



## Mapping Mediterranean seagrasses with Sentinel-2 imagery

Dimosthenis Traganos<sup>a,\*</sup>, Peter Reinartz<sup>b</sup>

<sup>a</sup> German Aerospace Center (DLR), Remote Sensing Technology Institute, Rutherfordstraße 2, 12489 Berlin, Germany

<sup>b</sup> German Aerospace Center (DLR), Earth Observation Center (EOC), 82234 Weßling, Germany



### ARTICLE INFO

#### Keywords:

*Posidonia oceanica*  
*Cymodocea nodosa*  
 Mediterranean seagrasses  
 Thermaikos Gulf  
 Sentinel-2  
 Water column correction

### 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<sup>2</sup> 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.

### 1. Introduction

Despite only covering 0.82% of the ocean surface, the Mediterranean Sea supports a great biodiversity of up to 18% of all known marine species (~17,000) (Coll et al., 2010). These clear waters are home to one of the most productive and valued, yet threatened, ecosystems worldwide, seagrasses. Seagrasses are marine flowering plants that can form extensive seafloor, monospecific meadows limited to the coastal continental shelf (0–45 m), where there is enough light for their growth via photosynthesis. Seven of the total sixty seagrass species are found in the Mediterranean Sea (Short et al., 2007). Endemic to this enclosed sea, *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 of 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). Furthermore, the economic value of *P. oceanica* meadows in relation to their protection from coastal erosion and sediment retention services has been estimated between 51.5 thousand €/km<sup>2</sup>/yr and 1.72 million €/km<sup>2</sup>/yr (Campagne et al., 2014; Vassallo et al., 2013).

*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 (MSFD; 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 increased 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,

\* Corresponding author.

E-mail addresses: [dimosthenis.traganos@dlr.de](mailto:dimosthenis.traganos@dlr.de) (D. Traganos), [peter.reinartz@dlr.de](mailto:peter.reinartz@dlr.de) (P. Reinartz).

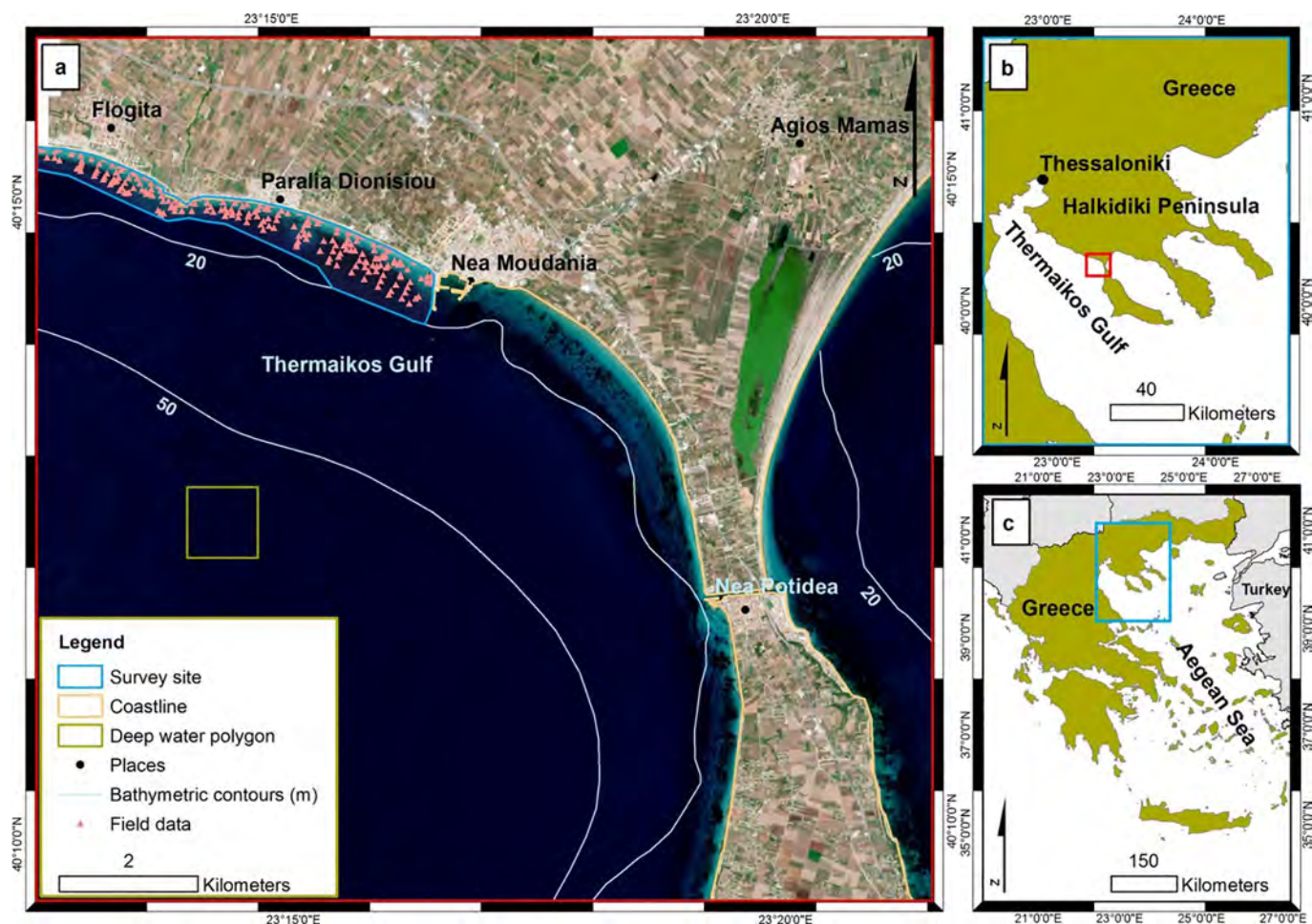


Fig. 1. Location of survey site within (a) Thermaikos Gulf, (b) Halkidiki Peninsula, (c) Aegean Sea, Greece. The depicted Sentinel-2A satellite image in (a) is an atmospherically corrected, true color (band 2 as blue, band 3 as green, band 4 as red) composite in UTM (zone 34) system/WGS84 projection. The outline colors of the left and upper right frame indicate their exact location. The deep water polygon is a selected  $115 \times 115$  pixel window used in the water column correction step as it has very little water leaving radiance values in all visible bands. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 part of the Mediterranean basin (Telesca et al., 2015). Present data in this part of the Mediterranean Sea are based on experts' personal 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.

Currently, optical satellite remote sensing comprises one of the most important methods to detect, map and monitor seagrass ecosystems due to its time- and cost-effectiveness over large areas as well as remote locations (Hossain et al., 2014; Knudby and Nordlund, 2011; Mumby et al., 1999). Analysis of multidecade spaceborne remote sensing data allows retrospective quantitative assessment of seagrass meadows (Dekker et al., 2006). Focusing on the Mediterranean seagrass ecosystems, several studies have utilized satellite imagery to map the distribution of the dominant *P. oceanica* (Borfecchia et al., 2013; Fornes et al., 2006; Matta et al., 2014; Pasqualini et al., 2005). The meadows of *P. oceanica* were studied in an IKONOS image to occupy depths of up to 15 m in the Magalluf embayment, SW Mallorca, Spain with an observed accuracy of 92% (Fornes et al., 2006). Matta et al. (2014) processed RapidEye imagery to map *P. oceanica* beds that spanned a depth range of 0–10 m in the Gulf of Oristano, W Sardinia, Italy, with an overall

accuracy of 84%. Using two SPOT 5 images, Pasqualini et al. (2005) mapped *P. oceanica* habitats up to 20 m of depth in the Laganas Bay, located in the Marine National Park of Zakynthos, Greece (eastern Mediterranean) with an accuracy between 73% and 96%. Last but not least, Borfecchia et al. (2013) implemented an artificial neural network using high-resolution, multispectral Quickbird data to chart *P. oceanica* beds in the coastal Lazio region of the middle Tyrrhenian Sea, Italy with an overall accuracy of 84%. The aforementioned studies were local and since their publication, recent advances in satellite remote sensing technology have emerged. Therefore, a need arises for the evaluation of the potential of new satellite-derived data on detecting and monitoring Mediterranean seagrasses.

In this study, our main objective is to test the suitability of Sentinel-2 imagery for the mapping of the distribution of seagrass meadows and broadly shallow, coastal habitats in the Mediterranean. We processed and analyzed imagery from Sentinel-2A (S2A), the first satellite of the twin polar-orbiting, multispectral satellites of the European Commission's Copernicus program, launched on 23.06.2015. Our study site was the SE Thermaikos Gulf, NW Aegean Sea, Greece, an area with high water transparency and gentle depth slope where seagrass meadows are thriving (Traganos and Reinartz, 2016). The Thermaikos Gulf has been identified as one of the greatest biodiversity hot spots in the entire Mediterranean (Coll et al., 2010). Former studies in the Thermaikos Gulf have mainly focused on the oceanography and hydrography of its waters (Karageorgis and Anagnostou, 2001; Krestenitis et al., 2007; Poulos et al., 2000; Tsimplis et al., 1995; Zervakis et al.,

2005). The remote sensing of shallow aquatic regions can be considerably obstructed by scattering, reflection and absorption of light in the atmosphere, air-water interface, and water column (Dekker et al., 2006). Therefore, to improve the accuracy of our methodology, we developed a site-specific, satellite-derived bathymetry model in addition to applying an analytical-based inverse model to correct water column effects on the satellite signal. We applied three different supervised classifiers with the aid of field data that were collected almost concurrently to the satellite sensor passing. The use of the three different supervised classifications acted as a comparison between them for larger statistical benefits of our methodology and as an indicator of the need for atmospheric and water column corrections prior to mapping seagrass distribution. Overall, we envisage that based on its spatial, temporal, spectral and radiometric resolution, Sentinel-2 will cover important gaps in terms of existing data on the Mediterranean seagrass beds, and especially of *P. oceanica*, and lead to improved understanding and management practices concerning these highly valuable ecosystems.

## 2. Materials and methods

### 2.1. Study site

The study site is a  $\sim 3.3 \text{ km}^2$  area located within the SE part of the Thermaikos Gulf, NW Aegean Sea, Greece (Fig. 1; eastern Mediterranean Sea). With a coastline that exceeds 350 km in length and  $52,300 \text{ km}^2$  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  $9^\circ \text{C}$  and  $17.5^\circ \text{C}$ , 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  $9^\circ \text{C}$  and summer values of  $25^\circ \text{C}$ , 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 gulf. In fact, this part of the shelf shows particulate matter concentrations (PMC) of  $< 1 \text{ mg/l}$ , 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<sup>1</sup> 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.<sup>2</sup> In total, 19.3% of the population of this neighboring region to our study area is employed in the primary sector<sup>1</sup>, while there are 91 hotels and 66 factories operating in the entire coastal region<sup>1</sup>.

### 2.2. Data and methodology

#### 2.2.1. Satellite data

The satellite imagery used in this study was acquired by Sentinel-2A, 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 1). The imagery was downloaded from the Sentinels Scientific Data Hub<sup>3</sup> which provides unlimited open access to the Sentinel-2 Level-1C (L1C) user products. L1C products, also called granules or tiles, consist of  $100 \text{ km}^2$  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 1) for further analysis as it contains our study site (Fig. 1).

#### 2.2.2. 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).

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. 1) were visited, photographed from directly above the seabed (Fig. 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

<sup>1</sup> Source: National Statistics Service of Greece (census 2011).

<sup>2</sup> Source: National Statistics Service of Greece (census 2015).

<sup>3</sup> Sentinels Scientific Data Hub, <https://scihub.copernicus.eu/>

**Table 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 altitude	786 km

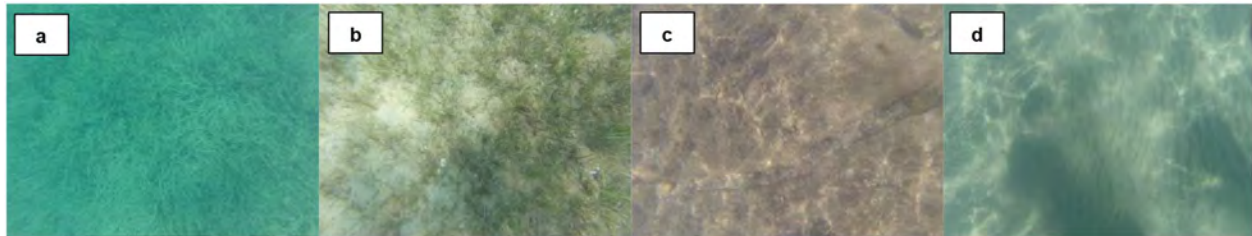


Fig. 2. Habitats exhibited in the survey site: a. dense *Posidonia oceanica* meadow, b. shallow *Cymodocea nodosa* habitat, c. photophilous algae on rocks, d. sand.

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: a) *Posidonia oceanica* seagrass, b) *Cymodocea nodosa* seagrass, c) photophilous algae on rocks and d) sand (Fig. 2).

### 2.2.3. 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.

## 2.3. Methodology

A schematic representation of the adapted methodology herein is shown in Fig. 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.

### 2.3.1. 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  $R_w$ ) are expressed in Integers, thus we divided them by a special quantification value included in the Sentinel-2A metadata (10000) to derive  $R_w$  values in the range 0–1.

### 2.3.2. 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. 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.

### 2.3.3. 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:

$$R_w = R_\infty + (R_b - R_\infty)e^{-2K_d Z} \quad (1)$$

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,  $R_\infty$ , plus substrate contrast (bottom reflectance,  $R_b - R_\infty$ ) after correction of the water depth effect (the term  $e^{-2K_d Z}$ ), where  $K_d$  is the light attenuation coefficient in the water and  $Z$  is the water depth. Maritorena et al.'s

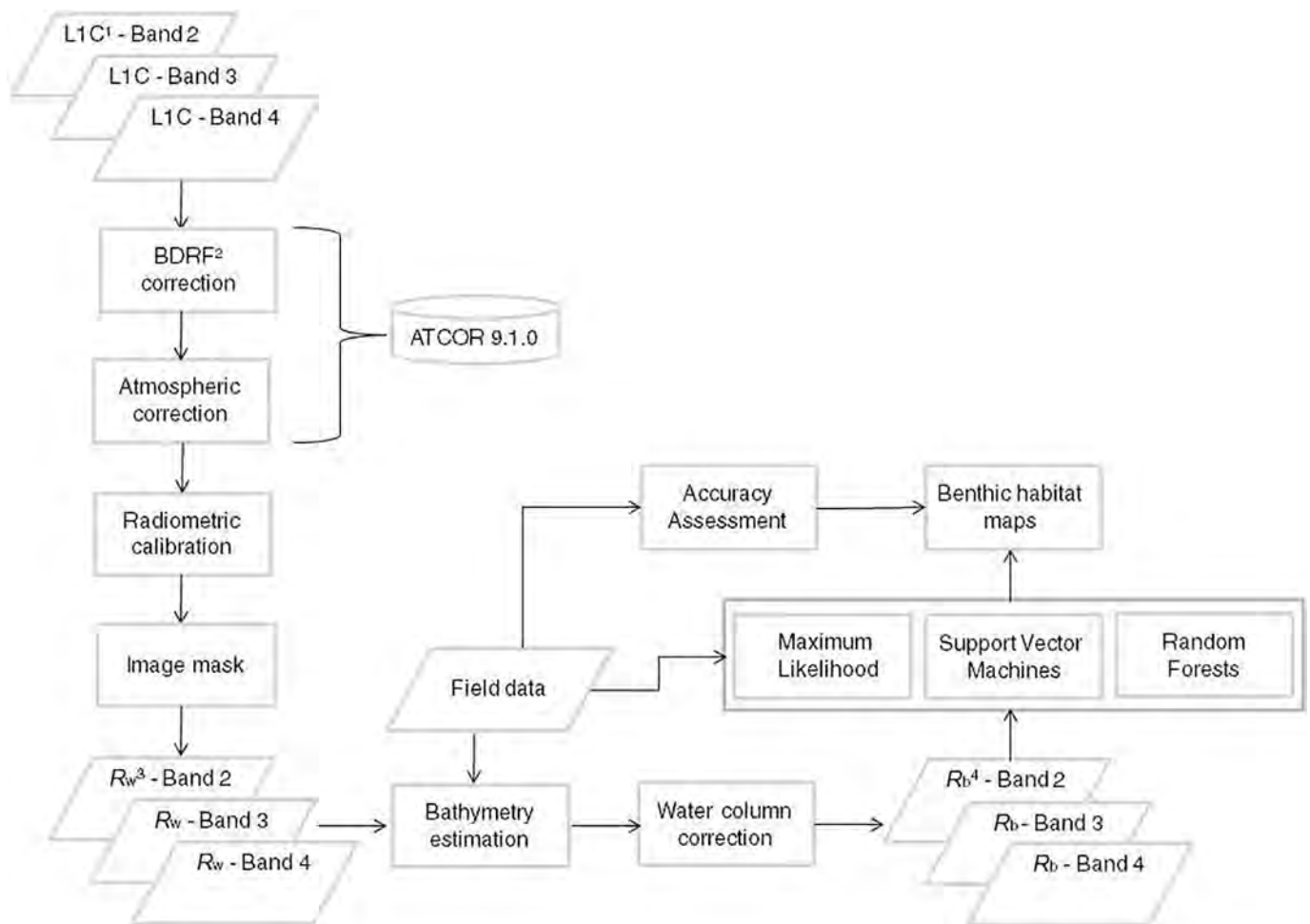


Fig. 3. Schematic presentation of the methodology. L1C<sup>1</sup> products are the initial orthorectified and spatially registered Sentinel-2A granules in UTM/WGS1984 projection, BDRF<sup>2</sup> is the bidirectional reflectance distribution function that defines how light is reflected at an opaque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

algorithm, therefore, can be inverted to derive bottom reflectances,  $R_b$ , from measurements of  $R_w$ ,  $R_\infty$ ,  $Z$  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.

**2.3.3.1. 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). The reflectance ratio was furthermore filtered with a median  $5 \times 5$  cell to reduce the resulting noise from the ratio calculation. Median filtering ranks

neighboring pixels (in this case a  $5 \times 5$  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

$$Z = 178.22x^2 - 428.78x + 259.17 \quad (2)$$

$$x = \frac{\ln[R_w(490)]}{\ln[R_w(560)]} \quad (3)$$

which explained > 96% of the variation ( $p < 0.001$ ) in measured bathymetry in 15 points (Fig. 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. 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. 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.

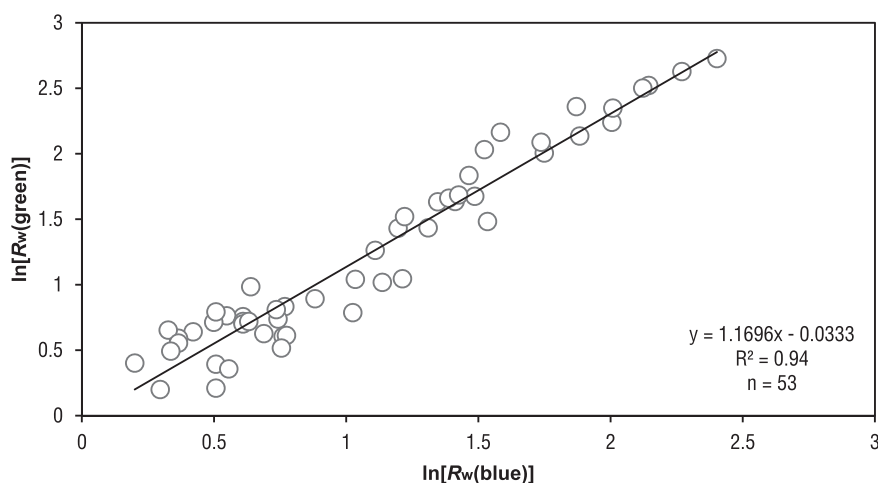


Fig. 4. Regression between log-transformed water-surface Sentinel-2A blue (band 2) versus Sentinel-2A green (band 3) reflectances over various habitats. An offset is applied to both log-transformed reflectances to facilitate presentation.

2.3.3.2. *Infinitely deep water reflectance estimation.* Water column correction following Maritorena et al. (1994) requires estimation of the parameter  $R_{\infty}$ , namely the reflectance measured over an infinitely deep water column. Here, we followed an image-based calculation of  $R_{\infty}$  from water surface reflectances,  $R_w$  by extracting the values of an  $115 \times 115$  pixel polygon (indicated with a gold frame in Fig. 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).

2.3.3.3. *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.

2.3.4. *Supervised classifications*

An important step in the majority of remote sensing studies of

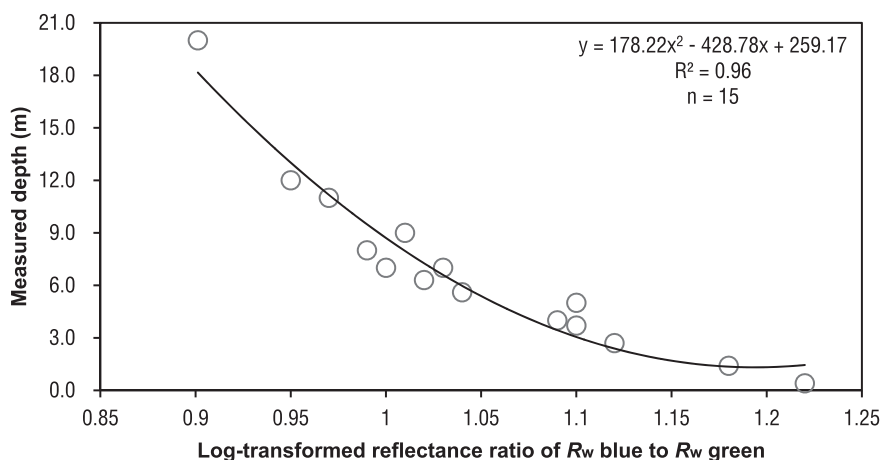


Fig. 5. Polynomial regression between log-transformed and passed with a  $5 \times 5$  median filter water surface blue to green reflectance ratio and *in situ* depth measurements from the Thermaikos survey site. The displayed polynomial equation was used to measure bathymetry in the study site.

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. 6A), b) after applying atmospheric correction ( $R_w$ , water surface reflectances; Fig. 6B), and c) after applying atmospheric and water column corrections ( $R_b$ , bottom reflectances; Fig. 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

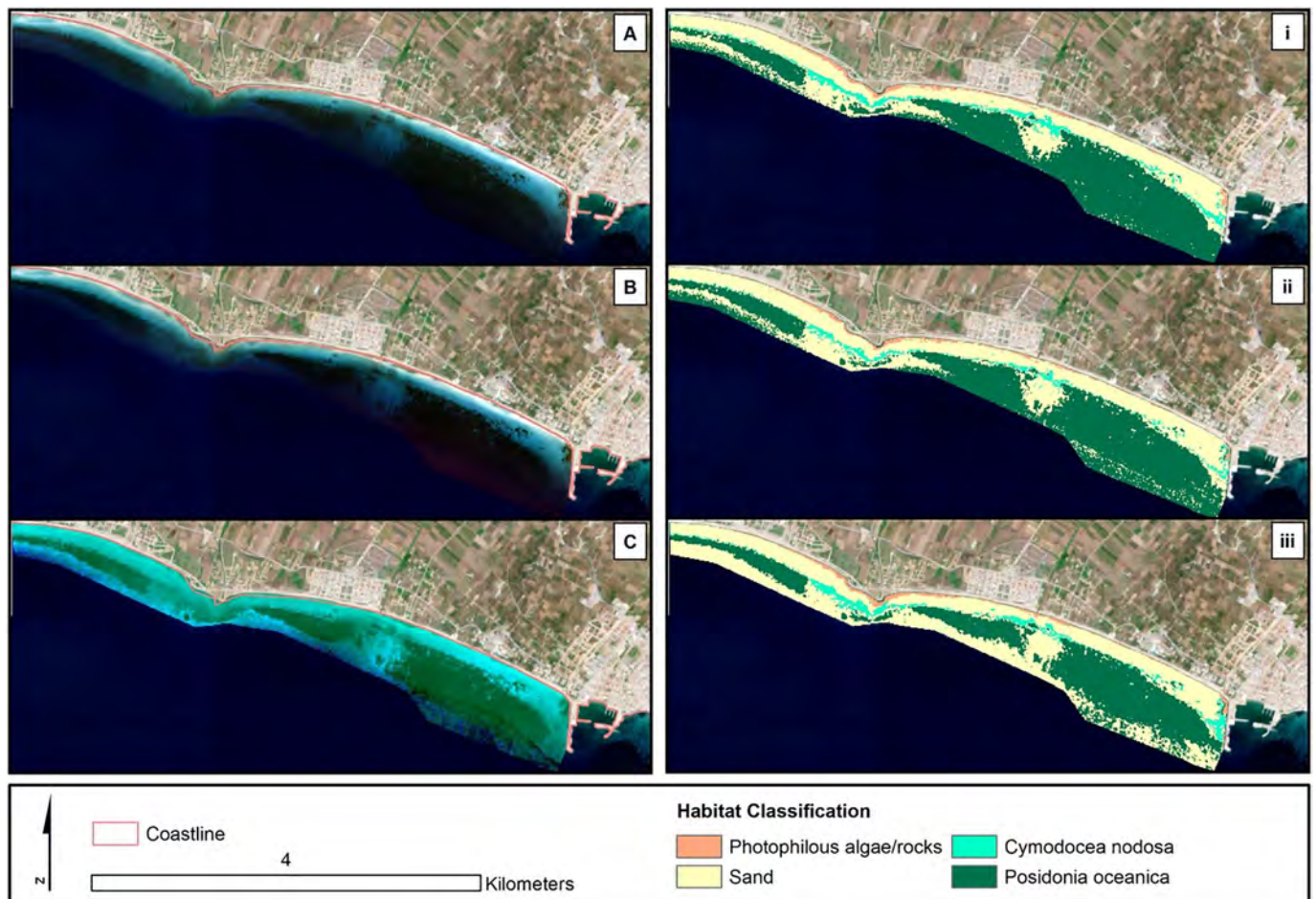


Fig. 6. Sentinel-2A RGB (10-m bands 2, 3, 4) composites of the various methodological steps and corresponding best performing classifiers in the Thermaikos Gulf study site. A. Masked LIC-composite (no correction applied), B.  $R_w$ , water surface reflectance composite (with atmospherically and geometric 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.

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.

**2.3.4.1. 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. 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).

**2.3.4.2. 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)

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

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 pre-defined 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).

**2.3.4.3. Random forests.** Random Forests (RF) is an ensemble supervised classification algorithm that combines multiple self-learning 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

$$n_r = \sqrt{n_a} \quad (5)$$

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.

**2.3.4.4. 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.

### 3. Results

#### 3.1. Pre-classification steps

Differences between the various methodological approaches that we followed in this study are displayed in succession in Fig. 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. 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. 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  $R_b$  for the blue, green and red bands of Sentinel-2A. The equation takes  $R_w$ ,  $Z$  (bathymetry),  $R_\infty$  (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. 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. 1), for a sum of 40 depth points (Fig. 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 a  $r$ -squared 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. 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. 9).

Furthermore, the optically deep water column reflectances were estimated using the median values of 13,287  $R_w$  pixels of the  $\sim 115 \times 115$  gold-framed polygon in Fig. 1. These values were determined as 0.0172, 0.0122 and 0.0092 for  $R_w$  490,  $R_w$  560 and  $R_w$  665, respectively. The diffuse attenuation coefficient for the three Sentinel-2A bands was calculated plotting the modeled depths versus the logarithms of surface reflectances (Fig. 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. 10 which represent the parameter  $-2K_d(\lambda)$  ( $\lambda$  denominates each of the three used bands). Fig. 11 displays the spectral dependence and increasing pattern of attenuation coefficient  $K_d$  with  $R_w$  490,  $R_w$  560 and  $R_w$  665. Values of attenuation coefficient are in meters because they are depth-specific, as the reflectances are also unitless.

#### 3.2. 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. 6A), the second was the  $R_w$  composite (Fig. 6B) and the third was the  $R_b$  composite (Fig. 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 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. 6 (i, ii, iii). SVM classifier performed better than MLC and RF in classifying the non-atmospherically corrected data (Fig. 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 2). RF produced the best results classifier-wise in the classification of the atmospherically corrected reflectances (Table 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. 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 2; Fig. 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. 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. 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 and qualitatively, we considered the SVM-derived habitat map of the

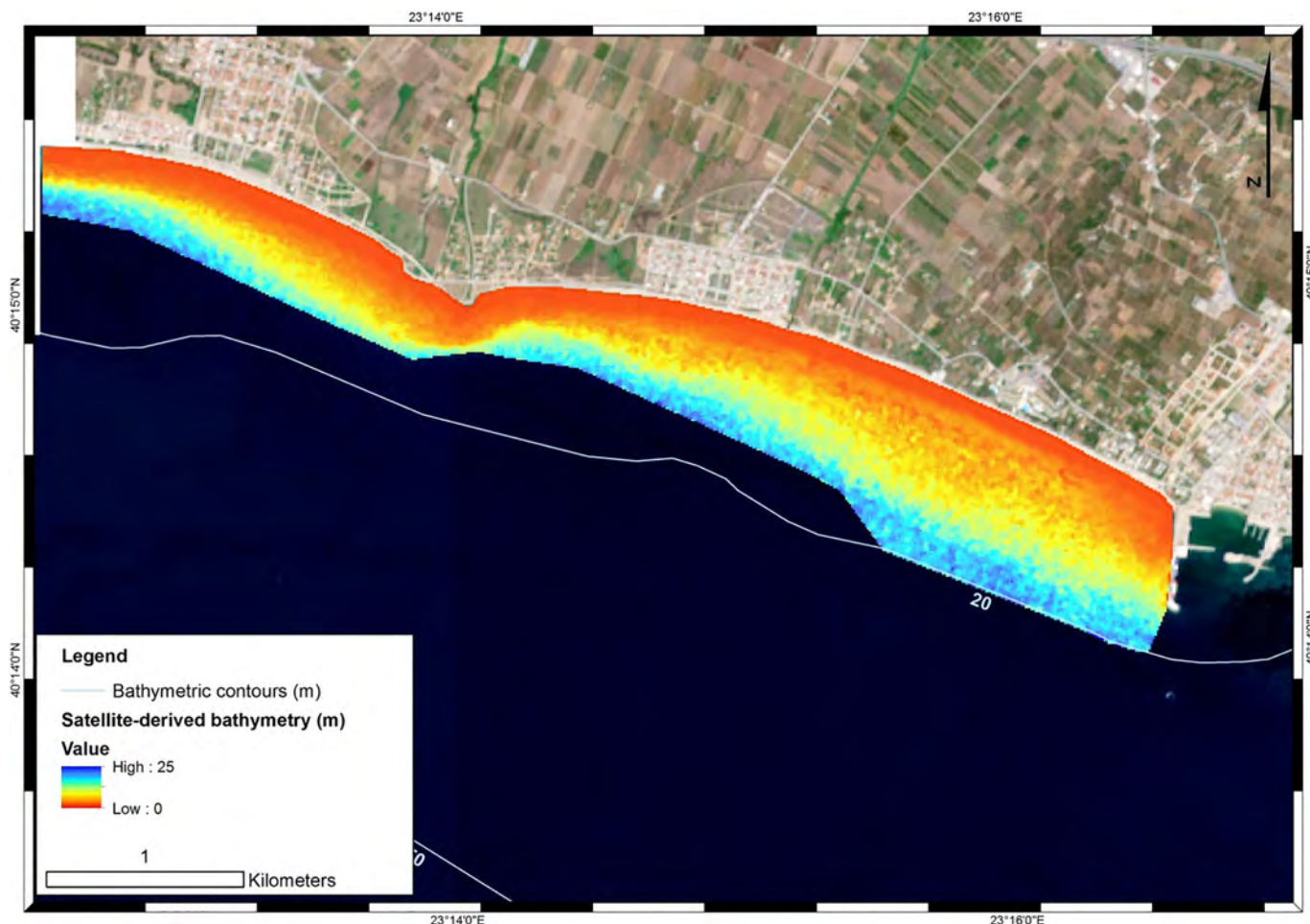


Fig. 7. Bathymetry map of the survey site draped over an atmospherically corrected S2A RGB composite. Projection is UTM (zone 34) system/WGS84.

bottom reflectances (Fig. 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<sup>2</sup> 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<sup>2</sup>) 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 south-easternmost part of the Thermaikos Gulf, extending to a length and width of 1.96 and 0.53 km, respectively (Fig. 6 - iii). Generally, the two seagrass species were distributed in the 45% of the total surveyed area of 3.3 km<sup>2</sup>.

#### 4. Discussion

##### 4.1. Sentinel-2 application to Mediterranean seagrass meadows

The main aim of this study was to test the efficiency of the multi-spectral 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

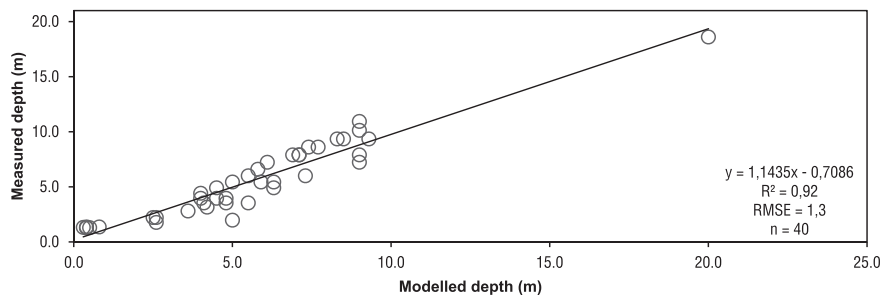


Fig. 8. Plot of modeled depth versus in situ measured depth for the validation of the bathymetry map of the survey site. Depth model was derived from the relationship of Fig. 5.

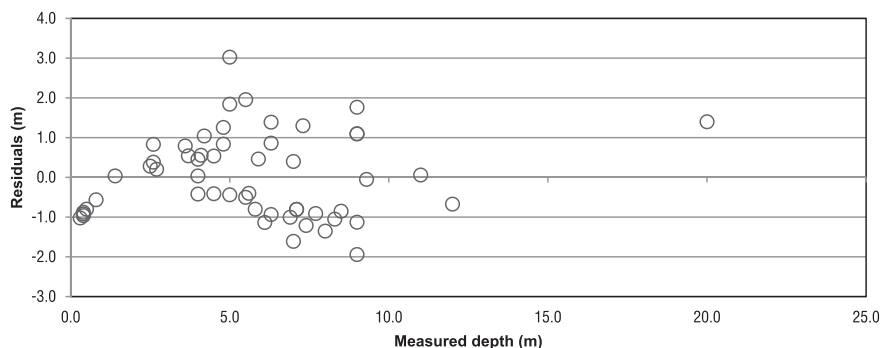


Fig. 9. Plot of the depth residuals from the regression of Fig. 7, between modeled depth and *in situ* measured depth. Note that RMSE is 1.3 m.

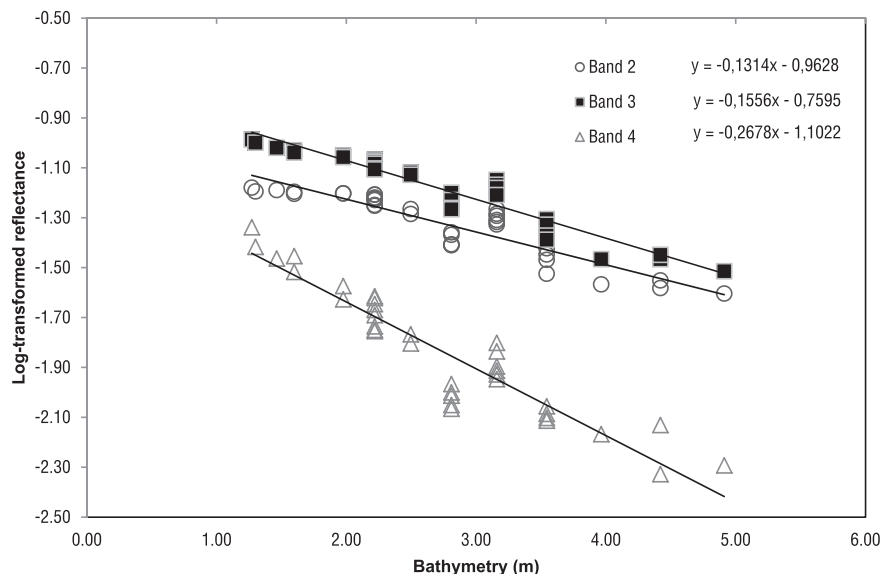


Fig. 10. Modeled bathymetry versus log-transformed water surface reflectances  $R_{w,490}$ ,  $R_{w,560}$  and  $R_{w,665}$ . The surface reflectance pixels were extracted from a sub-area on the satellite image that exhibited a constant sandy substrate but a depth variance from 1 to 5 m. Attenuation coefficients were estimated using the slope of each regression line which represents the quantity  $-2K_i$  (Bierwirth et al., 1993).

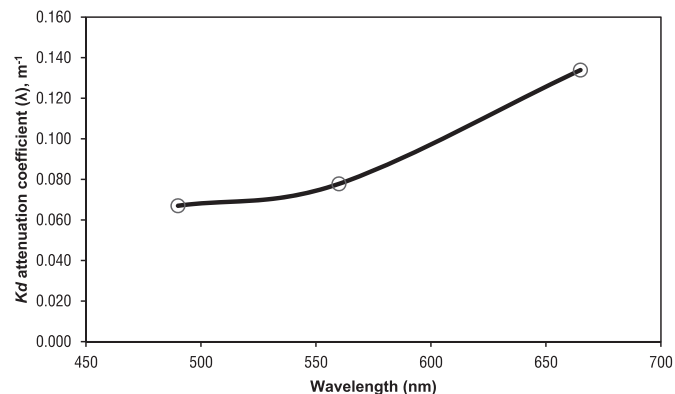


Fig. 11. Spectral dependence of diffuse attenuation coefficient for water surface reflectances  $R_{w,490}$ ,  $R_{w,560}$  and  $R_{w,665}$ . Diffuse attenuation coefficients estimations were derived from the regressions of Fig. 9.

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<sup>2</sup>, 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.

#### 4.2. 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*

**Table 2**  
Best error matrixes and statistical tests for accuracy assessment using different combination of datasets and classifiers.

Classes	Rocky algae	Sand	<i>Cymodocea nodosa</i>	<i>Posidonia oceanica</i>	Total	User accuracy (%)
Rocky algae	18	1	0	0	19	94.7
Sand	3	45	6	3	57	79
<i>Cymodocea nodosa</i>	0	2	8	0	10	80
<i>Posidonia oceanica</i>	0	5	0	52	57	91.2
Total	21	53	14	55	143	
Producer accuracy	85.7	84.9	57.1	94.6		
Overall accuracy: 86%						
Notes: Non-atmospherically corrected image (L1C) - Support Vector Machine classifier						
Rocky algae	20	0	0	0	20	100
Sand	1	51	3	2	57	89.5
<i>Cymodocea nodosa</i>	0	0	11	0	11	100
<i>Posidonia oceanica</i>	0	2	0	53	55	96.4
Total	21	53	14	55	143	
Producer accuracy (%)	95.2	96.2	78.6	96.4		
Overall accuracy: 94.4%						
Notes: Atmospherically-corrected water surface image ( $R_w$ ) - Random Forest classifier (100 trees)						
Rocky algae	20	0	0	0	20	100
Sand	0	48	2	0	50	96
<i>Cymodocea nodosa</i>	0	3	12	0	15	80
<i>Posidonia oceanica</i>	0	2	0	55	58	94.8
Total	21	53	14	55	143	
Producer accuracy (%)	95.2	90.6	85.7	100		
Overall accuracy: 94.4%						
Notes: Water-column corrected bottom image ( $R_b$ ) - Support Vector Machine classifier						

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 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 so-called ‘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 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 multi-species 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).

### 4.3. Pre-classification steps

#### 4.3.1. 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 (Lyzena, 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* extends. 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. 5) and 66 points (used in the validation of the bathymetric algorithm; Fig. 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. 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. 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.

#### 4.3.2. 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. 10) for the calculation of water attenuation coefficients. The principal reason is that attenuation of light increases when wavelength increases (Fig. 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. 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, field-based 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.

## 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 pre-processing, 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.

## Acknowledgements

Dimosthenis Traganos is supported by a DLR-DAAD Research Fellowship (No. 57186656). Sentinel-2A imagery was acquired from the ESA's Sentinels Scientific Data Hub, which provides open data access from all Sentinel missions. ATCOR® software was used thanks to an entitled license of DLR-Berlin provided by ReSe Applications Schläpfer. EnMap-Box software is provided for free under the EnMAP Open Source License. We would also like to thank Bogdanos boats (Nea Potidea, Halkidiki, Greece) for the provision of the boat for the field surveys.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:<http://dx.doi.org/10.1016/j.marpolbul.2017.06.075>. These data include the Google map of the most important areas described in this article.

## References

- 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. <http://dx.doi.org/10.1371/journal.pone.0030454>.
- 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.

- Bonacorsi, M., Pergent-Martini, C., Bréand, N., Pergent, G., 2013. Is *Posidonia oceanica* regression a general feature in the Mediterranean Sea? *Mediterr. Mar. Sci.* 14, 193–203. <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 sea-truth calibration. *Int. J. Remote Sens.* 34, 4680–4701. <http://dx.doi.org/10.1080/01431161.2013.781701>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <http://dx.doi.org/10.1023/A:1010933404324>.
- 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. *Mar. Pollut. Bull.* 97, 391–400. <http://dx.doi.org/10.1016/j.marpolbul.2015.05.061>.
- 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>.
- Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Lasram, F.B.R., Ballesteros, E., Bianchi, C.N., Corbera, J., Dailianis, T., Danovaro, R., Estrada, M., Frogliola, C., Galil, B.S., Gasol, J.M., Gertwagen, R., Gil, J., Guilhaumon, F., Kesner-Reyes, K., Kitsos, M.S., Koukouras, A., Lampadariou, N., Laxamana, E., Lopez-Fe de la Cuadra, C.M., Lotze, H.K., Martin, D., Mouillot, D., Oro, D., Raicevich, S., Rius-Barile, J., Saiz-Salinas, J.I., San Vicente, C., Somot, S., Templado, J., Turon, X., Vafidis, D., Villanueva, R., Voultsiadou, E., 2010. The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. *PLoS One* 5. <http://dx.doi.org/10.1371/journal.pone.0011842>.
- 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](http://dx.doi.org/10.1016/0034-4257(91)90048-b).
- 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. [http://dx.doi.org/10.1007/978-1-4020-2983-7\\_15](http://dx.doi.org/10.1007/978-1-4020-2983-7_15).
- 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. *Limnol. Oceanogr.* 48, 444–455. [http://dx.doi.org/10.4319/lo.2003.48.1\\_part\\_2.0444](http://dx.doi.org/10.4319/lo.2003.48.1_part_2.0444).
- ESA, 2015. SENTINEL-2 User Handbook. 64 pp. doi:GMES-S10P-EOPG-TN-13-0001.
- 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. <http://dx.doi.org/10.1016/j.isprsjprs.2006.04.002>.
- Fourqurean, J.W., Duarte, C.M., Kennedy, H., Marbà, N., Holmer, M., Mateo, M.A., Apostolaki, E.T., Kendrick, G.A., Krause-Jensen, D., McGlathery, K.J., Serrano, O., 2012. Seagrass ecosystems as a globally significant carbon stock. *Nat. Geosci.* 5, 505–509. <http://dx.doi.org/10.1038/ng eo1477>.
- 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>.
- 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. 577 pp.
- Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. *Pattern Recogn. Lett.* 27, 294–300. <http://dx.doi.org/10.1016/j.patrec.2005.08.011>.
- 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.
- 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. <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](http://dx.doi.org/10.1016/S0278-4343(01)00048-6).
- Knudby, A., Nordlund, L., 2011. Remote sensing of seagrasses in a patchy multi-species environment. *Int. J. Remote Sens.* 32, 2227–2244. <http://dx.doi.org/10.1080/01431161003692057>.
- 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>.
- Lyzenga, D.R., 1978. Passive remote sensing techniques for mapping water depth and bottom features. *Appl. Opt.* 17, 379–383. <http://dx.doi.org/10.1364/AO.17.000379>.
- 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. <http://dx.doi.org/10.4319/lo.1994.39.7.1689>.
- 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, pp. 5152–5155. <http://dx.doi.org/10.1109/IGARSS.2014.6947658>.
- 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>.
- Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. *ISPRS J. Photogramm. Remote Sens.* 66, 247–259. <http://dx.doi.org/10.1016/j.isprsjprs.2010.11.001>.
- 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>.
- Nordlund, L.M., Koch, E.W., Barbier, E.B., Creed, J.C., 2016. Seagrass ecosystem services and their variability across genera and geographical regions. *PLoS One* 11, 1–23. <http://dx.doi.org/10.1371/journal.pone.0163091>.
- 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 Sens. Environ.* 94, 39–45. <http://dx.doi.org/10.1016/j.rse.2004.09.010>.
- Pergent, G., Semroud, R., Djellouli, A., Langar, H., Duarte, C., 2010. *Posidonia oceanica*. The IUCN Red List of Threatened Species. <http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T153534A4516034.en>.
- Poulos, S.E., Chronis, G.T., Collins, M.B., Lykousis, V., 2000. Thermaikos Gulf Coastal System, NW Aegean Sea: an overview of water and 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](http://dx.doi.org/10.1016/S0924-7963(00)00008-7).
- Pu, R., Bell, S., Meyer, C., 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. <http://dx.doi.org/10.1016/j.ecss.2014.07.014>.
- 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 Thermaikos Gulf, north-western Aegean Sea. *Estuar. Coast. Shelf Sci.* 16, 363–378. [http://dx.doi.org/10.1016/0272-7714\(83\)90100-2](http://dx.doi.org/10.1016/0272-7714(83)90100-2).
- Short, F., Carruthers, T., Dennison, W., Waycott, M., 2007. Global seagrass distribution and diversity: a bioregional model. *J. Exp. Mar. Biol. Ecol.* 350, 3–20. <http://dx.doi.org/10.1016/j.jembe.2007.06.012>.
- Stumpf, R.P., Holderied, K., Sinclair, M., 2003. Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnol. Oceanogr.* 48, 547–556. [http://dx.doi.org/10.4319/lo.2003.48.1\\_part\\_2.0547](http://dx.doi.org/10.4319/lo.2003.48.1_part_2.0547).
- Telesca, L., Belluscio, A., Criscoli, A., Arduzzone, G., Apostolaki, E.T., Fraschetti, S., Gristina, M., Knittweis, L., Martin, C.S., Pergent, G., Alagna, A., Badalamenti, F., Garofalo, G., Gerakaris, V., Pace, M.L., Pergent-Martini, C., Salomidi, M., 2015. Seagrass meadows (*Posidonia oceanica*) distribution and trajectories of change. *Nat. Publ. Group* 3303, 1–14. <http://dx.doi.org/10.1038/srep12505>.
- 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>.
- 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>.
- 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>.
- 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. *Mar. Pollut. Bull.* 75, 157–167. <http://dx.doi.org/10.1016/j.marpolbul.2013.07.044>.
- Waycott, M., Duarte, C.M., Carruthers, T.J., Orth, R.J., Dennison, W.C., Olyarnik, S., Calladine, A., Fourqurean, J.W., Heck, K.L., Hughes, A.R., Kendrick, G.A., Kenworthy, W.J., Short, F.T., Williams, S.L., 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12377–12381. doi: 10.1073/pnas.0905620106.
- 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>.