after running several experiments, we decided to modify Stumpf's algorithm to derive a bathymetry map for our study site as a result of negative values over *P. oceanica* beds, when using the initial algorithm. Although Stumpf et al. (2003) argued the effectiveness of the ratio algorithm for variable bottom type and albedo differences, they tested and tuned it in shallow-water regions that had higher reflectance than deep water, unlike our study site which hosts extensive dense *P. oceanica* meadows that are less reflective than deep water. Our modified bathymetry algorithm implemented the ratio between blue reflectance at 490 nm and green reflectance at 560 nm. Half of the depth points used for tuning the bathymetry algorithm were over *P. oceanica* beds that exhibit higher reflectance in the green than in the blue wavelength. The magnitude of this reflectance difference was greater than the magnitude difference in water column attenuation between the two wavelengths with increasing depth, resulting in a polynomial relationship between the reflectance ratio and in situ depth. Generally, our study lacked a larger availability of depth points, especially in the optically deeper regions to which *P. oceanica* extents. Nevertheless, bathymetric estimations provided adequate results to model water column effects.

It is also worthwhile to discuss the influence of the outlying data points in Figs. 5 and 8 on resulting regression statistics. These data points are the median values of 41 points (used in the estimation of the site-specific bathymetric algorithm; Fig. 4.1.5) and 66 points (used in the validation of the bathymetric algorithm; Fig. 4.1.8), respectively. They were extracted from the spatial concurrence of the 20-m contour with the pixels of the atmospherically corrected Sentinel-2A scene (Fig. 4.1.1). Although they are outliers, we decided to use them to ensure that the satellite-derived bathymetric map will have accurate values over the whole depth range of seagrasses. The initial field depth measurements covered a depth range between 0 and 12 m while *Posidonia oceanica* seagrass was mapped up to 16.5 m. Exclusion of the outlier in Fig. 4.1.8 would decrease r-squared value from 0.92 to 0.86, while the RMSE would decrease by 0.022. It is also worth noting that we did not run classifications on the water column corrected data that are derived following the exclusion of this particular depth point to quantitatively compare the differences in resulting habitat maps and accuracies.

Water column correction

The herein applied analytical model of Maritorena et al. (1994) was based on a clear discussion of the physics pertaining optically shallow waters, whose reflectance contain a benthic signal. In general, analytical models untangle the weak bottom signal from satellite imagery following compensation for the water column attenuation (Dekker et al., 2006). However, these models are always subjected to simplifications inherent in their analytical nature. One of these simplifications concerns light attenuation, which is naturally split into four vertical attenuation coefficients that describe the wavelength dependent absorption and scattering of light due to water components (Maritorena et al., 1994). Given the lack of field measurements or radiative transfer modeling, the diffuse attenuation coefficient is simplified into one parameter.

Maritorena et al. (1994) successfully validated the retrieval of water surface reflectances, but neither the initial study nor other studies have tested the model in an inverse manner to retrieve bottom reflectances since then, to the best of our knowledge. In one of its few applications in an inverse manner, Dierssen et al. (2003) argued that the model yields good agreement with in situ bottom reflectances over dense seagrass (Thalassia testudinum) cover, while, in contrast, overestimated bottom reflectances over intermediate and sparse cover. Deepest areas for their study site were 9 m in contrast to 15 m in our study. In another application of Maritorena et al.'s model, Pu et al. (2014) retrieved bottom reflectances from four 30-m Landsat TM images to map three seagrass species in Florida, USA. They demonstrated that the water column correction (alongside other image pre-processing steps) improved overall accuracies up to 14% in comparison to studies that utilized similar imagery. However, the depth range of seagrass beds in their site was 0–4 m with the majority of seagrass species in < 3 m, nearly 20% of our observed depth range of seagrass presence. All in all, Maritorena et al. (1994) assume a high water transparency, and a horizontal and vertical homogeneity which are in line with the assumption for a uniform mixed character of waters that the bathymetry estimation dictates. Here, both diffuse attenuation coefficient and infinitely deep water reflectance were empirically calculated based on the atmospherically corrected Sentinel-2A image. As regards to the calculation of diffuse attenuation coefficient, we considered only the first 5 m of water column (Fig. 4.1.10) for the calculation of water attenuation coefficients. The principal reason is that attenuation of light increases when wavelength increases

(Fig. 4.1.11), thus rendering the satellite signal (reflectance) weaker and "immersed in sensor noise for deeper waters" (Bierwirth et al., 1993). Therefore, in the red part of the spectrum, even a highly reflective submerged sandy substrate, which was selected here to measure diffuse attenuation coefficients, in optically clear waters (as is the case of the Thermaikos Gulf) is difficult to be detected in depths over 5 m due to very low reflectance values (Maritorena et al., 1994). Extension of the plot in Fig. 4.1.10 in deeper waters than 5 m would hence result in a not so well-defined axis due to the lack of an elongated data cloud after this depth. Generally, fieldbased or radiative transfer-based estimations of both attenuation coefficients and infinitely deep water reflectance are expected to further improve the calculation of bottom reflectances, however, they also raise the cost of the given remote sensing approach of coastal waters.

4.1.5 Conclusions

This study demonstrates the suitability of multispectral Sentinel-2 imagery for the mapping of the distribution of Mediterranean seagrass, namely the dominant *Posidonia oceanica* and *Cymodocea nodosa*. With the launch of the European Union's Sentinels, there is now an extensive fleet of satellites providing systematic coverage of the Earth's coastal areas including seagrass habitats. This new wealth of open, high spatial and temporal resolution remotely sensed data requires suitable preprocessing, processing and analysis to be exploited in an efficient and practical manner. In this study, we utilized atmospheric and water column correction in addition to three different supervised classifiers (MLC, RF and SVM) to accurately map the distribution of *P. oceanica* and *C. nodosa* in the clear and homogeneous, optically shallow waters of the Thermaikos Gulf, Aegean Sea, Greece (eastern Mediterranean Sea). We also created a bathymetry map of the SE Thermaikos Gulf through the development of a site-specific algorithm which uses Sentinel-2A blue and green bands. Together, the methods presented in our study provide an off-the-shelf solution to the mapping of seagrasses and coastal submarine habitats in a range of coastal waters. In the near future, we will focus on the development of NDVI (Normalized Difference Vegetation Index)-like indices using the essential Sentinel-2 data to map the distribution, cover and other important biophysical parameters of seagrasses, namely biomass, leaf area index, carbon sequestration, using readily available in situ data. In addition, we will attempt to map and monitor changes in Mediterranean seagrass dynamics, from seasonal to decadal scales. Quantitative knowledge of seagrass spatial distribution, biophysical parameters and dynamics will resolve existing trends, unravel data issues, and allow better management and conservation practices in regard to a better future for these immensely valuable ecosystems and for their numerous ecosystem services.

4.2 Cubesat-derived Detection of Seagrasses using Planet Imagery following Unmixing-based Denoising: Is Small the Next Big?

4.2.1 Abstract

Seagrasses are one of the most productive and widespread yet threatened coastal ecosystems on Earth. Despite their importance, they are declining due to various threats, which are mainly anthropogenic. Lack of data on their distribution hinders any effort to rectify this decline through effective detection, mapping and monitoring. Remote sensing can mitigate this data gap by allowing retrospective quantitative assessment of seagrass beds over large and remote areas. In this paper, we evaluate the quantitative application of Planet high resolution imagery for the detection of seagrasses in the Thermaikos Gulf, NW Aegean Sea, Greece. The low Signal-to-noise Ratio (SNR), which characterizes spectral bands at shorter wavelengths, prompts the application of the Unmixing-based denoising (UBD) as a pre-processing step for seagrass detection. A total of 15 spectral-temporal patterns is extracted from a Planet image time series to restore the corrupted blue and green band in the processed Planet image. Subsequently, we implement Lyzenga's empirical water column correction and Support Vector Machines (SVM) to evaluate quantitative benefits of denoising. Denoising aids detection of *Posidonia oceanica* seagrass species by increasing its producer and user accuracy by 31.7% and 10.4%, correspondingly, with a respective increase in its Kappa value from 0.3 to 0.48. In the near future, our objective is to improve accuracies in seagrass detection by applying more sophisticated, analytical water column correction algorithms to Planet imagery, developing time- and cost-effective monitoring of seagrass distribution that will enable in turn the effective management and conservation of these highly valuable and productive ecosystems.

4.2.2 Materials and Methods

Field data

In this study, we employ field data that we collected during a boat-, snorkelling- and free-diving-based survey in between 10 – 13.07.2016. The temporal difference of four months between the field survey and satellite data acquisition, as noted in 1.2, is acceptable due to the nature of Mediterranean seagrass meadows which undergo changes in greater time scales. Field data consist of points related to specific habitat and coordinates. Four different substrates were observed: a) *Posidonia oceanica* seagrass, b) *Cymodocea nodosa* seagrass, c) Sand, d) Algae and Rocks.

Satellite data

The satellite data used in our study are acquired by Planet's PlanetScope (PS) Earth-Imaging multispectral CubeSat constellation (Planet Team, 2017). Here, we process the Ortho Scene products which are one of the three product lines for PlanetScope imagery[7]. PS Ortho Scenes are orthorectified, scaled Top of Atmosphere Radiance image products (Level 3B) and are delivered as analytical (4-band) and visual (RGB) products. PS CubeSats captures imagery at a ground sampling distance of 3.7 m at a reference altitude of 475 km and the imagery is then orthorectified to a pixel size of 3 m. The analytical Ortho Scene imagery used in processing and analysis steps herein was acquired on 17/11/2016 and information about its specific attributes is shown in Table 4.2.1. The signal-to-noise ratio (SNR) of the selected image is 68.8.

Unmixing-based Denoising

A schematic workflow of our adapted methodology is presented in Fig. 4.2.1. The low Signal-to-noise Ratio of (SNR) of the selected PS image prompts the application of the Unmixing-based denoising (UBD) as a Pre-processing step. UBD has been recently proposed to selectively retrieve spectral bands characterised by a low SNR by exploiting their correlation with non-corrupted pixels across the whole spectral dimension in hyperspectral images (Cerra et al., 2014). Spectral unmixing is the process which decomposes a hyperspectral image element into a linear combination of signals representing the backscattered solar radiation in each spectral band. Considering the physical properties of a mixed spectrum, UBD assumes the residual vector derived from the unmixing process to be mostly composed of noise and more relevant in spectral bands where atmospheric absorption effects are stronger, and therefore ignored in the reconstruction. In spite of adopting a linear unmixing model, which in theory does not adapt well to water, UBD achieved satisfactory results when applied to spectral bands in the blue range for scenes acquired over coastal waters in previous works (Cerra et al.,

---

[7] https://www.planet.com/products/satellite-imagery/files/Planet_Imagery_Product_Specs.pdf.

2013). Recently, it has been successfully applied to multitemporal stacks of multispectral images (Cerra et al., 2016a, 2016b).

Here, a total of 15 spectral-temporal patterns is extracted from a multitemporal Planet image time series, selected among the purest image elements in the stack. In addition to the PS image of the 17/11/2016, two other PS images from 30/09/2016 and 03/12/2016 with SNRs of 90.1 and 71.8, respectively, comprise this multispectral image time series. Subsequently, UBD is applied to restore the corrupted blue and green band in the PlanetScope image of the 17/11/2016.

Empirical water column correction

To display the quantitative benefits of UBD on the Planet high spatial resolution image, we implement Lyzenga's empirical water column correction (Lyzenga, 1981, 1978) on the initial

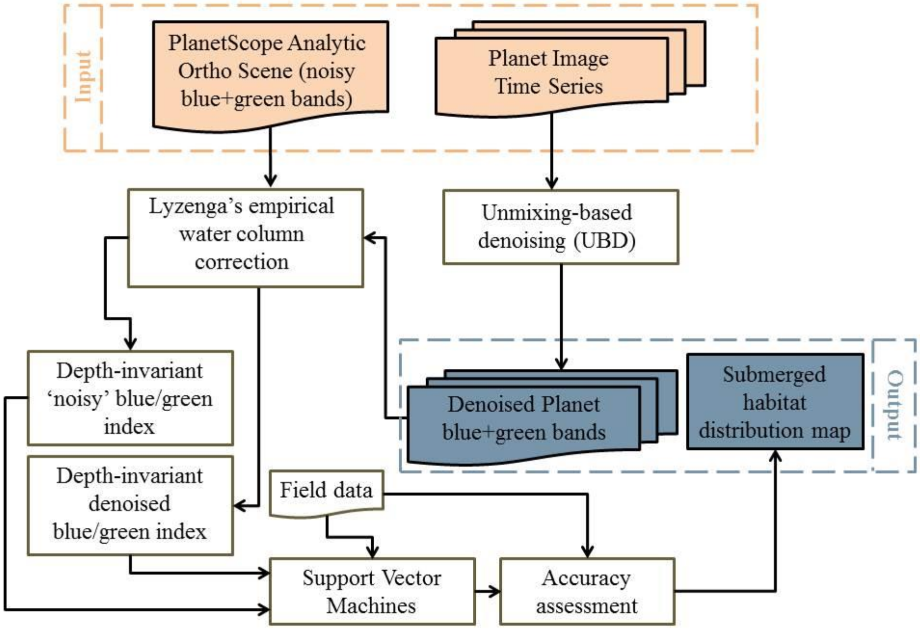


Fig. 4.2.1. Workflow of the herein processed methodology. The Planet Image Time Series consists of three different PlanetScope Analytic Ortho Scenes. Lyzenga's empirical water column correction is adapted from Lyzenga (1978, 1981).

Table 4.2.1. PlanetScope Analytic Ortho Scene Product Attributes.

Product Attribute	**Description**
Instrument	PS2 (2nd generation)
Acquisition date	17/11/2016
Acquisition time	9:27:08 AM GMT+1
Pixel Size	3 m
Spectral bands	Blue: 455 - 515 nm Green: 500 - 590 nm Red: 590 – 670 nm NIR: 780 – 860 nm
Pixel depth	Analytic (DN): 12-bit Analytic (Radiance): 16-bit
Orbit altitude	475 km

'noisy' and denoised (restored) blue and green bands. These bands are selected on the basis of their great water penetration in contrast with the red and NIR bands. The two main assumptions behind Lyzenga's algorithm are that: a) radiance differences between pixels of the same bottom are due to different depth and b) water column attenuation coefficient is constant for each band. We select sandy bottoms in the use of the empirical water column correction as they are the easiest to identify and cover large areas over different depths. The empirical algorithm derives one depth-invariant bottom index before and one after the denoising. These indices are not bottom reflectances, but rather represent a relation between radiances in the blue and green spectrum without a depth effect.

We choose Lyzenga's empirical algorithm due to its practicality (it does not require field or bathymetry data) and efficiency in improving mapping accuracies in waters that feature horizontal and vertical homogeneity of their optical properties of up to 15 m which are the case in the Thermaikos survey site. Furthermore, this survey site features same substrates (sand and seagrass) over a wide depth range (0 – 16 m), inducing the implementation of Lyzenga's algorithm.

Support Vector Machines

The Support Vector Machines (SVM) comprise a group of theoretically superior, supervised classifiers based on the statistical learning theory of Vapnik (1995). These essentially machine learning algorithms solve classification problems by fitting an optimal hyperplane, which comprises the decision boundary, to the training samples. The samples that fall closer to this hyperplane are the so-called support vectors which are the only data used by SVM in the classification step. We apply SVM utilising the universal Gaussian radial basis function kernel (RBF):

$$K(x, x_i) = \exp(-g\|x - x_i\|^2) \quad (1)$$

where g is the parameter that defines the width of the RBF. In addition, we set a regularisation parameter, C which controls the degree of acceptable misclassification through limiting the influence of individual training samples. Furthermore, we select a Cross validation parameter which monitors the accuracy of results during the grid search for the selection of the best pair of values for g and C. Specific model parameters to run both SVM experiments are shown in Table 4.2.2. The RBF kernel is employed on the acquired field data to classify both depth-invariant indices before and after the application of UBD. The SVM experiments are run using the open-source LIBSVM library (Chang and Lin, 2013).

Table 4.2.2. Model parameters used to run Support Vector Machines utilising the universal Gaussian radial basis function kernel (RBF) for both experiments; before and after the application of Unmixing-based Denoising (UBD). Min and Max report the minimum and maximum values which define the grid range (g and C dimension). A multiplier of 10 is set for grid range for both experiments.

Model parameters	**Before UBD**	**After UBD**
Support vectors	43	31
RBF width, g	100	1000
Regularization parameter, C	0.01	10
Min (g), Max (g)	0.01	1000
Min (C), Max (C)	0.01	1000
Cross validation	3	3

Termination criterion for grid search	0.1	0.001
Termination criterion for final training	0.1	0.001

The quantitative evaluation of the potential benefits of UBD on Planet data for detection of Mediterranean seagrasses and other shallow submerged habitats is performed through the accuracy assessment of the SVM-derived results. The accuracy assessment employs field data from Thermaikos survey site using the error matrix and Kappa value (Congalton, 1991).

4.2.3 Results

Our experimental results are displayed in methodological succession in Fig. 4.2.2 and 4.2.3. Visual inspection of Fig. 4.2.2A and 4.2.2B, which show the initial 'noisy' blue and green bands of PlanetScope image over our test site, reveals several artefacts, generated and enhanced in the acquisition and pre-processing of the acquired images. While we have not been able to determine to which extent these artefacts are introduced for each processing step, they appear to be significantly decreased following application of the Unmixing-based denoising (Fig. 4.2.3A, 4.2.3B). The depth-invariant index of 'noisy' blue and green bands (Fig. 4.2.2C) indicates that Lyzenga's empirical water column correction fails to eliminate the effect of water column attenuation on bottom reflectances. In contrast, applying Lyzenga's algorithm on the denoised depth-invariant index (Fig. 4.2.3C) yields a seemingly more realistic representation of the bottom. The SVM-derived classification results comprise submerged habitat distribution maps of our survey site in the Thermaikos Gulf (Fig. 4.2.2D, 4.2.3D). Accuracy assessment (Table 4.2.3, 4.2.4) of the classifications of the resulting 'noisy' and denoised depth-invariant indices following application of UBD and empirical water column correction reflects the visual inspections. The denoised depth-invariant index of blue and green Planet reflectances improves overall accuracy and Kappa value from 53.2% and 0.26 to 68.1% and 0.5, respectively. Denoising aids detection of *Posidonia oceanica* seagrass species by increasing its producer and user accuracy by 31.7% and 10.4%, correspondingly, with a respective increase in its Kappa value from 0.3 to 0.48. On the other hand, both classified depth-invariant indices, before and after denoising, fail to detect *Cymodocea nodosa* seagrass species. Moreover, the error matrix indicates that *P. oceanica* seagrass class is mainly confused with sand class. Generally, denoising unveils fine, smaller patches of sand within the homogeneous *P. oceanica* meadows, an essential feature of the high spatial resolution Planet imagery towards finer scale seagrass detection. Both classification results lack clear delineation of the seaward limit of *P. oceanica* seagrass.

4.2.4 Discussion

Using high spatial resolution CubeSat-derived Planet imagery, we detect *Posidonia oceanica,* the dominant Mediterranean seagrass species, with acceptable accuracies following application of UBD, empirical water column correction and SVM. On one hand, we expect Planet's high spatio- temporal resolution imagery to cover gaps in finer scale seagrass

detection of temporal cover of remote sensing data provided by bigger multispectral satellites like Sentinel-2 and Landsat-8. On the other hand, remote sensing of coastal ecosystems like seagrasses requires sensors with higher SNR than the ones applied in terrestrial ecosystems. Planet's low SNR for pixel-based coastal and aquatic, generally, applications may bring degradations in the image quality. In our study, these are shown to hinder the production of informative habitat maps following Lyzenga's empirical water column correction. We attribute this problem to the selection of pixels that represent sandy bottoms in different depths but suffer from notable degradations.

As we show here, the low SNR drawback can be balanced, first, by the application of UBD which significantly improves seagrass detection amongst other submerged habitats including sand and algae/rocks. A second solution for the SNR deficit for water applications can be possibly provided by applying data fusion between Planet high spatio-temporal resolution images and Sentinel-2 and/or Landsat-8 lower spatial, but higher spectral resolution. A third and last, possible solution for the noise issue can be provided through the use of Object-based Image Analysis (OBIA) which in contrast to pixel-based approaches, first analyse image scenes to relatively homogeneous areas by decomposition and then produce object-based instead of pixel-based classifications (Blaschke, 2010). In contrast to its increasing popularity and efficiency to terrestrial approaches, however, applications of OBIA to coastal areas are scarce.

In the near future, our objective is to improve accuracies in seagrass detection by applying more sophisticated, analytical water column correction algorithms to Planet imagery, developing time- and cost-effective mapping and monitoring of seagrass distribution. Effective exploitation of this new kind of high spatio-temporal resolution multispectral dataset can meet the scale and aim of the majority of future seagrass mapping and monitoring projects, from a small meadow to a whole coastline, allowing for improved management and conservation of these significant coastal ecosystems.

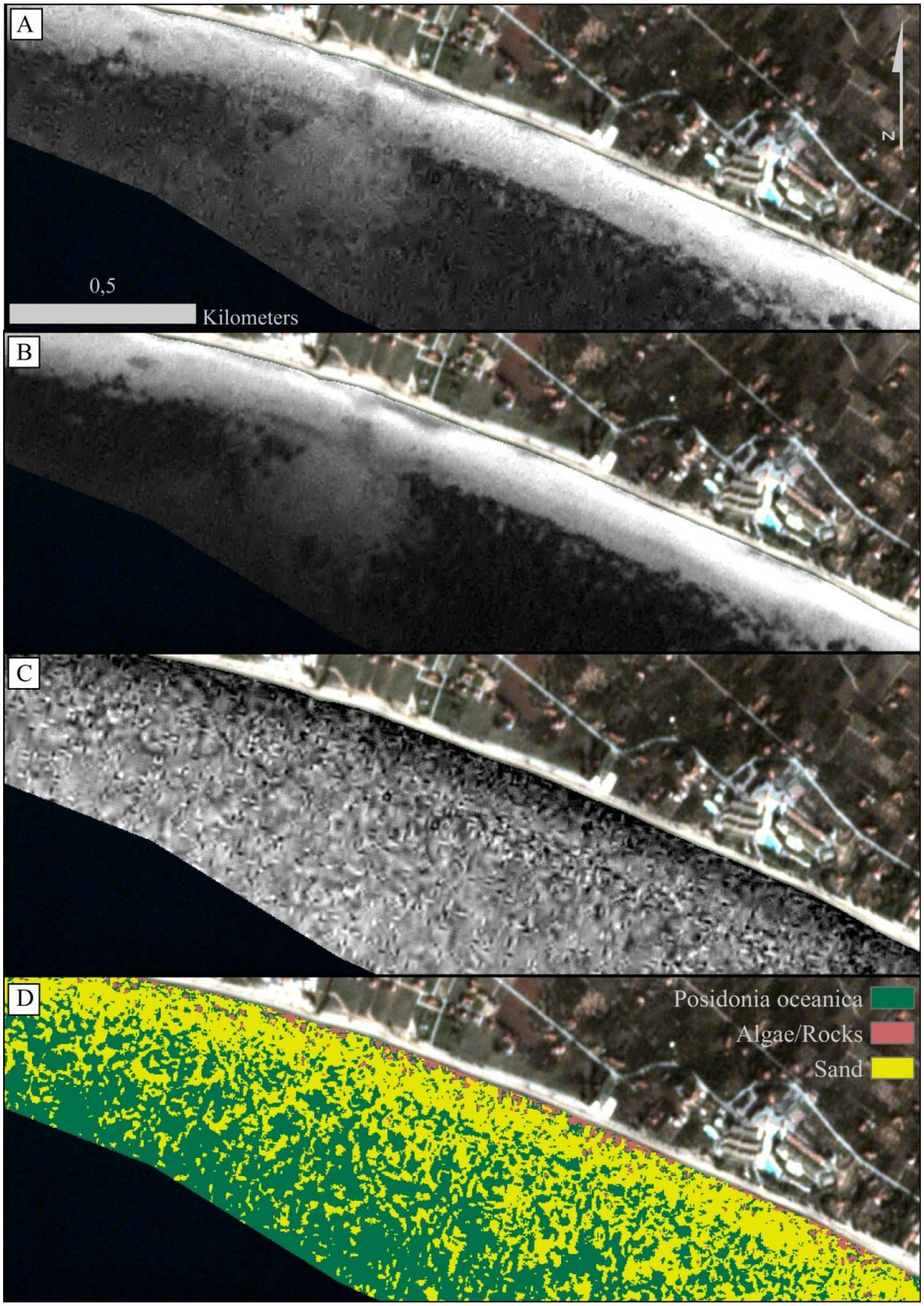


Fig. 4.2.2. Experimental steps without the application of Unmixing-based denoising. A and B display the initial, 'noisy' blue and green Planet reflectance, respectively. C is the depth-invariant index (Lyzenga's empirical water column correction) of A and B before applying the Unmixing-based denoising (UBD). D shows Support Vector Machine classification result of C. All rasters are draped over a PlanetScope analytic Ortho RGB Scene of 17/11/2016 over our test site in Thermaikos Gulf, NW Aegean Sea, Greece, projected in UTM (zone 34) system/WGS84.

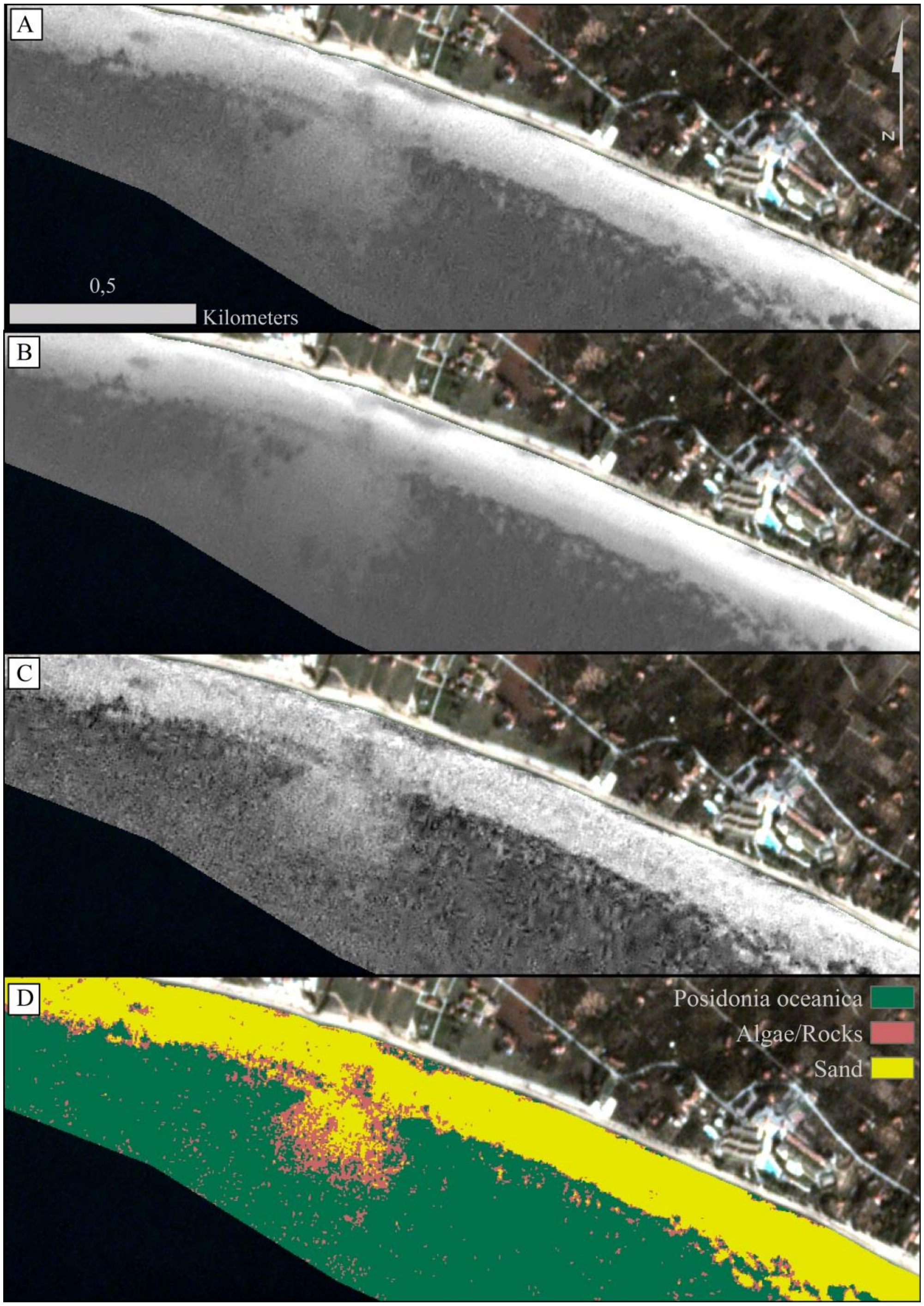


Fig. 4.2.3. Experimental steps with the application of Unmixing-based denoising. A and B display the denoised blue and green Planet reflectance following the UBD application on the A and B of Fig. 4.2.2., respectively. Differences in overall brightness from A and B of Fig. 4.2.3 are due to their different histogram stretch. C is the depth-invariant index of A and B after the UBD application. D shows Support Vector Machine classification result of C. All rasters are draped over a PlanetScope analytic Ortho RGB Scene of 17/11/2016 over our test site in Thermaikos Gulf, NW Aegean Sea, Greece, projected in UTM (zone 34) system/WGS84.

Table 4.2.3. Error matrix of Fig. 4.2.2D before the application of UBD on Planet imagery. CN: *Cymodocea nodosa* seagrass, PO: *Posidonia oceanica* seagrass, A/R: Algae and Rocks, S: Sand

Before UBD						
Class	CN	PO	A/R	S	Total	User accuracy
CN	0	0	0	0	0	0
PO	3	26	0	14	43	60.5
A/R	0	1	6	1	8	75
S	6	14	5	18	43	41.9
Total	9	41	11	33	94	
Producer accuracy	0	63.4	54.6	54.6		
Kappa value	0	0.3	0.72	0.1		
Overall accuracy: 53.2; Kapp value: 0.26						

Table 4.2.4. Error matrix of Fig. 4.2.3D after the application of UBD on Planet imagery. CN: *Cymodocea nodosa* seagrass, PO: *Posidonia oceanica* seagrass, A/R: Algae and Rocks, S: Sand

After UBD						
Class	CN	PO	A/R	S	Total	User accuracy
CN	0	0	0	0	0	0
PO	2	39	2	12	55	70.9
A/R	3	1	7	3	14	50
S	4	1	2	18	25	72
Total	9	41	11	33	94	
Producer accuracy	0	95.1	60.1	81.1		

Kappa value	0	0.48	0.43	0.57		
Overall accuracy: 68.1; Kapp value: 0.5						

4.3 On the use of Sentinel-2 for Coastal Habitat Mapping and Satellite-derived Bathymetry Estimation using Downscaled Coastal Aerosol Band

4.3.1 Abstract

Coastal habitats provide a plethora of ecosystem services, yet they undergo continuous pressure and degradation due to the human-induced climate change. Conservation and management imply continuous monitoring and mapping of their spatial distribution at first. The present study explores the capabilities of the Copernicus Sentinel-2 mission and the contribution of its coastal aerosol band 1 (443 nm) for the mapping of the dominant Mediterranean coastal marine habitats and the bathymetry in three survey sites in the East Mediterranean. The selected sites have shallow to deep habitats and a high variability of oceanographic and seabed morphological conditions. The major findings of our study demonstrate the advantages of the downscaled Sentinel-2 coastal aerosol band 1 for both optically shallow habitat and satellite-derived bathymetry mapping due to its great water penetration. The use of Sentinel-2 band 1 allows detection of *Posidonia oceanica* seagrass beds down to 32.2 m of depth. Sentinel-2 constellation with its 10-m spatial resolution at most of the spectral bands, 5-day revisit frequency and open data policy can be an important tool to provide crucial missing information on the spatial distribution of coastal habitats and on their bathymetry distribution, especially in data-poor and/or remote areas with large gaps in a retrospective, rapid and non-intrusive manner. As such, it becomes a crucial ally for the conservation and management of coastal habitats globally.

4.3.2 Materials and Methods

Study areas

The selected sites are in water bodies with different oceanographic and topographic characteristics (SoHelME, 2005). This variety allows comparisons for the capabilities of Sentinel-2 in seabed mapping in areas with medium and low productivities. The National Park of Marathon-Schinias (hereafter Schinias) covers a landscape of 56 km² in the marine body of the Petalioi Gulf, at the outlay of the South Evoikos Gulf. The studied area is the coastal zone of the park where seagrass meadows are present between 2 and 20 m (Fig. 4.3.1a). Samaria National Park (hereafter Samaria) is located in the West part of the island of Crete; it was declared as a National Park in 1962. The selected area lies up to 38 m deep and features mainly patchy seagrasses (Fig. 4.3.1b). Gavdos Island (hereafter Gavdos) is in the southernmost point of Europe and Greece. The studied area concerns the coastal part of the island where seagrass meadows thrive between 5 and 25 m of depth (Fig. 4.3.1c).

Sentinel-2 data

Copernicus Sentinel-2A (S2-A) L2A satellite data from the first of the polar-orbiting satellites, have been selected and downloaded from the Sentinel Scientific Data Hub for the present study (ESA, 2015). L2A data have been corrected from L1C using the rural (continental) aerosol type, a mid-latitude summer atmospheric profile, a value of 331 Dobson Units for ozone concentration, a land-average value for water vapor over water, a visibility of 40 km and an adjacency range of 1 km. Here, we have processed and analyzed two S2-A images (Table 4.3.1); one for Schinias, and one for Samaria and Gavdos, as one S2 tile covers both areas.

In situ data

Field data for the three distinctive Mediterranean coastal habitats in the selected three sites (Fig. 4.3.2) have been collected with various approaches. These are: a) *P. oceanica* seagrass meadows (Fig. 4.3.2a, c); b) rocky bottoms covered by photophilous algae (Fig. 4.3.2b); c) soft/sandy bottoms (Fig. 4.3.2d). For the sites located in Crete, we collected the data in summer 2013 during research activities for the ecological status of the seagrass meadows (Poursanidis et al., 2014) in Samaria and Gavdo Natura 2000 site, and updated them in summer 2017.

We collected the data by snorkeling between 0–45 m and SCUBA diving between 5–45 m. In both cases, the field scientist used a water- proof GPS unit (GARMIN GPSMAP series) either on a buoy or on hand and collected the data. In each site (Samaria and Gavdos), we acquired 20 POIs per habitat for use in image classification. Bathymetric data were collected only in Samaria by using a Lowrance HDS5 gen2 with a single beam sonar head placed 15 cm below the water surface. We collected 107 field points in total in Gavdos and 79 field points in Samaria. In the case of Schinias, we extracted 64 field points by interpreting high-resolution aerial images by the National Cadaster & Mapping Agency S.A. from 2010 and open access high resolution images between 2010–2016 in Google Earth.

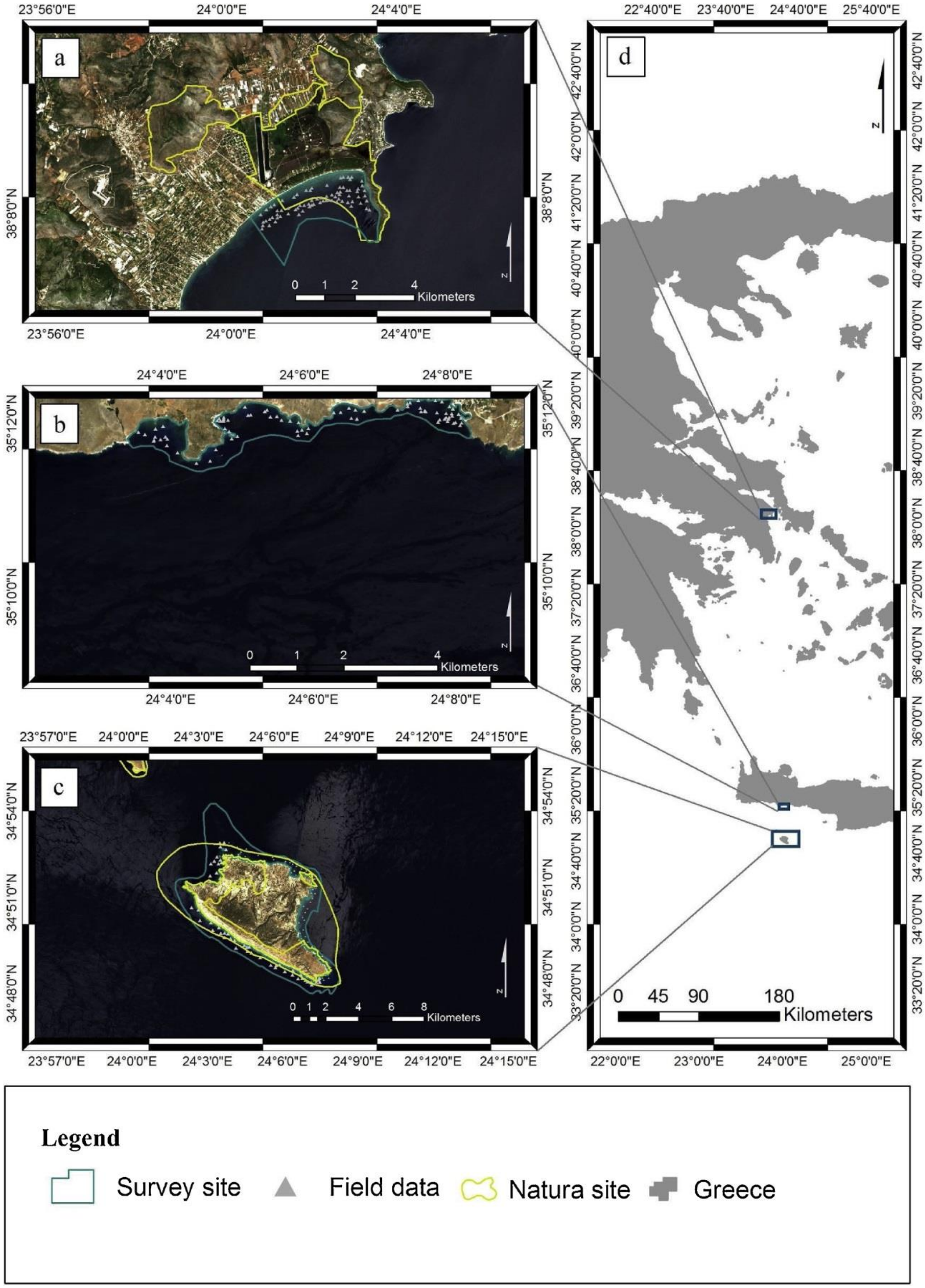


Fig. 4.3.1. Location of survey site in (a) Schinias Marathon National Park, (b) Samaria National Park, (c) Gavdos Island, and (d) Greece.

Table 4.3.1
Characteristics of Sentinel-2A images.

Survey site	**Schinias**	**Samaria & Gavdos**
Granule	34SGH	34SGD
Acquisition data	15 June 2017	26 April 2017
Acquisition time	09:05 am UTC	09:10 am UTC
Mean solar zenith angle	20.4°	25.3°
Mean solar azimuth angle	132.6°	142.7°

Fig. 4.3.2. The habitats of the selected sites: (a) *Posidonia oceanica* meadows at 30 m, (b) Rocky reef covered by photophilic algae, (c) *Posidonia oceanica* meadows at 5 m, and (d) Soft sandy bottom.

Pre-processing steps

Fig. 4.3.3 provides an overview of the designed methodology for the processing and classification of S2-A images. As a panchromatic band is absent in Sentinel-2, a different approach than pan-sharpening is needed to increase the spatial and spectral information of the coastal aerosol band 1 (443 nm). Band 1 is of significant importance in the remote sensing of coastal waters (Mouw et al., 2015) due to its deeper penetration than the other visible bands. Brodu (2017) developed a downscaling method, namely super-resolution, for multispectral, multi- resolution imagery as the Sentinel-2 satellites. We run the downscaling algorithm using the Sen2Res plugin in ESA'S Sentinel Application Platform (SNAP) Version 5.0 for all survey sites (Fig. 4.3.4) in order to downscale all spectral bands at 10 m. We use the downscaled Short-wave Infrared SWIR band 11 at 1610 nm to mask all terrestrial features in all survey sites by classifying the image into water and land with an Iso Cluster Unsupervised Classification. We correct sunglint using the method of Hedley et al. (2005). We also implement the Lyzenga depth- invariant indices (DIV) (Lyzenga, 1978, 1981) to empirically address the influence of variable depth on optically shallow seabed. We estimate three DIV (b1-b2, b2-b3, b1-b3) for all three survey sites.

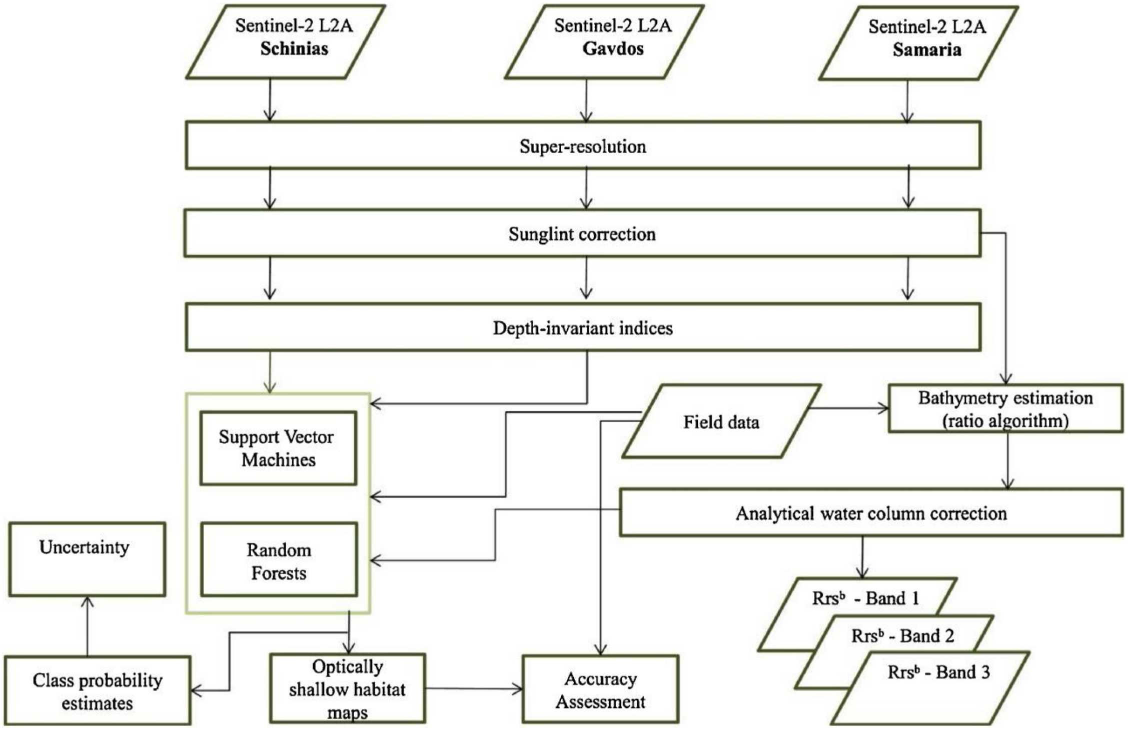


Fig. 4.3.3. Schematic methodological workflow.

Analytical water column correction – Bathymetry estimation

In contrast to the empirical image-based technique of Lyzenga (1978, 1981), semi-analytical and analytical water column correction are expected to increase classification accuracies by minimizing the water column interference on the signal of the submerged habitats in a more efficient manner. We implement the analytical water column correction model of Maritorena et al. (1994), while for the bathymetry estimation we employ the ratio method of Stumpf et al. (2003).

Classifications and uncertainty estimation

We utilize the machine learning classifiers of Support Vector Machines (SVM) and Random Forests (RF) - widely used in the remote sensing of optically shallow habitats (Eugenio et al., 2015; Traganos and Reinartz, 2017a, 2017b; Zhang, 2015). We employ both classifiers on all types of reflectance composites - the first three Sentinel-2A bands in 10-m resolution (bands 1–3) in all three survey sites: a) Atmospherically corrected L2A imagery (R); b) deglinted L2A imagery (R'); c) DIV; d) substrate remote sensing reflectance composites ($R\text{rs}^{b}$) in the case of Samaria. SVM are a group of universal machine learning algorithms based on the statistical learning theory of Vapnik (1995), while RF comprise an ensemble method for supervised classification based on classification and regression trees (CART) and were developed by Breiman (2001). To increase understanding as well as interpretation of our classification results, we decided to output them as class probability estimates (soft classification) instead of the typical multi-class hard classification outputs. To estimate per pixel probabilities, binary RF- and SVM-derived decision values are first transformed into binary probabilities according to Platt (2000) and then into class probabilities following the pair-wise coupling approach (one-against-one) of Wu et al. (2004). All experiments with SVM/RF and probability estimates were run in EnMAP-Box (van der Linden et al., 2015). We utilise Shannon entropy *(H)* (Shannon, 1948) to output per-pixel uncertainty as a function of class probabilities. We evaluate the hard classification results (the coastal habitat maps) by calculating overall (OA), producer (PA) and user accuracy (UA) of every class per survey site.

4.3.3 Results

Pre-processing results

Fig. 4.3.4 displays Sentinel-2 A L2A 4-3-2 composites from our three survey sites along with the coastal aerosol band (443 nm) over each site in its initial and super-resolved resolution (60 and 10-m correspondingly). The pink and yellow polygons in 4(e) and (f) display selected super-resolved pixels (sandy and rocky seabed) with same reflectance as the initial 60-m initial pixel— 0,0489 ± 0,0009 (rocks) and 0,0629 ± 0003 (sands)— demonstrating that super-resolution pre-processing does not impact the sensor radiometry. Deglinted L2A 3-2-1 composites are depicted in Fig. 4.3.5. Especially in the Schinias (Figs. 5a, b) and Gavdos (Figs. 5e, f, g, h), benthic features were spectrally enhanced following the sunglint correction of Hedley et al. (2005). Fig. 4.3.6 shows the image-based DIV (Lyzenga, 1978, 1981) in all three field sites. Light blue polygons in the insets of Fig. 4.3.6 indicate pixel areas of sandy sub- strate used in the estimation of the indices. Negative DIV values were observed and were corrected to cover positive ranges, incorporating an offset. Essential to both the estimation of accurate depth-invariant indices as well as to the analytical water column correction, the ratios of diffuse attenuation coefficient $k1/k2$, $k2/k3$, and $k1/k3$ were estimated as 0.64, 0.53, and 0.33 for Schinias; 0.77, 0.53, and 0.37 for Samaria 1.18, 0.28, and 0.05 for Gavdos.

Analytical water column correction and bathymetry estimation

Available field depth data allowed the estimation of substrate remote sensing reflectances (R_{rs}^{b}) (Fig. 4.3.7a). Wavelength-dependent attenuation coefficients (k) of the Sentinel-2 b1, b2 and b3 bands in Samaria were estimated as 0.03 m^{-1}, 0.04 m^{-1} and 0.08 m^{-1}. These values were incorporated in Equation (3) to derive R_{rs}^{b} . Furthermore, utilising k, we estimated the 'effective penetration depth' (EPD) in the first three visible bands; 74 m in 443 nm, 57.5 m in 490 nm, and 30.2 m in 560 nm. Employing Equations (5) and (6), we produced a 10-m bathymetry (Z) map of the Samaria site (Fig. 4.3.8). The implemented ratio of Equation 6 exploited the different attenuation of the b1 and b3 to retrieve water depths. Additionally, we derived a second bathymetry map (not shown here), utilising Equation (7) and (8) where we used the ratio of the b2 to b3 band. The validation of both SDBs was assessed with 1658 *in situ* depth measurements from the total of 2569 (Fig. 4.3.9a). Linear re- gressions between SDBs and field measurements revealed a superior r- squared value of 0.85 and a smaller root-mean-square error (RMSE) of 4.31 m following use of the b1-b3 ratio (Fig. 4.3.9) in comparison to the r-squared value of 0.68 and the RMSE of 5.46 m of the b2-b3 ratio. Blending knowledge of the bathymetry and effective penetration depth of Sentinel-2 imagery, we masked waters below 30.2 m (Figs. 7a). Past the EPD in 560 nm, the analytical water correction model might be possibly compromised or not be able to correct the depth effect (e^{-2kZ}) in Equation (3). We also masked waters below 43 m following the findings of Poursanidis et al. (2018) which show that the deep limits of *P. oceanica* seagrass lies on this depth.

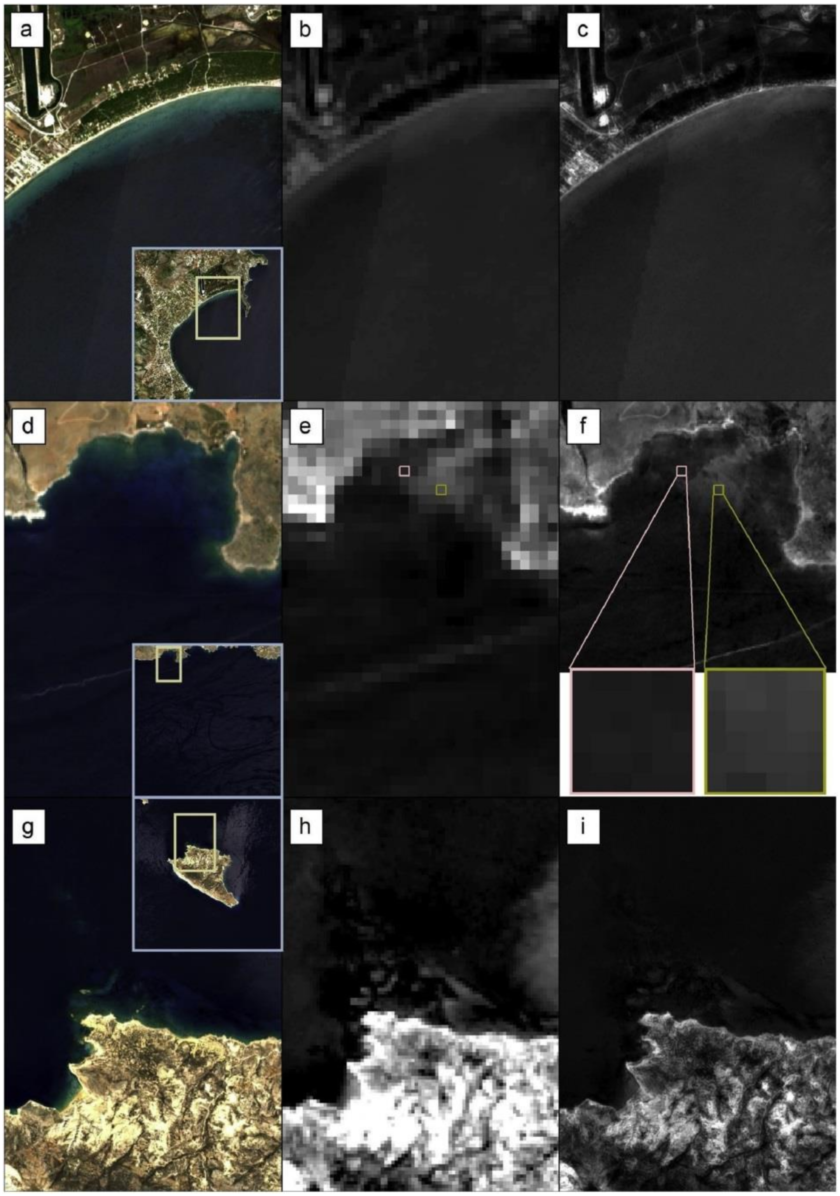


Fig. 4.3.4. Downscaling of Sentinel-2 band 1 following Brodu (2017). (a), (d), (g): 4-3-2 composite images of Schinias, Samaria, and Gavdos Island. (b), (e), (h): Original band 1 (443 nm, 60-m resolution). (c), (f), (i): Downscaled band 1 at 10-m resolution. The pink and yellow squares in Fig. 4.3.4(f) show the downscaled 10-m pixels over a rocky and sandy seabed, respectively. The polygon in the three insets of 4(a), (d) and (g) indicates the location of the displayed panels within the extent of each survey site.

Classifications and uncertainty estimation

Two machine learning classifiers, SVM and RF, were applied to three different composites in three survey sites for a total of 24 classified habitat maps. Except from the case of $R_{rs}{}^{b}$ composites in the Samaria site, the other classified composites were the S2-A L2A, the deglinted L2A and the DIV. In Samaria, *P. oceanica* seagrass occupy depths up to 32.2 m in the eastern part of this site with a 96.8% probability and 14.7% uncertainty as revealed by the SVM-classified full-depth $R_{rs}{}^{b}$ composite (OA: 75%, PA: 100%, UA: 72.5%) and SDB (Fig. 4.3.8). Generally, SVM exhibited a marginal higher overall accuracy on classifying the L2A and deglinted L2A composites (+2.5%) than the $R_{rs}{}^{b}$ imagery. In addition, the highest uncertainties are depicted in deeper areas (> 20 m) which are immersed in noise (Fig. 4.3.7e).

In Schinias, the SVM-classified deglinted L2A composite (Fig. 4.3.12a) yielded the best OA of 87.2%. This classified composite featured similar PA of seagrasses (82.6%) to the SVM L2A and RF L2A composites. On the other hand, the RF-classified DIV composite was the most accurate product in both PA and UA. Classifier-wise, SVM were slightly more accurate than Random Forests in the two-class experiments here, exhibiting 0.9% better OA. They also displayed 2.9% greater PA but 1% lower UA in classifying seagra Fig. 4.3.11d indicates a lower uncertainty in areas shallower than 10 m.

In Gavdos, the RF-derived deglinted L2A imagery (Fig. 4.3.11a) showed the best OA (58.2%) in classifying the benthic habitats of the island. More particularly, as regards to the class of *P. oceanica* and overall classifier performance, RF featured a 10.9% better OA than SVM, a 2.7% lower PA and a greater UA of 13.6%. *P. oceanica* class probability seems to be overestimated over optically deep areas (Fig. 12b), while uncertainty measure reveals a higher certainty on the eastern and southern part of Gavdos Island in contrast to its western part (Fig. 4.3.12e). While SVM exhibited a marginal difference from RF in OA by 2.7%, it demonstrated noteworthy greater PA and UA of *P. oceanica* seagrass class of 25.9% and 21.1%.

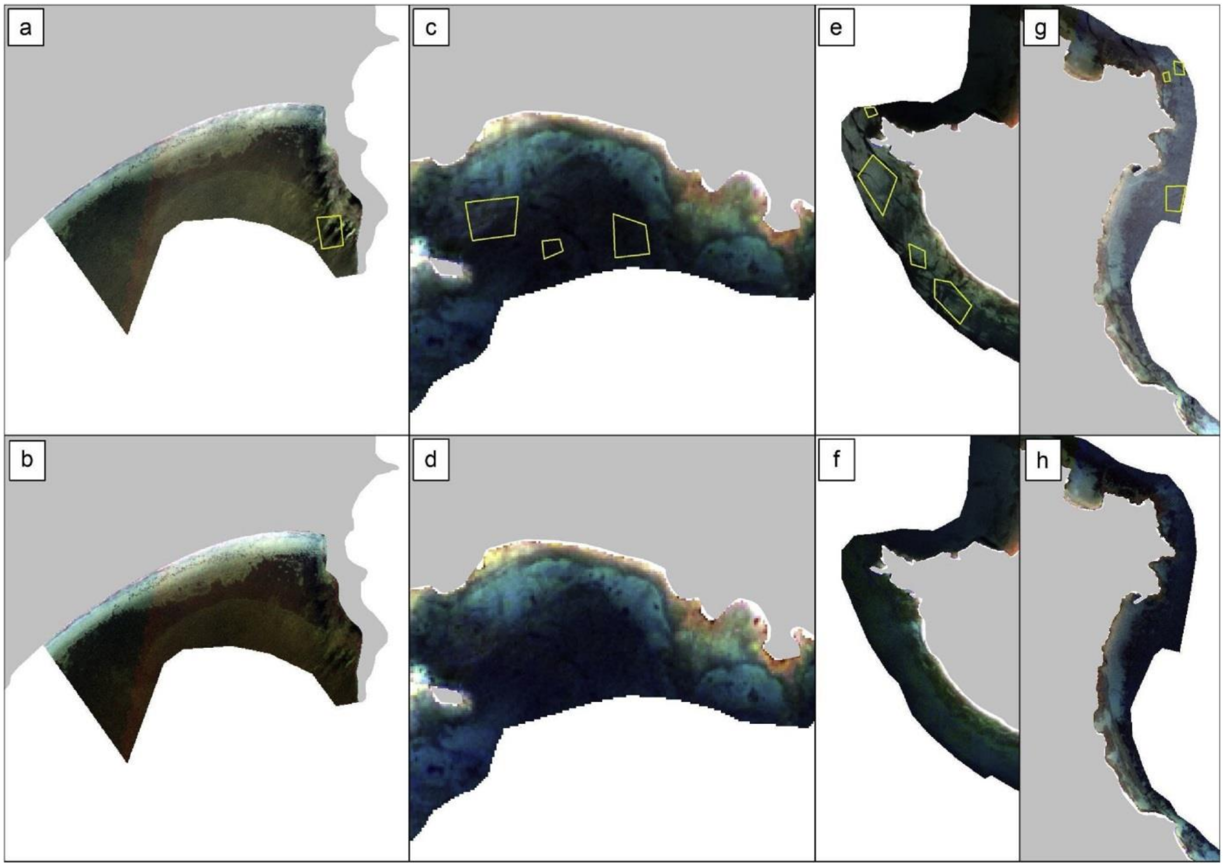


Fig. 4.3.5. Sunglint correction of Sentinel-2 A images following Hedley et al. (2005). All images are 3-2-1 L2A composites with downscaled band 1 at 10-m resolution. Yellow polygons indicate pixels selected for the regression. (a), (b): Initial and deglinted image of Schinias. (c), (d): Initial and deglinted image of the centre o Samaria. (e), (g), (f), (h): Initial and deglinted images of NW and E Gavdos.

4.3.4 Discussion

The present study selected three protected, eastern Mediterranean sites to evaluate the potential of the Sentinel-2 sensor in coastal habitat mapping and satellite-derived bathymetry over a variety of seabed morphologies and oceanographic characteristics. The coastal aerosol band 1 of S2 in 443 nm is the common denominator in the two major findings of our study. Leveraging empirical to analytical corrections and machine, we first found that S2 can detect *P. oceanica* seagrasses 32.2-m deep into the South Cretan water column (Fig. 4.3.10). We mainly attribute this finding to the high oligotrophic Cretan Sea which features a very low primary production (Psarra et al., 2000) and a strong pelagic microbial loop which subsequently reduces particle flux to the sediment (Chronis et al., 1996; Danovaro et al., 1996). Previous S2-based SDB estimations in the same basin reached an optical depth limit of 16.5 m (Traganos and Reinartz, 2018a).

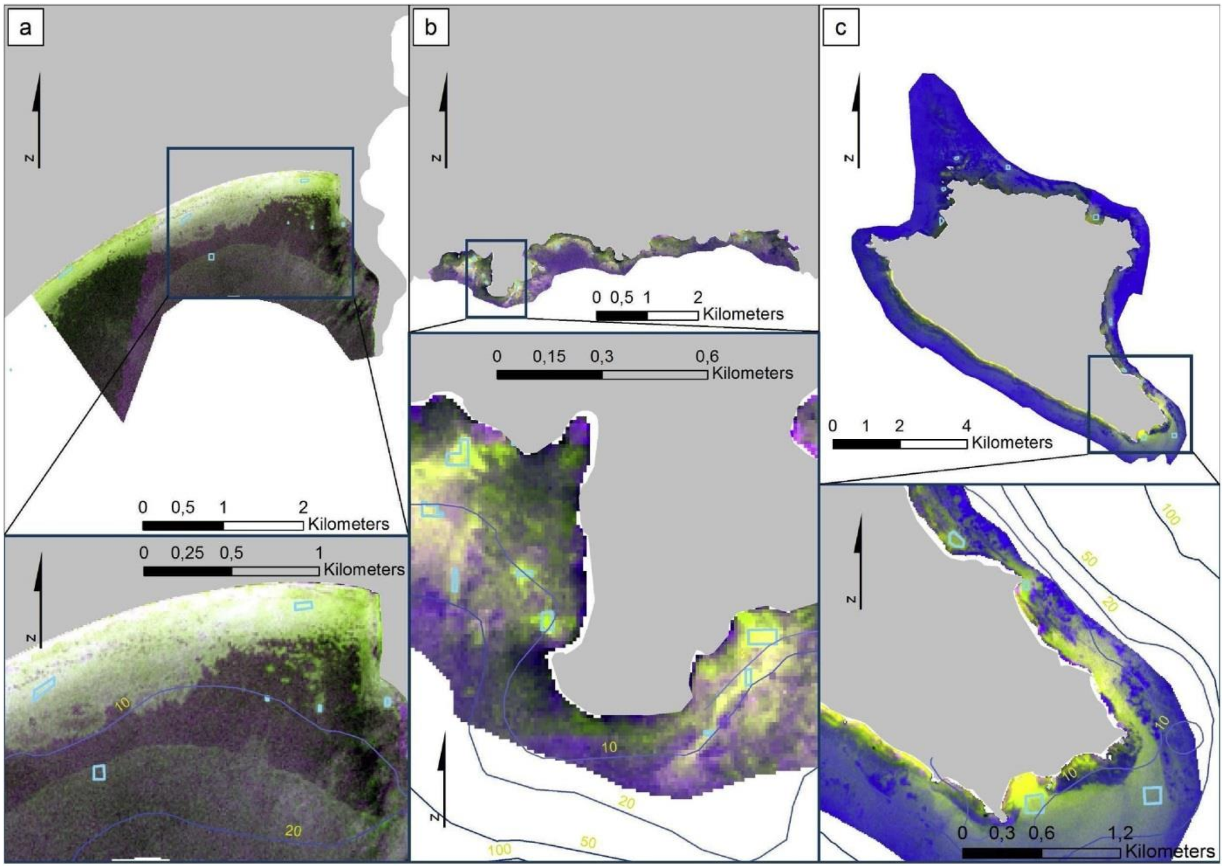


Fig. 4.3.6. Depth-invariant bottom indices (b1-b2, b1-b3 and b2-b3) following Lyzenga (1978, 1981). Light blue polygons indicate regression pixels that represent the same substrate, sand, in different depths. (a) Schinias, (b) Samaria, (c) Gavdos.

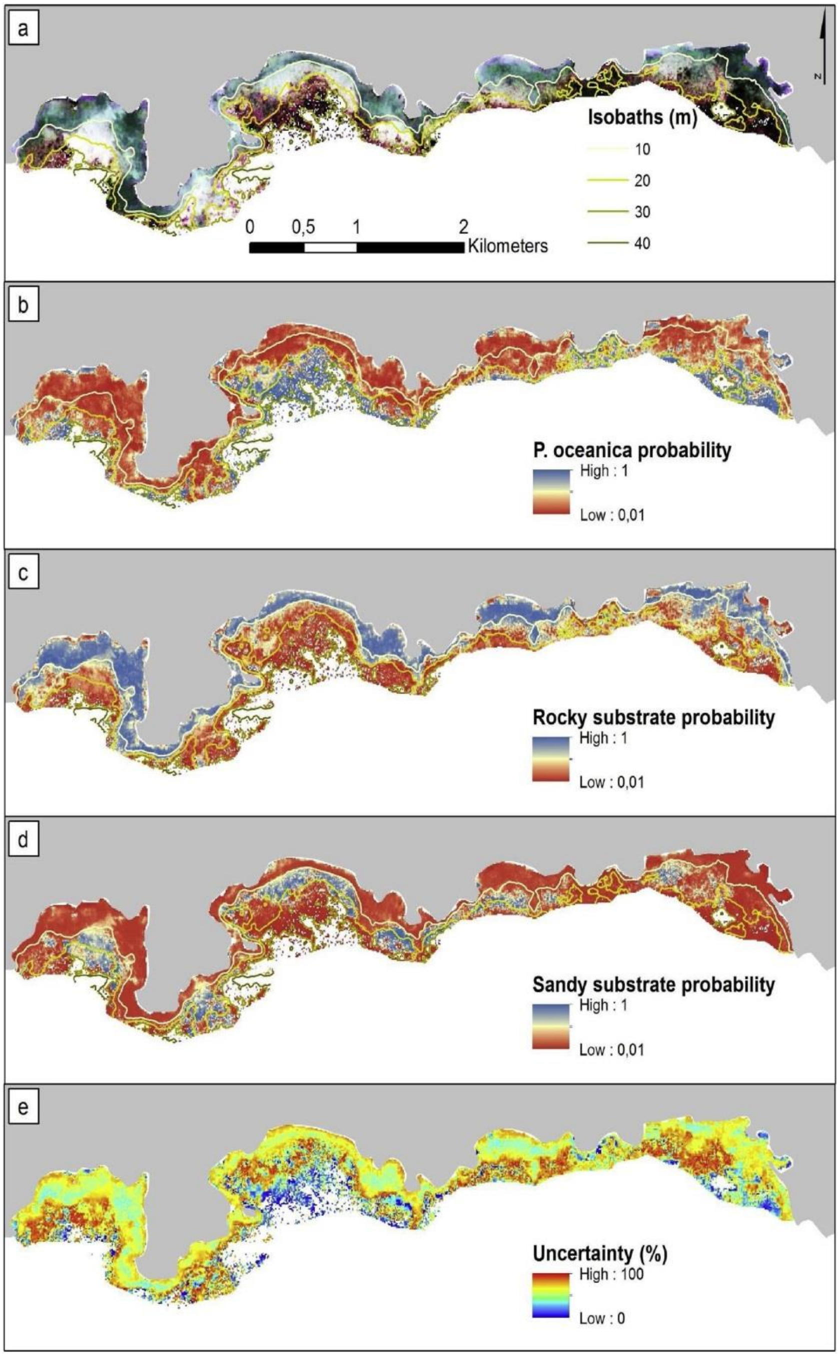


Fig. 4.3.7. Water column corrected S2-A 3-2-1 composite of Samaria survey site, and related probability and uncertainty of habitat distribution. Isobaths are based on a 5 × 5 smoothing of the SDB. (a) Substrate reflectance ($R_{rs}{}^{b}$) image up to 30.2 m of depth. (b), (c), (d) SVM-derived probability of *P. oceanica*, rocky and sandy substrate (%), respectively.

Related errors to the observed detection depth of the deepest *P. oceanica bed* could arise due to two factors. The first concerns the effective penetration depth of imagery which exploits information on the diffuse attenuation coefficient to estimate the maximum depth for which an underwater habitat could be identified by optical remote sensing. A combination of an image-based (Lyzenga, 1978, 1981) and a semi-analytical approach (Lee et al., 2000) yielded a depth penetration of S2 imagery at 30.2 m in 560 nm. Therefore, the observed detection depth lies 2 m beyond the EPD of the green band (560 nm); past the EPD in 560 nm, water column correction could possibly produce erroneous results as light would be reflected only by the water column and not by the substrate. The second factor is related to the SDB estimation using an empirical ratio-based approach. Fig. 4.3.9 indicates that the implemented SDB approach here—coastal aerosol to green ratio algorithm—saturates in depths over 22 m in addition to providing a RMSE of 4.31 m (b1/b3), respectively. This could produce an over-estimation of SDB due to greater water penetration in 443 nm in contrast to 560 nm. The second major finding concerns the statistical advantage of forming the ratio of the b1 with the green band to derived bathymetry in comparison to the more widely implemented blue to green band ratio (Stumpf et al., 2003; Traganos and Reinartz, 2018a).

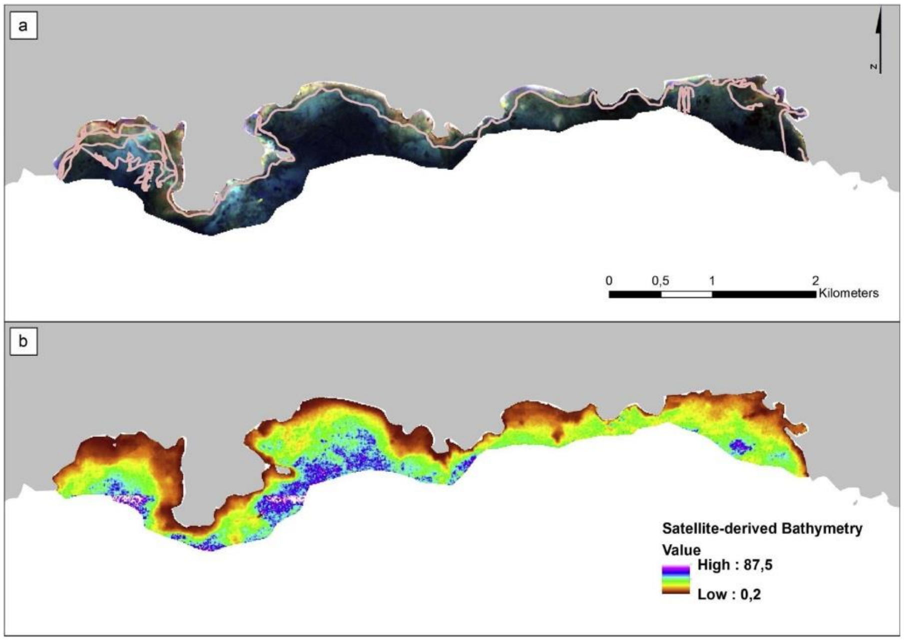


Fig. 4.3.8. Satellite-derived bathymetry (SDB) of the Samaria survey site based on the ratio of coastal aerosol to green band. (a) Deglinted L2A 3-2-1 composite of the survey site. *In situ* depth data (n = 2569) are shown in light pink. (b) SDB for the whole depth range.

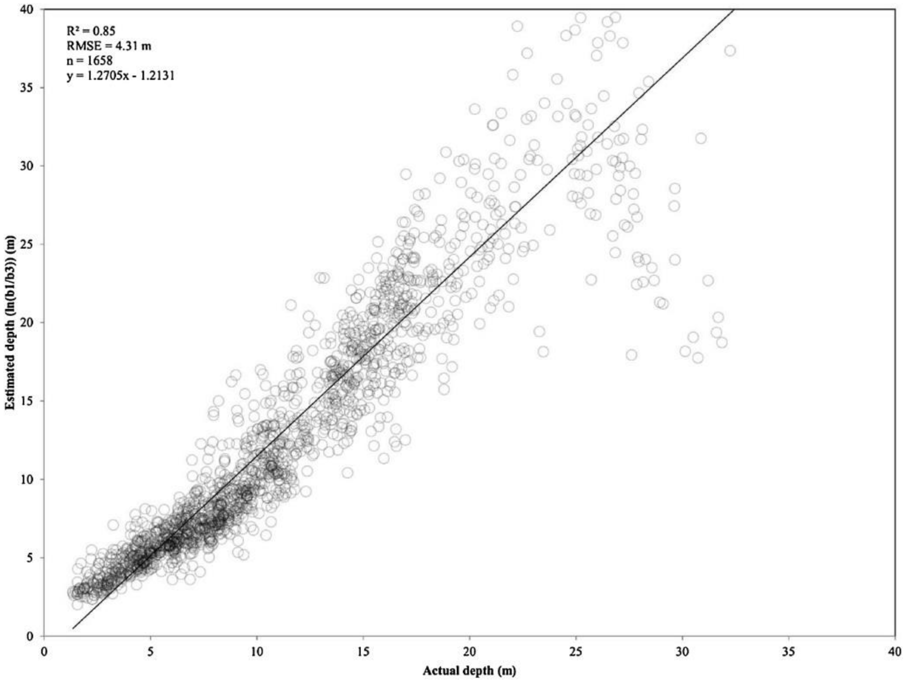


Fig. 4.3.9. Linear regression between SDB using the b1-b3 ratio and the actual depth measurements from the Samaria site. The SDB was implemented in the analytical water column corrections.

This is justified by the improved r-squared value, lower RMSE, and the smaller sensitivity of the former to variation in bottom reflectance than the latter (Fig. 4.3.9). The statistical gains of b1/b3 are clearly attributed to the greater water column penetration of b1 than b2. The errors in both ratio approaches derive from the highly steep and diverse underwater seascape morphology in Samaria. Studies on SDB (Bramante et al., 2013; Collin and Hench, 2012; Pacheco et al., 2015) have shown that sensors with coastal aerosol bands perform better in comparison to those lacking one. Further investigation in different underwater seascapes (low slopes, homogeneous seabed cover) could yield more insights into the capability of S2 for coastal SDB, the requirements on the *in situ* data acquisition, and the exploitation of an operational semi-automated, cloud-based method for basin- to global-scale coastal SDB (Traganos et al., 2018a, 2018b).

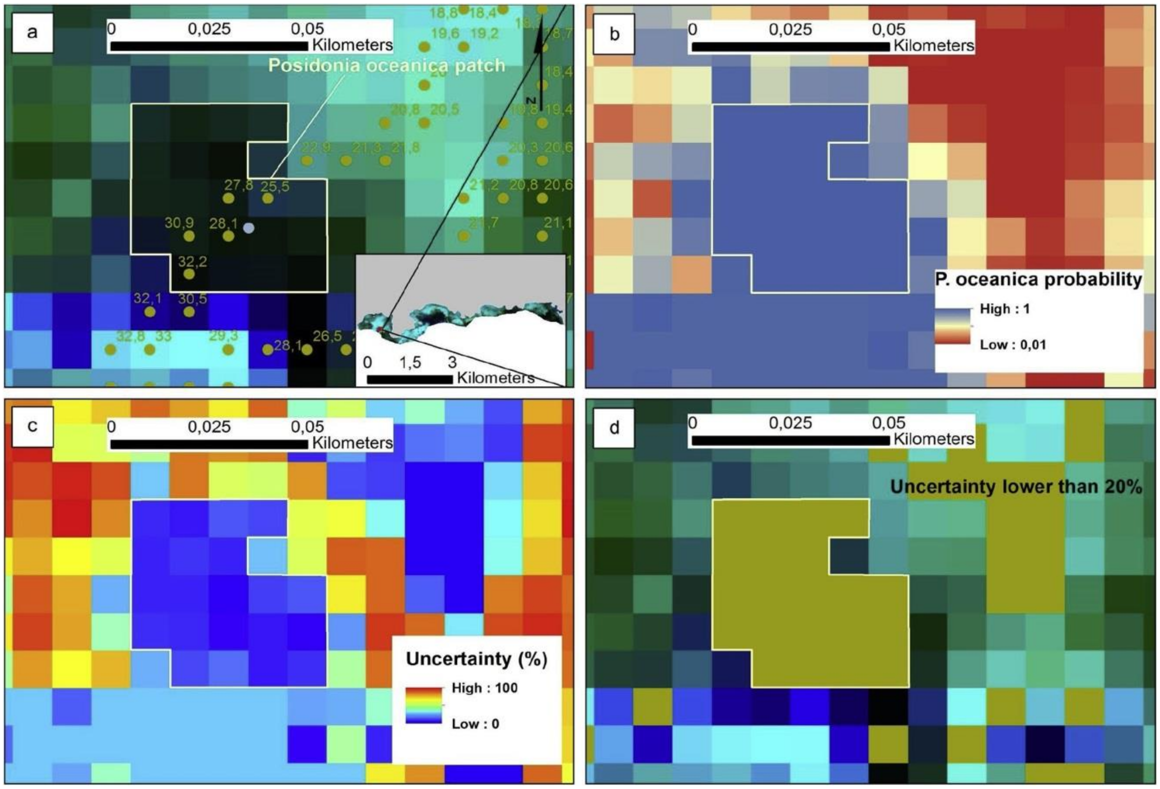


Fig. 4.3.10. Probability and uncertainty of *P. oceanica* seagrass distribution in the Samaria site (south Crete). (a) Rrs^b S2-A 3-2-1 composite with deep dense *P. oceanica* seagrass patch within yellow polygon. The light blue circle indicates the presence of *P. oceanica* seagrass as observed *in situ.* The golden circles depict mean depths of the *in situ* depth measurement points that fall within the same pixels. (b) Probability of occurrence of *P. oceanica* seagrass following SVM-derived classification. (c) General uncertainty (%). (d) Uncertainties of *P. oceanica* seagrass distribution lower than 20% are displayed with golden polygons.

We summarize this section by reporting some additional sensor- and method-wise issues which could affect the accuracy of near-future applications of Sentinel-2 in coastal waters of varying water surface conditions, water column composition, depth, and seabed reflectance. First, regarding corrections, the applied atmospheric correction to derive the L2A product, Sen2cor has been designed for land-based applications and does not estimate water aerosols.

It has demonstrated, however, its efficiency in correcting the atmospheric interference over water in areas without sunglint (B. Pflug, personal communication; 06/09/2017). In the present study, two sites, Schinias and Gavdos, are heavily 'polluted' with sunglint (Fig. 4.3.5). Additionally, the different angle acquisitions of S2 cause granule limits to be visible (Figs. 4.3.1a, c, Fig. 4.3.4a, b, c). This appears to be a S2-specific issue and can hinder image statistical analysis, especially in change detection studies due to the difficulty behind radiometrically-corrected composites. Moreover, the sunglint technique following Hedley et al. (2005) can over-correct very bright reflectances (*e.g.* shallow sands) due to the contribution of substrate NIR reflectance to the water-leaving NIR reflectance. We observe this issue in Fig. 4.3.11d with higher uncertainty (yellow values) over the brighter, shallow sands. Furthermore, the utilised adjacency effect corrections of Sen2Cor perform in- adequately for water surface pixels near the shoreline (Dörnhöfer et al., 2016). Adjacency effect could, therefore, produce

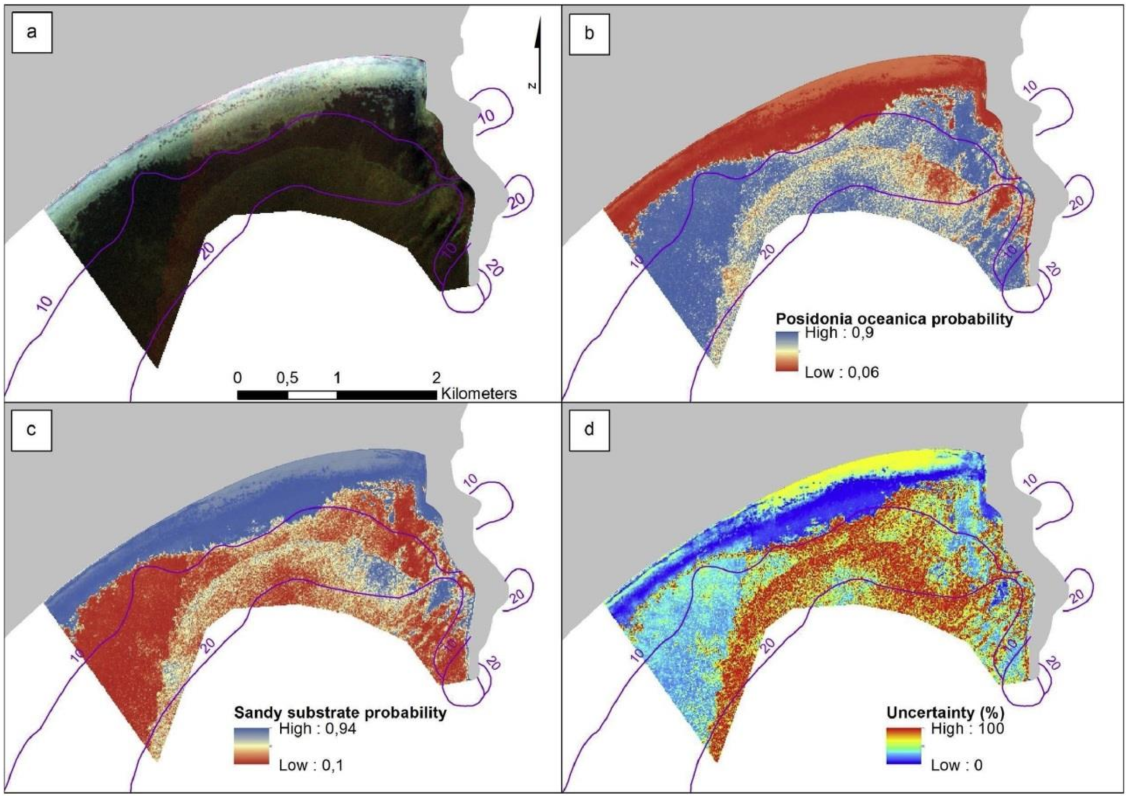


Fig. 4.3.11. Probability and uncertainty of habitat distribution in Schinias. Displayed probabilities are the most accurate results (overall accuracy) of machine learning- based classification. Bathymetric contours derived from the Hellenic Navy Hydrographic Office are displayed in purple. Masked land is displayed in grey colour. (a) Deglinted S2-A 3-2-1 composite of the survey site. (b) SVM-derived P. oceanica probability. (c) SVM-derived sandy substrate probability. (d) Uncertainty (%).

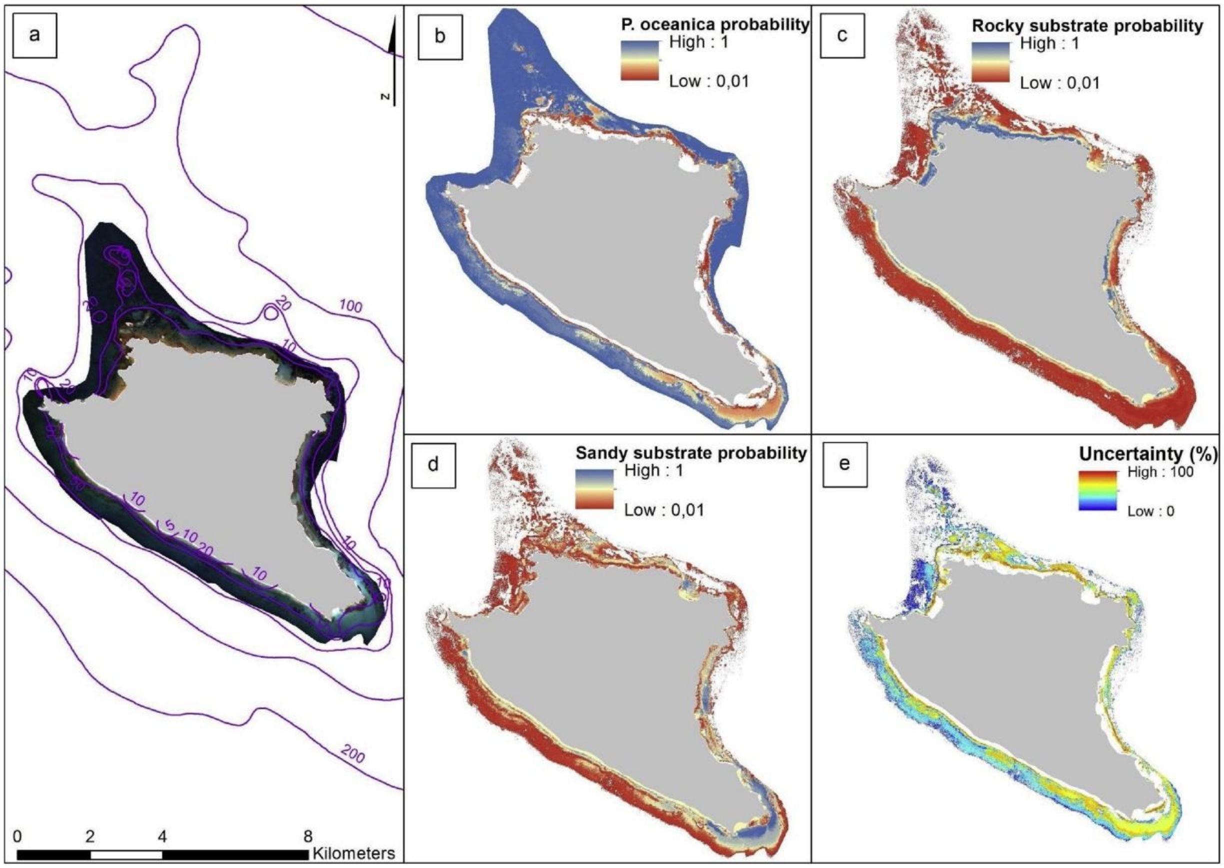


Fig. 4.3.12. Probability and uncertainty of habitat distribution in Gavdos. Displayed probabilities are the most accurate results (overall accuracy) of machine learning-based classification. Bathymetric contours derived from the Hellenic Navy Hydrographic Office are displayed in purple. Masked land is displayed in grey colour. (a) Deglinted S2-A 3-2-1 composite of the survey site. (b) RF-derived P. oceanica probability. (c) RF-derived rocky substrate probability. (d) RF- derived sandy substrate probability. (e) Uncertainty (%).

erroneous reflectance values in both sites in the south of Crete due to their large seabed slopes. As for 15/10/2017, there is a reported 3% error in the spectral response function of Sentinel-2 A over coastal aerosol and blue wave- lengths (both in use here), which can potentially further affect reflectances in coastal water applications, wherein these bands are highly exploited (B. Pflug, personal communication; 06/09/2017). Last but not least, downscaling b1 (Brodu, 2017) enhances its resolution from 60- to 10-m/pixel; this exploits the spatial information of b1 but introduces relationships between it and the other visible bands which could possibly obstruct subsequent corrections.

Random Forests and Support Vector Machines algorithms have been used to derive the probability of occurrence of the three distinctive habitats in the protected areas, while uncertainty of each class has been calculated for each site based on Shannon index (Shannon, 1949) to unveil regions where the presented S2-A-based methodological work- flow may have succeeded or failed. Both machine learning algorithms have been extensively used in image classification under several approaches (Belgiu and Dragut, 2016; Maulik and Chakraborty, 2017). The use of Shannon's entropy as a metric of uncertainty is also not common in the domain of satellite remote sensing. The two algorithms provide different accuracies in each site. In Schinias, SVM provide the highest overall accuracy (S2 file) at 87.18%, with a small mis- classification between the two classes. This misclassification can be possibly attributed to the mixed pixels which occur in the area between seagrass and sand. These can occur mainly in the borders of the meadow as the meadow has a compact form and few patches exit out of it or after a short storm that has transferred fine sand particles on the leafs. In Gavdos, RF produces the highest accuracy of 58.18% which is somewhat poor. This is maybe due to the fact that Gavdos coastal habitats are characterized by a highly fragmented seascape where all three mentioned habitats can exist within one S2 pixel (100 m^2). In Samaria, the highest accuracy in the classification of the seagrass class is produced from SVM classifier at 100% accuracy with an overall ac- curacy of 80% (S2 file). There, mixed habitats also exist in a S2 pixel while the rocky substrate is temporarily covered by sand or small grain of pebbles which amplifies spectral confusion.

The selected field data were collected from field activities and/or from the interpretation of high-resolution aerial (1-m/pixel) and satellite images (2-m/pixel). By using them in an image classification with 10-m/pixel, we expect a mixed-pixel problem—inhomogeneity is introduced from neighboring areas which will be included in the pixel, thus a random pixel will contain spectral information from different types of cover. Per pixel level unmixing can be a solution to overcome such issues (Poursanidis et al., 2015). Another approach could be the collection of large homogeneous datasets corresponding to the same pixel size of Sentinel-2 in the form of polygons or the selection of pixels directly from the image corresponding to homogeneous areas; a comparison with high resolution imagery can support the selection of such areas. Also, the use of object-oriented analysis can be advantageous where spatial and texture infomation is utilized in addition to spectra information (Ma et al., 2017).

All in all, open and free archive of the Sentinel-2 satellite con stellation allows time- and cost-efficient, highly accurate coastal habitat mapping and satellite derived bathymetry. This new wealth of high-quality remote sensing data raises the need for suitable pre-processing, processing and analysis to extract appropriate quantitative and qualitative information. Standardizing methodologies for the two herein explored mapping efforts will enable monitoring from seasonal to decadal scales at regional to global spatial scales; this, in turn, will allow the development of spaceborne quantitative indicators for effective ecosystem management and conservation through the utilization of newly developed cloud environment (Google Earth Engine, DIAS etc).

4.4 Machine Learning-based Retrieval of Benthic Reflectance and *Posidonia oceanica* Seagrass Extent using a Semi-Analytical Inversion of Sentinel-2 Satellite Data

4.4.1 Abstract

In the epoch of the human-induced climate change, seagrasses can mitigate the resulting negative impacts due to their carbon sequestration ability. The endemic and dominant in the Mediterranean *Posidonia oceanica* seagrass contains the largest stocks of organic carbon among all seagrass species, yet it undergoes a significant regression in its extent. Therefore, suitable quantitative assessment of its extent and optically shallow environment are required to allow good conservation and management practises. Here, we parameterise a semi-analytical inversion model which employs above-surface remote sensing reflectance of Sentinel-2A to derive water column and bottom properties in the Thermaikos Gulf, NW Aegean Sea, Greece (eastern Mediterranean). In the model, the diffuse attenuation coefficients are expressed as functions of absorption and backscattering coefficients. We apply a comprehensive pre-processing workflow which includes atmospheric correction using C2RCC (Case 2 Regional CoastColour) neural network, resampling of the lower spatial resolution Sentinel-2A bands to 10m/pixel, as well as empirical derivation of water bathymetry and machine learning-based classification of the resulting bottom properties using the Support Vector Machines. SVM-based classification of benthic reflectance reveals ~300 ha of *P. oceanica* seagrass between 2 and 16 m of depth, and yields very high producer and user accuracies of 95.3% and 99.5%, respectively. Sources of errors and uncertainties are discussed. All in all, recent advances in Earth Observation in terms of optical satellite technology, cloud computing and machine learning algorithms have created the perfect storm which could aid high spatio-temporal, large-scale seagrass habitat mapping and monitoring, allowing for its integration to the Analysis Ready Data era and ultimately enabling more efficient management and conservation in the epoch of climate change.

4.4.2 Materials and Methods

Study site

The broader study site here contains ~12657 submerged hectares within the southeast Thermaikos Gulf, NW Aegean Sea, Greece (Fig. 4.4.1). The prevailing climatic, oceanographic, hydrographic and sedimentological conditions of this gulf, as well as the various socioeconomic activities in its waters have been thoroughly summarised elsewhere (Traganos and Reinartz, 2018a). The Thermaikos Gulf is thriving in two types of seagrasses in depths between 1.4 and 16.5m: the endemic and dominant in the Mediterranean, *Posidonia oceanica* and the *Cymodocea nodosa* species (Traganos et al. 2017; Traganos and Reinartz, 2018a; Traganos

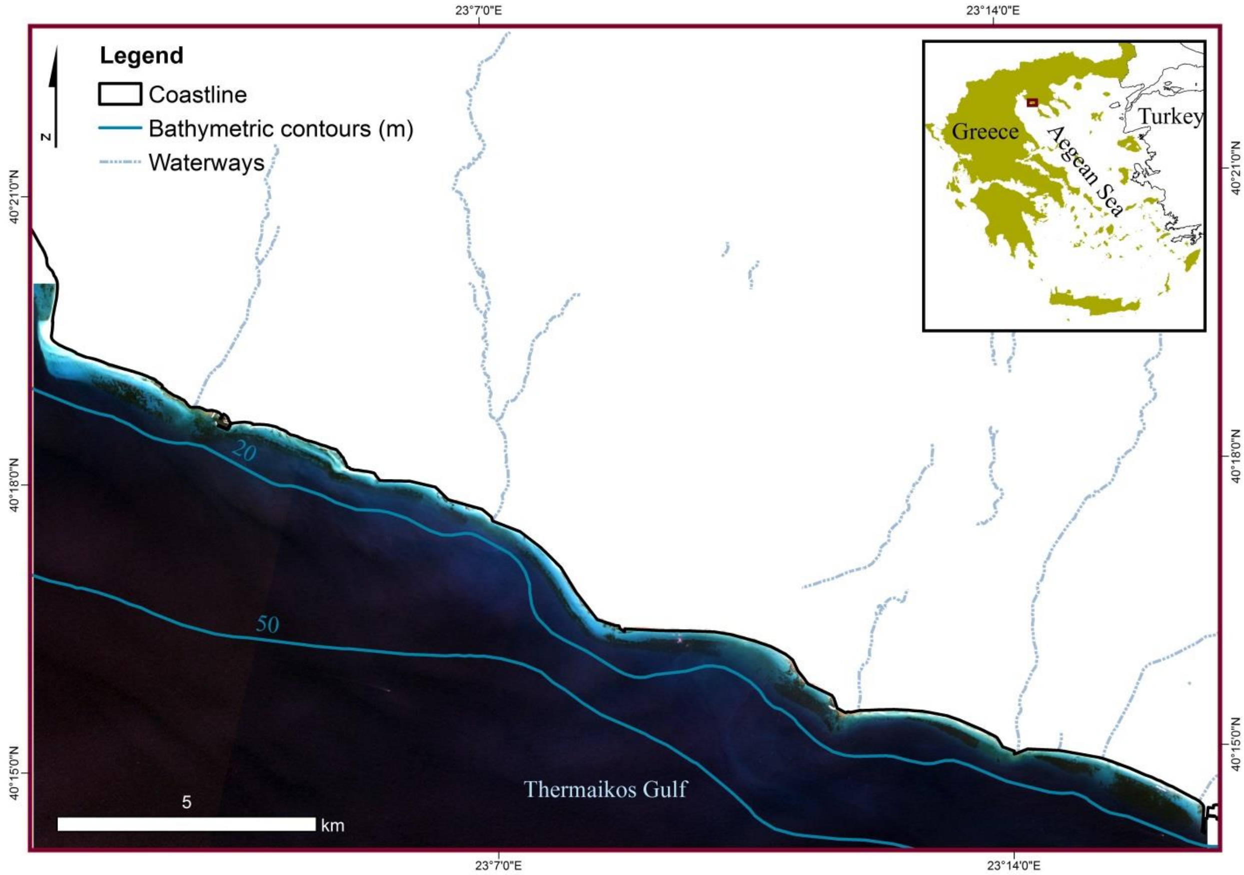


Fig. 4.4.1. Location of survey site within the Thermaikos Gulf and the Aegean Sea, Greece (inset map). The displayed Sentinel-2A image is a top-of-atmoshpere (toa) 4-3-2 composite in UTM (zone 34) system/WGS84 projection. The depicted bathymetric contours are provided by the Hellenic Navy Hydrographic Service (HNHS). The displayed waterways are provided by the OpenStreetMap (http://download.geofabrik.de/europe/greece.html).

and Reinartz, 2018).

Data

Satellite data

We used Copernicus Sentinel-2A (S2-A) satellite data in the present study. Along with Sentinel-2B, they comprise a system of twin-polar orbiting satellites, Sentinel-2. The first one was launched on 23 June 2015, while the second one on 7 March 2017. European Union's Copernicus programme is the most comprehensive Earth Observation (EO) programme up-to-date. With a free, full and open data policy, this EO programme envisages to tackle global challenges in its six thematic areas: land, marine, atmosphere, climate change, emergency management and security. The Sentinel-2 constellation carries the MultiSpectral Imager (MSI) instrument which images the Earth's surface from a mean altitude of 786 km, in 13 spectral bands, from the visible-near infrared (VNIR) to the short wave infrared (SWIR) spectrum at 3 spatial resolutions (10, 20, 60 m) with a swath width of 290 km. The combined temporal resolution of the twin satellites is 5 days. Sentinel-2's radiometric resolution of 12-bit allows for 4096 brightness levels, which is expected to be beneficial in the delineation of the lower limit of *Posidonia oceanica* seagrass and optically deep water.

The characteristics of the processed and analysed S2-A image are displayed in Table 4.4.1. We downloaded the L1C product of the S2A image from the Sentinel Scientific Data Hub (https://scihub.copernicus.eu/). The selected L1C product consists of 100 km^2 orthorectified and spatially co-registered images in Universal Transverse Mercator (UTM)/ World Geodetic System 1984 (WGS94) projection. The atmospheric correction is conducted with the C2RCC neural network processor through ESA's Sentinel toolbox SNAP. This processor depends on a large database of toa radiances and simulated water leaving reflectances to invert the atmospheric correction spectrum and to retrieve inherent optical properties (IOP) following the training of neural networks. It employs the Hydrolight model for the in-water modelling (Mobley, 1994) which derives all relevant IOPs at 443-nm wavelengths (S2 coastal aerosol wavelength). As C2RCC requires all bands in the same resolution, we also used the S2 resampling processor within SNAP, as well, to resample the coarser resolution bands of 60- and 20-m to 10-m pixel with a bilinear interpolation. In addition to the upsampling, this approach takes into account the particularities of the Sentinel-2 acquisition angles which can degrade S2-based image analysis by rendering visible the steps between neighboring granules (Gascon et al. 2017).

Table 4.4.1. Sentinel-2A image characteristics

Tile ID	S2A_MSIL1C_20170330T092021_N0204_R093_T34TFK_20170330T092849
Acquisition date	30 March 2017
Acquisition time	09.20 am UTC
Solar zenith angle	38.75°
Sensor incidence angle	4.29
Sensor altitude	786 km

Field and auxiliary data

The field data collection is described in Traganos and Reinartz (2018a). We collected points associated with habitats and corresponding coordinates as well as bathymetric data during a 4-day survey in July 2016. Based on these data and our survey experience, we manually digitised polygons on the L1C image, displaying each of the four habitats present in our survey site (*Posidonia oceanica* and *Cymodocea nodosa* seagrasses, sand and photophilous algae-rocks). As we will discuss in section 2.3.3., these polygons were used to train, cross-validate and test the herein implemented machine learning classifier of Support Vector Machines (SVM). In addition to our own acquisition, we utilised bathymetric data from the Hellenic Navy Hydrographic Service (HNHS) to aid the training and validation of the Satellite-derived Bathymetry (SDB) of the Thermaikos Gulf.

Methodology

Here, we utilised the semi-analytical inversion method of Lee et al. (1998, 1999) and adapted it to the implemented satellite data and survey site. Based on the quasi-single scattering theory (Gordon, Brown and Jacobs, 1975), the Hyperspectral Optimisation Process Exemplar (HOPE) model has served, in turn, as the basis on which several optically shallow water inversion algorithms have been developed (Brando et al. 2009; Hedley, Roelfsema and Phinn, 2009; Klonowski, Fearns and Lynch, 2007; Mobley et al. 2005). In HOPE, the remote sensing signals of the bottom and water column are mathematically separated and expressed by a few parameters (e.g. bottom reflectance, water depth, and water absorption and backscattering at a single wavelength). The accuracy of these radiative transfer equations (RTE) always depends on the required simplifications which are inherent in their nature.

Initially, HOPE was developed using Hydrolight (Mobley, 1994) to simultaneously derive water depth, bottom reflectance and the various optical water column constituents from

R_{rs} (above-surface remote sensing reflectance) spectra via optimisation. R_{rs}, defined as the ratio of the water leaving radiance to downwelling radiance just above the water surface, is the fundamental parameter upon which all subsequent inversions are based.

In the following paragraphs, we analytically describe how to derive water column and bottom properties from the atmospherically corrected Sentinel-2 visible bands before mapping the *Posidonia oceanica* seagrass meadows of our survey site. Naturally, in contrast to the 16 bands between 400 - 700 nm used as an environmental input in the development of HOPE, we adopted a simplified approach, tailored to the multispectral band set of S2; five S2-A bands in the visible range were utilised (coastal aerosol, blue, green, red and vegetation red edge bands at central wavelengths of 443.9 ($\lambda_{1)}$, 496.6 (λ_2), 560 (λ_3), 664.5 (λ_4) and 740.2 (λ_5) nm, respectively. The first four bands are the usual suspects in the remote sensing of optically shallow, coastal waters, as they penetrate deeper into the water column, while the vegetation red edge band is used to compensate to a degree for under- or over-correction of the atmospheric interference. Additionally, we employed the Shortwave Infrared (SWIR) band 11 (central wavelength: 1613.7 nm) to mask out all terrestrial features.

Terrestrial mask

To further enhance aquatic features, we masked all terrestrial pixels using the resampled SWIR band 11 at 1610 nm. Xu (2006) showed that the use of bands in the SWIR wavelength retrieves information from water bodies with greater accuracy than the use of Near Infrared (NIR) bands. We applied an Iso Cluster Unsupervised Classification on BOA band 11 to separate terrestrial and aquatic pixels into two classes and subsequently retrieved the latter.

Semi-analytical inversion of water column properties and benthic reflectance

General description

Remote sensing reflectance above surface, R_{rs}, forms the basis for the inversion of sub-surface properties like the water column constituents, depth and bottom reflectance. In optically shallow waters, numerous parameters contribute to the remotely sensed from a spaceborne sensor R_{rs}, such as the absorption and backscattering coefficients of suspended and dissolved matter, water depth and spectral bottom albedo. For vertically and horizontally homogeneous waters and assuming zero contribution from the water column inelastic scattering to the water leaving radiance, R_{rs} is conceptually summarised here as:

$$R_{\mathrm{rs}}(\lambda) = f[a(\lambda), b_{\mathrm{b}}(\lambda), \rho(\lambda), H, \theta_{\mathrm{w}}, \theta_{\mathrm{v}}, \varphi] \quad (1)$$

where $a(\lambda)$ is the absorption coefficient, $b_b(\lambda)$ is the backscattering coefficient, $\rho(\lambda)$ is the bottom albedo, H is the water depth, θ_w is the sub-surface solar zenith angle, θ_v sub-surface sensor viewing angel from nadir, and ϕ is the viewing azimuth angle from the solar plane. For brevity, we will omit the wavelength notation for each wavelength-dependent parameter unless when it is required for clarity.

The resulting reflectance Sentinel-2 composite of the atmospheric correction and resampling is not R_{rs}, rather than the normalised water leaving reflectance, R_{hown} which we converted to R_{rs} using:

$$R_{\mathrm{rs}} = \frac{R_{\mathrm{hown}}}{\pi} \quad (2)$$

Afterwards, we followed the two-fold pre-processing method of Lee et al. (2001), first, to minimise the possible produced noise from low signals in longer wavelengths than 600 nm (Red and Vegetation red edge bands for the herein used Sentinel-2 band set) and, second, to compensate for under- or over-correction of the atmospheric interference by C2RCC. The first part was performed by applying a 3x3 low-pass filter in the two aforementioned bands. For the second part, if we assume that the product of (2) is $R_{rs}{}^{raw}$:

$$R_{\mathrm{rs}}^{*}(\lambda) = R_{\mathrm{rs}}^{\mathrm{raw}}(\lambda) - R_{\mathrm{rs}}^{\mathrm{raw}}(\lambda_5)$$

$$\Delta = 0.0001 + 0.02 R_{\mathrm{rs}}^{*}(\lambda_4)$$

$$R_{\mathrm{rs}}(\lambda) = R_{\mathrm{rs}}^{*}(\lambda) + \Delta \quad (3)$$

This corrected R_{rs} is then used to invert sub-surface properties.
Following Hydrolight simulations, Lee et al. (1999) approximated optically shallow R_{rs} with $\phi = 90°$ as:

$$R_{\mathrm{rs}} = \frac{\zeta r_{\mathrm{rs}}}{1 - \Gamma r_{\mathrm{rs}}} \quad (4)$$

where r_{rs} is the sub-surface remote sensing reflectance, that is the ratio of the water leaving radiance to downwelling radiance just below the water surface. The denominator $1-\Gamma r_{rs}$ accounts for the internal reflection of the air-water interface, while ζ is the water-to-air divergence factor (Mobley, 1994). The effects of the internal reflection can be large in turbid or very shallow waters and thus are retained here due to the existence of shallow waters (<1 m) in our survey site. We used $\zeta = 0.5$ and $\Gamma = 1.5$ after the findings of Morel and Gentili (1993) for remote-sensing observation angles.

$$r_{\mathrm{rs}} = r_{\mathrm{rs}}^{\mathrm{c}} + r_{\mathrm{rs}}^{\mathrm{b}}$$

$$\approx r_{\mathrm{rs}}^{\mathrm{dp}}\left[1-\exp(-\left(\frac{1}{\cos(\theta_{\mathrm{w}})}+\frac{D_{\mathrm{u}}^{\mathrm{C}}}{\cos(\theta_{\mathrm{v}})}\right)\kappa H)\right]+\frac{1}{\pi}\rho\exp\left(-\left(\frac{1}{\cos(\theta_{\mathrm{w}})}+\frac{D_{\mathrm{u}}^{\mathrm{b}}}{\cos(\theta_{\mathrm{v}})}\right)\kappa H\right) \tag{5}$$

with

$r_{\mathrm{rs}}^{\mathrm{dp}}$, the sub-surface remote sensing reflectance for optically deep waters equaling here:

$$r_{\mathrm{rs}}^{\mathrm{dp}} \approx (0.084+0.170u)u \tag{6}$$

with

$$u=\frac{b_{\mathrm{b}}}{a+b_{\mathrm{b}}} \tag{7}$$

$$D_{\mathrm{u}}^{\mathrm{C}} \approx 1.03(1+2.4u)^{0.5}, \text{and } D_{\mathrm{u}}^{\mathrm{b}} \approx 1.04(1+5.4u)^{0.5} \tag{8}$$

and,

$$\kappa=\alpha+b_{\mathrm{b}} \tag{9}$$

$$a=a_{\mathrm{w}}+a_{\mathrm{phy}}+a_{\mathrm{g}} \tag{10}$$

$$b_{\mathrm{b}}=b_{\mathrm{bw}}+b_{\mathrm{bp}} \tag{11}$$

In the above-given set of equations, D_u^c and D_u^b are the optical path-elongation factors for water column and bottom photons, respectively, and are approximated as in Lee et al. (1999). Additionally, κ is the attenuation coefficient, a_w is the absorption coefficient of pure seawater, a_{phy} is the absorption coefficient of phytoplankton pigments, and a_g is the absorption coefficient of gelbstoff and detritus. Additionally, b_{bw} is the backscattering coefficient of pure seawater and b_{bp} is the backscattering coefficient of suspended particles. Both u in Equation (7) and κ in Equation (9) are inherent optical properties and altogether, the combination of Equations (5)-(11) provides the expression for R_{rs}.

The left term in Equation (6) comprises the path radiance contribution originating in optically deep waters without any bottom signal, while the right term expresses the bottom contribution following attenuation by the two-way path throughout the water column.

Parameterization

We employed the described model of Equation (5) to solve for the benthic reflectance, r_{rs}^b. The retrieval of such sub-surface properties requires the spectral decomposition of Equation (5). As we utilised Sentinel-2A's bands 1-4 in 10-m resolution for all subsequent calculations, Equation (5) is:

$$Rrs(\lambda_1) = f[a_{\mathrm{w}}(\lambda_1), a_{\mathrm{phy}}(\lambda_1), \alpha_{\mathrm{g}}(\lambda_1), b_{\mathrm{bw}}(\lambda_1), b_{\mathrm{bp}}(\lambda_1), r_{rs}^{b}(\lambda_1), H\,]$$

$$Rrs(\lambda_2) = f[a_{\mathrm{w}}(\lambda_2), a_{\mathrm{phy}}(\lambda_2), \alpha_{\mathrm{g}}(\lambda_2), b_{\mathrm{bw}}(\lambda_2), b_{\mathrm{bp}}(\lambda_2), r_{rs}^{b}(\lambda_2), H\,]$$

$$Rrs(\lambda_3) = f[a_{\mathrm{w}}(\lambda_3), a_{\mathrm{phy}}(\lambda_3), \alpha_{\mathrm{g}}(\lambda_3), b_{\mathrm{bw}}(\lambda_3), b_{\mathrm{bp}}(\lambda_3), r_{rs}^{b}(\lambda_3), H\,] \qquad (12)$$

$$Rrs(\lambda_4) = f[a_{\mathrm{w}}(\lambda_4), a_{\mathrm{phy}}(\lambda_4), \alpha_{\mathrm{g}}(\lambda_4), b_{\mathrm{bw}}(\lambda_4), b_{\mathrm{bp}}(\lambda_4), r_{rs}^{b}(\lambda_4), H\,]$$

To invert Equation (12) and obtain r_{rs}^{b}, we implemented an additional series of empirical and semi-analytical equations to mathematically solve for the unknowns and ultimately derive the bottom properties. It is important to note here that we slightly adapted equations, when needed, to match the central wavelengths of Sentinel-2A bands.

The absorption coefficient of phytoplankton pigments, $a_{\mathrm{phy}}(\lambda)$, is estimated using the single-parameter model of Lee (1998):

$$a_{\mathrm{phy}}(\lambda) = [\alpha_{\mathrm{o}}(\lambda) + \alpha_1(\lambda)\ln\big(a_{\mathrm{phy}}(\lambda_1)\big]\, a_{\mathrm{phy}}(\lambda_1) \qquad (13)$$

with $a_{phy}(\lambda_1)$ yielded by C2RCC processor. Empirically-derived values of a_{o} and a_1 were taken from Table 2 of Lee et al. (1999).

Following Bricaud, Morel and Prieur (1981) and Carder et al. (1999), the absorption coefficient of gelbstoff and detritus, $a_{\mathrm{g}}(\lambda)$ is:

$$a_{\mathrm{g}}(\lambda) = a_{\mathrm{g}}(\lambda_1)e^{-0.015(\lambda-440)} \qquad (14)$$

with $a_g(\lambda_1)$ also estimated by C2RCC.

The backscattering coefficient of suspended particles, $b_{\mathrm{bp}}(\lambda)$ equals:

$$b_{\mathrm{bp}}(\lambda) = b_{\mathrm{bp}}(\lambda_2)\left(\frac{550}{\lambda}\right)^{Y} \qquad (15)$$

with

$$b_{\mathrm{bp}}(560) = 8R_{\mathrm{rs}}(\lambda_4) \qquad (16)$$

after Dekker et al. (2011). Y expresses the spectral shape parameter of particle backscattering. We used the default value of 0.5 for all the pixels here (Lee et al. 2001).

Satellite-derived Bathymetry (SDB)

The final unknown to be determined before the inversion of bottom reflectance is the water depth, *H*. Estimations of bathymetry in optically shallow waters is vital to decrease reflectance changes which are attributed to variable depth and water column attenuation. Derivation of bathymetry over dense *P. oceanica* seagrass beds can be challenging due to their small reflectance, caused by shading within the canopy (Dekker et al. 2006), which often render the beds to reflect less light than the neighbouring optically deep waters. In addition, these beds may appear deeper than a seabed of higher reflectance (e.g. sand) at the same true depth. Researchers have traditionally derived bathymetry using empirical regression of satellite band ratios (Dierssen et al. 2003; Lyzenga, 1978; Mishra et al. 2005; Stumpf et al. 2003; Traganos and Reinartz, 2018a). Empirical SDB approaches are valid only within the extent of their applied survey site, requires less data processing, however they are expected to be less accurate than semi-analytical and analytical approaches (Brando et al. 2009; Hedley, Roelfsema and Phinn, 2009; Klonowski, Fearns and Lynch, 2007; Lee et al. 1999; Mobley et al. 2005).

Site-specific, empirical SDB algorithms exploit the assumption that the ratio of attenuation coefficients for a pair of two bands remains the same independent of the bottom type and water depth (Lyzenga, 1978). Here, we regressed the log-transformed ratio of sub-surface remote sensing coastal aerosol and green reflectance, after Equation (5), against *in situ* measured depths to develop an exponential relationship:

$$H = 0.4102e^{1.3814x} \tag{17}$$

$$x = \ln\left(\frac{r_{\mathrm{rs}}(\lambda_1)}{r_{\mathrm{rs}}(\lambda_3)}\right) \tag{18}$$

which explained >86% of the variation ($p < 0.001$) in estimated bathymetry in 43 points (Fig. 4.4.2) between 0.8 – 25 m with a root-mean-square error (RMSE) of 2.2 m. It is noteworthy that from the 43 points, 2 were over rocky seabed with algae, 4 over *Cymodocea nodosa* seagrass, 18 over sand, and 19 over *Posidonia oceanica* seagrass. We employed Equation (17) to estimate *H* pixel-wise for the optically shallow extent of

our survey site, an extent taken from the observations of Traganos and Reinartz (2018a). To remove any potential noise resulting from the band rationing, we applied a 5x5 median filter on the SDB result. Finally, we validated the median filtered SDB using 22 additional *in situ* depth estimations.

Machine learning-based classification of Posidonia oceanica seagrass beds

Following the correction of Sentinel-2A bands 1-4 to the benthic reflectance as if the satellite MSI instrument was directly sensing the seabed, we applied the machine learning classifier of Support Vector Machines (SVM) at all six stages of reflectance correction (R_{toa}, R_{hown}, R_{rs}^{raw}, R_{rs}, r_{rs}, r_{rs}^{b}) to map the per-pixel distribution of *P. oceanica* seagrass meadows in our survey site and test the contribution of each one of the correction steps. SVM are a group of theoretically superior, universal machine learning algorithms based on the statistical learning theory (Vapnik, 1995). The aim of SVM is to differentiate two classes by fitting an optimal hyperplane to the training data of each class in a multi-dimensional space. Support vectors are the data points closer to the separating hyperplane and they are the only ones which define it. Researchers have employed efficiently SVM to solve both linear and, more significantly, non-linear classification problems. In the latter case, the use of a kernel function in a higher dimensional feature space classifies more accurately the training data by the better fit of a linear hyperplane to the new data distribution. In this study, we utilised the Gaussian radial basis function kernel (RBF):

$$K\left(x_{\mathrm{i}}, x_{\mathrm{j}}\right) = \exp(-g|x_{\mathrm{i}} - x_{\mathrm{j}}|^2) \tag{19}$$

where g defines the width of the RBF and x_i and x_j represent feature vectors. Moreover, one has to also specify the regularisation parameter C, which regulates the degree of acceptable misclassification errors, hence limiting the influence of particular support vectors. We also conducted a two-dimensional grid search inside a preset range to find an optimal pair of values for g and C. The accuracy of this grid search is monitored internally by the k-fold cross-validation.

To classify, validate and test the presence of *P. oceanica* seagrass, we split our initially chosen 2510 pixels (in the forms of polygons), comprising 1.45% of the whole amount of image pixels, into three parts: a) the training set, where we parameterised multiple SVM algorithms, b) the cross-validation set, where the parameterised algorithms of (a) were cross-validated based on their performance, with the best performing being chosen, and finally, c) the test set, where the best performing SVM algorithm was tested for its accuracy on the whole image. Both the training and cross-validation set contained 80% of the 2510 pixels with the number of cross-validation folds, k, determining the final amount of both sets. For example, if k is 4, then with the 80% of the pixels allocated for the training and cross-validation phase, the latter will employ 20% (80/4) for every internal cross-validation, while the former will employ the remaining 60%. The test set used the remaining 20% of the pixels. Finally, we ran all SVM experiments using the LIBSVM library (Chang and Lin, 2013) in EnMAP-Box (van der Linden et al. 2015).

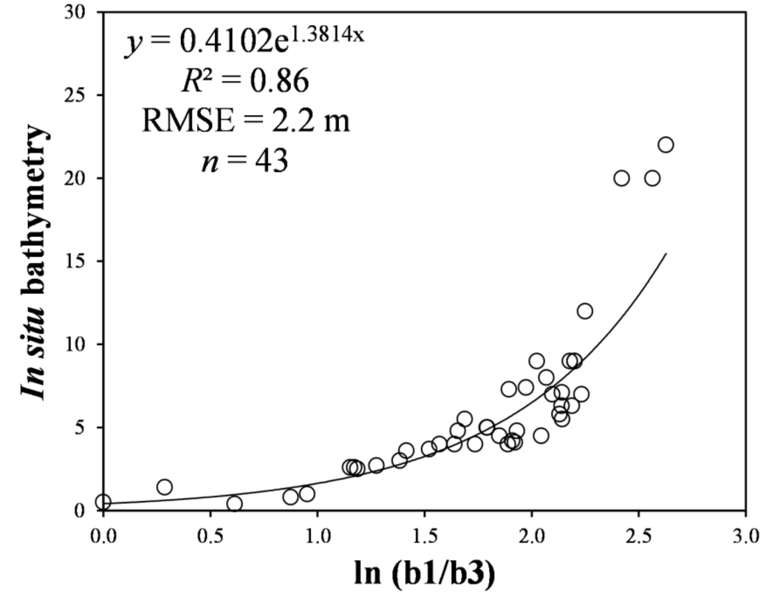


Fig. 4.4.2. Exponential regression between the log-transformed ratio of sub-surface remote sensing coastal aerosol (443.9 nm) and green (560 nm) reflectances, and in situ measured bathymetry from the Thermaikos survey site. The displayed exponential algorithm was used to empirically derived bathymetry in the present study.

We evaluated the SVM classification results (presence/absence of *P. oceanica* seagrasses) using overall (OA), producer (PA) and user accuracies (UA) per class. While PA is an important statistical metric for the scientist who creates the habitat map, UA is more vital towards the management of the region under study as it provides a quantitative probability of the actual presence of a mapped habitat in this region *i.e. P. oceanica* seagrass beds.

4.4.3 Results

Water column properties

The first step towards the retrieval of bottom properties and the mapping of the extensive *Posidonia oceanica* seagrass meadows in the waters of the Thermaikos Gulf was the pixel-wise inversion of water column properties from a Sentinel-2A satellite image, sensed on the 30 March 2017. In the implemented semi-analytical inversion for optically shallow waters (Lee et al. 1998, 1999), both downward and upward diffuse attenuation coefficients, are explicitly expressed as functions of the absorption, a, and backscattering, b_b, coefficients.

As regards to individual absorption coefficients, for a solar zenith angle of 38.75° and a wind speed of approximately 0.9 m s^{-1}, $a_{\text{phy}}(\lambda_1)$, the absorption coefficient of phytoplankton pigments at the coastal aerosol wavelength of S2-A, employed to estimate a_{phy} at all other herein used wavelengths, ranged between 0 – 1.84 m^{-1} in the optically shallow waters of the Thermaikos Gulf (Fig. 4.4.3a). The absorption coefficient of gelbstoff and detritus at 443.9 nm, $a_{\text{g}}(\lambda_1)$, used to derive a_{g} at all other used wavelengths (Equation 14), spanned the 0 – 5.52 m^{-1} range (not shown here). Along with a_{w}, absorption coefficient of pure seawater, values of 0.0075 m^{-1}, 0.0191 m^{-1}, 0.0619 m^{-1} and 0.429 m^{-1} at the respective, first four S2-A band wavelengths, we calculated the total absorption coefficients, a after Equation (9), which were 0.01 – 5.53 m^{-1} at 443.9 nm, 0 – 6.5 m^{-1} at 496.6 nm, 0.062 – 1.09 m^{-1} at 560 nm, and, finally, 0.43 – 1.94 m^{-1} at 664.5 nm.

On the other hand, concerning individual backscattering coefficients, $b_{\text{bp}}(\lambda_3)$, the backscattering coefficient of suspended particles at the green S2-A wavelength, which was utilised in Equation (15) to derive all other b_{bp} (Fig. 4.4.3b; Equation (16)) ranged between 0– 0.12 m^{-1} . With estimated b_{bw} values of 0.0037 m^{-1}, 0.0023 m^{-1}, 0.0014 m^{-1} and 0.0006 m^{-1} at respective S2-A coastal aerosol, blue, green and red wavelengths (Morel, 1974), we derived the total backscattering coefficients, b_{b} after Equation (11), between 0.01 – 0.14 m^{-1} at 443.9 nm, 0 – 0.13 m^{-1} at 496.6 nm, 0 – 0.12 m^{-1} at 560 nm, and, finally, 0– 0.11 m^{-1} at 664.5 nm.

Following estimations of both absorption and backscattering coefficients, we derived their pixel-wise sum, κ, attenuation coefficient in the optically shallow extent of the Thermaikos Gulf, which ranged between 0.01 – 5.58 m^{-1} at 443.9 nm (Fig. 4.4.3c), 0 – 6.49 at 496.6 nm (produced by C2RCC), 0.06 – 1.2 nm at 560 nm, and 0.43 – 2.04 at 664.5 nm.

In addition to κ, calculation of a and b_{b}, yielded u, an important inherent optical parameter, after Equation (7), towards the derivation of both the ranges of optical path-elongation factors for water column and bottom photons, D_{u}^{c} and D_{u}^{b}, respectively, (Equation 8) and $r_{\text{rs}}^{\text{dp}}$, the sub-surface remote sensing reflectance for optically deep waters (Equation 6). The values of D_{u}^{c} and D_{u}^{b} were 1.03 – 2.14 and 1.4 – 3.03 in all bands, correspondingly. Median $r_{\text{rs}}^{\text{dp}}$ values over 96581 optically deep water pixels were 0.05 sr^{-1}, 0.007 sr^{-1}, 0.004 sr^{-1} and 0.0006 sr^{-1} at coastal aerosol, blue, green and red central wavelengths, respectively.

Satellite-derived Bathymetry

We used the site-specific exponential algorithm of Equations (17) & (18) to empirically derive pixel-wise bathymetry, H, of the optically shallow extent of our survey site in the Thermaikos Gulf (Fig. 4.4.4); the mean depth of the Gulf is 10.98 m. We validated the 5x5 median filtered SDB using 22 *in situ* measurements (of the initial 65 points). These measurements revealed a coefficient of determination (R^2) value of 0.48 with an RMSE of 3.67 m over the whole optically shallow extent (Fig. 4.4.5).

Benthic reflectance

Employing Equations (5)-(18), we performed the pixel-wise, semi-analytical inversion of bottom properties, videlicet, water column correction, of the 10m/pixel coastal aerosol, blue, green and red S2-A bands in the optically shallow waters of the Thermaikos Gulf. This mixture of empirical and semi-analytical equations derived optical water column constituents and water depth per pixel which were used in turn as inputs to output pixel-wise bottom reflectance, r_{rs}^{b}, in 173559 pixels (Fig. 4.4.6a). The retrieved r_{rs}^{b} were 0– 0.02 sr^{-1} at 443.9 nm, 0.001 – 0.04 sr^{-1} at 496.6 nm, 0.005 – 0.076 sr^{-1} and 0.01 – 0.04 sr^{-1} at 664.5 nm over all optically shallow habitats.

Machine learning-based classification of Posidonia oceanica seagrass beds

Fig. 4.4.6 shows the classified by SVM reflectances (3-band composites, 6-fold cross validation) of *P. oceanica* seagrass meadows at the bottom reflectance stage. The respective accuracy assessment is provided in Table 4.4.2. In terms of all stages of reflectance correction, however, the best performing classification was surprisingly the R_{toa} composite with an overall accuracy of 98.8% (0.55% standard error), and PO producer and user accuracy of 97.4% and 99.47%, respectively. It was followed by r_{rs}^{b} which featured decreased OA, PA and UA of 0.79% (lower 0.15% standard error), 2.09% and 0.01%, respectively. We observe that the raw above-surface remote sensing reflectance, R_{rs}^{raw} exhibited the lowest accuracies, while in comparison, the corrected for atmospheric interference one, R_{rs}, and the sub-surface remote sensing reflectance, r_{rs} were both 0.59%, 0.52% and 1.1% more accurate in terms of OA, PA and UA, correspondingly.

Following the elaboration on the split between training and cross-validation data in the second paragraph of 2.3.3., the 6-fold cross-validation means that from the 80% of the data which were allocated for training and cross-validating the individual SVM algorithm, 67% was used for the former and the remaining 13% for the latter step.

Optically shallow remote sensing of the Thermaikos Gulf using Sentinel-2A imagery revealed eight dense *P. oceanica* seagrass meadows which cover an area of ~300 ha between 2 – 16 m. PO beds, therefore, occupied 17.3% of the optically shallow extent of the survey site.

OA is overall accuracy, SE is standard error, PA is producer accuracy and UA is user accuracy. PO represents *P. oceanica* seagrass, while NON-PO represents all other merged classes.

Given herein implemented polygons for the machine learning-based classification of *P. oceanica* seagrass habitat, we extracted 320 pixel values over this particular habitat from all the produced reflectance composites of the present study (e.g. R_{toa}, R_{hown}, $R_{\mathrm{rs}}^{\mathrm{raw}}$, R_{rs}, r_{rs}, $r_{\mathrm{rs}}^{\mathrm{b}}$) (Fig. 4.4.7). We show Fig. 4.4.7 to aid understanding of the spectral behaviour of *P. oceanica*-vegetated seabed across the relevant to optically shallow remote sensing spectrum and relevant corrections. There are several notable characteristics in Fig. 4.4.7:

1. All corrections between the toa and bottom layers featured a reflectance peak at 496.6 nm and decay towards the 664.5-nm wavelength.
2. The benthic reflectance of PO peaked at 560 nm, while was nearly equal at the S2-A blue and red central wavelengths.
3. $r_{\mathrm{rs}}^{\mathrm{b}}$ ranged from 0.03% of the R_{toa} at the coastal aerosol wavelength to 2.1% of it at the green wavelength.
4. PO reflectances at all levels were lower than 12.7% of the initial remote sensing signal.

We also show the mean reflectance of *Posidonia oceanica* seagrass and other co-existing habitats (*Cymodocea nodosa* seagrass, photophilous algae and deep water) for the produced reflectance composite of $r_{\mathrm{rs}}^{\mathrm{b}}$) (Fig. 4.4.8). Habitats in our survey site exhibited different peaks, troughs and variable reflectance across the Sentinel-2 bands which allow their more accurate machine learning-based discrimination.

4.4.4 Discussion

Our main objective in this study was to evaluate the suitability of the Sentinel-2 satellite to quantitatively derive benthic reflectance as well as the distribution of the endemic and dominant in the Mediterranean *P. oceanica* seagrass meadows between 0 – 16m in 1736 optically shallow hectares of the Thermaikos Gulf, NW Aegean Sea, Greece. To reach our

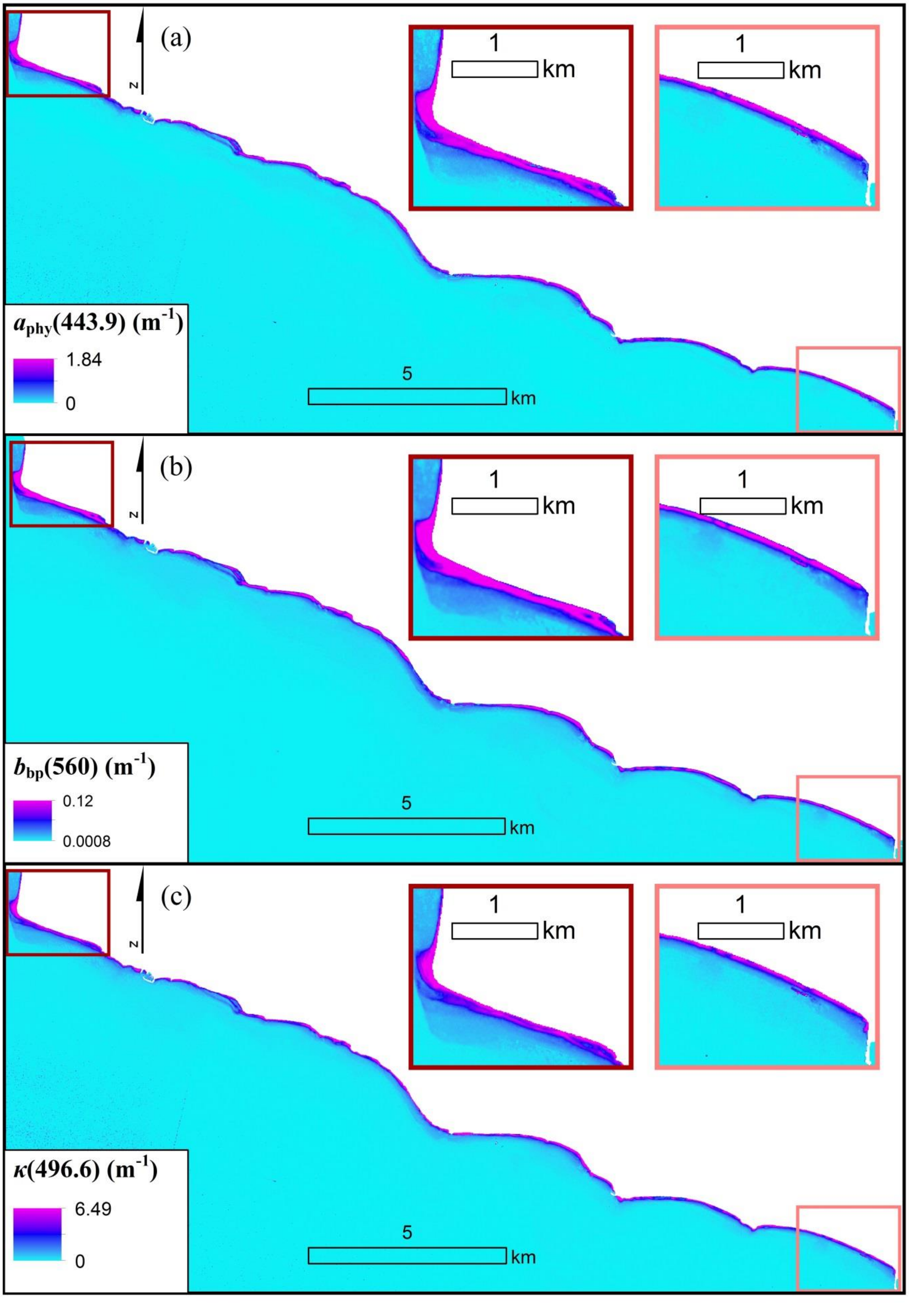

(a)
N
1
km
1
km
a_{phy}(443.9) (m^{-1})
1.84
0
5
km
(b)
N
1
km
1
km
b_{bp}(560) (m^{-1})
0.12
0.0008
5
km
(c)
N
1
km
1
km
κ(496.6) (m^{-1})
6.49
0
5
km

Fig. 4.4.3. Examples of retrieved inherent parameters from both C2RCC and the herein parameterised HOPE (Lee et al. 1999) model in the Thermaikos Gulf (see Fig. 4.4.1). (a) Absorption coefficient of phytoplankton pigments at 443.9 nm, (b) Backscattering coefficient of suspended particles at 560 nm, (c) Attenuation coefficient at 496.6 nm.

objective, we employed a methodological workflow which included atmospheric and air-water interface correction, resampling of the implemented lower resolution Sentinel-2 bands, satellite-derived bathymetry using an empirical site-specific ratio, semi-analytical inversion of optical water column constituents and bottom reflectance (Lee et al. 1998, 1999) and, finally, the machine learning algorithm of Support Vector Machines (Vapnik, 1995).

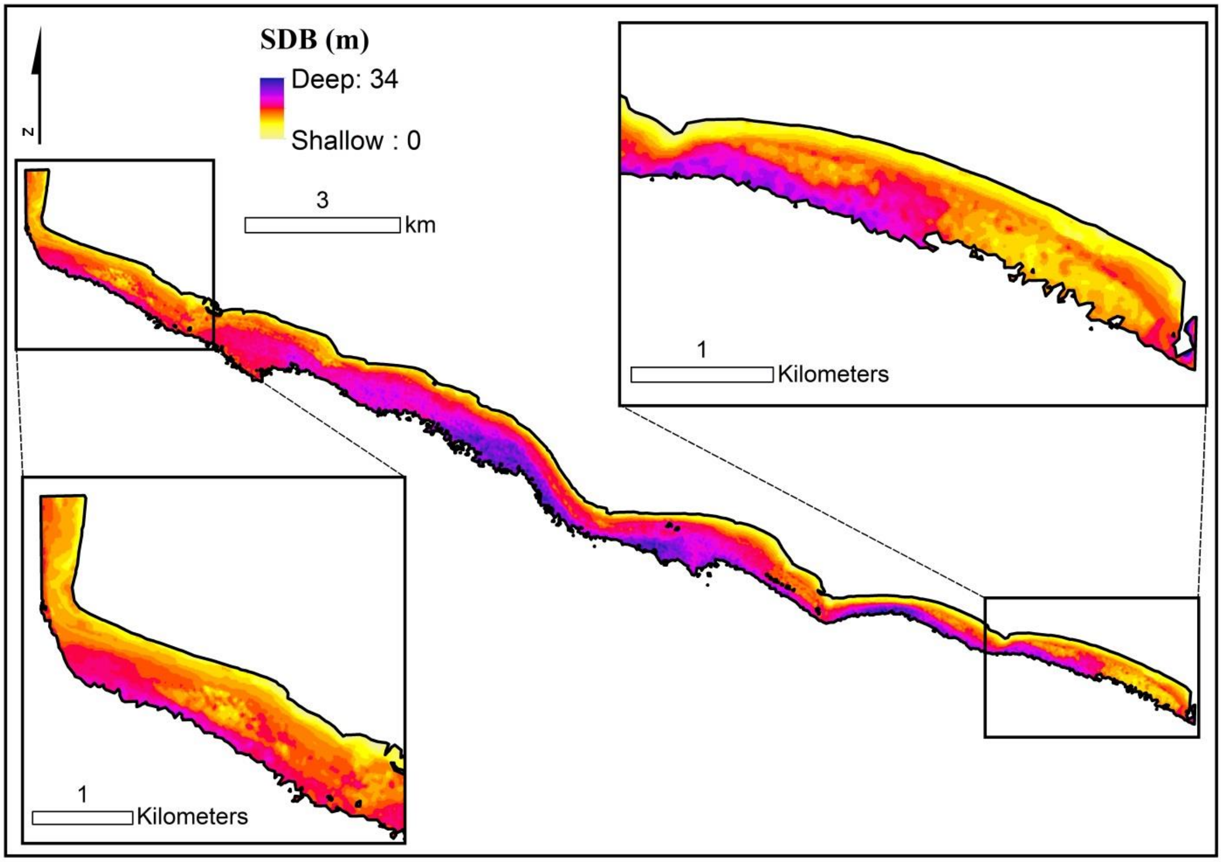


Fig. 4.4.4. 5x5 median-filtered, Satellite-derived bathymetry (H) of the Thermaikos Gulf (see Fig. 4.4.1) based on the empirical regression of Equations (17)-(18) shown in Fig 2. The displayed extent follows the application of an optically shallow mask until 16 m after Traganos and Reinartz (2018a).

Sentinel-2 suitability for P. oceanica seagrass mapping

Ideally, mapping approaches in optically shallow waters, as the herein studied site, should involve suitable atmospheric, air-water interface and water column corrections to derive bottom reflectance prior to the classification step. Here, the SVM classifier on S2-A coastal aerosol, blue, green red bottom reflectances showed a very high producer accuracy of 97.4% and user accuracy of 99.47% on a two-class problem (presence or absence of *P. oceanica* beds) with an RBF kernel and a 6-fold cross-validation. SVM yielded more accurate results following classification of the R_{toa} composite, namely prior to any correction of the reflectance, while the classified benthic reflectance composite, R_{rs}^{b} yielded the second best observed accuracies. This observation raises questions about the necessity of correcting the top-of-atmosphere to the benthic reflectance when one does not need accurate retrieval of the latter—in contrast to the case of forward modelling or existence of accurate field data for optimisation approaches— rather than an accurate classification of the extent of benthic habitat. We assume that application of the same workflow on a different Sentinel-2 or other multispectral/hyperspectral image would produce increasing efficiency from stages 1 to 6 of the successive reflectance corrections. Nevertheless, the herein utilised workflow for *P. oceanica* mapping produced accuracies (OA, PA, UA) greater than 93% at all stages of reflectance correction (Table 4.4.2).

The small misclassification between the two classes could be attributed to mixed pixels of any of the three following types: a) *P. oceanica* seagrass and sand (near or on top of seagrass canopy), b) *P. oceanica* and *Cymodocea nodosa* seagrass, which is also studied near the meadows of the former (Traganos and Reinartz (2018a), and c) *P. oceanica* seagrass and optically deep seabed due to the low reflectance of the former and the inclusion of the latter in the implemented optically shallow mask here. We applied the optically shallow mask following previous findings concerning the depth range of *P. oceanica* presence (Traganos and Reinartz, 2018a). More accurate approaches as the Substratum Detectability Index implemented in the SAMBUCA semi-analytical model (Brando et al. 2009) could have provided better discrimination between optically shallow and deep pixels, and thus mapping of the aquatic habitats in study. Polygons utilised for training, cross-validation and test of SVM were extracted by interpretation of the presence/absence of *P. oceanica* seagrass meadows on the processed S2-A imagery.

In this study, we paired the machine learning of SVM with the semi-analytical model of HOPE to retrieve the benthic extent of seagrass in a Mediterranean Gulf with neighbouring habitats of high variation in their reflectance (shallow sands of high reflectance to deep seagrasses of low reflectance). Modifying the benthic reflectance parameterisation of HOPE, BRUCE semi-analytical model (Klonowski, Fearns and Lynch, 2007) retrieves benthic cover of three substrates (seagrass, sediment and brown algae) along with bathymetry by spectral matching of modelled and image reflectances. While we did not test the applicability of BRUCE model and/or compare it to the SVM approach here, we expect the adaption and estimation of four-component bottom reflectance fractions to improve accuracies of our methodological workflow.

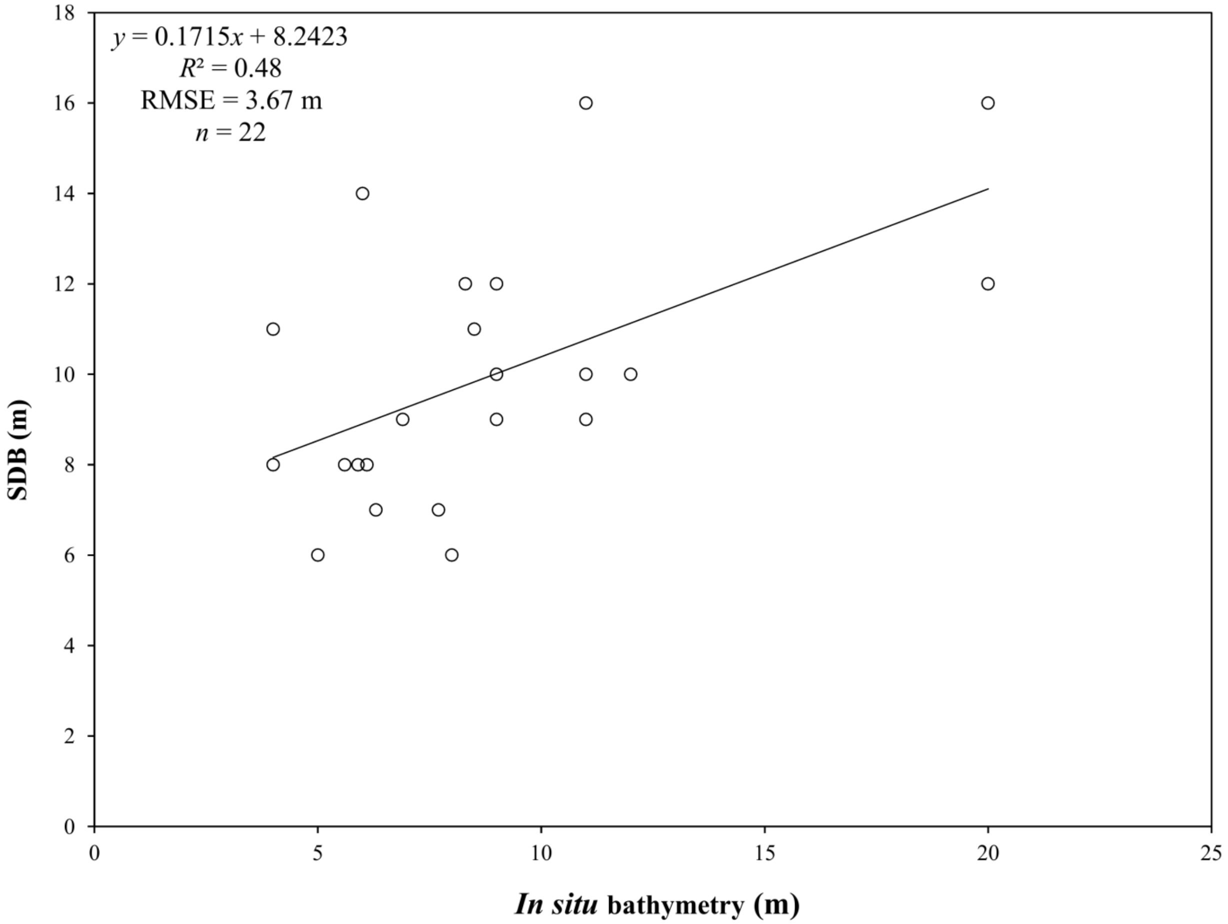


Fig. 4.4.5. Validation of the Satellite-derived Bathymetry (H) shown in Fig. 2 and 4.4.4. H was utilised in Equation (5) to invert bottom reflectances.

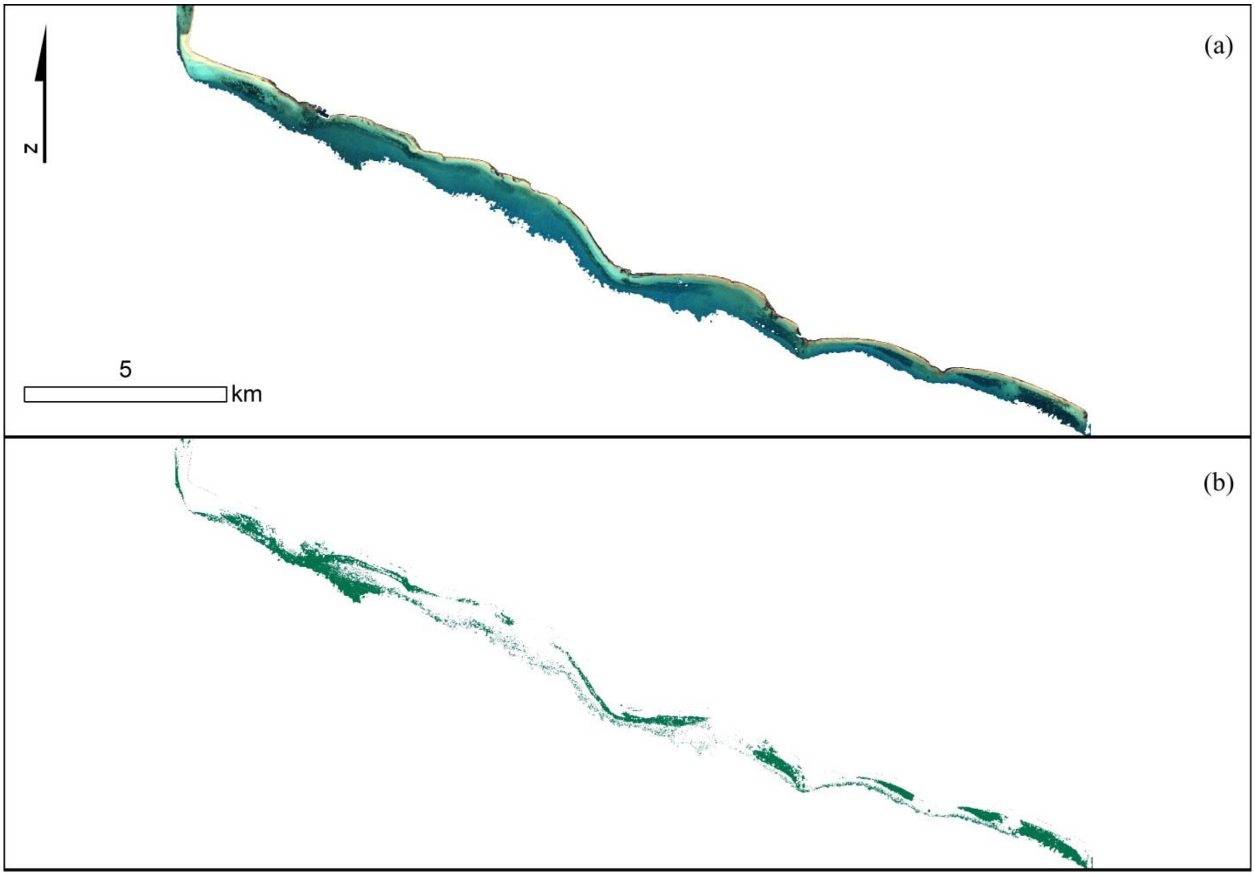


Fig. 4.4.6. Machine-learning-based classification of P. oceanica seagrass—shown in green in panel (b) —bottom reflectances. Accuracy assessment is shown in Table 2.

Table 4.4.2. Accuracies of classified reflectance composites.

Reflectance	Number of folds	Number of bands	OA (%)	SE (%)	Habitat	PA (%)	SE (%)	UA (%)	SE (%)
R_{toa}	6	3	98.81	0.55	NON-PO	99.68	0.22	98.42	0.56
					PO	97.40	6.15	99.47	0.32
R_{hown}	6	3	97.42	0.84	NON-PO	99.68	0.24	96.28	0.84
					PO	93.75	7.43	99.45	0.33
R_{rs}^{raw}	6	3	96.83	0.87	NON-PO	99.04	0.31	95.96	0.88
					PO	93.23	7.76	98.35	0.57
R_{rs}	6	3	97.42	0.84	NON-PO	99.68	0.24	96.28	0.84
					PO	93.75	7.47	99.45	0.33
r_{rs}	6	3	97.42	0.84	NON-PO	99.68	0.24	96.28	0.84
					PO	93.75	7.43	99.45	0.33
r_{rs}^{b}	6	3	98.02	0.73	NON-PO	99.68	0.23	97.19	0.74
					PO	95.31	6.58	99.46	0.33

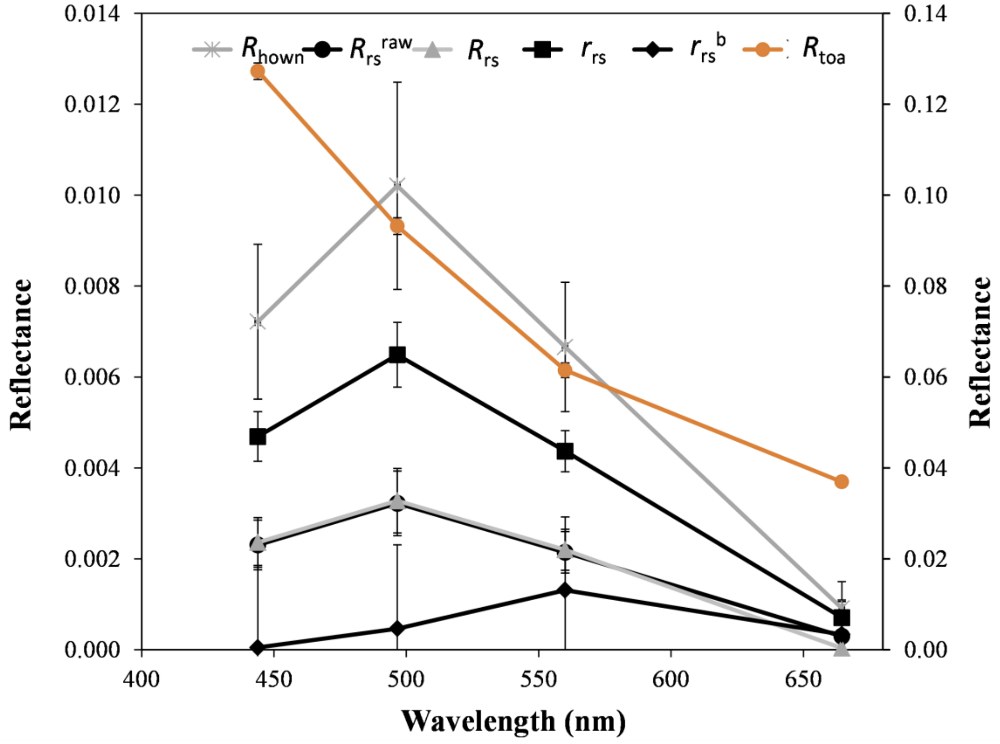


Fig. 4.4.7. Mean spectral reflectances (plus standard deviation) over 320 pixels of classified as P. oceanica seagrass beds for the four first bands of Sentinel-2A (all in 10m/pixel resolution) for all the estimated products in the present study. Description of reflectance denominators are provided in section 2.3.2. The secondary y axis displays reflectance values only for the R_{toa}.

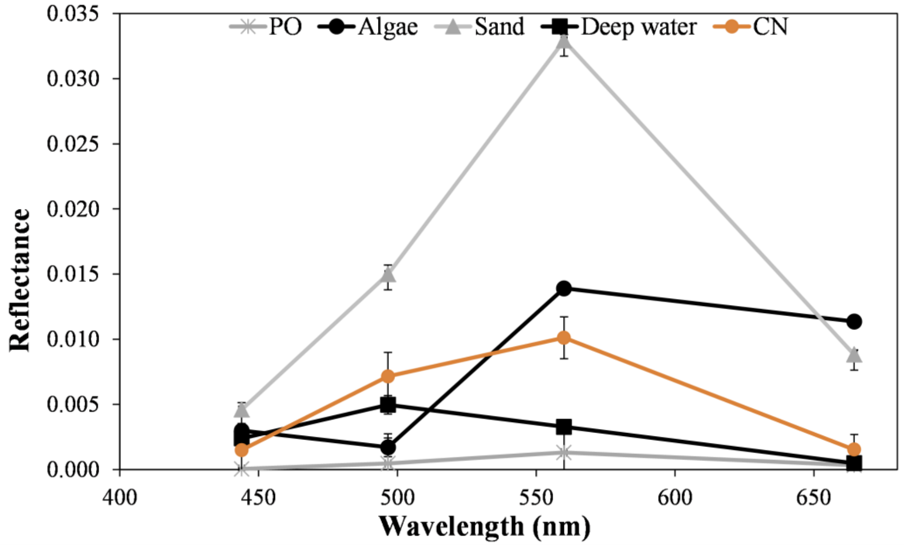


Fig. 4.4.8. Mean spectral reflectances of the classified habitats for the four first bands of Sentinel-2A (all in 10m/pixel resolution) for bottom reflectance product. PO and CN represent Posidonia oceanica and Cymodocea nodosa seagrasses, respectively.

In general, machine learning-based mapping approaches in the remote sensing of optically shallow habitats are rare. In one of these SVM and RF(Random Forests)-based approaches, also directly comparable to the present study, due to the same underwater setting (Thermaikos Gulf) and satellite imagery, the best observed accuracies were somewhat similar (94.4% overall accuracy, 100% producer accuracy, 94.8% user accuracy), albeit using an almost 10-fold smaller set of training and test pixels (228 vs 2510 pixels here) in a smaller optically shallow extent (three vs eight meadows here). This last observation about the amount of training data comes in line with the findings that SVM define well the hyperplane even from a small number and/or data quality (Mountrakis, Im and Ogole, 2011). SVM defines the hyperplane allocating only the support vectors from the whole amount of the training data. To achieve higher accuracies, therefore, one has to select suitable data points which are likely to be support vectors rather than a large data set. This is exemplified by the observed ratios of support vectors to the total amount of training data in use in our study and the study in comparison (Traganos and Reinartz, 2018a). In our approach, only 4.7% of the training data in use are support vectors, while in the latter the respective ratio is 45.9% on a nearly 25 times smaller training set. It is noteworthy that the training data set in our study consisted of polygons (set of pixels) rather than individual pixels as in Traganos and Reinartz (2018a).

In another application in the same coastal environment of the Thermaikos Gulf, SVM yielded an overall accuracy of 68.1%, producer accuracy of 95.1%, and a user accuracy of 70.9% using Planet's multispectral CubeSat-derived imagery and the empirical water column correction of Lyzenga (1978) (Traganos et al. 2017). This highlights the issue of SVM with noisy data, such as the Planet's CubeSat-derived imagery over coastal areas, as they are algorithms not optimised to deal with this problem (Mountrakis, Im and Ogole, 2011). SVM performance in Traganos et al. (2017) increased by the use of the Unmixing-based Denoising (UBD) which has also shown remarkable improvements in hyperspectral scenes over coastal waters (Cerra et al, 2013, 2017). Elsewhere, Eugenio, Marcello and Martin (2015) employed SVM algorithm with an RBF kernel in a similarly parameterised radiative transfer model to ours to map benthic habitats on a WorldView-2 with an overall accuracy of 73% in a depth range between 0 – 25m.

Sentinel-2 suitability for inverting water column properties, bathymetry and benthic reflectance in optically shallow waters

The initial model-driven spectral optimisation technique of HOPE (Lee et al. 1998, 1999) simultaneously retrieves optical water column properties and water depth. Here, we simplified its radiative transfer equations to derive water column and bottom properties from a Sentinel-2 image, while we estimated bathymetry using an empirical regression. The simplification is required due to the limited spectral bands of S2 in comparison to the initially 16 bands between 400-700 nm implemented in the development of HOPE. Moreover, the central wavelengths of S2 bands are slightly different than the utilised wavelengths in the initial radiative transfer equations (e.g. 560 nm for S2-A band 3 vs 550 nm in the $b_{bp}(\lambda_3)$ coefficient) which is expected to be a source of error here. Lack of relevant *in situ* data constrained the assessment of the retrieval error of resulting water column and bottom properties—produced either by C2RCC or the empirical equations. As regards to the atmospheric and in-water processor C2RCC, it is mainly used as the ground segment processor of the Sentinel 3 OLCI for generating Case 2 water-related products and also in the MERIS 4th reprocessing. Toming et al. (2017) has utilised it to map water leaving reflectance, IOPs and water quality parameter (e.g. chlorophyll *a,* suspended matter and dissolved organic matter) in the optically complex Baltic Sea from Sentinel-3 data; they reported realistic spectra for the yielded water leaving reflectance in contrast to bad correlations of the IOPs and water quality parameters with *in situ* data.

Concerning the optical water column properties, derived $\kappa(\lambda_2)$ (Fig. 4.4.3c) by C2RCC displayed higher values (maximum of 6.49 m^{-1}) and much greater variation in the optically shallow waters of the Thermaikos Gulf near the terrestrial part than the bottom portion of the Sentinel-2 image which comprise optically deep waters. This behavior is normal as κ is the sum of a, the total absorption coefficient and b_b, the total backscattering coefficient, and these optical parameters tend to be higher and to strongly vary near the terrestrial side due to the river runoff (Fig. 4.4.1) and higher backscattering of possibly suspended sand particles of the underlying seabed. Additionally, the mean $\kappa(\lambda_2)$ of 0.044 m^{-1} is somewhat lower than the image-derived 0.067 m^{-1} of Traganos and Reinartz (2018a) for the same waters. All in all, existence of relevant field data could further resolve the accuracy of the produced parameters by C2RCC.

Regarding the satellite-derived bathymetry, H, we estimated a 10-m pixelwise bathymetry map of our survey site using an empirical regression between the super-resolved to 10m/pixel coastal aerosol and the green sub-surface remote sensing reflectance, r_{rs}. We utilised the coastal aerosol instead of the previously implemented blue band (Traganos and Reinartz, 2018a) to exploit its greater penetration in the optically shallow water column. Studies related to SDB approaches have shown that sensors with bands centred around 443 nm yield more accurate results in contrast to those lacking a similar spectral band (Collin et al. 2012; Pacheco et al. 2015). On the other hand, we should note two inherent issues in empirical SDB approaches over optically shallow environments with dense seagrass meadows as in the Thermaikos Gulf. The first issue is related to the contribution of the underlying low reflectance of seagrass to the SDB which is apparently large. This issue is two-fold: a) First, it causes an overestimation of SDB due to the somewhat failure of the herein implemented ratio to minimise seagrass contribution, and b) As we implemented an optically shallow mask between 0 – 16m, following the findings of Traganos and Reinartz (2018a), this is expected, according to point (a), to include optically deep pixels in the mask which could cause misclassification in the detection of *P. oceanica* seagrass due to the marginal difference between the beds of the latter and neighbouring optically deep waters. The second issue is more general and is related to the inherent errors due to the simple nature of the empirically-derived bathymetry in comparison to more sophisticated semi- or analytical derivations. Nonetheless, the empirical approach provides a time- and cost-efficient way for estimating SDB given existence of a set of accurate *in situ* bathymetry measurements (Dierssen et al. 2003; Stumpf et al. 2003).

Finally, concerning the suitability of Sentinel-2A to invert benthic reflectances, quantitative assessment (Fig. 4.4.7) reveals that our processing chain succeeded in retrieving the characteristic spectrum shape for seagrass species presented elsewhere (Fyfe, 2003; Thornhaug, Richardson and Berlyn, 2005). In contrast to the correct spectral shape, the magnitude at all four herein wavelengths was nearly ten-fold less at 443.9 nm, 50% lower at 496.6 nm, seven times less at 560 nm, and nearly twice as low at 664.5 nm in comparison to the reported magnitudes of Thornhaug, Richardson and Berlyn (2005). Inversion methods like the algorithm of Lee et al. (1998, 1999) may produce incorrect results due to uncertainties related to the estimation of certain parameters, uncertainties that propagate to the output. Moreover, the assumption of horizontal homogeneity of optical water column parameters by HOPE could produce inaccurate results (Lee et al. 2007). Also, pre-processing steps like the selected atmospheric and sun glint corrections could possibly impact HOPE results (Goodman, Lee and Ustin, 2008).

A few studies have used the inversion scheme of HOPE to accurately derive water column and bottom properties, and water depth simultaneously (Goodman, Lee and Ustin, 2008; Lee et al, 2001, 2007; Mishra et al. 2005). Goodman, Lee and Ustin (2008) employed a hyperspectral AVIRIS imagery of Kaneohe Bay, Hawaii to examine impacts of pre-processing methods (atmospheric and sun glint correction) on a HOPE-based retrieval of four bottom types (sand, coral, algae and flat spectrum). Lee et al. (2001) used the same type of hyperspectral imagery over Tampa Bay, Florida in a HOPE implementation over two types of bottom: sand and seagrass. Finally, Lee et al. (2007) successfully mapped optical parameters, bottom reflectance and bathymetry over a single sandy bottom using EO-1 Hyperion data over Looe Key, Florida, while Mishra et al. (2005) mapped sand, seagrass and corals at the Roatan Island, Honduras using a simplified version of HOPE adapted to IKONOS data to derive an overall accuracy of 80.6%.

Seagrasses in the epoch of Analysis Ready Data (ARD) and climate change

In the human-induced climate change era, seagrasses should be awarded a more central role mainly due to their ability to act as a natural carbon capture technology (Fourqurean et al. 2012). There are existing gaps as regards to the large-scale quantification of the CO_2 absorption by seagrasses. This is immediately related to SDG 13 ('Take urgent action to combat climate change and its impacts') and SDG 14 ('Conserve and sustainably use the oceans, seas and marine resources for sustainable development') (UN, 2017). To achieve both these goals, and thus to derive quantitative information on the carbon absorption and other biophysical parameters indicative of seagrass ecosystem services and health, remote sensing can be one of the most vital elements in the toolbox of ocean science. This is attributed to the capability of the remote sensing technologies to time- and cost-efficiently measure numerous parameters in a non-intrusive manner and high spatio-temporal scale. Recent advances in the development of new multispectral satellites (i.e. Sentinel series, Landsat-8, WorldView series, Planet's Doves), cloud computing platforms (the already launched Google Earth Engine and the soon-to-be-launched DIAS-Copernicus Data and Information Services), machine and deep learning algorithms could all aid high spatio-temporal, large-scale mapping and monitoring of seagrass habitats (Traganos et al. 2018a) which would galvanise two things: a) better management and conservation practices, and b) the integration of seagrasses to the epoch of *Analysis Ready Data* (ARD) and sensor agnostic approaches where satellite data will be already pre-processed to an already high scientific standard and offered as science-ready data immediately after being imaged by a single sensor (Catapult, 2017; Collison and Wilson, 2017; Landsat, 2017) or by multiple sensors (Descartes Labs, 2017). These approaches will decrease the required processing time of satellite data, will increase their usability by a greater number of non-remote sensing experts and will finally enable physical change of relevant parameters to seagrass habitats and coastal ecosystems to be indexed much alike how Google has indexed the Internet.

4.5 Interannual Change Detection of Mediterranean Seagrasses using RapidEye Image Time Series

4.5.1 Abstract

Recent research studies have highlighted the decrease in the coverage of Mediterranean seagrasses due to mainly anthropogenic activities. The lack of data on the distribution of these significant aquatic plants complicates the quantification of their decreasing tendency. While Mediterranean seagrasses are declining, satellite remote sensing technology is growing at an unprecedented pace, resulting in a wealth of spaceborne image time series. Here, we exploit recent advances in high spatial resolution sensors and machine learning to study Mediterranean seagrasses. We process a multispectral RapidEye time series between 2011 and 2016 to detect interannual seagrass dynamics in 888 submerged hectares of the Thermaikos Gulf, NW Aegean Sea, Greece (eastern Mediterranean Sea). We assess the extent change of two Mediterranean seagrass species, the dominant *Posidonia oceanica* and *Cymodocea nodosa*, following atmospheric and analytical water column correction, as well as machine learning classification, using Random Forests, of the RapidEye time series. Prior corrections are necessary to untangle the initially weak signal of the submerged seagrass habitats from satellite imagery. The central results of this study show that *P. oceanica* seagrass area has declined by 4.1%, with a trend of −11.2 ha/yr, while *C. nodosa* seagrass area has increased by 17.7% with a trend of +18 ha/yr throughout the 5-year study period. Trends of change in spatial distribution of seagrasses in the Thermaikos Gulf site are in line with reported trends in the Mediterranean. Our presented methodology could be a time- and cost-effective method toward the quantitative ecological assessment of seagrass dynamics elsewhere in the future. From small meadows to whole coastlines, knowledge of aquatic plant dynamics could resolve decline or growth trends and accurately highlight key units for future restoration, management, and conservation.

4.5.2 Materials and Methods

Study site

The study site is a submerged area of 888 hectares in the eastern Thermaikos Gulf, NW Aegean Sea, Greece (eastern Mediterranean Sea; Fig. 4.5.1). The climate, oceanography, and hydrography of its water have been comprehensively described elsewhere (Poulos et al., 2000; Traganos and Reinartz, 2018a). Satellite-derived mapping has revealed that the Thermaikos Gulf contains extensive beds of varying density of two seagrass species, *P. oceanica* and *C. nodosa*, between 1.4 and 16.5 m of depth (Traganos et al., 2017; Traganos and Reinartz, 2018a).

The coastal system of the eastern Thermaikos Gulf administratively belongs to the Municipality of Nea Propontida. With a population of 36,500 (National Statistics Service of Greece, 2011), Nea Propontida hosts numerous socioeconomic activities including fishing, aquaculture, tourism, agriculture, industry, and trade directly or indirectly influence the marine environment. Furthermore, the administrative region of Central Macedonia, which contains Nea Propontida, features a total fishing catch of 11,869 t (18.5% of the total Greek fishing catch) (National Statistics Service of Greece, 2015). Last but not least, 19.3% of the total citizens of the municipality are employed in the primary sector, while 66 factories and 91 hotels are running in the coastal region (National Statistics Service of Greece, 2011).

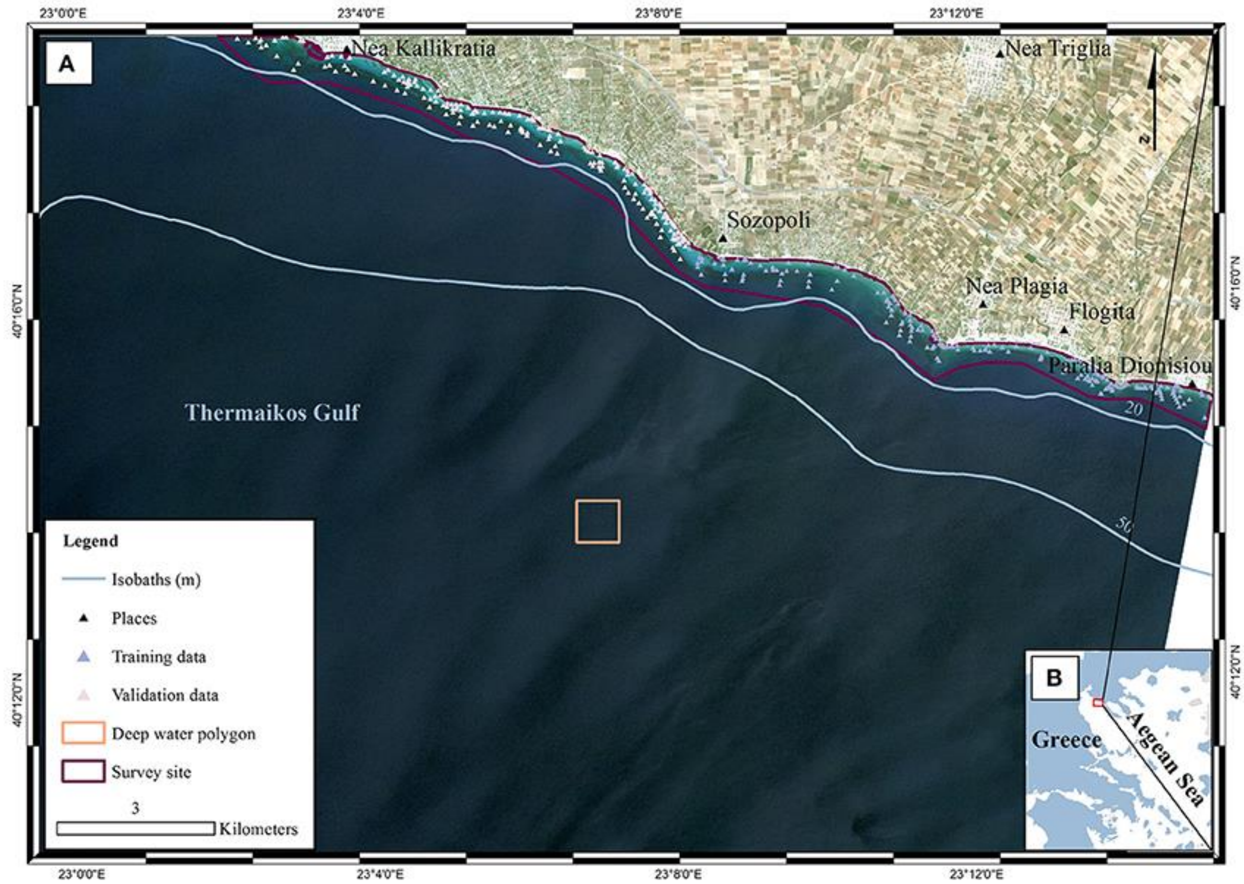


Fig. 4.5.1. Location of survey site within (A) Thermaikos Gulf, (B) Aegean Sea, Greece. The displayed RapidEye imagery is a non-atmospherically corrected, true color (band 1 as

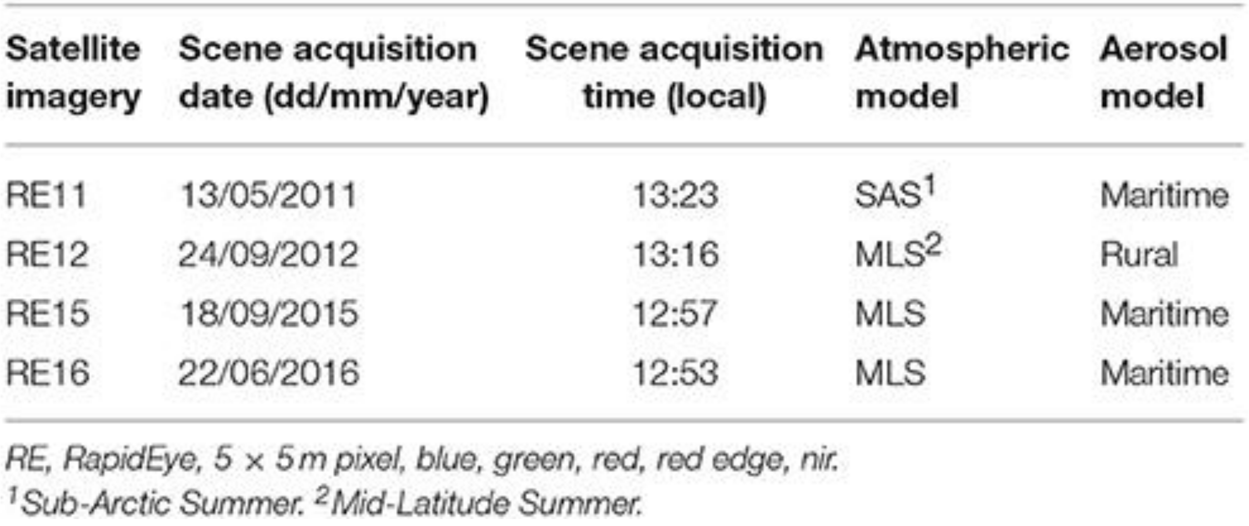
blue, band 2 as green, band 3 as red) composite in UTM (zone 34) system/WGS84 projection. The imagery was acquired on 22/06/2016 (RE16 in text). The red polygon in (B) depicts the location of (A) within the Thermaikos Gulf. The deep water polygon represents a ~160 × 160 pixel window implemented in the water column correction of the image time series as it represents an area with very little water leaving radiance values in all three bands.

Satellite Data

Planet's RapidEye constellation consists of five multispectral (five bands between 440 and 850 nm) satellites which collect high- spatial (5-m orthorectified pixel size) and temporal (daily off- nadir and 5.5 days at nadir) imagery. Although designed to operate for a minimum of seven years, RapidEye satellites have already collected an 8-year image archive. Fourty-seven Level 3A image tiles fell within the extent of our study site. These tiles are individual 25 × 25 km orthorectified imagery products with applied geometric, radiometric, and sensor corrections.

Based on a preliminary visual examination, we selected four from these image tiles (Table 4.5.1) which satisfied optimum conditions for remote sensing of optically shallow environment (e.g., cloud-, sunglint- and skyglint-free, no or low concentration of water column constituents, same season). The four images comprise a time series which spans the years between 2011 and 2016. For ease, we will refer to the four images with the abbreviation RE (stands for RapidEye) and the two last digits from the year of acquisition; RE11, RE12, RE15, RE16.

Table 4.5.1. Characteristics of the high spatial resolution satellite imagery and respective input parameters for running FLAASH module



Satellite imagery	Scene acquisition date (dd/mm/year)	Scene acquisition time (local)	Atmospheric model	Aerosol model
RE11	13/05/2011	13:23	SAS[1]	Maritime
RE12	24/09/2012	13:16	MLS[2]	Rural
RE15	18/09/2015	12:57	MLS	Maritime
RE16	22/06/2016	12:53	MLS	Maritime

RE, RapidEye, 5 × 5 m pixel, blue, green, red, red edge, nir.

[1] Sub-Arctic Summer. [2] Mid-Latitude Summer.

Field Data

The field data collection is described in Traganos and Reinartz (2018a). We collected these field data, namely habitat-related points with associated coordinates and bathymetry data, during a boat-based survey between 10 and 13 July 2016. Furthermore, we added more data points following interpretation of the high resolution RapidEye imagery. We selected data points that have indicated the same habitat within the 5-year span of our time series analysis. Four-hundred data points (Fig. 4.5.1) were used for both training and validation of the machine learning classifier implemented here. In the bathymetry estimation step, we employed the image chronologically closest to the field data collection, RE16, to develop a pixel-based bathymetry map for subsequent use in the water column correction step.

Methodology

To derive quantitative information from coastal image time series using remote sensing, the analyst has to address the interference of the atmospheric, air-water interface, and water column by applying the same processing protocol on all satellite images which comprise the image time series. Fig. 4.5.2, 4.5.3 show a schematic and visual representation of the processed protocol herein, respectively, until the classification step. The pre-classification steps which we followed in the present study included: (1) atmospheric correction to derive at-water surface reflectances without atmospheric interference (Fig. 4.5.4B), (2) bathymetry estimation for use in the water column correction step (Fig. 4.5.4C), and (3) water column correction to derive bottom reflectances without water column interference (Fig. 4.5.4D). The classification step concerned the use of Random Forests (RF), an ensemble supervised classification algorithm which has received small attention in the remote sensing of optically shallow environments. The accuracy assessment of the RF-derived results was performed using the traditional error matrices reporting overall, producer, user, and kappa accuracies (Table 4.5.2). Finally, we conducted the interannual change detection of the two Mediterranean seagrass species, *P. oceanica* and *C. nodosa*, on the basis of area change and related trend throughout the 4 years.

Atmospheric Correction

The first step of the pre-classification procedure was the atmospheric correction. We implemented the Fast Line-of- sight Atmospheric Analysis of Spectral Hypercubes

(FLAASH) algorithm to correct the atmospheric interference on all RE images. The input parameters to run the FLAASH module are described in Table 4.5.1. All aerosol models were set as Maritime type except from the RE12 imagery for which experiments using the same type resulted to negative reflectances, therefore, we used the Rural type. The FLAASH module resulted to at-water surface reflectances, *R,* of all 5 RE bands (Fig. 4.5.4B). The positional accuracy on the initial Level 3A RE image tiles was found to be adequate, hence we performed no additional co-registration on the four RE images which is a necessary step otherwise due to the pixel-based approach of the present study. In addition, due to our preliminary visual examination of the Level 3A products, no evident sunglint was found in the at-water surface reflectance composites.

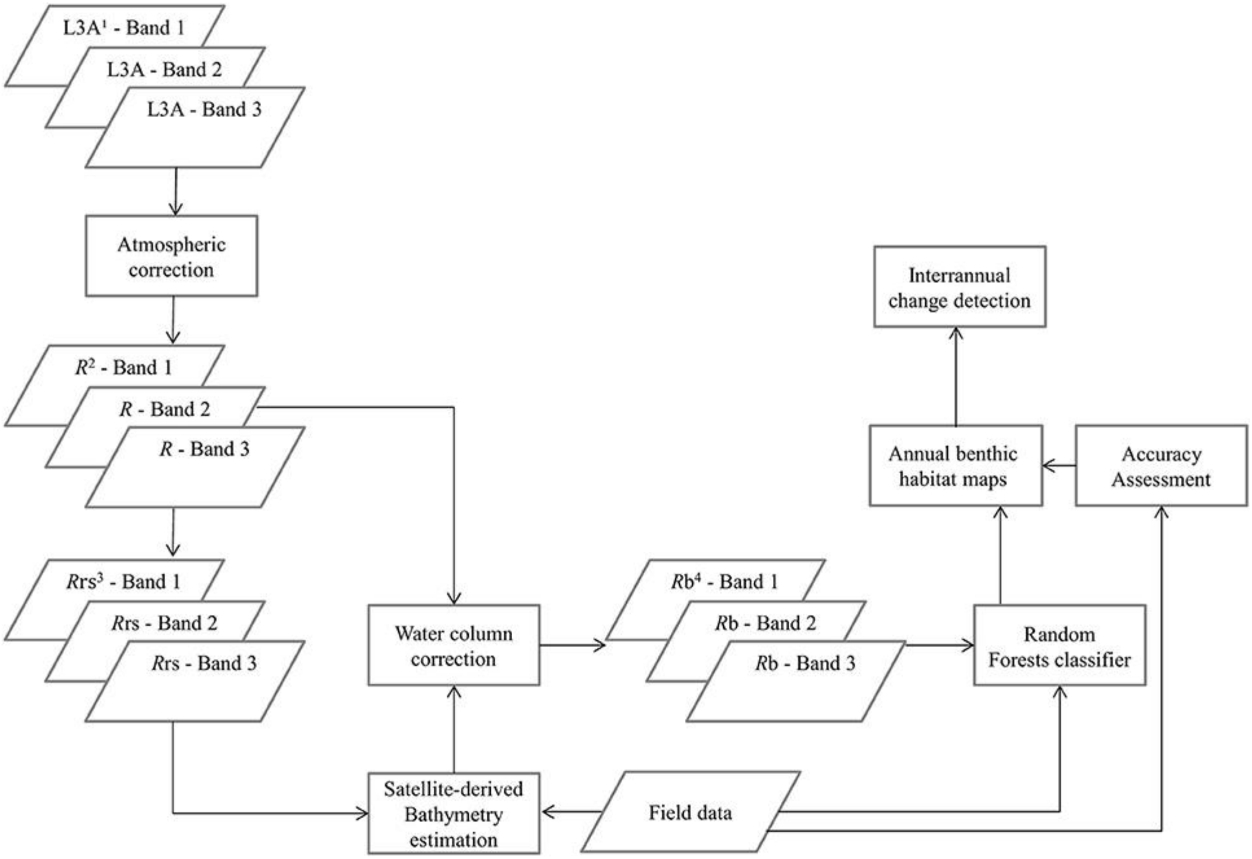


Fig. 4.5.2. Schematic representation of the methodology. [1]L3A ortho products are the initial radiometric, sensor, and geometrically corrected RapidEye images in UTM/WGS1984 projection, [2]*R* represent atmospherically-corrected (FLAASH module), at-water surface reflectances, [3]R_{rs} are remote sensing reflectances, transformed from *R* using Equation (1), [4]*Rb* are water-column-corrected, bottom reflectances using the analytical model of Maritorena et al. (1994).

Bathymetry estimation

Bathymetry knowledge of a coastal site is crucial to reduce reflectance changes due to water column attenuation and variable depth. Differences between reflectances of coexisting submerged habitats can hinder their detection through remote sensing. As a result of the shading which occurs within the canopy, *P. oceanica* seagrass exhibit lower reflectances than its seaward neighbor, optically deep water (Dekker et al., 2006). Coupled with the high reflectances of a submerged sandy substrate, *P. oceanica* would look deeper than the sandy substrate at the same true depth (Traganos and Reinartz, 2017). Remote sensing researchers have developed band ratios to tackle the aforementioned issues and to measure bathymetry (Lyzenga, 1978; Dierssen et al., 2003; Stumpf et al., 2003). The basic assumption of band ratios is that the reflectance ratio in these bands remains constant irrespectively of the submerged environment.

To further eliminate interference at air-water interface, we chose to retrieve pixel-based bathymetry from the remote sensing reflectance, R_{rs}, which is also less sensitive to water column properties (Mobley, 1994). We derived R_{rs} from at-water surface reflectances, R, using

$$R_{rs} = \frac{tR}{Q} \tag{1}$$

where t is the transmittance of E_d, spectral downwelling plane irradiance, and Lu, spectral upwelling radiance, through the air- water interface and was calculated as 0.54 by Mobley (1994). Q factor is the ratio of E_u, spectral upwelling plane irradiance, and L_u just beneath the water surface and depends on the type and depth of the bottom and the wavelength. We chose the π value for Q factor for the calculations of R_{rs}, which is the theoretical value for Lambertian surfaces (Dierssen et al., 2003). We mapped the bathymetry of our survey site using the log-ratio of R_{rs} blue to R_{rs} green at 475 and 555 nm from RE16, respectively. We used RE16 because it is closer chronologically to our *in situ* data acquisition of July 2016. The first two RapidEye bands are attenuated less in the water column than the ones in the red, red edge and NIR. Thus, they comprise the ideal contestants to develop a second-order polynomial after plotting their ratio against *in situ* estimated bathymetry, Z

$$Z = 24.135x^2 - 70.038x + 51.571 \tag{2}$$

$$x = \ln\left(\frac{Rrs(475}{Rrs(555)}\right) \tag{3}$$

which explained >91% of the variation ($p < 0.001$) in estimated bathymetry in 32 points (Fig. 4.5.3) which spanned the whole depth range of habitat presence in our survey site. The site- specific algorithm of (2) was subsequently employed to create a pixel-based bathymetry map (Fig. 4.5.4C). This bathymetry was further smoothed with a 5 × 5 low pass filter to reduce local variation and unwanted noise which would impede water column correction and possibly decrease classification accuracy. Based on the satellite-derived bathymetry and the findings of Traganos and Reinartz (2018a), we applied an optically deep water mask utilizing the depth limit of 16.5 m to enhance submerged features in the classification step.

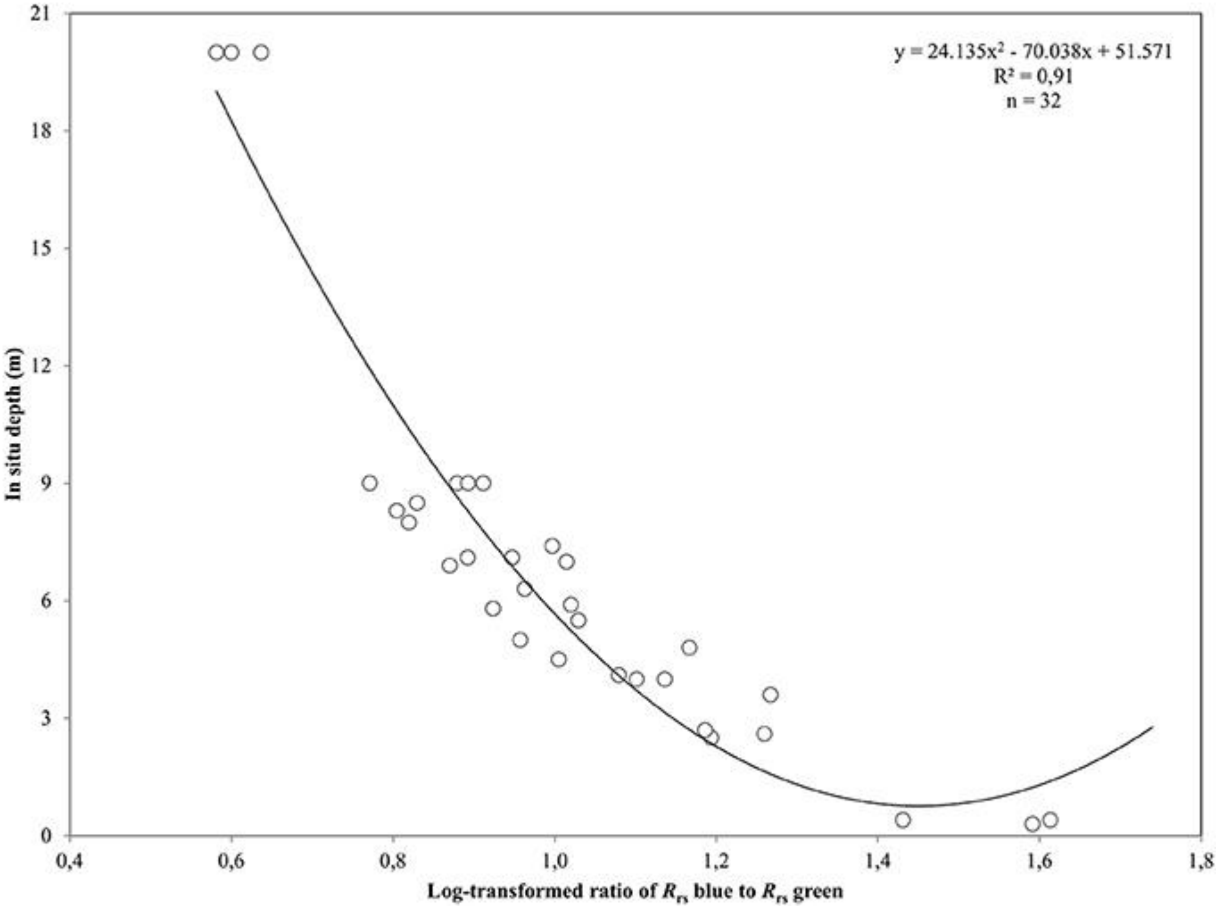


Fig. 4.5.3. Polynomial regression between the log-transformed ratio of blue and green remote sensing reflectances, R_{rs}, and *in situ* depth measurements from the Thermaikos survey site. The shown polynomial equation was implemented to estimate the bathymetry map displayed in (C) of Fig. 4.5.4.

Water Column Correction

The water column correction step is vital to retrieve bottom reflectances from at-water surface reflectances. Here, we employ the approximate analytical solution of Maritorena et al. (1994) for optically shallow water.

$$R(\lambda) = R_{\infty}(\lambda) + (R_b(\lambda) - R_{\infty}(\lambda))\exp(-2K_d(\lambda)Z) \tag{4}$$

$R_{\infty}(\lambda)$ is the reflectance over an infinitely deep water column; $R_b(\lambda)$ is the bottom reflectance (Fig. 4.5.4D); $K_d(\lambda)$ is the operational attenuation coefficient which expresses the attenuation of both upwelling and downwelling stream as these are originating from the seabed and from the water column (Maritorena et al., 1994). Ideally, $R_{\infty}(\lambda)$ and $K_d(\lambda)$ are estimated using *in situ* optical measurements and radiative transfer simulations. In the absence of these, we used image- based techniques and existing measurements (Traganos and Reinartz, 2018a). For the infinitely deep water column reflectance, $R_{\infty}(\lambda)$, we extracted mean values from the deep water polygon of Fig. 4.5.1A from the at-water surface reflectance RE16 composite, mean values which composed $R_{\infty}(\lambda)$ for all four images of the RE time series. For the $K_d(\lambda)$, we used image-based calculated values for the water column corrections of all four dates from Traganos and Reinartz (2018a) who used Bierwirth et al. (1993) approximations in the same area. Bottom reflectances (Fig. 4.5.4D) were calculated for every pixel of the first three bands of the atmospherically corrected, at-water surface reflectance composites of all four RE images. We selected RE bands 1, 2, and 3 at 475, 555, and 658 nm since seagrasses and, generally, underwater habitats cannot be detected by wavelengths past 680 nm due to the great attenuation of pure water (Kirk, 1994). Last but not least, to save valuable space in the remaining text, we will refer to each quantity which is wavelength-dependent without its wavelength notation except when it is needed i.e., to discriminate between two quantities.

Random Forest Classification

The machine learning approach of Random Forests (RF) comprises an ensemble supervised classification algorithm that implements multiple self-learning decision trees to handle collinearity and, more significantly, non-linearity between predictor variables. Developed by Breiman (2001), RF are based on the assumption that different independent tree predictors give wrong predictions in different regions. By combining the results of the predictions, RF improve the efficiency of the model. Every decision tree in the implemented RF algorithm here is trained with a bootstrapped sample of the training data and at every split node, a subset of randomly selected features is utilized. The outputs are then combined by a simple majority vote. Generally, RF are robust against overtraining and noisy data in addition to providing good results with relatively small datasets (Gislason et al., 2006). Three parameters must be set before running the RF classifier: (a) the number of decision trees (k), (b) the number of randomly selected features (n_r), and (c) the split selection. We selected 100 trees to run all RF experiments as they featured the best results out of a plethora of runs with different number of trees. In addition, we chose two for the number of n_r as well as the Gini Index for the measurement of the best split selection. We parameterized and ran all RF experiments using the EnMAP- Box software (van der Linden et al., 2015). All the experiments were performed using the bottom reflectance composites, R_b and 400 training and validation data for all four classes: (a) *C. nodosa* seagrass, (b) *P. oceanica* seagrass, (c) Rocky algae, and (d) Sand. The training and validation data were split equally into 50 data points per class for both training and subsequent validation through accuracy assessment.

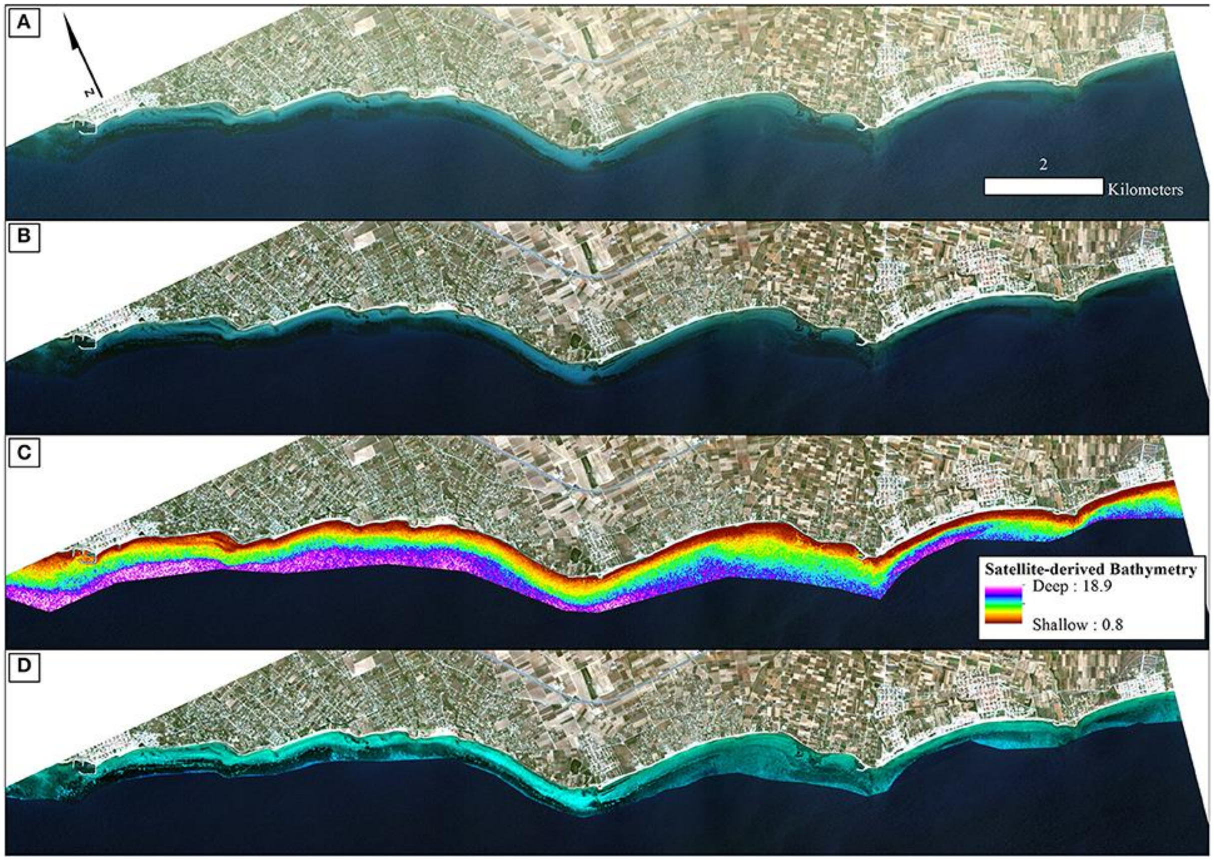


Fig. 4.5.4. Methodological steps from atmospheric to water column correction in order of successive processing. All four panels are true color RapidEye image composites (22/06/2016; RE16 in text) projected in UTM (zone 34) system/WGS84. (A) Non-atmospherically corrected composite. (B) Atmospherically-corrected composite using the FLAASH module. (C) Satellite-derived Bathymetry map of the survey site draped over

the atmospherically-corrected composite of (B) using the site-specific polynomial algorithm of Equation (2) as shown on Fig. 4.5.3. We applied a 5 × 5 low-pass filter on the initial ratio-derived bathymetry (not shown here) to reduce potential noise which would be transferred to the water-column corrected product. (D) Water-column corrected composite following application of the water column correction algorithm of Maritorena et al. (1994) draped over the atmospherically-corrected composite of (B) and masked using the optically deep limit of 16.5 m to enhance bottom features and potentially increase classification accuracies

Accuracy Assessment

We used the error matrices (Table 4.5.2) to validate the results of the Random Forest classifications. As discussed in Random Forest Classification section, 50 data points per class were used to validate the RF classifier. The error matrix contains a square array of rows and columns where each of them represents one habitat class in the classification. Each cell in this matrix is the number of classified training samples, while the rows comprise classified training data and the columns are validation data for the assessment of the classified data. The error matrix outputs the overall, producer, and user accuracy (Congalton, 1991). The overall accuracy is the ratio of the number of correctly classified validation samples to the total number of validation data (200 in our study). On one hand, the producer accuracy expresses the number of correctly classified validation data in one class divided by the total number of validation data in the same class (50 in our study). On the other hand, the user accuracy corresponds to the number of correctly classified validation in one class divided by the total number of validation data that were classified in the same class. Although the producer accuracy is a solid statistical value for the creator of the habitat map (the remote sensing scientist as the case in point), the user accuracy is more vital from a management point of view as it reports the quantitative probability for the tangible presence of the habitat in the studied region i.e., *P. oceanica* and *C. nodosa* seagrass meadows.

4.5.3 Results

Pre-Classification Steps

All pre-classification steps are displayed schematically in Fig. 4.5.2 and visually in Fig. 4.5.4. It is visually apparent that both the atmospheric and water column correction in Fig. 4.5.4B,D, respectively, enhance bottom features following the increase of the seabed spectral variability from Fig. 4.5.4A (initial top-of-the- atmosphere reflectance composite).

After converting the atmospherically-corrected at-water surface reflectances, R to R_{rs} (1), we developed a site-specific polynomial algorithm (2) using the log-transformed ratio of the blue to green RapidEye bands of the RE16 image (3) to map bathymetry, Z, in our site (Fig. 4.5.4C). The depth of the eastern coast of the Thermaikos Gulf spanned the depth range between 0.8 and 18.9 m with a mean depth of 7.7 m and a mean slope of 5.4°. The validation of the Satellite-derived Bathymetry was conducted using 14 *in situ* depth points and revealed an r-squared value of 0.86 with a root mean square error (RMSE) of 2.6 m (Fig. 4.5.5). It is worth noting that from the 32 points used in the bathymetry estimation, 15 were measured above *P. oceanica* beds, 3 over *C. nodosa* beds, 11 over sandy seabed, and 3 over rocky seabed with photophilous algae. On the other hand, from the 14 points used in the bathymetry validation, 8 were over *P. oceanica* beds and 6 over sandy seabed.

Employing Equation (4), we performed water column correction for every pixel in all RE images. Equation (4) takes the at-water surface reflectance, R, the per-pixel Satellite-derived bathymetry, Z, the reflectance of an optically deep column, R_∞, and, finally, the diffuse attenuation coefficient, K_d, as inputs and outputs bottom reflectance, R_b. The reflectances of an optically deep column were determined based on 25,599 pixels within the deep water polygon of Fig. 4.5.1; R_∞ (475) = 0.033, R_∞(555) = 0.024, and R_∞ (658) = 0.017. As mentioned in section Water Column Correction, K_d values were selected for the whole time-series from Fig. 4.1.11 in Traganos and Reinartz (2018a; section 4.1 here); K_d(475) = 0.067, K_d (555) = 0.078, and K_d (658) = 0.134 (in m−1 as they are calculated based on the unitless R and depth).

Random Forest Classification

We employed the Random Forest machine learning classifier on all bottom reflectance images which comprised the studied time series (Fig. 4.5.4D). The results of the random forest classifications are presented in Fig. 4.5.6. The accuracy assessment of the classification results for all images and habitats is presented in Table 4.5.2 with the form of four error matrices, one for each image from RE11 to RE16. All experiments were run using 100 trees. The two Mediterranean seagrasses under study here,
P. oceanica and *C. nodosa*, showed a mean producer accuracy of 89 and 50%, respectively, with a mean user accuracy of 91.6 and 63.1%, correspondingly. From the whole time-series, *P. oceanica* seagrass was

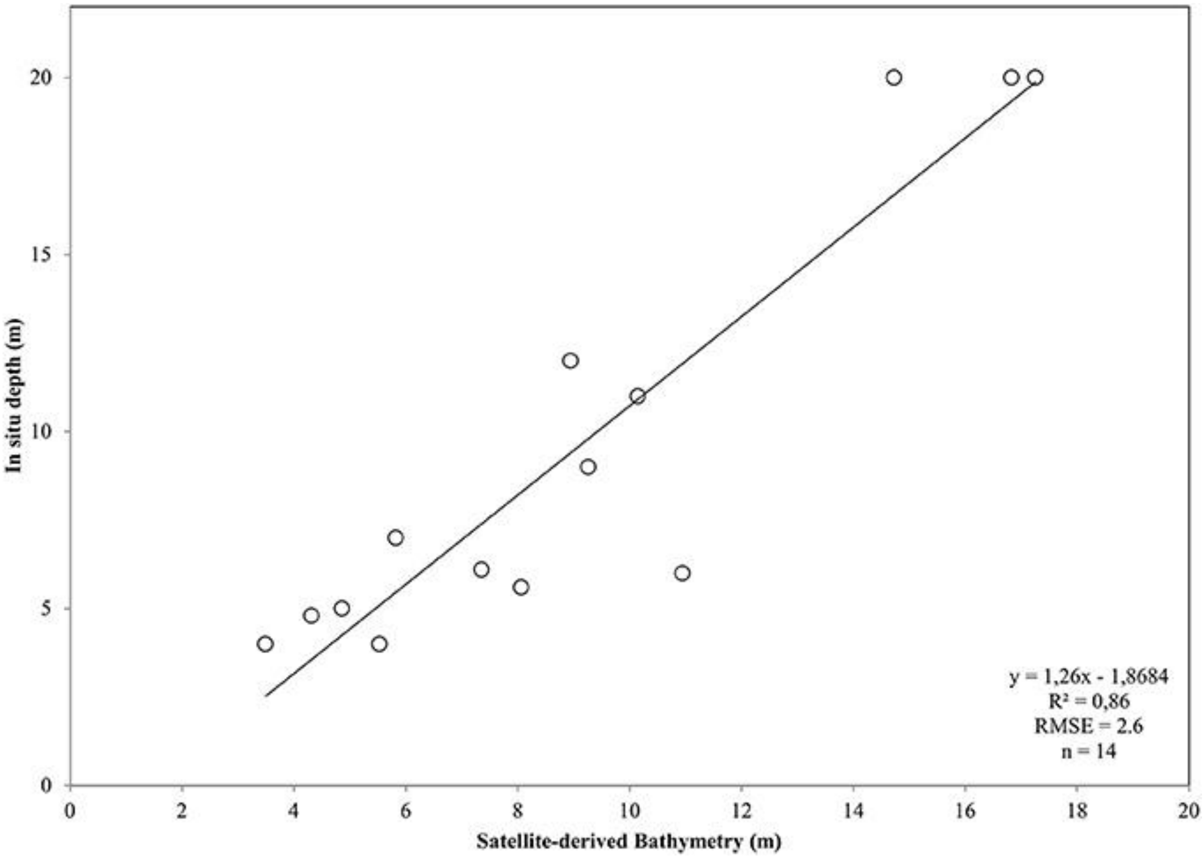


Fig. 4.5.5. Plot of Satellite-derived Bathymetry (SDB) vs. *in situ* measured depth for the validation of the bathymetry map of the Thermaikos Gulf (Fig. 4.5.4C). SDB was derived from Equation (2). Regressed SDB have been previously smoothed with a 5 × 5 low pass filter to reduce unwanted noise.

more accurately classified in RE16 (Fig. 4.5.6D; 98% producer accuracy) and RE11 (Fig. 4.5.6A; 92% producer accuracy), but less accurately identified in the same images according to the user accuracy of 84.5% of the former and 86.8% of the latter. The best user accuracies concerning *P. oceanica* were produced for the RE12 (Fig. 4.5.6B; 100%) and the RE15 (Fig. 4.5.6C; 95%).

As regards to *C. nodosa* species, RF correctly classified it to 62% and 60% producer accuracies in RE12 (Fig. 4.5.6B) and RE15 (Fig. 4.5.6C), while the former exhibited the second best user accuracy of 86.1% following the 92% of RE16 (Fig. 4.5.6D). Generally, RE12 featured the second best overall accuracy (81%), marginally behind RE16 (82%), but possessed the best mean producer and user accuracies of the two seagrass habitats (76 and 93%, correspondingly). In the contrary, RE11 revealed the worst results with the worst overall accuracy of 73.5%, worst mean producer accuracy (62%), worst mean user accuracy (81.5%).

Generally, the error matrices indicate that errors in both producer and user accuracies in all four images are mainly attributed to confusion between the two seagrasses and less with sandy or rocky seabed.

Based on the classified water column corrected RE16 composite, *P. oceanica* seagrass meadows covered an area of 264 ha in depths between 0.8 and 17.9 m, with an average depth presence of 8 m. On the other hand, *C. nodosa* beds covered 242 ha and were spread between depths of 0.8 and 16.1 m, with a mean depth presence of 5.8 m.

Change Detection

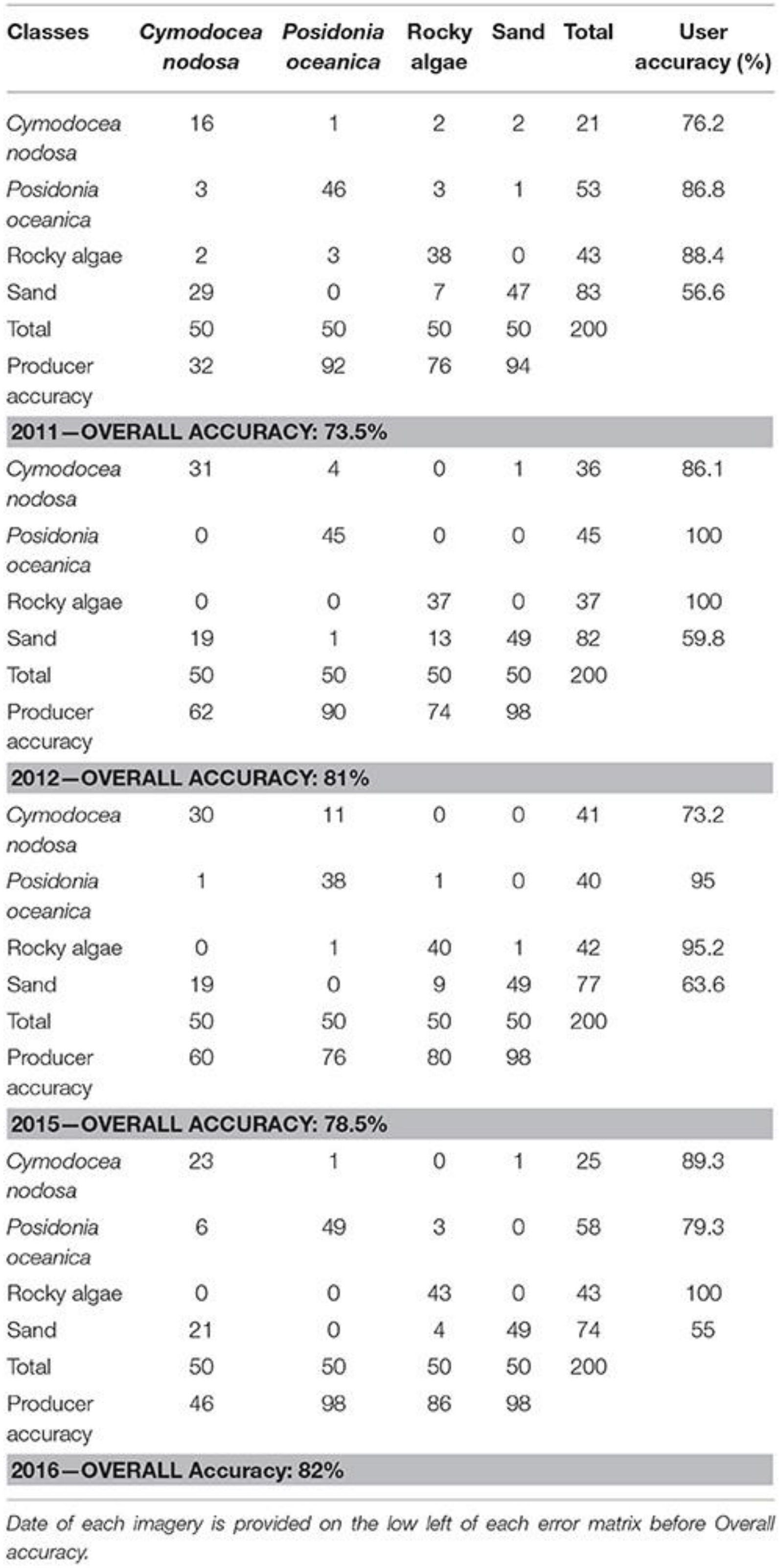
We report the interannual change detection of *P. oceanica* and *C. nodosa* seagrasses here as change of their extent (Fig. 4.5.7, 4.5.8) following random forest classification of all four RE images (Fig. 4.5.6). Fig. 4.5.7 shows the areas of both seagrasses and total seagrass area in each of the four studied years in addition to indicating change trends at species and total level (black lines). We observe that *P. oceanica* area declined by 4.1% (from 275 to 264 ha) between 2011 and 2016, while its declining trend was 11.2 ha/yr. On the contrary, *C. nodosa* area increased by 17.7% (from 199 to 242 ha), while its increasing trend was 18 ha/yr. Overall, therefore, the area of seagrasses in the Thermaikos Gulf increased by 6.8% (474 to 506 ha) between 2011 and 2016, with an increasing trend of 6.8 ha/yr. The highest decrease of *P. oceanica* seagrass was displayed between 2012.

4.5.4 Discussion

Change Detection

The main objective of this study was to evaluate whether the application of machine learning algorithms, namely Random Forests, on a time-series of high resolution satellite images, namely RapidEye, was effective for mapping the interranual change detection of two Mediterranean seagrasses, *P. oceanica* and *C. nodosa,* in 888 submerged hectares of the Thermaikos Gulf, NW Aegean Sea, Greece between 2011 and 2016 (total of

Table 4.5.2. Error matrices of the four water-column corrected bottom reflectance images.



Classes	*Cymodocea nodosa*	*Posidonia oceanica*	Rocky algae	Sand	Total	User accuracy (%)
Cymodocea nodosa	16	1	2	2	21	76.2
Posidonia oceanica	3	46	3	1	53	86.8
Rocky algae	2	3	38	0	43	88.4
Sand	29	0	7	47	83	56.6
Total	50	50	50	50	200	
Producer accuracy	32	92	76	94		
2011—OVERALL ACCURACY: 73.5%						
Cymodocea nodosa	31	4	0	1	36	86.1
Posidonia oceanica	0	45	0	0	45	100
Rocky algae	0	0	37	0	37	100
Sand	19	1	13	49	82	59.8
Total	50	50	50	50	200	
Producer accuracy	62	90	74	98		
2012—OVERALL ACCURACY: 81%						
Cymodocea nodosa	30	11	0	0	41	73.2
Posidonia oceanica	1	38	1	0	40	95
Rocky algae	0	1	40	1	42	95.2
Sand	19	0	9	49	77	63.6
Total	50	50	50	50	200	
Producer accuracy	60	76	80	98		
2015—OVERALL ACCURACY: 78.5%						
Cymodocea nodosa	23	1	0	1	25	89.3
Posidonia oceanica	6	49	3	0	58	79.3
Rocky algae	0	0	43	0	43	100
Sand	21	0	4	49	74	55
Total	50	50	50	50	200	
Producer accuracy	46	98	86	98		
2016—OVERALL Accuracy: 82%						

Date of each imagery is provided on the low left of each error matrix before Overall accuracy.

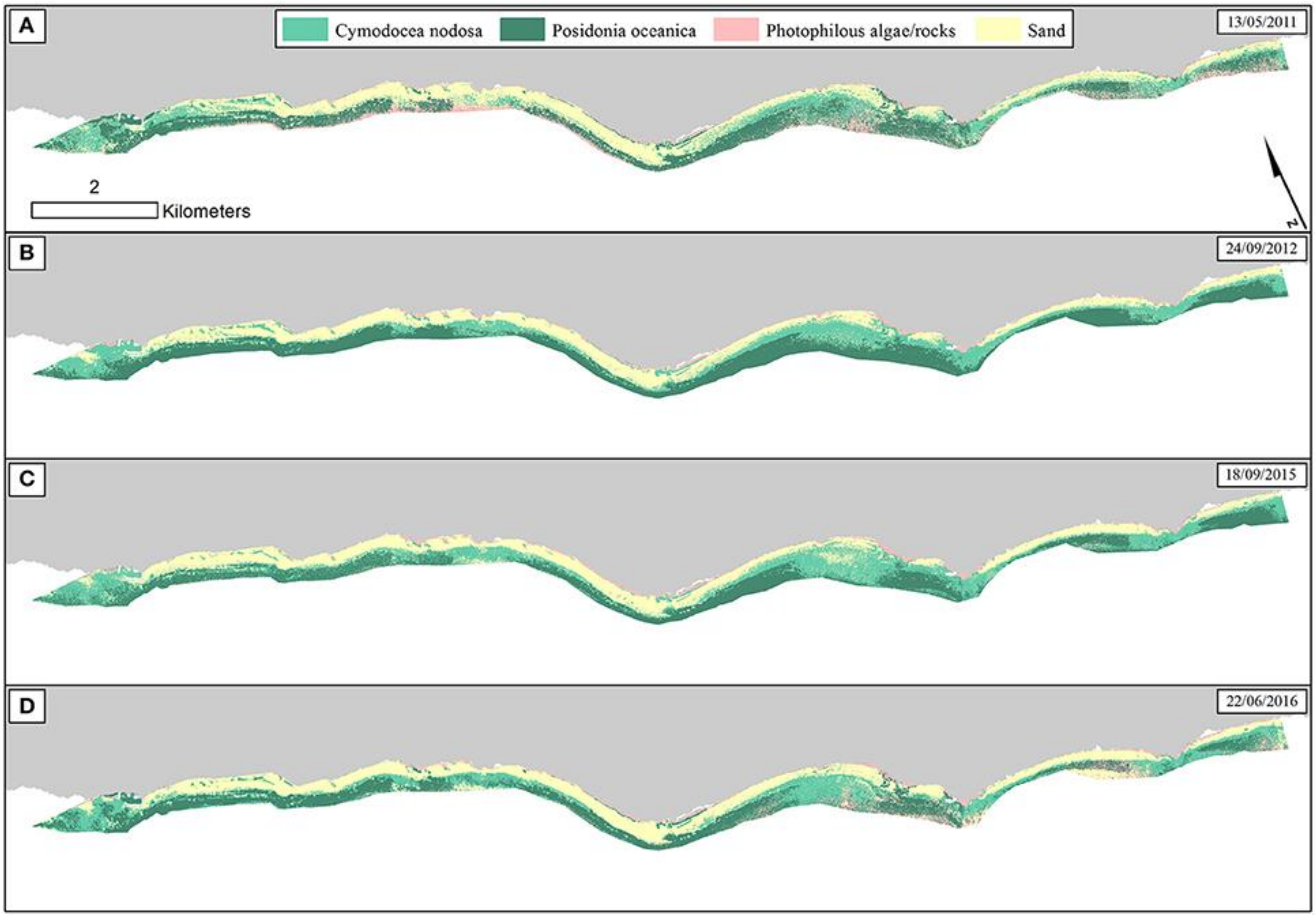


Fig. 4.5.6. Classified water-column-corrected RapidEye composites from the 4 years using Random Forest machine learning classifier (100 trees). The frames on the upper right of each panel indicate the date of each RapidEye image. (A) RE11—Overall accuracy: 73.5%. (B) RE12—Overall accuracy: 81%. (C) RE15—Overall accuracy: 78.5%. (D) RE16—Overall accuracy: 82%.

four images). As attested by Fig. 4.5.7, our main findings reveal that the distribution of *P. oceanica* seagrass meadows have declined by 4.1% with a decreasing trend of 11.2 ha/yr, while *C. nodosa* beds have increased by 17.7% with a gain trend of 18 ha/yr. Generally, total seagrass area increased by 6.3% at a +6.8 ha/yr trend throughout the 5 years. Approximate trends of seagrass distribution change are indicated by the slope coefficient of a standard linear regression between seagrass area and related years (Fig. 4.5.7). This method of presenting remotely sensed time series of seagrass was utilized efficiently in 200 km2 of the Eastern Banks, Moreton Bay, Australia (Lyons et al., 2013).

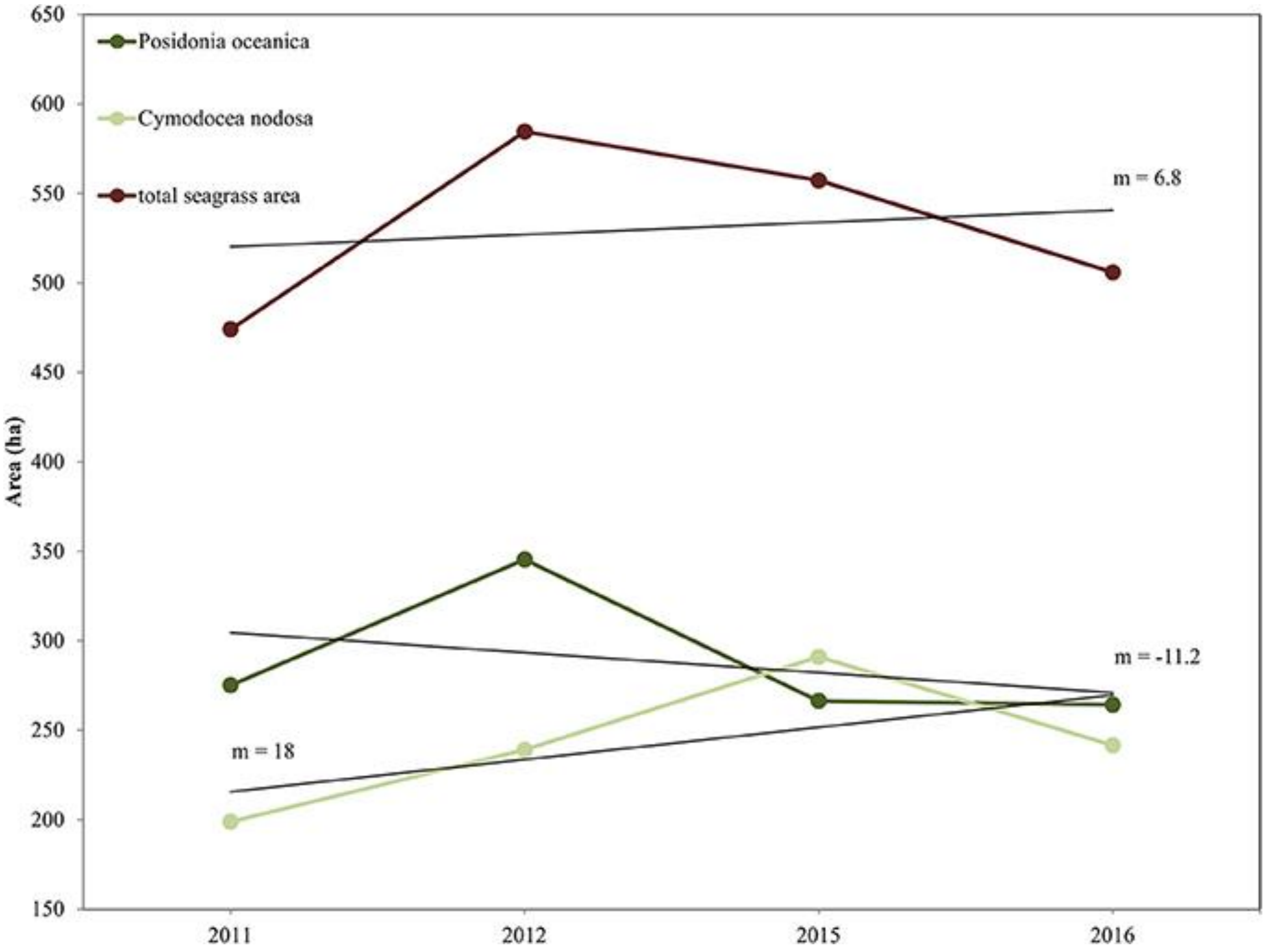


Fig. 4.5.7. Interranual change detection of seagrasses in the Thermaikos Survey site between 2011 and 2016 using RapidEye satellite images. The trajectory plot displays change of area (in hectares; y-axis) over the years (x-axis) of *Posidonia oceanica* and *Cymodocea nodosa* species, and of total seagrass area. Linea regression black lines (m = slope) show approximate trend in area between 2011 and 2016. *Posidonia oceanica* seagrass is decreasing at 11.2 ha/yr, *Cymodocea nodosa* seagrass is increasing at 18 ha/yr, while total seagrass area is expanding at 6.8 ha/yr.

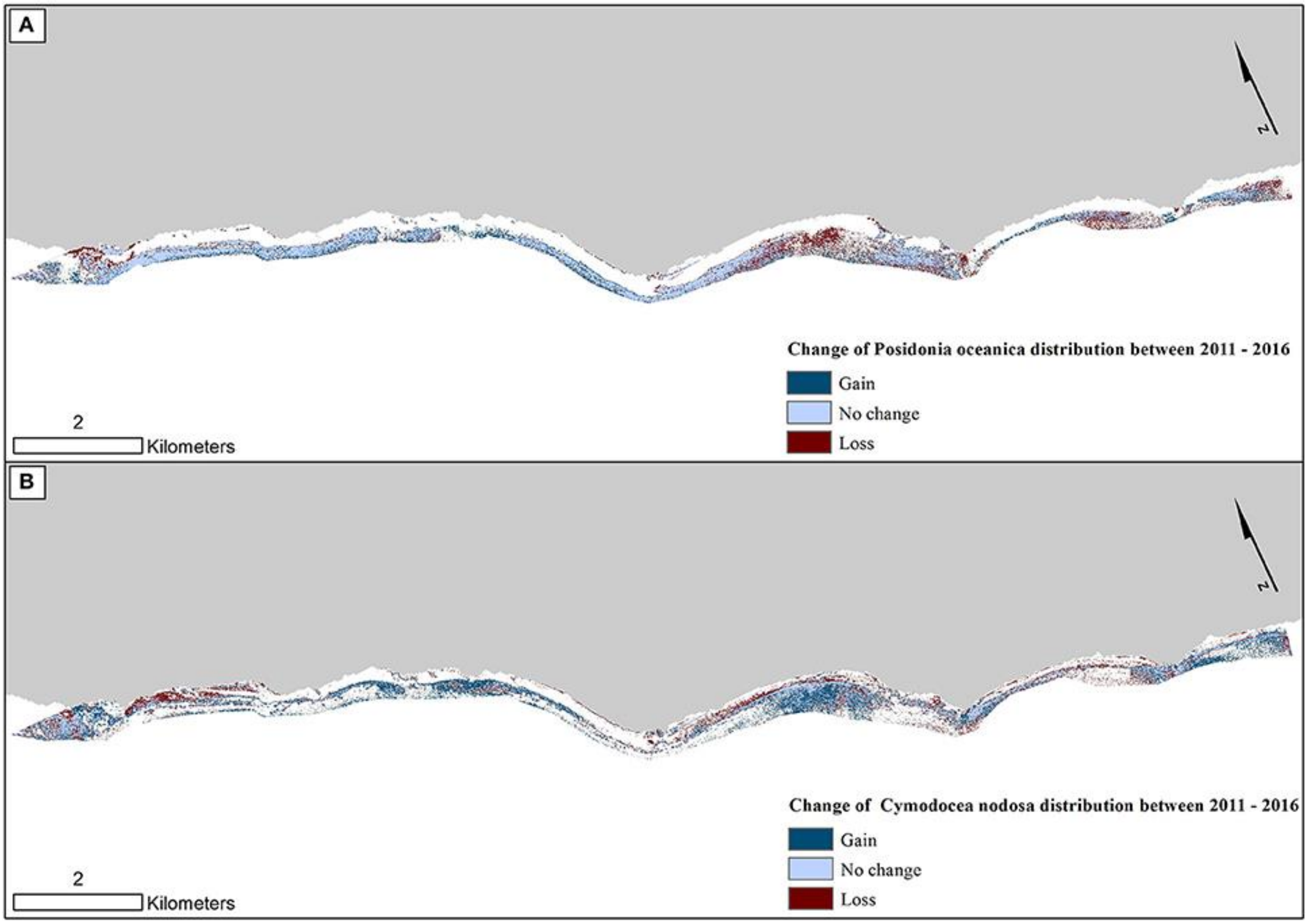


Fig. 4.5.8. Change in seagrass distribution in the Thermaikos survey site between 2011 and 2016 for (A) *Posidonia oceanica* and (B) *Cymodocea nodosa*. Between 2011 and 2016, *P. oceanica* seagrass meadows have declined by 4.1%, while *C. nodosa* seagrasses have increased by 17.7%.

To the best of our knowledge, the studied regression of 4.1% of *P. oceanica* meadows between 2011 and 2016 is the first report of regression of this particular seagrass species in the Greek seas and one of the first reports in the whole Eastern Mediterranean, a poorly mapped area. It is also in line with the reported trends in the Mediterranean and globally. Telesca et al. (2015) estimated an average regression of 10.1% for the whole extent of the Mediterranean basin during the past 50 years, which further increased to 33.6% for areas with existing historical information; Greece lacks this significant information. Marbà et al. (2014) further estimated that between 13 and 38% of initial *P. oceanica* meadows have been lost since 1960, with a decreasing trend of 1.74%/yr. On assessing 215 studies worldwide, Waycott et al. (2009) has shown that since 1990, seagrass grounds are disappearing at a median rate of 7%/yr, a 7-fold increase from the median rate of 0.9%/yr before 1940.

Based on the ecological value of *P. oceanica*, its reported regression of 4.1% in the Thermaikos Gulf translates into loss of the relevant ecosystem services which it provides in the broader region, including protection from coastal erosion, carbon sequestration, nursery grounds, and nutrient cycling among others. More specifically, the related economic loss to the declining rate of *P. oceanica* of 11.2 ha/yr is 19.264 million e/yr (Vassallo et al., 2013). This regression is more alarming due to the slow growth of *P. oceanica* meadows and the existing pressure from climate change. Mortality rates of *P. oceanica* seagrass are expected to increase 3-fold with an increase of 3◦C in maximum annual seawater surface temperature (SSTmax) (Marbà and Duarte, 2010). In addition, the temperature of 28◦C is the critical SSTmax, above which *P. oceanica* functional losses accelerate. Study of the specific drivers of the observed *P. oceanica* regression is out of our scope in the present study. Future studies, however, of the temporal dynamics of SSTmax and the extent of *P. oceanica* seagrass meadows could unravel the underlying causes of their regression in the Thermaikos Gulf and elsewhere in the Mediterranean.

On the other hand, *C. nodosa* seagrass area faced a 17% increase of its extent, gaining 43 hectares between 2011 and 2016. We could attribute part of this increase to a combination of two physical factors. First, *C. nodosa* is a fast-growing seagrass with a reported rate of horizontal growth of up to 2 m/yr. Second, the decline in terms of *P. oceanica* seagrass allowed the fast recolonization of its regressed beds by the former seagrass, causing the expansion of its area. This substitution of *P. oceanica* by *C. nodosa* between 2011 and 2016 is particularly observed in the shallower parts of the northwesternmost, the middle and the southeasternmost regions of the Thermaikos Gulf (Fig. 4.5.8).

Temporal dynamics of *C. nodosa* distribution have been sparsely studied elsewhere. *C. nodosa* populations have shown a progression of up to 42% between 1994 and 2011 in the Western Mediterranean, following a regression of 49% between 1973 and 1994 (Garrido et al., 2013). Furthermore, *C. nodosa* seagrass has been found to re-colonize the shallower regressed beds of *P. oceanica* seagrass (Montefalcone et al., 2007). All in all, *C. nodosa* seagrass competes with *P. oceanica* in terms of its expansion, while it is also considered as a significant step in the ecological succession prior to beds of the latter. Further research efforts are needed to increase the body of literature on *C. nodosa*, in terms of its spatio-temporal dynamics, associated drivers, and especially potential links with the ongoing climate change, and its provided ecosystem services.

As regards to the harness of the variety of existing satellite data, numerous studies have assessed seagrass dynamics in a plethora of spatial and temporal scales. In one of the first and most important seagrass change detection assessments, Dekker et al. (2005) exploited four Landsat 5 and 7 images spanning a total of 14 years to map four seagrass species, including ones with great ecologically sensitivity, in Wallis Lake, an estuarine lake in Australia. Pu et al. (2014) also used Landsat 5 data to evaluate seagrass dynamics between 2003 and 2005 in Florida coast. While the two studies monitored change detection of seagrass extent and cover, respectively, Roelfsema et al. (2014) mapped seagrass species, cover and above ground biomass processing a 142-km2 time-series of high spatial resolution WorldView-2, IKONOS and Quickbird imagery between 2004 and 2013 with an object-based approach in Moreton Bay, Australia. In another yet exploitation and longest, to the best of our knowledge, of the Landsat archive, Lyons et al. (2012) developed an object-based approach to assess seagrass extent between 1972 and 2010 in Moreton Bay as well. The common denominator of the aforementioned four time- series studies is that they assessed shallow waters up to 7 m in contrast to the 16.5-m deep limit in our study. They also all highlight the significance of remote sensing time series of seagrass habitats for seagrass ecology. Roelfsema et al. (2014) argued that the study of the correlation of seagrass-related physicochemical parameters like water quality and temperature with seagrass distribution and composition is vital. Future approaches to this direction could benefit time-series studies and more broadly the management and conservation of seagrasses.

Generally, there is a need for the development of an automated workflow which would exploit the great quantity of remote sensing information and develop time series of seagrass distribution and other management-related parameters in a time- and cost-efficient as well as accurate fashion. This workflow would enable fast assessment of problematic areas (areas of existing or ongoing regression) and raise the need for appropriate management and conservation measures. As mentioned in the Introduction section, however, to achieve accurate time series of seagrass and broadly coastal submerged habitats, one has to select suitable classifier(s) to the subject in study, as classification of these habitats always precede time series analysis.

Classification of Submerged Habitats

We selected Random Forests to solve the classification problem of discriminating between four habitats in the optically shallow waters of the Thermaikos Gulf, NW Aegean Sea. We run the classification experiments on atmospherically and water-column corrected RapidEye reflectance composites using 100 trees which yielded better quantitative results than other numbers of trees. We also used 50 data for each habitat for both the classification and validation to avoid possible overestimation of any of the classes (Traganos and Reinartz, 2018a). It is noteworthy, as Fig. 4.5.1 shows, that we chose training data only from the southeastern part of our survey site and validation data only from the northwestern part of our survey site. This could have led to biased classified results.

Overall, as reported in the error matrices of Table 4.5.2, RF exhibited high accuracies in classifying and identifying both seagrass species and especially *P. oceanica* species, up to 98% and 100% producer and user accuracy, respectively. The 5-m pixel-based random forest classification of *P. oceanica and C. nodosa* species displayed slightly worse producer accuracies, but higher user accuracies than similar efforts using Sentinel- 2A 10-m imagery in the same waters (Traganos and Reinartz, 2018a). Particularly for the latter species, its sparse and mixed nature with sandy beds inhibits classification and identification approaches causing the so-called "mixed" pixels. Higher than 5-m resolution approaches employing linear unmixing models and/or object-based classifications could solve this classification issue.

We chose RF to classify and identify Mediterranean seagrasses to achieve more accurate results than simpler, theoretically, algorithms like Maximum Likelihood (Traganos and Reinartz, 2018a) which has had a wide application history in the literature of both single- and multi-date studies of seagrass ecosystems (Dekker et al., 2005; Pasqualini et al., 2005; Pu and Bell, 2013; Pu et al., 2014). MLC assumes a normal distribution of classes which is rare in the nature of the examined classes, thus producing inferior results to the more sophisticated machine learning classifiers (Traganos and Reinartz, 2018a). RF produced promising results concerning classification of seagrasses recently in two studies (Zhang, 2015; Traganos and Reinartz, 2018a). (Zhang et al., 2013) demonstrated the advantage of RF over MLC using hyperspectral imagery. Utilizing 150 trees to run the RF-based experiments, Zhang (2015) achieved better accuracies than the machine learning classifiers of Support Vector Machines (SVMs) and *k*-Nearest Neighbor (*k*-NN) in identifying patchy seagrass in a 40-km2 area in lower Florida Keys, but slightly lower accuracy for continuous seagrass. Traganos and Reinartz (2018a) compared RF, SVMs and MLC classifiers in a small section of the surveyed site in the present study. They showed that both RF and SVMs performed evidently better than MLC on classifying *P. oceanica* and *C. nodosa* habitats. In the same study, both RF and SVMs displayed lower accuracies on the classification of *C. nodosa* than *P. oceanica* due to the smaller number of field data for the former in addition to its mixed ground with sand.

In summary, machine learning classifiers like RF and SVMs gain more and more interest in coastal habitat remote sensing and, more broadly, in the remote sensing literature

(Gislason et al., 2006; Mountrakis et al., 2011). Deep learning techniques concerning submerged habitats are still in their infancy (Call et al., 2003; Calvo et al., 2003) and it is still unknown whether the extra processing power and time to design the experiments are worthy for the potentially better identification that they would offer than machine learning classifiers.

Pre-classification Steps in Detection, Mapping, and Time Series of Submerged Habitats

The steps which precede the classification and subsequent development and analysis of the time series of submerged habitats include geometric, atmospheric, and water column corrections of satellite data in addition to developing satellite- derived bathymetry. In the present study, geometric corrections were already done in the Level 3A RapidEye imagery which we processed in the time series. The FLAASH module, implemented for the atmospheric correction, has been already deployed by several studies for studying the change detection of underwater habitats (Lyons et al., 2010, 2011; Pu and Bell, 2013; Roelfsema et al., 2014).

Regarding the water column correction, the analytical model of Maritorena et al. (1994) accurately retrieved bottom reflectances of both *P. oceanica* and *C. nodosa* seagrass species in the RE time series of the present study. In the same area, Traganos and Reinartz (2018a) employed successfully the same model to perform water column corrections for the mapping of the two same species. Dierssen et al. (2003) discussed the good agreement of the model's calculated bottom reflectances with *in situ* ones over dense *Thalassia testudinum* beds in contrast to overestimated seabed reflectances over beds of the same species of sparse to intermediate density up to depths of 9 m in Lee Stocking Island, Bahamas. In addition, Pu et al. (2014) conducted water column corrections following the same analytical model to identify three seagrass species (*T. testudinum*, *Syringodium filiforme,* and *Halodule wrightii*) in depths up to 4 m in Florida, USA. A similar to ours image pre-processing and processing methodological approach to seagrass change detection led to 14% improved overall accuracies than studies which used analogous data. In another application of Maritorena et al. (1994) water column correction model, Dekker et al. (2005) mapped the change detection of *Posidonia australis*, *Halophila ovalis*, *Zostera capricorni,* and *Ruppia megacarpa* in the waters of Wallis Lake in Australia, in depths of <3 m. In antithesis to the field optical measurements of the latter study, we used and developed image-based estimations of both the diffuse attenuation coefficient and infinitely deep water column reflectance. Future *in situ* optical measurements are expected to increase accuracies in water column corrections and succeeding classifications and time series analyses, however, these measurements would also raise the cost of the given study.

Regarding the satellite-derived bathymetry, we created a 5-m resolution bathymetry map of the Thermaikos Gulf using RapidEye imagery to aid the RE-based time series of the two Mediterranean seagrass beds. We should note here that we employed the closest imagery to our *in situ* depth estimations, RE16, to estimate bathymetry for all four images. Traditionally, remote sensing scientists have calculated depth in optically shallow regions by using the band ratio (Lyzenga, 1978). Employing the ratio of blue to green, but in the different central wavelengths of Sentinel-2, 490 and 560 nm, than in the present study, Traganos and Reinartz (2018a) developed a bathymetry model for a subsection of the survey site of the present study. They also applied a low pass filter before the estimation of the bathymetry in contrast to the present study where we applied a low pass filter after mapping depth. The lower wavelength of the blue band of RapidEye at 475 nm, however, in comparison to the 490 m of Sentinel 2 is anticipated to produce more accurate bathymetry estimation due to the higher penetration of the blue band in the water column in this case. Moreover, we chose to convert at-water surface reflectances to the remote sensing reflectances for the development of the bathymetry model as the latter are considered more robust to interactions in the air-water interface and water column constituent composition (Dierssen et al., 2003; Dekker et al., 2011).

The dense canopy and the incidental shading produce the often lower than the adjacent optically deep waters reflectance of *P. oceanica* seagrass (Dekker et al., 2006). This issue imposes a problem to accurate bathymetry estimations over this type of seabed. Traganos and Reinartz (2018a) overcome this problem by modifying the widely utilized bathymetry algorithm of Stumpf et al. (2003) which displayed negative values over *P. oceanica* beds in relevant experiments. It is worth noting that due to the lack of extensive *in situ* depth data, we tuned our polynomial algorithm using data, mainly from the southeasternmost part of the Thermaikos Gulf. Nevertheless, as we presented in section Pre-classification Steps, we chose *in situ* depth data over all four habitats, namely the two seagrasses, sand and rocky seabed with algae, achieving an accuracy of 91% in the development of the site-specific depth algorithm and a r-squared value of 0.86 with a RMSE of 2.6 m in the validation of this site-specific bathymetry.

Other studies have used either existing pixel-based depth maps produced with acoustic equipment (Pu and Bell, 2013; Pu et al., 2014), have developed their own satellite-derived bathymetry maps implementing either linear or ratio algorithms (Lyons et al., 2011) or have run simulation experiments of the bathymetry effects using HYDROLIGHT, a robust radiative transfer model. It would be interesting to compare in the future bathymetries derived from all of the above sources to study how accuracies deviate in turn.

4.5.5 Conclusions

The present study demonstrates an off-the-shelf methodology to quantitatively assess the spatio-temporal dynamics of seagrasses and other submerged habitats in clear and homogeneous optically shallow waters using Planet's RapidEye time series of four 5-m satellite images. The methodology includes three stages: (a) pre-processing including atmospheric and water column correction of the satellite data along with satellite-derived bathymetry, (b) machine learning classification using the Random Forest algorithm, and (c) interannual change detection which is presented here as a change of area and associated trend. We applied these three steps to study the dynamics of two Mediterranean seagrasses, *P. oceanica* and *C. nodosa*, in the waters of the Thermaikos Gulf (NW Aegean Sea, Greece) between 2011 and 2016. Total seagrass area has increased by 6.3% at a rate of +6.8 ha/yr, while *P. oceanica* seagrass has regressed by 4.1% at a rate of −11.2 ha/yr and *C. nodosa* seagrass has progressed by 17.7% at a rate of +18 ha/yr throughout the 5 years. In some occasions, *C. nodosa* has been studied to substitute the regressed beds of *P. oceanica.* The aforementioned trends, especially in terms of the regression of the *P. oceanica*, are in line with the reported regression of this valuable seagrass elsewhere in the Mediterranean. This study is the first to report spatio-temporal dynamics of both seagrasses in large scales using remotely sensed data. The remote sensing of seagrasses lying in optically shallow waters (where the observed surface reflectance contains signal from the bottom in contrast to an optically deep column) faces a plethora of inherent obstacles due to the complex nature of the media above the seagrass beds themselves. Obstacles like water column constituents, sunglint, and skyglint presence, air-water interface interference could impede the detection of seagrasses and require, usually, consideration through relevant algorithms. The presented methodological workflow could act like an alternative ecological assessment showing current trends, revealing regressing seagrasses, and allowing better conservation of these complex but also significant ecosystems. Potential improvements in the given approach could be the existence of *in situ* optical measurements of several relevant parameters, broader bathymetry field data, advanced radiative transfer simulations, possible comparison of different machine learning algorithms for the improvement of classification and identification of seagrasses and better tuning of those algorithms. Currently, seagrasses are decreasing in alarming rates in a global scale. Linkage of this decreasing trend with the anthropogenic and natural interference through Earth observation of climate change, eutrophication, coastal development as well as temperature, salinity, and hydrodynamic change could develop and refine machine learning models to ecologically assess seagrass status worldwide. Harnessing the wealth of Earth observation data that this century offers and state-of-the-art machine learning algorithms, we could better understand the thresholds of different seagrass habitats in different aquatic environments and strengthen their conservation management, allowing a brighter future for these significant ecosystem service suppliers.

4.6 Towards Global-Scale Seagrass Mapping and Monitoring using Sentinel-2 on Google Earth Engine: The case study of the Aegean and Ionian Seas

4.6.1 Abstract

Seagrasses are traversing the epoch of intense anthropogenic impacts which significantly decrease their coverage and invaluable ecosystem services, necessitating accurate and adaptable, global-scale mapping and monitoring solutions. Here, we combine the cloud computing power of Google Earth Engine with the freely available Copernicus Sentinel-2 multispectral image archive, image composition and machine learning approaches to develop a methodological workflow for large-scale, high spatiotemporal mapping and monitoring of seagrass habitats. The present workflow can be easily tuned to space, time and data input; here, we show its potential, mapping 2,510.1 km^2 of *P. oceanica* seagrasses in an area of 40,951 km^2 between 0 and 40 m of depth in the Aegean and Ionian Seas (Greek territorial waters) after applying Support Vector Machines to a composite of 1,045 Sentinel-2 tiles at 10-m resolution. The overall accuracy of *P. oceanica* seagrass habitats features an overall accuracy of 72% following validation by an independent field data set to reduce bias. We envisage that the introduced flexible, time- and cost-efficient cloud-based chain will provide the crucial seasonal to inter-annual baseline mapping and monitoring of seagrass ecosystems in global scale, resolving gain and loss trends and assisting coastal conservation, management planning, and ultimately climate change mitigation.

4.6.2 Materials and Methods

Study site

The study area includes the Greek territorial waters—i.e. Ionian Sea and Greek part of the Aegean Sea—covering a total surface of 40,951 km² (Fig. 4.6.1). We define this area on the basis of an edited 3-km buffer zone along the Greek coastline of approximately 18,000 km, following the results of (Topouzelis et al., 2018) (see section 2.3.2). The Aegean Sea features an intricate geomorphology that reflects past geologic history and recent geodynamic processes. Shallow shelves, deep basins, and troughs alternate throughout the buffer-defined area whose deepest point is ~2,500 m (deepest area in the Aegean trench is ~ 5,000m). It has also a distinctive insular character with more than 1,400 islands or islets, while its extensive coastline consists of several landforms, including sandy beaches, rocky shores, cliffs, coastal lagoons and deltaic systems, as well as a notable variety of coastal and marine habitat types. On the other hand, the Ionian Sea is considered an oligotrophic area, based both on low nutrient concentrations and primary production. Biological data reflect a very oligotrophic area, dominated by a microbial food web, where new production mostly derives from limited events in space and time, mainly driven by climatological factors generating mesoscale instabilities. In terms of biology, the Ionian Sea is influenced by the Adriatic water in the northern part of the subbasin, with higher phytoplankton biomass (particularly diatoms), while water of Atlantic origin makes up its southern part with the exception of the whole eastern side (Casotti et al., 2003; Sini et al., 2017). Seagrass meadows exist in protected bays and gulfs while their mean maximum depths of presence are between 25 and 35 m depending on the local conditions (Gerakaris et al., 2014).

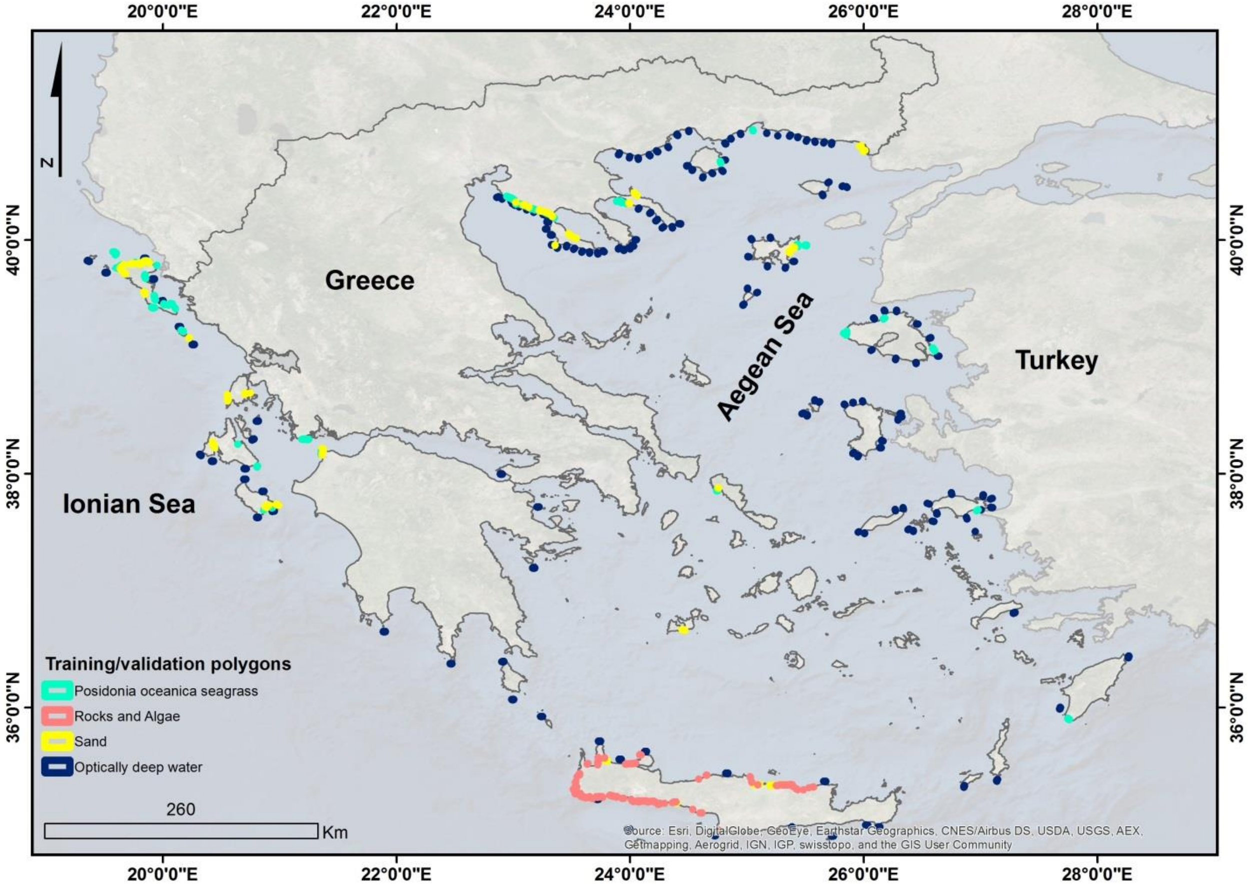
Satellite Data

We use Copernicus Sentinel-2 Level-1C (L1C) Top of Atmosphere (TOA) reflectance satellite data, the standard S2 archive in GEE. The available data extends from 23 June 2015 (date of launch of Sentinel-2A) to today with a 5-day temporal resolution (with the use of Sentinel-2B). For the present study, we choose a period between 1 September and 1 October 2017 which satisfies availability of both S2-A and S2-B imagery, and, more importantly, a better-stratified water column in the study areas of Aegean and Ionian seas (Traganos et al. 2018b). Our data input is the seven S2 bands: b1-coastal aerosol, b2-blue, b3-green, b4-red, b8-NIR, b11-SWIR1, and the QA60 band, the bitmask band which contains information on cloud mask. Bands b1 and QA60 are in 60-m spatial resolution, while b11 is in 20-m spatial resolution. GEE reprojects all to the 10-m native resolution of the b2, b3, b4, and b8. In total, 1,045 S2 tiles—100x100 km² sub-images— compose our initial image dataset. All available datasets, from satellite images to field and auxiliary data are projected in the GCS WGS84 World Geodetic System.

Field Data

Training and Validation Data

We manually digitised all training data on very high spatial resolution images using the ArcGIS World Imagery basemap (< 60-cm pixel) (Fig. 4.6.1). To ensure data quality, we selected areas in relative shallow waters where seagrass was easily interpreted. The digitisation took place in areas where we have a high level of knowledge on the seabed cover and composition due to past and ongoing fieldwork activities. The non-seagrass class contained rocky, sandy and deep-sea areas. To ensure consistency across all classes, we implement 4x4 homogeneous polygons (16 S2 pixels). In ArcGIS Desktop, the digitization tool allows to design polygon data with specific dimension and size. Thus, the size of each polygon is the same and the homogeneity is related to the seabed cover as this is identified by experienced image interpreters. In total we used 1,457 homogeneous polygons (23,312 pixels), for two classes (seagrass/non-seagrass) (Fig. 4.6.2; Table 4.6.1). The homogeneity of all polygons is vital because their absence may cause misclassification. The validation data consists of 322 independent field data points based on unpublished data provided by terraSolutions m.e.r. (http://www.terrasolutions.eu/) to reduce bias.



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

Fig. 4.6.1. Geographical location of survey site and training polygons of the herein considered classes. All polygons are in GCS WGS84 World Geodetic System.

Auxiliary Data

We utilise two additional auxiliary datasets to aid both the time (thus computing) efficiency of our processing chain and the distillation of its results. First, to mask out both land and deeper waters, we edit a shapefile of the Greek Coastline at scale 1:90,000 (HNHS, 2018) by the Hellenic Navy Hydrographic Service (HNHS) in a two-step way; a) we create a buffer of 3 km to encompass the whole optically shallow extent (where there is remote sensing signal from the seabed), b) we manually edit the buffer of 3 km to include the deeper and/or larger meadows (according to the seagrass polygons of Topouzelis et al. (2018) and to delete all the vertices over land. The resulted coastal area comprises our survey site of a total area of ~40,951 km². In addition, we use the bathymetry of the Aegean Sea for the post-classification stage (see section 2.4.3). The Aegean Sea Bathymetry depth zones (5-m intervals until 50-m depth) is a fused product of HNHS, EMODnet (the European Marine Observation and Data Network), and in situ data collected during MARISCA project (Sini et al. 2017).

Methodology

In addition to the easy and parallel access to the satellite image archives, GEE offers quick and adaptable computational tools for remotely sensed data processing and analysis. We exploit GEE tools to build our methodological chain which is divided into three parts: a) pre-classification, b) classification, c) post-classification. Fig. 4.6.3 displays this chain, while Fig. 4.6.4 depicts its various successive stages.

Table 4.6.1. Number of implemented polygons and pixels per class for our ~40,951 km^2 survey site.

Class	**Polygons**	**Pixels**	**%**
Seagrass	329	5,264	22.6
Non-seagrass	1,128	18,048	77.4
Sum	1,457	23,312	100

Pre-Classification

Our pre-classification part consists of six steps:

1. Cloud mask: We use the QA60 bitmask band to mask opaque and cirrus clouds and scale S2 L1C TOA images by 10,000 (Fig. 4.6.4a, b).
2. Land mask: Although here we utilise the buffered coastline shapefile of Greek waters to mask out terrestrial Greece, we include a CART classifier (Breiman et al., 1984) in the GEE code which future users could employ to mask out their terrestrial

part. The classifier is applied on a b3-b8-b11 composite and the user should train it with relevant pixels over land and water.

3. Image composition: We apply image composition which yields a new pseudo-image composite whose pixels are the first quartile (Q1) of the median values of the cloud corrected and masked for land images of step 2. The purpose of this approach is to decrease non-corrected image artefacts by the previous steps.

4. Atmospheric correction: We implement a modified dark pixel subtraction method following (Armstrong et al., 1993) to empirically address path radiance and noise in all bands; this method subtracts the average reflectance and two standard deviations of optically deep water (> 40 m) (Fig. 4.6.4b).

5. Sunglint correction: We further correct the atmospherically corrected image composites with the sunglint correction algorithm of Hedley et al. (2005). Following a user-defined set of pixels of variable sunglint intensity, the algorithms equals the corrected for sunglint composite to the initial first quartile composite minus the product of the regression slope of b8 against b1 - b4 and the difference between b8 and its minimum value (Fig. 4.6.4c).

6. Depth invariant indices calculation: To compensate the influence of variable depth on seabed habitats, we derive the depth invariant indices (Lyzenga, 1978, 1981) for each pair of bands with reasonable water penetration (b1-b2, b2-b3, b1-b3) with the statistical analysis of (Green et al. 2000). Prior to the machine learning-based classification, we apply a 3x3 low pass filter in the depth invariant as well as the sunglint-corrected input to minimise remaining noise over the optically deep water extent which would have caused misclassified seagrass pixels otherwise (Fig. 4.6.4d).

Classification

Although we experimented with three machine learning classifiers, Support Vector Machines (SVM) (Vapnik, 1995), Random Forests (RF) (Breiman, 2001), and CART (Breiman et al., 1984) we end up using only SVM due to their better yielded classification output, both qualitatively and quantitatively (Fig. 4.6.4e). Based on the statistical learning theory of SVM of Vapnik (1995), solve linear and non-linear classification problems by fitting a separating hyperplane to the training data of the studied classes; they take their name from the support vectors, namely the points closest to the hyperplane which are the only ones that ultimately define it. A small number of studies have utilised SVM to map optically shallow habitats (Zhang, 2015; Traganos and Reinartz, 2018a; Traganos et al., 2017; Poursanidis et al., 2018). Here, we run SVM with a Gaussian radial basis function kernel (RBF), a parameter *g*, the width of the RBF, of 1000, and a regularisation parameter, *C*, which governs the degree of acceptable misclassification, of 100. We empirically select the adequate pair values for *g* and *C* based on experiments that we run setting their range between a minimum of 0.01 and a maximum of 1000 using a multiplier of 10. The input to the classifier is the sunglint-corrected S2 composite of b1, b2, b3, b4 and the depth invariant index b2-b3.

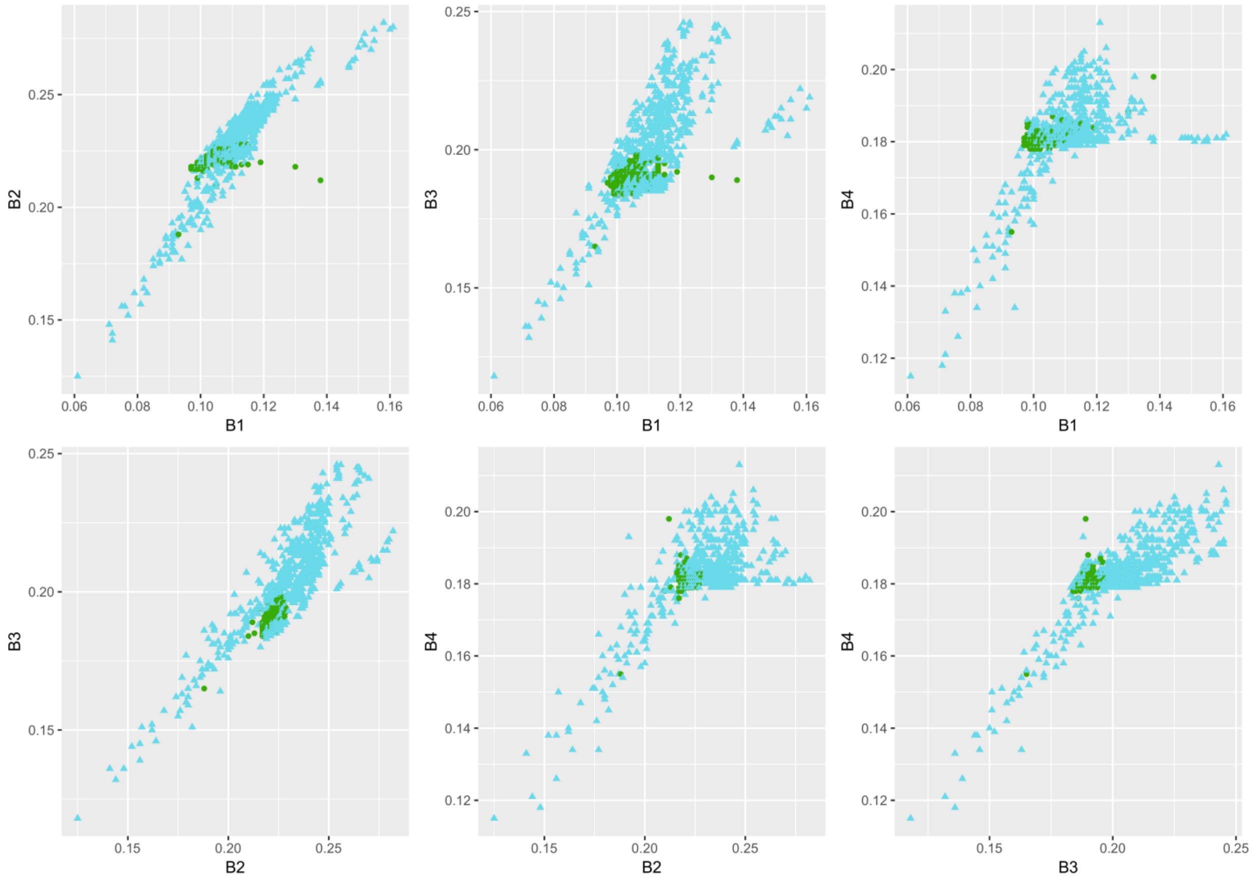


Fig. 4.6.2. Scatter plots of the first four, sunglint-corrected Sentinel-2 bands depicting waveband reflectivity of the herein 1,457 polygons for the whole extent of the study area. Seagrasses are in green circles and non-seagrasses are in light blue triangles.

Post-Classification

The post-classification part consists of two components: the editing of seagrass polygons due to misclassified pixels as seagrass in deep water and the accuracy assessment of the these machine learning derived edited seagrass polygons. The accuracy assessment employs an independent validation data set (unpublished data from terraSolutions m.e.r.) consisting of 322 data points to reduce general bias. We report the overall accuracy of seagrass habitats in the extent of the Aegean and Ionian seas which is the ratio of the entire number of correctly classified pixels to the total number of validation pixels. The post-classification correction of seagrass pixels over deep water serves a better visualisation of the classification output, but more significantly, it decreases the overprediction tendency of the classification results, namely seagrass area in the Aegean and Ionian seas; a single misclassified 10-m S2 pixel as *P. oceanica* seagrass would cause an over-prediction of 100 m^2.

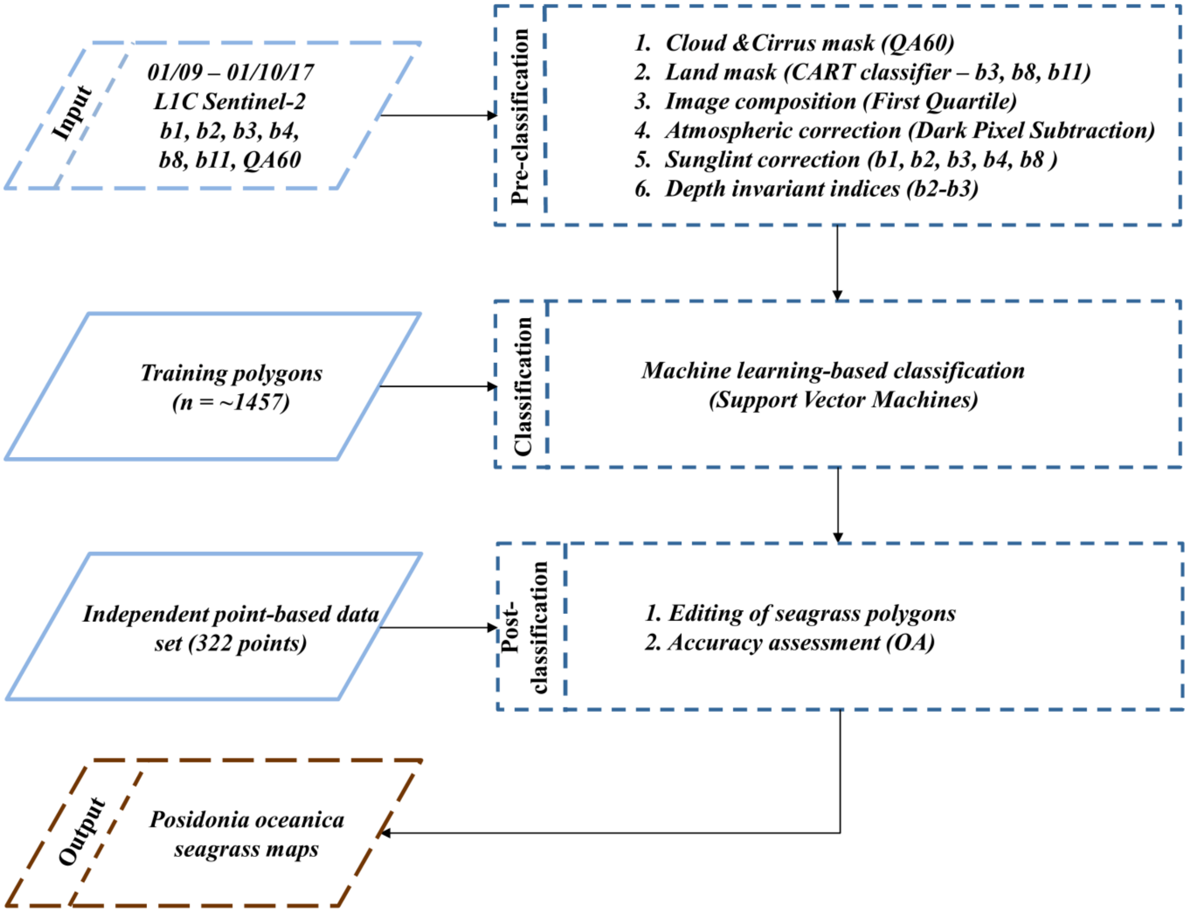


Fig. 4.6.3. Methodological workflow of the present study within Google Earth Engine. OA denotes overal accuracy. In the present study we did not implement step 2 (due to the use of a coastline buffer), however, we include it as it is an important component of the methodological chain.

4.6.3 Results

Pre-Classification

Regarding the pre-classification steps, r^2 of regression between sunglint-polluted b8 and b1 - b4 composites are b1: 0.3, b2: 0.3, b3: 0.39, and b4: 0.39. As regards to the only herein employed depth invariant index of b2-b3 bands, the ratio of water attenuation coefficient is 0.57. Machines-based classified product draped over the b2-b3 depth invariant index layer, (f) PlanetScope surface reflectance product (as imaged by Planet's Doves) in natural colour for high resolution reference (3 m) (ID: 20170828_084352_100e/20170828_084352_100e_3B_AnalyticMS_SR). The pink squares indicate sunglint presence in (b) and its correction in (c). The green and yellow polygons show employed seagrass and sand pixels in the machine learning classification. All panels are in GCS WGS84 World Geodetic System.

Classification

Our methodological workflow reveals that the seagrass area of the Greek Seas is 2,510.1 km^2; 1,885 km^2 in the Aegean Sea and 625 km^2 in the Ionian Sea (Fig. 4.6.5). As regards to individual geographical areas, the ones with the largest seagrass area (and maximum observed depth, where available, in parentheses) are Limnos (Fig. 4.6.5d) with 254 km^2 (40 m), NW Peloponissos (Fig. 4.6.5c) with 99 km^2, Corfu with 90 km^2, Crete with 70 km^2 (40 m), Thasos (Fig. 4.6.5b) with 53 km^2 (25 m), and the Thermaikos Gulf (Fig. 4.6.5a) with 49 km^2 (20 m). Across both seas, according to the employed auxiliary depth information, the depth of seagrass habitats is between 0 and 40 m.

Post-classification

The observed overall accuracy of *P. oceanica* seagrass in both the Aegean and Ionian seas is 72% based on the independent field data set.

4.6.4 Discussion

On Global Mapping and Monitoring of Seagrasses, and the Results of the Present Case Study

The exponential increase in cloud-based computing power, machine learning algorithms, and freely available satellite image archives has rendered the vision of global-scale mapping and monitoring of seagrasses and their neighbouring habitats more feasible than ever before. However, seagrass environments vary in regards to species

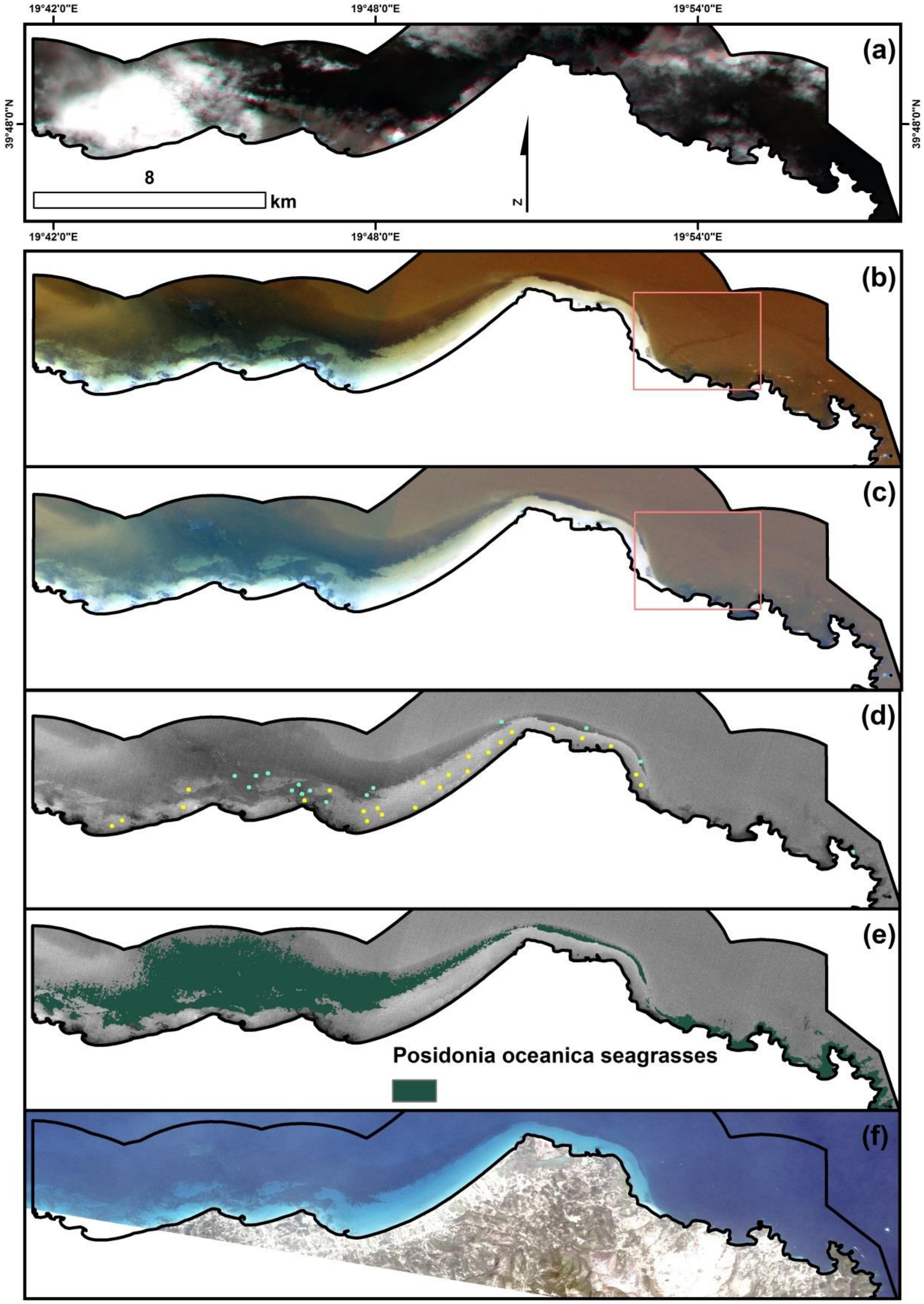

19°42'0"E
19°48'0"E
19°54'0"E
39°48'0"N
39°48'0"N
(a)
8
km
N
19°42'0"E
19°48'0"E
19°54'0"E
(b)
(c)
(d)
(e)
Posidonia oceanica seagrasses
(f)

Fig. 4.6.4. Successive stages of the developed workflow through the resulting false-colour (b1-b2-b3) Sentinel-2 composites. (a) Initial S2, (b) Cloud- and atmospherically-corrected, (c) Sunglint-corrected, (d) Depth invariant index b2-b3, (e) Support Vector

composition and abundance, but also water clarity (secchi depth) which will influence the capability to apply approach globally. At present, while there are tangible efforts for the global monitoring of coral habitats (Washington D.C., 2018) similar efforts focusing on seagrasses ecosystems have been yet not envisaged.

Here, we have developed a methodological workflow within Google Earth Engine which employs a plethora of universally used algorithms in coastal aquatic habitat remote sensing along with image composition and machine learning that could potentially be applied to map and monitor seagrasses globally. We demonstrate its power along with its issues in the Greek Seas, namely the Aegean and Ionian Sea; in a total coastline extent of 40,951 km^2, we map 2,510.1 km^2 of Greek seagrasses (*Posidonia oceanica* species) between 0 and 40 m of depth applying Support Vector Machines in a pixel-based fashion on 1,045 Sentinel-2 tiles. In comparison to existing mapping efforts and known distribution of seagrasses in our study site, our area findings are ~4.2% less than the respective coverage estimations of (Topouzelis et al. (2018) (2619.3 km^2). This can be attributed to the different methods and data in use: pixel vs object-based approach, different type of in situ data and difference in spatial resolution e.g. Sentinel-2 vs Landsat-8 spatial resolution (10 vs 30 m, respectively, resulting in a minimum mapping unit of 100 m^2 vs 900 m^2). A near-future regional comparison of pixel- to object-based approaches in the same context could shed further light upon the nature of their discrepancies—e.g. statistical vs environmental. On the other hand, in comparison to the UNEP WCMC Version 5 seagrass distribution, the herein seagrass area are nearly four times more (639.5 km^2) (UNEP-WCMC, Short FT, 2017).

The Good, the Bad and the Best Practices of the Proposed Cloud-based Workflow

The strength of our methodological chain lies mainly in the fact that it can be easily adjusted in space, time and data input. In comparison to the pre-processing chain of (Traganos et al. 2018b) and the therein use of the median value, we implement the first quartile here which yields less noisy image composites because it filters higher reflectances (=clouds and sunglint). The chain requires specific input to run:

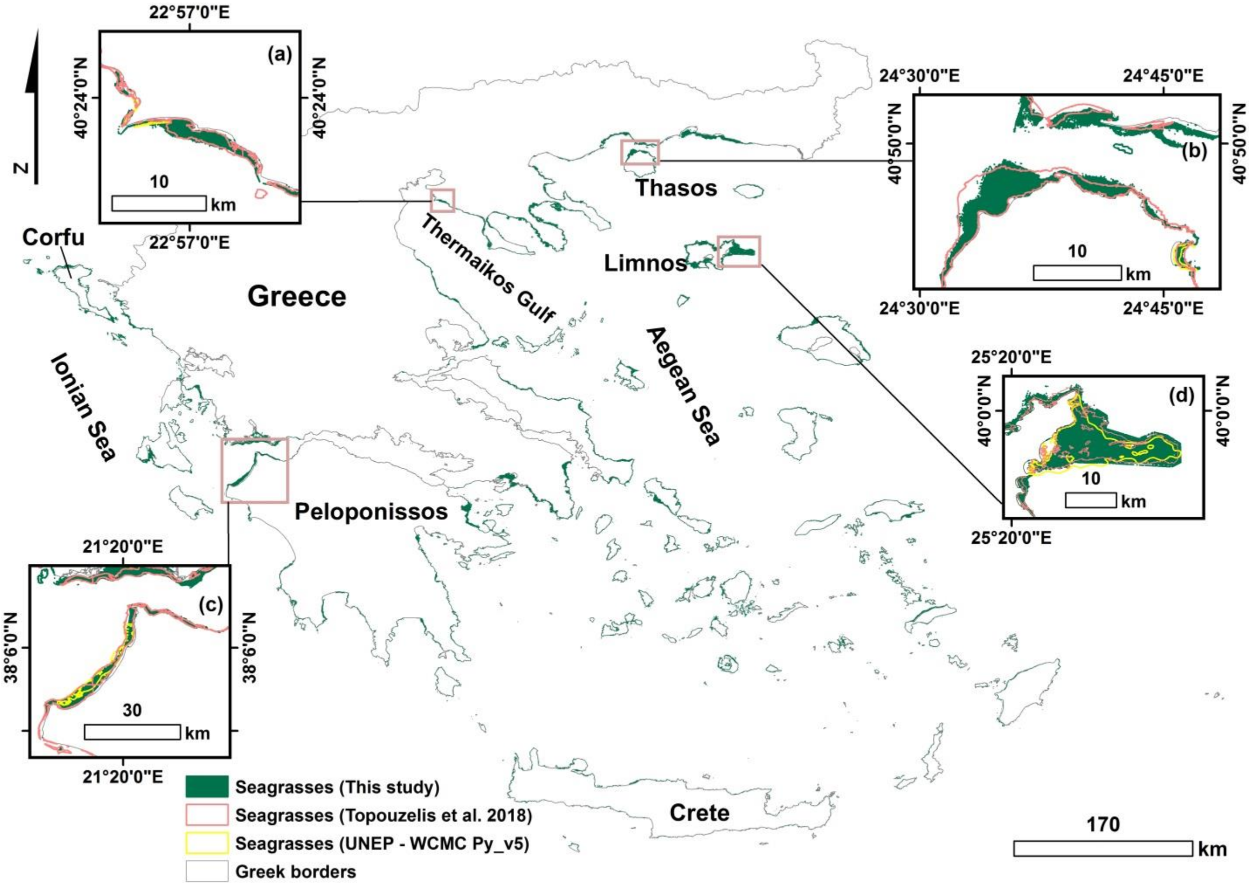


Fig. 4.6.5. Distribution of seagrasses in the Greek Seas. (a) Thermaikos Gulf, (b) Thasos Island, (c) NE Peloponissos, (d) Limnos Island. Inset maps contain results from Topouzelis et al. (2018) and UNEP-WCMC, Short FT (2017) for reference and further validation of our results. All panels are in GCS WGS84 World Geodetic System.

a. Selection of a suitable time range; the suitability relates to possible available in situ data to run the machine learning classifiers, the atmospheric, water surface and column conditions of the study area, but also the season of maximum growth of the seagrass species of interest, especially for change detection studies. Here, we have chosen one month of Sentinel-2 imagery within the period of better water column stratification of the Greek Seas.

b. Selection of suitable points that will represent land and water for land masking (if needed), polygons over deep water (for atmospheric correction), variable sunglint intensity (for sunglint correction), and sandy seabed of variable depth (for the depth invariant index calculation).

c. Accurate in situ data that will cover all the existing habitats within the study extent for training of the machine learning classifications and validation with an ideally independent data set to reduce potential bias; here, we design remotely sensed, homogenous 4x4 (1,600 m^2) polygons for the training of the machine learning model and employ an independent point-based data set for the validation. We also decided to design deep-water polygons to minimise possible misclassifications with seagrasses.

In contrast, the weaknesses of the present workflow are two:

1. Method-wise, the herein image-based, empirical algorithms (e.g. dark pixel subtraction, sunglint correction, depth invariant indices) contain inherent assumptions in their nature and necessitate a sufficient selection of pixels to produce valuable results. Concerning the sunglint correction, specifically, an image composition spanning a large period of time can amplify the artificiality of the produced pseudo-composite, causing the sunglint correction algorithm to be unable to capture any existing interference by this phenomenon.

2. Data-wise, there is a three-fold problem with Sentinel-2 applications in the remote sensing of optically shallow benthos and broadly aquatic extent. First, the tile limits of Sentinel-2 data are visible due to differences in viewing angles (odd and even detectors feature a different viewing angle) which produces striping. In turn, this artifact could severely impact classification output as it alters neighbouring reflectances. A first possible solution for striping could be the application of pseudo-invariant feature normalisation using a tile as a reference image and all the others as the slave ones—a theoretically, computationally expensive operation within GEE. A second solution is to split the initial study area into sub-areas—ultimately every tile within the visible stripes—where we could select polygons and run the classifier. The second data-related issue is the coastal aerosol band 1 which is originally in 60-m resolution in comparison to the 10-m resolution of all the other visible bands. Although on-the-fly reprojected to 10-m for visual purposes and integral towards coastal habitat mapping and SDB due to its great penetration, it causes artefacts upon application of the depth invariant indices of (Lyzenga, 1978, 1981). Therefore, we only utilise b2-b3 index during the classification step. This could be solved through the implementation of a

downscaling approach of band 1 (Brodu, 2017) into the existing workflow which is under exploration in terms of computation time efficiency. The third and last data-wise issue is the selection of training and validation data. We designed as homogeneous as possible polygons that represent seagrasses, sands, rocks and deep water based on very high spatial resolution images; however, these will be as accurate as our experienced eye will dictate us. Fig. 4.6.2 shows that at all band to band scatter plots, the designed polygons of seagrass and non-seagrass beds are not well differentiated that might have caused misclassifications. Generally, the collection of field data for the classification of remote sensing of aquatic habitats is expensive, time-consuming and sparse today. More efforts should be driven towards allocating funding for accurate and high resolution in situ data and/or advocating the sharing of open datasets that would permit regional to global projects. The search for open access data on seagrass from relevant data repositories reveals a high number, however a fraction of these are potentially suitable for use in the remote sensing domain. Therefore, it is mandatory to urge a collaborative action between seagrass and remote sensing scientists which will galvanise the development of a protocol that could be easily adapted in any seagrass bioregion for the designation of accurate and well documented with metadata, in situ data for seagrass mapping using the present workflow.

Future Endeavours

In addition to the much-needed availability of accurate in situ datasets suitable for image analysis of variable scales (from 50-cm to 30-m pixel size), we discuss three future endeavours following the use of the present cloud-based workflow:

1. Basin- (Mediterranean) to global-scale mapping and monitoring of seagrasses and related biophysical variables (specifically the climate change-related carbon sequestration): The expected lifespan of Sentinel-2 and its succeeding complementary twin mission (7.25+7.25 years) would unravel issues related to open and free, high spatial resolution data availability and allow intra-annual (seasonal) to inter-annual monitoring activities in the optically shallow grounds of seagrasses for 14.5 years by 2029 which marks the end of the announced UN decade of ocean science (UNESCO, 2018).

2. Improvement of certain stages of the present workflow: a) Incorporation of a more sophisticated atmospheric correction algorithm like Py6S (Wilson, 2012), b) Implementation of optimisation approaches for simultaneous derivation of benthic reflectance and bathymetry based on the semi-analytical inversion model of (Lee et al., 1998, 1999), c) Inclusion of Best available pixel (BAP) approach within span of

well-stratified column period which use pixel-based scores, according to both atmospheric-, season- and sensor-related issues, to produce a composite with the *Best Available Pixel* (White et al., 2014), d) Incorporation of object-based segmentation and classification methods to improve classified outputs. The main drawback of the first three improvements is that they would possibly lower the time efficiency of the present version of the chain due to the higher demand in computational power based on the need to implement look up tables and/or run radiative transfer codes.

3. Integration of seagrasses and other coastal habitats to the Analysis Ready Data (ARD) era: Recent advances in optical multispectral remote sensing (e.g. Sentinel-2, Landsat 8, Planet's Doves), cloud computing and machine learning classifiers can enable multi-scale, multi-temporal and sensor-agnostic approaches where all the aforementioned data will be pre-processed to a high scientific standard (Cloud Optimised GeoTIFF (Cloud Optimised GeoTIFF, 2018); further harnessing past, present and future remotely sensed Big Data and facilitating the near real-time measurements of physical changes of these immensely valuable habitats for Earth.

4.6.5 Conclusions

The present study introduces a complete methodological workflow for large scale, high spatial and temporal mapping and monitoring of seagrasses and other optically shallow habitats. The workflow can be easily tuned to spatial, timely and data input; here, we showcase its large spatiotemporal and time efficiency, mapping 2,510.1 km^2 of *P. oceanica* seagrasses in 40,951 km^2 of the Greek Seas utilising a 10-m Sentinel-2 based composite of 1,045 tiles in seconds. The workflow could also ingest the freely available image archive of Landsat-8 surface reflectance as input. We envisage that the herein adaptable, accurate, and time- and cost efficient cloud-based workflow will provide the vital seasonal to inter-annual baseline mapping and monitoring of seagrass ecosystems in a global scale, identifying problematic areas, resolving current trends, and assisting coastal conservation, management planning, and ultimately climate change mitigation.

5 General Discussion

Building on the individual discussions of the six published contributions of the present PhD Thesis, I herein discuss their connectivity and relevance in the light of the most significant findings of my PhD, the pertinent natural and technological environment, and the emerging needs for the future of seagrass and seascape mapping using remote sensing technology.

5.1 Important Findings and Innovations

As provided in Section 1.3. and following Chapter 4, the cohesion of this section needs to reinforce the eight most prominent and innovative findings and contributions of the present PhD:

a) The exploitation of new satellite sensors (e.g., Sentinel-2, RapidEye, and PlanetScope) for mapping and monitoring of seagrasses in multiple spatial and temporal scales and resolutions
b) The use of machine learning algorithms and frameworks for seagrass mapping and the demonstration of their increased accuracies (in comparison to simpler classifiers e.g., Maximum Likelihood)
c) The revelation of seagrass loss in northern Greece through a five-year satellite image time-series of RapidEye data
d) The identification of deep seagrass meadows in depths of 33 m in the south of Crete (southern Greece) on Sentinel-2 imagery
e) The designation and integration of the probability and uncertainty metrics for the validation of coastal aquatic remote sensing data products
f) The design and application of a semi-analytical inversion model to retrieve optical properties of the water column above and seabed of Mediterranean seagrasses from Sentinel-2 imagery
g) The introduction and exploration of multi-temporal approaches in the remote sensing of seagrasses (which address better usual optical interferences in the coastal Earth Observation e.g., clouds, sunglint, waves, turbidity)
h) The design and development of a cloud-native methodological workflow for large-scale seagrass mapping and monitoring of seagrasses, combining multi-temporal satellite image composites of Sentinel-2, independent image-annotated and field-

based training and validation data, and a machine learning framework within the cloud platform of the Google Earth Engine.

The timing of the start of this PhD was ideal to coincide with the exponential increase in the launch of Earth Observation missions (e.g. Sentinel series, Planet's Cubesats), the application, solidification and democratisation of artificial intelligence frameworks in the remote sensing domain, and the increase in computing power which has brought about the establishment and growth of cloud computing platforms. And the harmonious synergy and symbiosis of all these Earth Observation technological advances, and the resulting developments for seagrass mapping, as presented in 4.6, are a result of the research, experiments, and applications—within local servers—of the chronologically prior publications presented in 4.1 – 4.5 within local servers; and, naturally, follow, are inspired by, and build on the theories, applications and observations of dozens of scientists in the last 50 years.

The cloud-based mapping workflow of section 4.6 could offer the yeast to optimise the remote sensing-based detection of seagrass habitats and their natural seascape environment at a planetary scale. The serverless machine learning, analytics and satellite data refineries will allow this optimisation through the ensuing gains in scalability, cost and time-optimisation, adaptability, and automation.

Cloud-native, large-scale satellite-based mapping requires large-scale, well-distributed, and independent training and validation datasets for nonpartisan mapping and monitoring of seagrass and their neighbouring habitats. And the subsequent validation of the mapping data requires, as well, the development, application, and standardisation of standard and new metrics in the cloud e.g., Error Matrices, Confidence, Probability, Uncertainty (similar to the ones implemented within a local environment in 4.3)

5.2 Standardise It Now

Following the written contributions of *The Good, the Bad and the Best Practices of the Proposed Cloud-based Workflow* Discussion section of 4.6, there are two emerging needs due to the transition from the local to the cloud environment and the resulting scalability of seagrass mapping; and it is vital to meet them for improved transparency, accuracy, efficiency, and ultimately understanding.

The first concerns the need for global-scale reference data to train and validate the cloud-native workflow and the resulting mapping products. We can meet this need by designing a central repository where all the available or new field data will be collected in a standardised fashion. These field datasets can be derived from:

a) New field campaigns using optical (e.g., drones) and acoustic means (e.g., side-scan sonars), and GPS or customized mobile applications (e.g., Collect.Earth - http://www.openforis.org/tools/collect-earth.html).
b) Existing open field data found in online inventories (e.g., PANGAEA - https://www.pangaea.de/; ZENODO - https://zenodo.org/; UNEP-WCMC - https://www.unep-wcmc.org/resources-and-data)
c) Design (i.e. annotation) of new reference data based on human photo-interpretation of high-resolution satellite base maps. Such base maps are offered within cloud computational platforms, easing the photo-interpretation.

The aforementioned photo-interpretation would still necessitate experienced analysts—with image analysis and the present seagrass species in a given region—to differentiate and annotate non-biased reference data. In this case, biases arise from the annotation of mixed classes within each digitised pixel block (representative of a certain class). Photo-interpretation could be implemented to annotate the drone-based high-resolution imagery (as explained in point (a) above). This would minimise the cost of collecting reference seagrass data in situ by means of snorkelling and diving and establish a yet unexplored synergy between drones and satellites in the coastal aquatic remote sensing domain.

Machine and/or deep learning-based image analysis tools could be also used to standardise and automate the collection and annotation of such field data inventories. Already designed and implemented approaches, mainly for semi-automated and automated interpretation of coral reefs, like the ones presented in Beijbom et al. (2015), Gonzalez-Rivero et al. (2016), and Williams et al. (2019) provide inspiration and technical capacity for the employment and/or development of similar systems for seagrass photo-interpretation.

It is also vital that the described envisaged central repository of seagrass and seascape-related data should be open, public, and free. Only this way, there will be advances into the research and non-research incentives of the involved stakeholders in the planetary seagrass environment.

The second need relates to the development, integration and standardisation of quantitative validation metrics in the cloud. Cloud-native mapping approaches will benefit from the universal implementation of standard accuracy assessment approaches like the error matrices (including the metrics of Overall, Producer, User Accuracy and F-measure score) and alternative ones like the metrics of Confidence, Probability, and Uncertainty (section 4.3). Such implementation and standardisation will highlight potential bottlenecks and biases of the applied algorithmical components; will reveal possible over and under-estimations of seagrass extent and of the ranges of related biophysical parameters; and will dictate current and future sampling designs, minimum mapping and spatial assessment units, and the required amount of cal/val data in a given seagrass mapping and monitoring exercise (FAO, 2016).

Although developed for terrestrial habitats and related environment variables, the practices for effective and accurate calibration and validation described in FAO (2016) and TERN (2018) could compose a roadmap for similar standardised practices, adaptations and efforts in a coastal aquatic context. And the theories, approaches, and metrics elaborated in Roelfsema and Phinn (2008, 2013), Tulloch et al. (2017), and section 4.3 could complement the two aforementioned practices towards standardising the validation component in cloud-native mapping solutions.

5.3 Seeing the Forest for the Trees

Seagrasses are core ecosystems within the global coastal seascape which also includes corals, mangroves, tidal flats and salt marshes. These vital seascape components are geographically and functionally interconnected; recession of any of the individual component habitats of the seascape will have negative consequences for the other components (Pendleton et al. 2012) e.g. the degradation of mangroves can induce sedimentation in the water column which will, in turn, decrease the light availability for the photosynthesis of seagrasses, stressing and degrading their meadows.

The majority of seascape mapping efforts to date has focused on specific habitats of interest, being unable to see the forest for the trees. In addition, we are lacking accurate estimates of the global extent of seagrasses, with current estimates featuring a geographical and historical bias (Macreadie et al. 2019). Near-future nonpartisan data-driven mapping of the global spatial extent of seagrasses will complement existing and current mapping efforts of the neighboring seascape habitats (Giri et al. 2010; Hamilton and Casey, 2016; Murray et al. 2019; Allen Coral Atlas, 2019). The cloud-native mapping solution presented in 4.6 could offer such global-scale mapping data in the next three years, following the necessary and aforementioned standardisations and automations in artificial intelligence frameworks and reference data.

I envisage that the planetary-scale baseline mapping of seagrass ecosystems should and will be followed and/or integrated into a comprehensive, long-term cloud-based mapping and monitoring service for the entire coastal seascape, similar to the ones which exist for terrestrial ecosystems e.g., Global Forest Watch. I also envisage that with the current availability and advances in open satellite data archives (e.g., Sentinel-2, Landsat series), cloud computing, artificial intelligence and the necessary reference data, such potential service could achieve annual to inter-decadal monitoring of the global seascape through the next decades. Future holistic seascape mapping approaches will unveil related ecosystem tipping points or regime shifts, covering significant gaps in seascape knowledge, science, mapping, management, protection and conservation.

5.4. Leave No One Behind

A previously unexplored mapping endeavor at a global scale is the quantification of the carbon stocks of seascape habitats (incl. seagrasses, mangroves and tidal flats). Spatial and temporal knowledge of carbon sequestration trends and current carbon stocks could improve current climate change mitigation schemes and global carbon budgeting, as they present a “no regrets” option for climate change mitigation and adaptation. As agreed in the 2018 UN climate conference in Katowice, Poland (COP24, 2019), from 2024, a new international climate regime will oblige all countries to report both their emissions and the related progress in cutting them biennially. A total of 151 countries contain at least one blue carbon habitat, while 71 countries contain all three; these countries could certainly benefit from spatio-temporal intelligence of seascape habitats. These envisaged incentives will be better facilitated by the inclusion of these habitats in carbon trading programs and payments for ecosystem services (PES, 2019) like REDD (Reduced emissions from deforestation and degradation) and REDD+ (REDD, 2019). Through these incentives and schemes, protection, management and subsequent decrease of the ongoing degradation of seascape habitats could help all the involved countries and communities to have access to relevant climate funds and carbon trading programs.

The aforementioned climate change mitigation schemes are part of the Sustainable Development Goals—especially Goal 13: Climate Action and 14: Life below water in the context of seascape ecosystems (PES, 2019; SDG, 2019). These goals address specifically the problem of recession and loss of coastal seascapes via efforts towards reducing climate-related risk, while improving economic growth and sustainable ocean management. The advances in Earth Observation technology and the resulting wealth of global seagrass mapping data can certainly assist national, international and global institutions not only to integrate SDGs into their policies and strategies but also to achieve them. This integration and achievement is expected to trigger $26 trillion in economic benefits by 2030 (SDG, 2019).

5.5 Essential Variables

An additional near-future priority in spaceborne seagrass and seascape mapping can be identified in the assessment of the Essential Variable frameworks—the Essential Ocean Variables (EOVs) of the IOC-UNESCO Global Ocean Observing System and the Essential Biodiversity Variables (EBVs) of the Group on Earth Observations Biodiversity Observation Network (GEO BON). The variables associated to seagrasses are their cover and composition, at both total and species level percentages, and have been designed following their importance to science and society, and viability for planetary-scale realisation (Miloslavich et al. 2018). To date, the use of remotely sensed data for the measurement of EOVs and EBVs has been largely unproven or demonstrated in limited cases (Muller-Karger et al. 2018). Globally coordinated monitoring networks and respective standardised practices and assessments could create routine, multi-purpose, and interoperable data of these essential variables; and in turn, such efforts and data could generate further interest in the operational and commercial sector, and attract supplementary funding for seagrass ecosystems (Duffy et al. 2019).

6 Conclusions and Outlook

From the dinosaurs and the Aristotle to the age of climate change and the Fourth Industrial Revolution, seagrasses have been an overlooked component of our Planet. They have been relentlessly contributing their ecosystem services for all these millions of years, but more recently they have experienced degradation and loss of their area due to the impacts of climate change. And these issues are exacerbated by the current spatial and temporal biases in the existing global seagrass dataset. My four-year PhD research aimed to build scalable tools and algorithms required for seagrass mapping exploiting the technological advances of the last five years in remote sensing and Earth Observation—the new satellite data, the democratisation of artificial intelligence frameworks, and the proliferation in cloud geospatial computing. These tools have been long needed to shed light into which seagrass species are where, how much they are, how much area they have lost (or gained), and the drivers of these trends.

Experimentation and experience with the use of new satellite image archives in local and cloud servers, in a single scene and multi-temporal fashion, in empirical to analytical solutions, in temperate and tropical seagrass regions, have more recently led to the design and development of an end-to-end, cloud-based workflow for seagrass mapping. Thus far, the cloud-native workflow has provided country-scale seagrass extent data (Greece – 41,000 km^2 of coastline) and in a yet unpublished effort, it will provide seagrass mapping data for the whole of the Mediterranean seagrass region (~180,000 km^2) through the employment of thousands of Sentinel-2 images.

Such data-driven, nonpartisan seagrass mapping approaches can be utilised to measure the global magnitude, condition and dynamics of seagrasses. And such efforts would require and bring about increased automation, scalability, democratisation, interoperability, standardisations in the cloud-native mapping and monitoring of seagrass and seascape ecosystems. Suitable and open reference data—both for training and validation of the seagrass mapping—for Earth Observation data are and will be a vital component for the realisation of planetary scalability in this context.

The (near) future of seagrass and seascape mapping should be built upon the developed momentum around the value and condition of seagrass habitats; the innovations and potential of the remote sensing technology; and the capacity of and in conjunction with similar global mapping services like the Allen Coral Atlas, the Global Mangrove Watch and the Global Forest Watch. Today's momentum and foreseen potential and benefits of seagrass mapping, science and policymaking can contribute to humanity's race to adapt to and mitigate climate change impacts within a sustainable 1.5 °C of global warming above pre-industrial levels. To avoid cross tipping points and irreversible changes in biodiversity, ecosystems and climate patterns past the sustainable boundaries of Earth, we need significant coordinated networks with and investments from the commercial and private sector, and more so from not so traditional players in the game of marine science (e.g., philanthropic foundations, online companies, technology informatics companies, insurance companies, global banks, gaming industry companies). Mapping, scientific evidence and risk assessment is one thing, but structurally changing our financial and economic incentives in a climate-changed world is another.

List of Figures

3.4 Example of leaf area index (LAI: ratio of leaf area to substrate area) mapping using two high-resolution PHILLS imagery at Lee Stocking Island, Bahamas. Histograms feature the distribution of LAI frequency within the outlined box of the respective left panels (from Dierssen et al. 2003).

3.5 Example of seagrass biomass mapping using IKONOS satellite imagery and field data in Zanzibar, Tanzania (from Knudby and Nordlund, 2011).

3.6 Example of leaf area index (LAI: ratio of leaf area to substrate area) mapping using a physics-based inversion model, PRISM hyperspectral imagery and field data in Florida Bay, USA (from Hedley et al. 2016). The transects of (b) represent the white lines of (a) while the black dots are quadrat-based in-situ data, lines are model inversion results for only clean leaves and 0-25% coverage by sediment. Grey areas represent the 90% confidence intervals based on uncertainty propagation.

4.1.1 Location of survey site within (a) Thermaikos Gulf, (b) Halkidiki Peninsula, (c) Aegean Sea, Greece. The depicted Sentinel-2A satellite image in (a) is an atmospherically corrected, true color (band 2 as blue, band 3 as green, band 4 as red) composite in UTM (zone 34) system/WGS84 projection. The outline colors of the left and upper right frame indicate their exact location. The deep water polygon is a selected 115 × 115 pixel window used in the water column correction step as it has very little water leaving radiance values in all visible bands.

4.1.2 Habitats exhibited in the survey site: (from left to right) Dense *Posidonia oceanica* meadows; Shallow *Cymodocea nodosa* habitat; Photophilous algae on rocks; Sand.

4.1.3 Schematic presentation of the methodology. L1C[1] products are the initial orthorectified and spatially registered Sentinel-2A granules in UTM/WGS1984 projection, BDRF[2] is the bidirectional reflectance distribution function that defines how light is reflected at an opaque surface, R_w[3] represents water surface reflectances and R_b[4] are bottom reflectances.

4.1.4 Regression between log-transformed water-surface Sentinel-2A blue (band 2) versus Sentinel-2A green (band 3) reflectances over various habitats. An offset is applied to both logtransformed reflectances to facilitate presentation.

4.1.5 Polynomial regression between log-transformed and passed with a 5 × 5 median filter water surface blue to green reflectance ratio and in situ depth measurements from the Thermaikos survey site. The displayed polynomial equation was used to measure bathymetry in the study site.

4.1.6 Sentinel-2A RGB (10-m bands 2, 3, 4) composites of the various methodological steps and corresponding best performing classifiers in the Thermaikos Gulf study site. A. Masked L1C-composite (no correction applied), B. R_w, water surface reflectance composite (with atmospherically and geometric correction applied), C. R_b, bottom reflectance composite (with water column correction applied), i. Support Vector Machine (SVM) classification of A-composite, ii. Random Forest classification of B-composite, iii. SVM classification of C-composite. All 6 images are draped over atmospherically corrected Sentinel-2A RGB imagery in UTM (zone 34) system/WGS84 projection.

4.1.7 Bathymetry map of the survey site draped over an atmospherically corrected S2A RGB composite. Projection is UTM (zone 34) system/WGS84.

4.1.8 Plot of modeled depth versus in situ measured depth for the validation of the bathymetry map of the survey site. Depth model was derived from the relationship of Fig. 4.1.5.

4.1.9 Plot of the depth residuals from the regression of Fig. 4.1.7, between modeled depth and in situ measured depth. Note that RMSE is 1.3 m.

4.1.10 Modeled bathymetry versus log-transformed water surface reflectances R_w490, R_w560 and R_w665. The surface reflectance pixels were extracted from a sub-area on the satellite image that exhibited a constant sandy substrate but a depth variance from 1 to 5 m. Attenuation coefficients were estimated using the slope of each regression line which represents the quantity $-2K_i$ (Bierwirth et al., 1993).

4.1.11 Spectral dependence of diffuse attenuation coefficient for water surface reflectances R_w490, R_w560 and R_w665. Diffuse attenuation coefficients estimations were derived from the regressions of Fig. 4.1.9.

4.2.1 Workflow of the herein processed methodology. The Planet Image Time Series consists of three different PlanetScope Analytic Ortho Scenes. Lyzenga's empirical water column correction is adapted from Lyzenga (1978, 1981).

4.2.2 Experimental steps without the application of Unmixing-based denoising. A and B display the initial, 'noisy' blue and green Planet reflectance, respectively. C is the depth-invariant index (Lyzenga's empirical water column correction) of A and B before applying the Unmixing-based denoising (UBD). D shows Support Vector Machine classification result of C. All rasters are draped over a PlanetScope analytic Ortho RGB Scene of

17/11/2016 over our test site in Thermaikos Gulf, NW Aegean Sea, Greece, projected in UTM (zone 34) system/WGS84.

4.2.3 Fig. 4.2.3. Experimental steps with the application of Unmixing-based denoising. A and B display the denoised blue and green Planet reflectance following the UBD application on the A and B of Fig. 4.2.2., respectively. Differences in overall brightness from A and B of Fig. 4.2.3 are due to their different histogram stretch. C is the depth-invariant index of A and B after the UBD application. D shows Support Vector Machine classification result of C. All rasters are draped over a PlanetScope analytic Ortho RGB Scene of 17/11/2016 over our test site in Thermaikos Gulf, NW Aegean Sea, Greece, projected in UTM (zone 34) system/WGS84.

4.3.1 Location of survey site in (a) Schinias Marathon National Park, (b) Samaria National Park, (c) Gavdos Island, and (d) Greece.

4.3.2 The habitats of the selected sites: (a) *Posidonia oceanica* meadows at 30 m, (b) Rocky reef covered by photophilic algae, (c) *Posidonia oceanica* meadows at 5 m, and (d) Soft sandy bottom.

4.3.3 Schematic methodological workflow.

4.3.4 Downscaling of Sentinel-2 band 1 following Brodu (2017). (a), (d), (g): 4-3-2 composite images of Schinias, Samaria, and Gavdos Island. (b), (e), (h): Original band 1 (443 nm, 60-m resolution). (c), (f), (i): Downscaled band 1 at 10-m resolution. The pink and yellow squares in Fig. 4.3.4(f) show the downscaled 10-m pixels over a rocky and sandy seabed, respectively. The polygon in the three insets of 4(a), (d) and (g) indicates the location of the displayed panels within the extent of each survey site.

4.3.5 Sunglint correction of Sentinel-2 A images following Hedley et al. (2005). All images are 3-2-1 L2A composites with downscaled band 1 at 10-m resolution. Yellow polygons indicate pixels selected for the regression. (a), (b): Initial and deglinted image of Schinias. (c), (d): Initial and deglinted image of the centre o Samaria. (e), (g), (f), (h): Initial and deglinted images of NW and E Gavdos.

4.3.6 Depth-invariant bottom indices (b1-b2, b1-b3 and b2-b3) following Lyzenga (1978, 1981). Light blue polygons indicate regression pixels that represent the same substrate, sand, in different depths. (a) Schinias, (b) Samaria, (c) Gavdos.
4.3.7 Water column corrected S2-A 3-2-1 composite of Samaria survey site, and related probability and uncertainty of habitat distribution. Isobaths are based on a 5 × 5

smoothing of the SDB. (a) Substrate reflectance (R_{rs}^{b}) image up to 30.2 m of depth. (b), (c), (d) SVM-derived probability of *P. oceanica*, rocky and sandy substrate (%), respectively.

4.3.8 Satellite-derived bathymetry (SDB) of the Samaria survey site based on the ratio of coastal aerosol to green band. (a) Deglinted L2A 3-2-1 composite of the survey site. *In situ* depth data (n = 2569) are shown in light pink. (b) SDB for the whole depth range.

4.3.9 Linear regression between SDB using the b1-b3 ratio and the actual depth measurements from the Samaria site. The SDB was implemented in the analytical water column corrections.

4.3.10 Probability and uncertainty of *P. oceanica* seagrass distribution in the Samaria site (south Crete). (a) Rrs^{b} S2-A 3-2-1 composite with deep dense *P. oceanica* seagrass patch within yellow polygon. The light blue circle indicates the presence of *P. oceanica* seagrass as observed *in situ.* The golden circles depict mean depths of the *in situ* depth measurement points that fall within the same pixels. (b) Probability of occurrence of *P. oceanica* seagrass following SVM-derived classification. (c) General uncertainty (%). (d) Uncertainties of *P. oceanica* seagrass distribution lower than 20% are displayed with golden polygons.

4.3.11 Probability and uncertainty of habitat distribution in Schinias. Displayed probabilities are the most accurate results (overall accuracy) of machine learning- based classification. Bathymetric contours derived from the Hellenic Navy Hydrographic Office are displayed in purple. Masked land is displayed in grey colour. (a) Deglinted S2-A 3-2-1 composite of the survey site. (b) SVM-derived P. oceanica probability. (c) SVM-derived sandy substrate probability. (d) Uncertainty (%).

4.3.12 Probability and uncertainty of habitat distribution in Gavdos. Displayed probabilities are the most accurate results (overall accuracy) of machine learning-based classification. Bathymetric contours derived from the Hellenic Navy Hydrographic Office are displayed in purple. Masked land is displayed in grey colour. (a) Deglinted S2-A 3-2-1 composite of the survey site. (b) RF-derived P. oceanica probability. (c) RF-derived rocky substrate probability. (d) RF- derived sandy substrate probability. (e) Uncertainty (%).

4.4.1 Location of survey site within the Thermaikos Gulf and the Aegean Sea, Greece (inset map). The displayed Sentinel-2A image is a top-of-atmoshpere (toa) 4-3-2 composite in UTM (zone 34) system/WGS84 projection. The depicted bathymetric contours are provided by the Hellenic Navy Hydrographic Service (HNHS). The displayed waterways are provided by the OpenStreetMap (http://download.geofabrik.de/europe/greece.html).

4.4.2 Exponential regression between the log-transformed ratio of sub-surface remote sensing coastal aerosol (443.9 nm) and green (560 nm) reflectances, and in situ measured bathymetry from the Thermaikos survey site. The displayed exponential algorithm was used to empirically derived bathymetry in the present study.

4.4.3 Examples of retrieved inherent parameters from both C2RCC and the herein parameterised HOPE (Lee et al. 1999) model in the Thermaikos Gulf (see Fig. 4.4.1). (a) Absorption coefficient of phytoplankton pigments at 443.9 nm, (b) Backscattering coefficient of suspended particles at 560 nm, (c) Attenuation coefficient at 496.6 nm.

4.4.4 5x5 median-filtered, Satellite-derived bathymetry (H) of the Thermaikos Gulf (see Fig. 4.4.1) based on the empirical regression of Equations (17)-(18) shown in Fig 2. The displayed extent follows the application of an optically shallow mask until 16 m after Traganos and Reinartz (2018a).

4.4.5 Validation of the Satellite-derived Bathymetry (H) shown in Fig. 2 and 4.4.4. H was utilised in Equation (5) to invert bottom reflectances.

4.4.6 Machine-learning-based classification of P. oceanica seagrass—shown in green in panel (b) —bottom reflectances. Accuracy assessment is shown in Table 2.

4.4.7 Mean spectral reflectances (plus standard deviation) over 320 pixels of classified as P. oceanica seagrass beds for the four first bands of Sentinel-2A (all in 10m/pixel resolution) for all the estimated products in the present study. Description of reflectance denominators are provided in section 2.3.2. The secondary y axis displays reflectance values only for the R_{toa}.

4.4.8 Mean spectral reflectances of the classified habitats for the four first bands of Sentinel-2A (all in 10m/pixel resolution) for bottom reflectance product. PO and CN represent Posidonia oceanica and Cymodocea nodosa seagrasses, respectively.

4.5.1 Location of survey site within (A) Thermaikos Gulf, (B) Aegean Sea, Greece. The displayed RapidEye imagery is a non-atmospherically corrected, true color (band 1 as blue, band 2 as green, band 3 as red) composite in UTM (zone 34) system/WGS84 projection. The imagery was acquired on 22/06/2016 (RE16 in text). The red polygon in (B) depicts the location of (A) within the Thermaikos Gulf. The deep water polygon represents a ~160 × 160 pixel window implemented in the water column correction of the image time series as it represents an area with very little water leaving radiance values in all three bands.

4.5.2 Schematic representation of the methodology. [1]L3A ortho products are the initial radiometric, sensor, and geometrically corrected RapidEye images in UTM/WGS1984 projection, [2]R represent atmospherically-corrected (FLAASH module), at-water surface reflectances, [3]R_{rs} are remote sensing reflectances, transformed from R using Equation (1), [4]Rb are water-column-corrected, bottom reflectances using the analytical model of Maritorena et al. (1994).

4.5.3 Polynomial regression between the log-transformed ratio of blue and green remote sensing reflectances, R_{rs}, and *in situ* depth measurements from the Thermaikos survey site. The shown polynomial equation was implemented to estimate the bathymetry map displayed in (C) of Fig. 4.5.4.

4.5.4 Methodological steps from atmospheric to water column correction in order of successive processing. All four panels are true color RapidEye image composites (22/06/2016; RE16 in text) projected in UTM (zone 34) system/WGS84. (A) Non-atmospherically corrected composite. (B) Atmospherically-corrected composite using the FLAASH module. (C) Satellite-derived Bathymetry map of the survey site draped over the atmospherically-corrected composite of (B) using the site-specific polynomial algorithm of Equation (2) as shown on Fig. 4.5.3. We applied a 5 × 5 low-pass filter on the initial ratio-derived bathymetry (not shown here) to reduce potential noise which would be transferred to the water-column corrected product. (D) Water-column corrected composite following application of the water column correction algorithm of Maritorena et al. (1994) draped over the atmospherically-corrected composite of (B) and masked using the optically deep limit of 16.5 m to enhance bottom features and potentially increase classification accuracies

4.5.5 Plot of Satellite-derived Bathymetry (SDB) vs. *in situ* measured depth for the validation of the bathymetry map of the Thermaikos Gulf (Fig. 4.5.4C). SDB was derived from Equation (2). Regressed SDB have been previously smoothed with a 5 × 5 low pass filter to reduce unwanted noise.

4.5.6 Classified water-column-corrected RapidEye composites from the 4 years using Random Forest machine learning classifier (100 trees). The frames on the upper right of each panel indicate the date of each RapidEye image. (A) RE11—Overall accuracy: 73.5%. (B) RE12—Overall accuracy: 81%. (C) RE15—Overall accuracy: 78.5%. (D) RE16—Overall accuracy: 82%.

4.5.7 Interannual change detection of seagrasses in the Thermaikos Survey site between 2011 and 2016 using RapidEye satellite images. The trajectory plot displays change of area (in hectares; y-axis) over the years (x-axis) of *Posidonia oceanica* and *Cymodocea nodosa* species, and of total seagrass area. Linea regression black lines (m = slope) show approximate trend in area between 2011 and 2016. *Posidonia oceanica* seagrass is decreasing at 11.2 ha/yr, *Cymodocea nodosa* seagrass is increasing at 18 ha/yr, while total seagrass area is expanding at 6.8 ha/yr.

4.5.8 Change in seagrass distribution in the Thermaikos survey site between 2011 and 2016 for (A) *Posidonia oceanica* and (B) *Cymodocea nodosa*. Between 2011 and 2016, *P. oceanica* seagrass meadows have declined by 4.1%, while *C. nodosa* seagrasses have increased by 17.7%.

4.6.1 Geographical location of survey site and training polygons of the herein considered classes. All polygons are in GCS WGS84 World Geodetic System.

4.6.2 Scatter plots of the first four, sunglint-corrected Sentinel-2 bands depicting waveband reflectivity of the herein 1,457 polygons for the whole extent of the study area. Seagrasses are in green circles and non-seagrasses are in light blue triangles.

4.6.3 Methodological workflow of the present study within Google Earth Engine. OA denotes overal accuracy. In the present study we did not implement step 2 (due to the use of a coastline buffer), however, we include it as it is an important component of the methodological chain.

4.6.4 Successive stages of the developed workflow through the resulting false-colour (b1-b2-b3) Sentinel-2 composites. (a) Initial S2, (b) Cloud- and atmospherically-corrected, (c) Sunglint-corrected, (d) Depth invariant index b2-b3, (e) Support Vector Machines-based classified product draped over the b2-b3 depth invariant index layer, (f) PlanetScope surface reflectance product (as imaged by Planet's Doves) in natural colour for high resolution reference (3 m) (ID: 20170828_084352_100e/20170828_084352_100e_3B_AnalyticMS_SR). The pink squares indicate sunglint presence in (b) and its correction in (c). The green and yellow polygons show employed seagrass and sand pixels in the machine learning classification. All panels are in GCS WGS84 World Geodetic System.

4.6.5 Distribution of seagrasses in the Greek Seas. (a) Thermaikos Gulf, (b) Thasos Island, (c) NE Peloponissos, (d) Limnos Island. Inset maps contain results from Topouzelis et al.

(2018) and UNEP-WCMC, Short FT (2017) for reference and further validation of our results. All panels are in GCS WGS84 World Geodetic System.

List of Tables

Bibliography

Allen Coral Atlas (2019). Available online: https://allencoralatlas.org/ (accessed on 09 September 2019)

Andréfouët S., Kraemer P., Torres-Puliza, D., Joyce K. E., Hochberg E. J., Garza-Perez J. R. (2003). Multi-site evaluation of IKONOS data for classification of tropical reef environments. Remote Sensing of Environment 88, 123-143. doi: https://doi.org/10.1016/j.rse.2003.04.005

Amazon AWS (2019). Available online: https://aws.amazon.com/(accessed on 03 August 2019)

Arnaud-Haond S., Duarte C.M., Diaz-Almela E., Marbà N., Sintes T., Serrão E.A. (2012). Implications of extreme life span in clonal organisms: millenary clones in meadows of the threatened seagrass Posidonia oceanica. PLoS One 7. doi: http://dx.doi.org/10.1371/journal.pone.0030454.

Armstrong, R.A. (1993). Remote sensing of submerged vegetation canopies for biomass estimation. Int. J. Remote Sens. 14, 621–627.

Barillé L., Robin M., Harin N., Bargain A., Launeau P. (2010). Increase in seagrass distribution at Bourgneuf Bay (France) detected by spatial remote sensing. Aquat. Bot. 92, 185–194. doi: 10.1016/j.aquabot.2009.11.006

Beijbom O. et al. (2015). Towards Automated Annotation of Benthic Survey Images: Variability of Human Experts and Operational Modes of Automation. PLoS ONE 10(7): e0130312. doi: https://doi.org/10.1371/journal.pone.0130312

Belgiu, M., Dragut, L., (2016). Random forest in remote sensing: a review of applications and future directions. Isprs J. Photogramm. Remote Sens. 114, 24–31.

Bierwirth, P.N., Lee, T.J., Burne, R.V., (1993). Shallow sea-floor reflectance and water depth derived by unmixing multispectral imagery. Photogramm. Eng. Remote Sens. 59, 331–338.

Blaschke, T. (2010). Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65(1), pp. 2-16.

Bonacorsi M., Pergent-Martini C., Bréand N., Pergent G. (2013). Is Posidonia oceanica regression a general feature in the Mediterranean Sea? Mediterranean Marine Science 14, 193–203. doi: http://dx.doi.org/10.12681/mms.334.

Borfecchia F., Cecco, L. De Martini S., Ceriola G., Bollanos S., Vlachopoulos G., Valiante L. M., Belmonte A., Micheli C. (2013). Posidonia oceanica genetic and biometry mapping through high-resolution satellite spectral vegetation indices and seatruth calibration. Int. J. Remote Sens. 34, 4680–4701. doi: http://dx.doi.org/10.1080/ 01431161.2013.781701.

Boström, C. and Bonsdorff, E. (1997). Community structure and spatial variation of benthic invertebrates associated with Zostera marina (L.) beds in the northern Baltic Sea. Journal of Sea Research 37 (1), pp. 153–166.

Boudouresque, C.F., et al. (2012). Protection and Conservation of Posidonia Oceanica Meadows. RAMOGE pub, Tunis.

Bramante, James F., Raju, Durairaju Kumaran, Sin, Tsai Min, (2013). Multispectral derivation of bathymetry in Singapore's shallow, turbid waters. Int. J. Remote Sens. Vol.34 (Iss. 6).

Brando V. E., J. M., Anstee M. Wettle A. G. Dekker S. R. Phinn, Roelfsema R. (2009). A physics based retrieval and quality assessment of bathymetry from suboptimal hyperspectral data. Remote Sensing of Environment 113: 755-770. doi:10.1016/ j.rse. 2008.12.003.

Brando, V. E., and A.G. Dekker. (2003). Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. IEEE Transactions on Geoscience and Remote Sensing 41: 1378–1387. doi: 10.1109/TGRS.2003.812907

Breiman, L., (2001). Random forests. Mach. Learn. 45, 5–32. http://dx.doi.org/10.1023/ A:1010933404324.

Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, (1984). C.J. Classification and Regression Trees. In Chapman and Hall/CRC; CRC Press: Boca Raton, FL, USA.

Bricaud, A., A. Morel, and L. Prieur. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. Limnology and Oceanography 26 (1): 43-53. doi:10.4319/lo.1981.26.1.0043.

Brockmann, C., R. Doerffer, P. Roland, M. Peters, K. Stelzer, S. Embacher, and A. Ruescas. (2016). Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. in Living Planet Symposium, Proceedings of the conference held in May, 9-13 May 2016 in Prague, Czech Republic. edited by L. Ouwehand. ESA-SP Volume 740.

Brodu, (2017). N. Super-Resolving Multiresolution Images with Band-Independent Geometry of Multispectral Pixels. IEEE Trans. Geosci. Remote Sens. 55, 4610–4617

Call, K. A., Hardy, J. T., and Wallin, D. O. (2003). Coral reef habitat discrimination using multivariate spectral analysis and satellite remote sensing. Int. J. Remote Sens. 24, 2627–2639. doi: 10.1080/0143116031000066990

Calvo, S., Ciraolo, G., and la Loggia, G. (2003). Monitoring Posidonia oceanica meadows in a Mediterranean coastal lagoon (Stagnone, Italy) by means of a neural network and (ISODATA) classification methods. Int. J. Remote Sens. 24, 2703–2716. doi: 10.1080/0143116031000066882

Campagne C.S., Salles J.M., Boissery P., Deter J. (2014). The seagrass Posidonia oceanica: ecosystem services identification and economic evaluation of goods and benefits. Marine Pollution Bulletin 97, 391–400. doi:http://dx.doi.org/10.1016/j.marpolbul.2015.05.061.

Carder, K. L., F. R. Chen, Z.P. Lee, and S.K. Hawes. (1999). Semi-analytic moderate resolution imaging spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures. Journal of Geophysical Research 104: 5403-5421. doi:10.1029/1998JC900082

Casotti, R.; Landolfi, A.; Brunet, C.; Ortenzio, F.D.; Mangoni, O.; Ribera d'Alcala, M. (2003). Composition and dynamics of the phytoplankton of the Ionian Sea (eastern Mediterranean). J. Geophys. Res. 108.

Catapult (2017). Sentinel-2 Analysis Ready Data: Standards and automated production. Available online: https://media.sa.catapult.org.uk/wpcontent/uploads/2017/07/09123212/Outputs-Sentinel-2-part-2-v0.2.pdf?lipi=urn%3Ali%3Apage%3Ad_flagship3_pulse_read%3BV1gVuKlFQbu0zMOmea6Qrw%3D%3D (accessed on 07 July 2019)

Cerra, D., Bieniarz, J., Beyer, F., Tian, J., Muller, R., Jarmer, T. and Reinartz, P. (2016). Cloud removal in image time series through sparse reconstruction from random measurements. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9(8), pp. 3615-3628.

Cerra, D., Bieniarz, J., Muller, R., Storch, T. and Reinartz, P. (2015). Restoration of simulated enmap data through sparse spectral unmixing. Remote Sensing 7(10), pp. 13190-13207.

Cerra, D., P. Gege, R. Müller, and P. Reinartz. (2013). Exploiting noisy hyperspectral bands for water analysis In 33th EARSeL Symposium Proceedings, 43-48.

Cerra, D., Muller, R. and Reinartz, P., 2014. Noise reduction in hyperspectral images through spectral unmixing. IEEE Geoscience and Remote Sensing Letters 11(1), pp. 109–113.

Cerra, D., D. Traganos, P. Gege, and P. Reinartz. (2017). Unmixing-based Denoising as a pre-processing step for Coral Reef Analysis. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-1/W1, 2017 ISPRS Hannover.doi: 10.5194/isprs-archives-XLII-1-W1-279-2017.

Chang, C., Lin, C., (2013). LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 1–39. http://dx.doi.org/10.1145/1961189.1961199.

Chronis, G., et al., 1996. Downward fluxes. In: Tselepides, A., Papadopoulou, K.N., Polychronaki, T. (Eds.), CINCS: Pelagic-Benthic Coupling the Oligotrophic Cretan Sea, MAST-II Mediterranean Targeted Project. pp. 82–111.

Cloud Optimized GeoTIFF. (2018). An Imagery Format for Cloud-Native Geospatial Processing. 2018. Available online: http://www.cogeo.org/ (accessed on 3 August 2019).

Coll M. et al. (2010). The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. PLoS One 5. doi: http://dx.doi.org/10.1371/journal.pone.0011842.

Collin, A., and J.L. Hench. (2012). Towards deeper measurements of tropical reefscape structure using the WorldView-2 spaceborne sensor. Remote Sensing 4 (5): 1425-1447. doi: 10.3390/rs4051425.

Collison, A., Wilson, N. (2017). Planet Surface Reflectance Product. Available online: https://assets.planet.com/marketing/PDF/Planet_Surface_Reflectance_Technical_White_Paper.pdf (accessed on 07 July 2019)

Congalton, R.G., (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46. http://dx.doi.org/10.1016/0034-4257(91)90048-b.

COP24: Key outcomes agreed at the UN climate talks in Katowice (2018). Available online: https://www.carbonbrief.org/cop24-key-outcomes-agreed-at-the-un-climate-talks-inkatowice?fbclid=IwAR3JnLrnlSkpMGRn26AltT5CJ_N4cjULbkffINN16ap3eXRbymypy8DOfhY. (accessed on 10 September 2019).

Copernicus DIAS (2019). Available online: https://www.copernicus.eu/en/access-data/dias (accessed on 03 August 2019).

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., et al. (1997). The value of the world's ecosystem services and natural capital. Nature 387, 253–260. doi: 10.1038/387253a0

Cullen-Unsworth, L., Unsworth, R., (2013). Seagrass meadows, ecosystem services, and sustainability. Environ. Sci. Policy Sustain. Dev. 55, 14–28. https://doi.org/10.1080/00139157.2013.785864.

Danovaro, R., Della Croce, N., et al., (1996). Seasonal and spatial distribution of pico-,-nano-, and microparticulate matter and bacterioplankton in the Eastern Mediterranean Sea. In: Tselepides, A., Papadopoulou, K.-N., Polychronaki, T. (Eds.), CINCS: Pelagic-Benthic Coupling the Oligotrophic Cretan Sea, MAST-II Mediterranean Targeted Project. pp. 69–81.

Dekker A. G., Brando V. E., Anstee J. M. (2005). Retrospective seagrass change detection in a shallow coastal tidal Australian lake. Remote Sensing of Environment 97, 415–433. doi: 10.1016/j.rse.2005.02.017

Dekker A., Brando V., Anstee J., Fyfe S., Malthus T., Karpouzli E. (2006). Remote sensing of seagrass ecosystems: use of spaceborne and airborne sensors. In: Larkum, A.W.D., Orth, R.J., Duarte, C.M. (Eds.), Seagrasses: Biology, Ecology and Conservation. Springer, Dordrecht, pp. 347–359. doi: http://dx.doi.org/10.1007/978-1- 4020-2983-7_15.

Dekker A. G. et al. (2011). Intercomparison of shallow water bathymetry, hydro-optics, and benthos mapping techniques in Australia and Caribbean coastal environments. Limnology and Oceanography: Methods 9: 396-425. doi: 10:4319/lom.2011.9.396

Descartes Labs (2017). A data refinery, built to understand our planet. Available online: https://medium.com/@DescartesLabs/announcing-our-30m-seriesb-e658d58b4024 (accessed on 07 July 2019)

Dierssen H. M., Zimmerman R. C., Leathers R. A., Downes T. V., Davis C. O. (2003). Ocean color remote sensing of seagrass and bathymetry in the Bahamas Banks by high-resolution airborne imagery. Limnology and Oceanography 48: 444–455. doi: 10. 4319/lo.2003.48.1_part_2.0444.

Dörnhöfer, K., A. Göritz, P. Gege, B. Pflug, N. Oppelt. (2016). Water constituents and water depth retrieval from Sentinel-2A - A first evaluation in an oligotrophic lake. Remote Sensing 8 (11): 941. doi: 10.3390/rs8110941

Duarte, C.M.; Borum, J.; Short, F.T.; Walker, D.I. (2008). Seagrass Ecosystems: Their Global Status and Prospects. In Aquatic Ecosystems: Trends and Global Prospects; Polunin, N., Ed.; Cambridge University Press: Cambridge, UK.

Duffy J.E. et al. (2019). Toward a coordinated global observing system for seagrasses and marine macroalgae. Frontiers in Marine Science. doi: https://doi.org/10.3389/fmars.2019.00317

Eugenio, F., Marcello, J., and Martin, J. (2015). High-resolution maps of bathymetry and benthic habitats in shallow-water environments using multispectral remote sensing imagery. IEEE Transactions on Geoscience and Remote Sensing 53 (7): 3539-3549.

ESA, (2015). SENTINEL-2 User Handbook. 64 pp. doi:GMES-S1OP-EOPG-TN-13-0001

FAO (2016). Map Accuracy Assessment and Area Estimation: A Practical Guide. Available online: http://www.fao.org/3/a-i5601e.pdf (accessed on 03 September 2019).

Fornes A., Basterretxea G., Orfila A., Jordi A., Alvarez A., Tintore J. (2006). Mapping Posidonia oceanica from IKONOS. ISPRS J. Photogramm. Remote Sens. 60, 315–322. doi: http://dx.doi.org/10.1016/j.isprsjprs.2006.04.002.

Fourqurean J.W. et al. (2012). Seagrass ecosystems as a globally significant carbon stock. Nature Geoscience 5, 505-509. doi: https://doi.org/10.1038/ngeo1477

Fritz, C., Dörmhöfer, K., Schneider, T., Geist, J., and Oppelt, N. (2017). Mapping submerged aquatic vegetation using RapidEye satellite data: the example of Lake Kummerow (Germany). Water 9:510. doi: 10.3390/w9070510

Fyfe, S.K., Dekker, A.G., (2001). Seagrass species: are they spectrally distinct? 2001 IEEE, Sydney. http://dx.doi.org/10.1109/IGARSS.2001.978147.

Garrido M., Lafabrie C., Torre F., Fernandez C., Pasqualini V. (2013). Resilience and stability of Cymodocea nodosa seagrass meadows over the last four decades in a Mediterranean lagoon. Estuar. Coast. Shelf Sci. 130, 89–98. doi: 10.1016/j.ecss.2013.05.035

Gascon F., C. Bouzinac, O. Thepaut, M. Jung, B. Francesconi, J. Louis, V. Lonjou et al. (2017). Copernicus Sentinel-2A calibration and products validation status. Remote Sensing 9 (6): 584. doi: 10.3390/rs9060584

Gerakaris, V.; Panayotidis, P.; Tsiamis, K.; Nikolaidou, A.; Economou-Amili, A. Posidonia oceanica meadows in Greek seas: Lower depth limits and meadow densities. In Proceedings of the 5th Mediterranean Symposium on Marine Vegetation, Portorož, Slovenia, 27–28 October 2014.

Gernez P, D. Doxaran and L. Barillé. (2017). Shellfish aquaculture from space: Potential of Sentinel2 to monitor tide-driven changes in turbidity, chlorophyll concentration and oyster physiological response at the scale of an oyster farm. Frontiers in Marine Science 4: 137. doi: 10.3389/fmars.2017.00137

Giannoulaki M., Belluscio A., Colloca F., Fraschetti S., Scardi M., Smith C., Panayotidis P., Valavanis V., Spedicato M.T. (2013). Mediterranean Sensitive Habitats. DG MARE Specific Contract SI2.600741, Final Report. 557 pp.

Giardino, C., Bresciani, M., Fava, F., Matta, E., Brando, V. E., and Colombo, R. (2015). Mapping submerged habitats and mangroves of Lampi Island Marine National Park (Myanmar) from in situ and satellite observations. Remote Sens. 8:2. doi: 10.3390/rs8010002

Giri C., Ochieng E., Tieszen L. L., Zhu Z., Singh A., Loveland T., Masek J., Duke N. (2010). Status and distribution of mangrove forests of the world using earth observation satellite data. Global Ecology and Biogeography 20, 154-159. doi: https://doi.org/10.1111/j.1466-8238.2010.00584.x

Gislason P. O., Benediktsson J. A., Sveinsson J. R. (2006). Random forests for land cover classification. Pattern Recognit. Lett. 27, 294–300. doi: 10.1016/j.patrec.2005.08.011

Githaiga MN, Kairo J.G, Gilpin L., Huxham M. (2017). Carbon storage in the seagrass meadows of Gazi Bay, Kenya. PLoS ONE 12(5): e0177001. doi: https://doi.org/10.1371/journal.pone.0177001

González-Rivero M., Beijbom O., Rodriguez-Ramirez A., Holtrop T., González-Marrero Y., Ganase A., Roelfsema C.; Phinn S., Hoegh-Guldberg O. (2016). Scaling up Ecological Measurements of Coral Reefs Using Semi-Automated Field Image Collection and Analysis. Remote Sens. 8, 30.

Goodchild M. et al. (2012). Next-generation Digital Earth. PNAS 109, 11088-11094. doi: https://doi.org/10.1073/pnas.1202383109

Goodman, J.A. and S.L. Ustin. (2007). Classification of benthic composition in a coral reef environment using spectral unmixing. Journal of Applied Remote Sensing 1. doi:10.1117/1.2815907.

Gordon, H.R., O. B. Brown, and M.M. Jacobs, (1975). Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. Applied Optics 14(2): 417-427. doi: 10.1364/AO.14.000417

Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27, doi: https://doi.org/10.1016/j.rse.2017.06.031

Green E.P., Mumby P.J., Edwards A.J., Clark C.D. (2000). Remote sensing handbook for tropical coastal management. In: Edwards, Alasdair J. (Ed.), Coastal Management Sourcebooks 3. UNESCO, Paris x + 316 pp.

Green E.P., and Short F.T. (2003). World atlas of seagrasses. Prepared by UNEP World Conservation Monitoring Centre. Berkeley (California, USA): University of California. 332 pp, url: http://data.unep-wcmc.org/datasets/9

Hamana, M., Komatsu, T., (2016). Real-Time Classification of Seagrass Meadows on Flat Bottom with Bathymetric Data Measured by a Narrow Multibeam Sonar System. Remote Sens. 8, 96.

Harmel, T., M. Chami, T. Tormos, N. Reynaud, and P.A. Danis. (2018). Sunglint correction of the Multi-Spectral Instrument (MSI)-SENTINEL-2 imagery over inland and sea waters from SWIR bands. Remote Sensing of Environment 204: 308-321. doi: 10.1016/j.rse.2017.10.022

Hamilton S., Casey D. (2016). Creation of a high spatio-temporal resolution global database of continuous mangrove forest cover for the 21st century (CGMFC-21). Global Ecology and Biogeography 25, 729-738. doi: https://doi.org/10.1111/geb.12449

Hansen M. C., Potapov P. V., Moore R., Hancher M., Turubanova S. A. A., Tyukavina A., Thau D., Stehman S. V., Goetz S. J., Loveland T.R. et al. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. Science, 342, 850–853, doi: 10.1126/science.1244693

Hedley, J.D.; Harborne, A.R.; Mumby, P.J. (2005). Technical note: Simple and robust removal of sun glint for mapping shallow-water benthos. Int. J. Remote Sens. 26, 2107–2112.

Hedley, J.D., Roelfsema, C.M., et al., (2016). Remote sensing of Coral Reefs for monitoring and management: a review. Remote Sens. 8 (2), 118

Hedley J., Russell B., Randolph K., Dierssen H. M. (2016). A physics-based method for the remote sensing of seagrasses. Remote Sensing of Environment 174, 134-147. doi: 10.1016/j.rse.2015.12.001

Hedley J., Roelfsema C., Phinn S.R. (2009). Efficient radiative transfer model inversion for remote sensing applications. Remote Sensing Environment 113: 2527-2532. doi:10.1016/ j.rse.2009.07.008.

HNHS. Greek Coastline at Scale 1:90,000. (2018). Available online: https://www.hnhs.gr/en/?option=com_ opencart&Itemid=268&route=product/product&path=86&producp_id=271 (accessed on 3 August 2019).

Hossain M.S., Bujang J.S., Zakaria M.H., Hashim M. (2014). The application of remote sensing to seagrass ecosystems: an overview and future research prospects. Int. J. Remote Sens. 36, 61–114. doi: http://dx.doi.org/10.1080/01431161.2014.990649

Karageorgis, A.P., Anagnostou, C.L., (2001). Particulate matter spatial–temporal distribution and associated surface sediment properties: Thermaikos Gulf and Sporades Basin, NW Aegean Sea. Cont. Shelf Res. 21, 2141–2153. http://dx.doi.org/10.1016/S0278-4343(01)00048-6.

Kirk, J. T. O. (1994). Light and Photosynthesis in Aquatic Ecosystems. Cambridge: University Press.

Klonowski, W., Fearns P. R. C. S., Lynch M. (2007). Retrieving key benthic cover types and bathymetry from hyperspectral imagery. Journal of Applied Remote Sensing 1(1): 011505. doi: 10.1117/1.2816113.

Knudby A., Newman C., Shaghude Y., Muhando C. (2010). Simple and effective monitoring of historic changes in nearshore environments using the free archive of Landsat imagery. Int. J. Appl. Earth Obs. Geoinf. 12, 116–122. doi: 10.1016/j.jag.2009.09.002

Knudby A., Nordlund L. (2011). Remote sensing of seagrasses in a patchy multi-species environment. Int. J. Remote Sens. 32, 2227–2244. doi: http://dx.doi.org/10.1080/01431161003692057.

Koedsin, W., Intararuang, W., Ritchie, R.J., Huete, A., (2016). An Integrated Field and Remote Sensing Method for Mapping Seagrass Species, Cover, and Biomass in Southern Thailand. Remote Sens. 8 (4), 292.

Krestenitis, Y.N., Kombiadou, K.D., Savvidis, Y.G., (2007). Modelling the cohesive sediment transport in the marine environment: the case of Thermaikos Gulf. Ocean Sci. 3, 91–104. http://dx.doi.org/10.5194/os-3-91-2007.

Lamb, J.B., Water, J.A.J.M. Van De, Bourne, D.G., Altier, C. (2017). Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. Science 355(6326), pp. 731-733.

Landsat (2017). U.S. Landsat Analysis Ready Data (ARD). Available online: https://landsat.usgs.gov/ard (accessed on 07 July 2019)

Lee, Z.P., Du, K.P., Arnone, R., (2000). A model for the diffuse attenuation coefficient of downwelling irradiance. J. Geophys. Res. 110.

Lee, Z., K.L. Carder, R.F. Chen, and T.G. Peacock. (2001). Properties of the water column and bottom derived from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data. Journal of Geophysical Research 106: 11639–11651. doi: 10.1029/2000JC000554

Lee Z., Carder K. L., Mobley C. D., Steward R. G., Patch J.S. (1998). Hyperspectral remote sensing for shallow waters. I. A semianalytical model. Applied Optics 37 (27): 6329-6338. doi: 10.1364/AO.37.006329.

Lee Z., Carder K. L., Mobley C. D., Steward R. G., Patch J.S. (1998). Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. Applied Optics 38(18): 3831–3843.

Lee, Z. et al. (2007). Water and bottom properties of a coastal environment derived from Hyperion data measured from the EO-1 spacecraft platform. Journal of Applied Remote Sensing 1: 011502. doi:10.1117/1.2822610.

Leriche, A., V. Pasqualini, C.F. Boudouresque, G. Bernard, P. Bonhomme, P. Clabaut, J. Denis. (2006). Spatial, temporal and structural variations of a Posidonia oceanica seagrass meadows facing human activities. Aquatic Botany 84: 287-293. doi: 10.1016/j.aquabot.2005.10.001

Levin, N., Coll, M., Fraschetti, S., Gal, G., et al., (2014). Biodiversity data requirements for systematic conservation planning in the Mediterranean Sea. Mar. Ecol. Prog. Ser. 508, 261–281.

Lyons, M. B., Phinn, S. R., and Roelfsema, C. M. (2010). Long term monitoring of seagrass distribution in Moreton Bay, Australia, from 1972–2010 using landsat, M. S. S., TM, ETM," in 2010 IEEE, IGARSS (Honolulu), 25–30.

Lyons, M., Phinn, S. R., and Roelfsema, C. M. (2011). Integrating quickbird multi-spectral satellite and field data. Remote Sens. 3, 42–64. doi: 10.3390/rs3010042

Lyons M. B., Phinn S. R., Roelfsema C. M. (2012). Long term land cover and seagrass mapping using Landsat and object-based image analysis from 1972 to 2010 in the coastal environment of South East Queensland, Australia. ISPRS J. Photogramm. Remote Sens. 71, 34–46. doi: 10.1016/j.isprsjprs.2012.05.002

Lyons, M., Roelfsema, C. M., and Phinn, S. R. (2013). Towards understanding temporal and spatial dynamics of seagrass landscapes using time-series remote sensing. Estuar. Coast. Shelf Sci. 120, 42–53. doi: 10.1016/j.ecss.2013.01.015

Lyzenga D.R.. (1978). Passive remote sensing techniques for mapping water depth and bottom features. Applied Optics 17: 379–383. doi: 10.1364/AO.17.000379

Lyzenga D.R. (1981). Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. International Journal of Remote Sensing 2 (1): 71-82. doi: 10.1080/01431168108948342.

Macreadie P. I. et al. (2019). The future of Blue Carbon science. Nature Communications 10, 3998. doi: https://doi.org/10.1038/s41467-019-11693-w

Marbà, N., Díaz-Almela E., Duarte C. M. (2014). Mediterranean seagrass (Posidonia oceanica) loss between 1842 and 2009. Biological Conservation 176: 183–190. doi: 10.1016/j.biocon.2014.05.024.

Marbà J. C. N, Duarte C. M. (2012). Mediterranean seagrass vulnerable to regional climate warming. Nature Climate Change 2: 821–824. doi: 10.1038/nclimate1533

Maritorena S., Morel A., Gentili B. (1994). Diffuse reflectance of oceanic shallow waters: Influence of water depth and bottom albedo. Limnol. Oceanogr. 39, 1689–1703. doi: http://dx.doi.org/10.4319/lo.1994.39.7.1689

Mateo, M.A., Romero J. (1997) Detritus dynamics in the seagrass Posidonia oceanica: elements for an ecosystem carbon and nutrients budget. Marine Ecology Progress Series 151: 43-53. doi: 10.3354/meps151043

Matta E., Aiello M., Bresciani M., Gianinetto M., Musanti M., Giardino C. (2014). Mapping Posidonia Meadow From High Spatial Resolution Images in the Gulf of Oristano (Italy). 2014 IEEE, Quebec City, 5152–5155. doi: http://dx.doi.org/10. 1109/IGARSS.2014.6947658.

Maulik, U., Chakraborty, D., (2017). Remote sensing image classification: a survey of support-vector-machine-based advanced techniques. IEEE Geosc. Remote Sens. Mag. 5, 33–52.

Microsoft Azure (2019). Available online: https://azure.microsoft.com/en-us/ (accessed on 03 August 2019)

Miloslavich, P. et al. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Chang. Biol. 105, 10456–10418. doi: 10.1111/gcb.14108

Mishra, D.R., Narumalani, S., Rundquist, D., Lawson, M., (2005). High-resolution ocean color remote sensing of benthic habitats: a case study at the Roatan island, Honduras. IEEE Trans. Geosci. Remote Sens. 43, 1592–1604. http://dx.doi.org/10.1109/TGRS.2005.847790.

Mobley, C.D. (1994). Light and Water: Radiative transfer in natural waters. Academic Press.

Mobley, C.D. et al. (2005). Interpreation of hyperspectral remote-sensing imagery by spectrum matching and look-up tables. Applied Optics 44 (17): 3576-3592. doi: 10.1364/AO.44.003576.

Montefalcone, M., Albertelli, G., Morri, C., and Bianchi, C. N. (2007). Urban seagrass: status of Posidonia oceanica facing the Genoa city waterfront (Italy) and implications for management. Mar. Pollut. Bull. 54, 206–2013. doi: 10.1016/j.marpolbul.2006.10.005

Morel, A. (1974). Optical Properties of Pure Water and Pure Seawater. In Optical Aspects of Oceanography; Jerlov, N.G., Steemann, N.E., Eds.; Academic Press: New York, NY, USA, 1974; pp. 1–24.

Morel, A. and B. Gentili. (1993). Diffuse reflectance of oceanic waters: II: Bidirectional aspects. Applied Optics 32(33): 6864-6879. doi: 10.1364/AO.32.006864

Mountrakis G., Im, J., Ogole C. (2011). Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 66, 247–259. doi: 10.1016/j.isprsjprs.2010.11.001

Mouw, C.B., Greb, S., Aurin, D., DiGiacomo, P.M., Lee, Z., Twardowski, M., Binding, C., Hu, C., Ma, R., Moore, T., Moses, W., Craig, S.E. (2015). Aquatic color radiometry remote sensing of coastal and inland waters: challenges and recommendations for future satellite missions. Remote Sens. Environ. 160, 15–30.

Mumby P. J., Edwards A. J. (2002). Mapping marine environments with IKONOS imagery: enhanced spatial resolution can deliver greater thematic accuracy. Remote Sensing of Environment 82, 248-257. doi: https://doi.org/10.1016/S0034-4257(02)00041-X

Mumby, P.J., Green, E.P., Edwards, A.J., Clark, C.D., (1999). The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. J. Environ. Manag. 55, 157–166. http://dx.doi.org/10.1006/jema.1998.0255.

Muller-Karger F. et al. (2018). Satellite sensor requirements for monitoring essential biodiversity variables of coastal ecosystems. Ecological Applications 28, 749-760. doi: https://doi.org/10.1002/eap.1682

Murray N. J., Phinn S. R., DeWitt M., Ferrari R., Johnston R., Lyons M. B., Clinton N., Thau D., Fuller R. A. (2019). The global distribution and trajectory of tidal flats. Nature, 565, 222-225, doi: https://doi.org/10.1038/s41586-018-0805-8

National Academies of Sciences, Engineering, and Medicine (2016). Achieving Science with CubeSats: Thinking Inside the Box. Washington, DC: The National Academies Press. pp. 130.

Nordlund L., Koch E.W., Barbier E.B., Creed J.C. (2016). Seagrass Ecosystem Services and Their Variability across Genera and Geographical Regions. PLoS ONE 11(10): e0163091. doi:10.1371/journal.pone.0163091

Nordlund, L.M, R.K.F. Unsworth, M. Gullström, and L.C. Cullen-Unsworth. (2017). Global significance of seagrass fishery activity. Fish and Fisheries 19 (3): 399-412. doi: 10.1111/faf.12259

Orth, R.J et al. (2016). A global crisis for seagrass ecosystems. Bioscience 56, 987–996.

Pacheco, A., J. Horta, C. Loureiro, and O. Ferreira. (2015). Retrieval of nearshore bathymetry from Landsat 8 images: A tool for coastal monitoring in shallow waters. Remote Sensing of Environment 159: 102-116. doi: 10.1016/j.rse.2014.12.004

Pahlevan, N., S. Sarkar, B.A. Franz, S.V. Blasubramarian, and J. He. (2017). Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. Remote Sensing of Environment 201: 47-56. doi: 10.1016/j.rse.2017.08.033.

Palandro D., Andréfouët S., Dustan P., Muller-Karger F. E. (2003). Change detection in coral reef communities using IKONOS satellite sensor imagery and historic aerial photographs. Int. J. Remote Sens. 24, 873–878. doi: 10.1080/0143116021000009895

Partnership to Develop First-Ever Global Map and Dynamic Map of Coral Reefs. (2018). Available online: https://carnegiescience.edu/node/2354 (accessed on 3 August 2019).

Pasqualini V., Pergent-Martini C., Pergent G., Agreil M., Skoufas G., Sourbes L., Tsirika A. (2005). Use of SPOT 5 for mapping seagrasses: an application to Posidonia oceanica. Remote Sensing of Environment 94, 39–45. doi: http://dx.doi.org/10.1016/j.rse.2004.09.010.

Pekel J. F., Cottam A., Gorelick N., and Belward A.S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540, 418–422, doi: 10.1038/nature20584

Pendleton L. et al. (2012). Estimating global "blue carbon" emissions from conversion and degradation of vegetated coastal ecosystems. Plos One 7, e43542. doi: https://doi.org/10.1371/journal.pone.0043542

Pergent G. et al. (2012). Mediterranean Seagrass Meadows: Resilience and Contribution to Climate Change Mitigation, A Short Summary. Gland, Switzerland and Málaga, Spain: IUCN. 40 pages.

Pergent G., Semroud R., Djellouli A., Langar H., Duarte C. (2010). Posidonia oceanica. The IUCN Red List of Threatened Species. doi: http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T153534A4516034.en.

PES (2019). Payments for Ecosystem Services. Available online: https://www.sdfinance.undp.org/content/sdfinance/en/home/solutions/payments-for-ecosystem-services.html (accessed on 10 September 2019)

Platt, J.C., (2000). Probabilistic outputs for Support Vector Machines and comparisons to regularized likelihood methods. In: Smola, A.J. (Ed.), Advances in Large Margin Classifiers, pp. 61–74.

Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com

Poulos, S.E., Chronis, G.T., Collins, M.B., Lykousis, V., (2000). Thermaikos Gulf Coastal System, NW Aegean Sea: an overview of water r sediment fluxes in relation to air – land – ocean interactions and human activities. J. Mar. Syst. 25, 47–76. http://dx.doi. org/10.1016/S0924-7963(00)00008-7.

Poursanidis, D., Barnias, A., Lymberakis, P., (2014). Assessment of the Conservation Status of Posidonia oceanica meadows in the Samaria National Park, an MPA in Crete, Greece. Proceedings of the 5th Mediterranean Symposium on Marine Vegetation 143–148p.

Poursanidis, D., Chrysoulakis, N., Mitraka, Z., (2015). Landsat 8 vs. Landsat 5: A comparison based on urban and peri-urban land cover mapping. Int. J. Appl. Earth Obs.Geoinf. 35, 259–269.

Poursanidis, D.; Topouzelis, K.; Chrysoulakis, N. (2018). Mapping coastal marine habitats and delineating the deep limits of the Neptune's seagrass meadows using VHR earth observation data. Int. J. Remote Sens.

Psarra, S., Tselepides, A., Ignatiades, L., (2000). Primary productivity in the oligotrophic Cretan Sea (NE Mediterranean): seasonal and interannual variability. Prog. Oceanogr. 46 (2–4), 187–204.

Pu R., Bell S. S., Meyer C. A. (2014). Mapping and assessing seagrass bed changes in Central Florida's West Coast using multitemporal Landsat TM imagery. Estuar. Coast. Shelf Sci. 149, 68–79. doi: 10.1016/j.ecss.2014.07.014

Purkis S., and Roelfsema C. M. (2015). Remote sensing of submerged aquatic vegetation and coral reefs, in Remote Sensing of Wetlands: Applications and Advances, ed R. W. Tiner, M. W. Lang, and V. V. Klemas (Boca Raton, FL: CRC Press), 223–242. doi: 10.1201/b18210-15

REDD (2019). Reducing Emissions from Deforestation and Forest Degradation in Developing Countries. Available online: https://www.sdfinance.undp.org/content/sdfinance/en/home/solutions/payments-for-ecosystem-services.html (accessed on 10 September 2019).

Richter, R., (2007). Atmospheric/Topographic Correction for Satellite Imagery (ATCOR - 2/3 User Guide). ATCOR-2/3 User Guide. Version 6.3. pp. 1–71. http://dx.doi.org/ 10.1017/CBO9781107415324.004.

Robles, F.L.E., Collins, M.B., Ferentinos, G., (1983). Water masses in Thermaicos Gulf, north-western Aegean Sea. Estuar. Coast. Shelf Sci. 16, 363–378. http://dx.doi.org/ 10.1016/0272-7714(83)90100-2.

Roelfsema C. M. et al. (2014). Multi-temporal mapping of seagrass cover, species and biomass: a semi-automated object based image analysis approach. Remote Sens. Environ. 150, 172–187. doi: 10.1016/j.rse.2014.05.001

Roelfsema, C.M.; Kovacs, E.; Phinn, S.R.; Lyons, M.; Saunders, M.; Maxwell, P. (2013). Challenges of Remote Sensing for Quantifying Changes in Large Complex Seagrass Environments. Estuar. Coast. Shelf Sci. 133, 161–171.

Roelfsema C., Phinn S. (2008). Evaluating eight field and remote sensing approaches for mapping the benthos of three different coral reef environments in Fiji. In: Robert J. Frouin, Serge Andrefouet, Hiroshi Kawamura, Mervyn J. Lynch, Delu Pan and Trevor Platt, Remote Sensing of Inland, Coastal, and Oceanic Waters. Remote Sensing of Inland, Coastal, and Oceanic Waters, Noumea, New Caledonia, (). 18 November 2008. doi:10.1117/12.804806

Roelfsema C., Phinn S. (2013). Validation. In James A. Goodman, Samuel J. Purkis and Stuart R. Phinn (Ed.), Coral Reef Remote Sensing: A Guide for Multi-level Sensing Mapping and Assessment (pp. 375-401) Dordrecht, Netherlands: Springer. doi:10.1007/978-90-481-9292-2

Roessler, S., Wolf, P., Schneider, T., and Melzer, A. (2012). “Multispectral remote sensing of invasive aquatic plants using RapidEye,” in Earth Observation of Global Changes (EOGC), eds J. M. Krisp, L. Meng, R. Pail, and U. Stilla (Berlin; Heidelberg: Springer), 109–123.

Sagawa, T., Boisnier, Etienne, Komatsu, Teruhisa, et al., 2010. Using bottom surface reflectance to map coastal marine areas: a new application method for Lyzenga's model. Int. J. Remote Sens. 31 (12), 3051–3064.

SDG (2019). Sustainable Development Goals. Available online: https://www.undp.org/content/undp/en/home/sustainable-development-goals.html (accessed on 10 September 2019).

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (3), 379–423.

Short F., Carruthers T., Dennison W., Waycott M. (2007). Global seagrass distribution and diversity: A bioregional model. Journal of Experimental Marine Biology and Ecology 1-2, 3-20. doi: https://doi.org/10.1016/j.jembe.2007.06.012

Sini, M. et al. (2017). Assembling Ecological Pieces to Reconstruct the Conservation Puzzle of the Aegean Sea. Front. Mar. Sci. 4.

SoHelME, 2005. In: Papathanassiou, E., Zenetos, A. (Eds.), State of the Hellenic Marine Environment. HCMR Publ., pp. 360.

Stumpf, R.P., Holderied, K., Sinclair, M., (2003). Determination of water depth with highresolution satellite imagery over variable bottom types. Limnol. Oceanogr. 48, 547–556. http://dx.doi.org/10.4319/lo.2003.48.1_part_2.0547.

Telesca L., et al. (2015). Seagrass meadows (Posidonia oceanica) distribution and trajectories of change. Nature Publishing Group 3303, 1–14. doi: http://dx.doi.org/10.1038/srep12505.

TERN (2018). Effective Field Calibration and Validation Practices. Available online: http://qld.auscover.org.au/public/html/LandscapeAssessmentPracticalHandbook_August2018.pdf (accessed on 03 September 2019)

Thorhaug, A., Richardson, A.D., Berlyn, G.P., (2015). Spectral reflectance of the seagrasses: Thalassia testudinum, Halodule wrightii, Syringodium filiforme and five marine algae. Int. J. Remote Sens. 1161, 1487–1501. http://dx.doi.org/10.1080/01431160600954662.

Tomas, F., X. Turon, and J. Romero. (2005). Seasonal and small-scale spatial variability of herbivory pressure on the temperate seagrass Posidonia oceanica. Marine Ecology Progress Series 301: 95-107. doi: 10.3354/meps301095

Toming, K., T. Kutser, R. Uiboupin, A. Arikas, K. Vahter, and B. Paavel. (2017). Mapping water quality parameters with Sentinel-3 Ocean and Land Colour Instrument imagery in the Baltic Sea. Remote Sensing 9: 1070. doi: 10.3390/rs9101070.

Topouzelis, K.; Makri, D.; Stoupas, N.; Papakonstantinou, A.; Katsanevakis, S. (2018). Seagrass mapping in Greek territorial waters using Landsat-8 satellite images. Int. J. Appl. Earth Obs. Geoinform 67, 98–113.

Topouzelis, K., C. Spondylidis, S. Papakonstantinou, and N. Soulakelis. 2016. The use of Sentinel-2 imagery for seagrass mapping: Kalloni Gulf (Lesvos Island, Greece) case study) in Proceedings of the SPIE, Volume 9688, id. 96881F 7 pp. doi: 10.1117/12.2242887

Traganos, D., Reinartz, P., (2016). Satellite-based Quantitative Assessment of Seagrass Habitats in Thermaikos Gulf, NW Aegean Sea, Greece [Version 1; Not Peer Reviewed]. F1000Research 2. http://dx.doi.org/10.7490/f1000research.1113300.1.

Traganos, D.; Cerra, D.; Reinartz, P. (2017). CubeSat-Derived Detection of Seagrasses Using Planet Imagery Following Unmixing-Based Denoising: Is Small the Next Big? Available online: https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-1-W1/283/2017/isprsarchives-XLII-1-W1-283-2017.pdf (accessed on 3 August 2019).

Traganos D., Aggarwal B., Poursanidis D., Topouzelis K., Chrysoulakis N., Reinartz, P. (2018a). Towards Global-Scale Seagrass Mapping and Monitoring Using Sentinel-2 on Google Earth Engine: The Case Study of the Aegean and Ionian Seas. Remote Sensing, 10:8, doi: https://doi.org/10.3390/rs10081227

Traganos, D.; Poursanidis, D.; Aggarwal, B.; Chrysoulakis, N.; Reinartz, P. (2018b). Estimating Satellite-Derived Bathymetry (SDB) with the Google Earth Engine and Sentinel-2. Remote Sens. 10, 859. [

Traganos, D. and Reinartz, P. (2018a). Mapping Mediterranean Seagrasses with Sentinel-2 Imagery. Available online: https://www.sciencedirect.com/science/article/pii/S0025326X17305726 (accessed on 3 August 2019).

Traganos, D. and Reinartz, P. (2018b). Machine learning-based retrieval of benthic reflectance and *Posidonia oceanica* seagrass extent using a semi-analytical inversion of

Sentinel-2 satellite data. International Journal of Remote Sensing 24: 39. doi: https://doi.org/10.1080/01431161.2018.1519289

Traganos, D. and P. Reinartz. (2018c). Interannual change detection of Mediterranean seagrasses using RapidEye image time series. Frontiers in Plant Science 9: 96. doi: 10.3389/fpls.2018.00096

Tsimplis, M.N., Proctor, R., Flather, R.A., (1995). A two-dimensional tidal model for the Mediterranean Sea. J. Geophys. Res. 100, 16223–16239. http://dx.doi.org/10.1029/95JC01671.

Tulloch V. J. et al. (2017). Trade-offs between data resolution, accuracy, and cost when choosing information to plan reserves for coral reef ecosystems. Journal of Environmental Management 188, 108-119. doi: https://doi.org/10.1016/j.jenvman.2016.11.070

UN General Assembly, Transforming Our World: The 2030 Agenda for Sustainable Development. 21 October 2015. Available online: http://www.refworld.org/docid/57b6e3e44.html (accessed on 3 August 2019).

UNEP-WCMC, Short F.T. (2018). Global distribution of seagrasses (version 6.0). Sixth update to the data layer used in Green and Short (2003). Cambridge (UK): UN Environment World Conservation Monitoring Centre, url: http://data.unep-wcmc.org/datasets/7

van der Linden, S., Rabe, A., Held, M., Jakimow, B., Leitão, P., Okujeni, A., Hostert, P., (2015). The EnMAP-Box - a toolbox and application programming interface for EnMAP data processing. Remote Sens. 7 (11), 249–11,266. http://dx.doi.org/10.3390/ rs70911249.

Vapnik, V., (1995). The Nature of Statistical Learning Theory. Springer-Verlag, New York. http://dx.doi.org/10.1007/978-1-4757-2440-0. 188 pp.

Vassallo P., Paoli C., Rovere A., Montefalcone M., Morri C., Bianchi C.N (2013). The value of the seagrass Posidonia oceanica: a natural capital assessment. Marine Pollution Bulletin 75, 157–167. doi: http://dx.doi.org/10.1016/j.marpolbul.2013.07.044.

Wabnitz, C.; Andréfouët, S.; Torres-Pulliza, D.; Müller-Karger, F.; Kramer, P. (2008). Regional-scale seagrass habitat mapping in the wider Caribbean region using Landsat sensors: Applications to conservation and ecology. Remote Sens. Environ. 112, 3455–3467.

Ward, L.G., Michael Kemp, W., Boynton, W.R. (1984). The influence of waves and seagrass communities on suspended particulates in an estuarine embayment. Marine Geology 59(1- 4), pp. 85–103.

Waycott M. et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. PNAS 106, 12377-12381. doi: 10.1073pnas.0905620106.

White, J.C et al. (2014). Pixel-based image compositing for large-area dense time series applications and science. Can. J. Remote Sens. 40, 192–212

Williams I. D., Couch C. S., Beijbom O., Oliver T. A., Vargas-Angel B., Schumacher B. D, Brainard R. E. (2019). Leveraging Automated Image Analysis Tools to Transform Our Capacity to Assess Status and Trends of Coral Reefs. Front. Mar. Sci. 6:222. doi: 10.3389/fmars.2019.00222

Wilson, R. Py6S—A Python Interface to 6S. (2012). Available online: https://py6s.readthedocs.io/en/latest/ (accessed on 3 August 2018).

Wu, T.F., Lin, C.J., Weng, R.C., (2004). Probability estimates for multi-class classification by pairwise coupling. J. Mach. Learn. Res. 5, 975–1005.

Xiong, J.; Thenkabail, P.S.; Gumma, M.K.; Teluguntla, P.; Poehnelt, J.; Congalton, R.G.; Yadav, K.; Thau, D. (2017). Automated cropland mapping of continental Africa using Google earth engine cloud computing. ISPRS J. Photogramm. Remote Sens. 126, 225–244.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing 27 (14): 3025-3033. doi: 10.1080/01431160600589179.

Zervakis, V., Karageorgis, A.P., Kontoyiannis, H., Papadopoulos, V., Lykousis, V., 2005. Hydrology, circulation and distribution of particulate matter in Thermaikos Gulf (NW Aegean Sea), during September 2001–October 2001 and February 2002. Cont. Shelf Res. 25, 2332–2349. http://dx.doi.org/10.1016/j.csr.2005.08.010.

Zhang, C. (2015). Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem. ISPRS J. Photogramm. Remote Sens. 104, 213–223. doi: 10.1016/j.isprsjprs.2014.06.005

Zhang, C., Selch, D., Xie, Z., Roberts, C., Cooper, H., and Chen, G. (2013). Object-based benthic habitat mapping in the Florida Keys from hyperpspectral imagery. Estuar. Coast. Shelf Sci. 134, 88–97. doi: 10.1016/j.ecss.2013.09.018

Zoffoli M. L., Frouin R., Kampel M. (2014). Water column correction for coral reef studies by remote sensing. Sensors 14, 16881-16931. doi: 10.3390/s140916881