NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF RURAL, SURVEYING &
GEOINFORMATICS ENGINEERING
Remote Sensing Laboratory

Marine Pollution Detection from Multispectral Remote Sensing Data

Doctoral Dissertation

Aikaterini Kikaki

Supervisor: Prof. Konstantinos Karantzalos

Athens,
October 2023

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ – ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
Εργαστήριο Τηλεπισκόπησης

Εντοπισμός Θαλάσσιας Ρύπανσης σε Πολυφασματικά Τηλεπισκοπικά Δεδομένα

Διδακτορική Διατριβή

Αικατερίνη Κικάκη

Επιβλέπων: Καθ. Κωνσταντίνος Καράντζαλος

Αθήνα,
Οκτώβριος 2023

PhD Committee

Advisory Committee

Konstantinos Karantzalos
Professor, National Technical University of Athens *(Supervisor)*

Vasilios Kapsimalis
Research Director, Hellenic Centre for Marine Research

Demetre Argialas
Emeritus Professor, National Technical University of Athens

Examination Committee

Konstantinos Karantzalos
Professor, National Technical University of Athens *(Supervisor)*

Vasilios Kapsimalis
Research Director, Hellenic Centre for Marine Research

Demetre Argialas
Emeritus Professor, National Technical University of Athens

Vasilia Karathanassi
Professor, National Technical University of Athens

Maria Papadopoulou
Professor, National Technical University of Athens

Aristomenis Karageorgis
Research Director, Hellenic Centre for Marine Research

Dionysios E. Raitsos
Assistant Professor, National and Kapodistrian University of Athens

Abstract

Monitoring marine pollutants, such as *Marine Debris* and *Oil Spills* is of fundamental importance for protecting the marine environment, public health and maritime activities. The deployment of the latest-generation Earth Observation satellite platforms has led to an increasing volume and variety of geospatial data, facilitating continuous monitoring of the marine ecosystem and operational systems development. However, the accurate detection and tracking of marine pollutants and other features on the sea surface from the currently broadly available satellite data is still a major challenge due to the properties of current sensors, the complex optical properties of specific floating objects, the variability of the sea conditions and the monolithic approaches employed so far (e.g., binary classification tasks). To this end, this dissertation is concerned with the mapping of marine pollution, i.e., *Marine Debris* and *Oil Spill*, and other sea surface features using multitemporal high-resolution multispectral data at a global scale.

The first topic of this dissertation is the investigation of the capabilities of the current Sentinel-2, Landsat-8 and Planet satellite sensors in detecting plastic debris. Since at the beginning of this thesis, the research on marine litter detection from satellite imagery was in its very infancy, the first step was to collect reliable in situ data of *Marine Debris* to verify multispectral satellite observations. Thus, focusing on Honduras Gulf (Caribbean Sea), a highly affected by plastic pollution region, significant insights about plastics' spectral behavior, sources, travelling trajectories and tracking were extracted, contributing to fundamental open issues regarding *Marine Debris* monitoring.

Secondly, this dissertation contributes to benchmarking *Marine Debris* detection using high-resolution multispectral Sentinel-2 satellite data. Well-established Random Forest classifier and Deep Learning U-Net model were developed and evaluated, and the labelled data were openly provided, triggering the research community to create and validate current and future *Marine Debris* detection solutions. The considered created database, named MARIDA (*Marine Debris* Archive), was the first open-access benchmark dataset that discriminates *Marine Debris* from other features that co-exist, such as floating macroalgae, ships, water-related classes, etc. A detailed spectral and statistical analysis of MARIDA was also offered through interactive online material, enabling the users to explore the spectral behaviour of *Marine Debris* and other recorded classes. MARIDA FAIR data and software are open and re-usable and can be integrated to several marine debris detection solutions based on artificial intelligence, or satellite pre-processing pipelines.

The third contribution of this dissertation is the development of a mapping framework for marine pollutants and sea surface features detection, including globally distributed Sentinel-2 satellite observations, aiming to support future

operational solutions development. Towards this direction, a new state-of-the-art convolutional model was proposed, named MariNeXt, which led to overall F1, mIoU and Overall Accuracy scores of 76%, 64% and 89%, outperforming all baselines by at least 12% in F1 and mIoU metrics. A *Marine Debris* and *Oil Spill* (MADOS) dataset was also offered, composed of 174 Sentinel-2 scenes captured between 2015-2022, with approximately 1.5M annotated pixels globally distributed and collected under various weather conditions, forming 15 different thematic classes. Finally, important findings regarding the discrimination potentials of the considered thematic classes are discussed in detail based on an extensive quantitative and qualitative evaluation.

To the best of our knowledge, there is no other available mapping framework and open-access dataset for marine pollutants and other sea surface features detection at 10m spatial resolution with such thematic analysis and quite promising mapping outcomes. By considering, without exceptions, the competing marine pollutants, other sea surface features and water-related classes, reliable operational monitoring solutions can be developed using the proposed framework and dataset.

Acknowledgements

First and foremost, I would like to express my appreciation to my supervisor, Konstantinos Karantzalos, for his trust, support, guidance and willingness to help during all stages of this PhD thesis. I would also like to express my gratitude to Vasilios Kapsimalis for introducing me to the HCMR team and for the valuable discussions and advice during this dissertation. Moreover, I thank Dionysios E. Raitsos for his precious time devoted to constructive discussions and valuable comments during this dissertation. I also thank the rest of the examination committee members for their comments and suggestions.

I am grateful to Vasilios Kapsimalis, Aristomenis Karageorgis, Soultana Zervoudaki Theodore Kanellopoulos, Athanassia Iona and Paraskevi Drakopoulou for their financial support under HCMR projects.

I thank all members of NTUA RSLab and especially my colleagues Ioannis Kakogeorgiou, Antonia Kournopoulou, Paraskevi Mikeli and Antonis Koutroumpas for our excellent collaboration. I also thank my HCMR colleagues and especially Erasmia Kastanidi and Spyros Chaikalis, for our collaboration and fruitful discussions about our research.

Additionally, I would like to thank my beloved family, Nikos, Sissy, Vassia, and my friends, who were always there to hear me and stand by me.

At last, I could not have undertaken this journey without the support of my beloved husband and friend Giannis, whose encouragement gave me the strength to keep going. Thank you!

For Giannis and little Konstantinos

"If your plan is for one year, plant rice. If your plan is for ten years, plant trees. If your plan is for one hundred years, educate children."
Confucius

Εκτεταμένη περίληψη

Η ζωή στον πλανήτη εξαρτάται σε σημαντικό βαθμό από τους ωκεανούς, που αποτελούν το μεγαλύτερο μέρος του νερού στη Γη. Επιπλέον, περίπου το ήμισυ του παγκόσμιου πληθυσμού ζει στην παράκτια ζώνη, ενώ οι επιχειρήσεις που βασίζονται στον ωκεανό συνεισφέρουν σημαντικά στην παγκόσμια οικονομία. Επομένως, η μελέτη του θαλάσσιου περιβάλλοντος αποτελεί σήμερα προτεραιότητα για μεγάλο αριθμό επιστημονικών κοινοτήτων, εθνικών και διεθνών φορέων και οργανισμών. Ιδιαίτερα η μελέτη της θαλάσσιας ρύπανσης (π.χ., θαλάσσια απορρίμματα, πετρελαιοκηλίδες) αποτελεί ένα από τα πιο πρόσφατα αλλά και πιο σημαντικά πεδία έρευνας λόγω των καταστροφικών επιπτώσεων για τη βιοποικιλότητα και τον άνθρωπο. Η έρευνα και η ακριβής καταγραφή και παρακολούθηση της θαλάσσιας ρύπανσης κρίνεται πλέον απαραίτητη τόσο για την προστασία του θαλάσσιου οικοσυστήματος και της δημόσιας υγείας όσο και για την υποστήριξη της γαλάζιας οικονομίας.

Τα δεδομένα παρατήρησης γης προσφέρουν τη δυνατότητα για καταγραφή και παρακολούθηση του θαλάσσιου περιβάλλοντος σε μεγάλες χωρικές και χρονικές κλίμακες με παγκόσμια κάλυψη. Ιδιαίτερα τα τελευταία χρόνια η διάθεση τηλεπισκοπικών δεδομένων από αισθητήρες δορυφόρων τελευταίας γενιάς και η ανάπτυξη μεθόδων που βασίζονται σε μοντέλα και μεθόδους μηχανικής μάθησης, έχει οδηγήσει σε εξαιρετικά αποτελέσματα σε σχέση με την παρακολούθηση του θαλάσσιου περιβάλλοντος. Παρόλο αυτά, ο ακριβής εντοπισμός των ρυπαντών και άλλων στοιχείων στην επιφάνεια της θάλασσας με τη χρήση των διαθέσιμων δορυφορικών δεδομένων αποτελεί μια ιδιαίτερα δύσκολη πρόκληση λόγω των φασματικών και χωρικών αναλύσεων των δορυφόρων και των πολύπλοκων οπτικών ιδιοτήτων ορισμένων αντικειμένων (π.χ., πλαστικά απορρίμματα). Επιπλέον, τα μοντέλα μηχανικής μάθησης που έχουν αναπτυχθεί έως και σήμερα για τον εντοπισμό της θαλάσσιας ρύπανσης τις περισσότερες φορές αφορούν σε πολύ συγκεκριμένες περιοχές (regional models), γεγονός που περιορίζει την ικανότητα γενίκευσης των μοντέλων, ενώ στοχεύουν κυρίως στον εντοπισμό μιας μόνο μορφής ρύπανσης (π.χ. πλαστικά/ όχι πλαστικά ή παρουσία/ απουσία πετρελαιοκηλίδας) και μάλιστα με όχι ικανοποιητικές επιδόσεις. Τέλος, τα περισσότερα συστήματα που έχουν αναπτυχθεί μέχρι σήμερα εστιάζουν κυρίως στην κατανόηση και το διαχωρισμό των φασματικών υπογραφών χωρίς να λαμβάνουν υπόψιν τη χωρική ή χρονική πληροφορία.

Προς την κατεύθυνση αυτή, στην παρούσα διδακτορική διατριβή μελετάται και αναλύεται διεξοδικά το αντικείμενο του εντοπισμού της θαλάσσιας ρύπανσης (θαλάσσια απορρίμματα και πετρελαιοκηλίδες) καθώς και άλλων αντικειμένων/ στοιχείων στην επιφάνεια της θάλασσας με χρήση διαχρονικών πολυφασματικών δορυφορικών δεδομένων υψηλής χωρικής ανάλυσης σε

παγκόσμια κλίμακα. Η βασική συνεισφορά της διατριβής είναι η ανάπτυξη και αξιολόγηση ενός μεθοδολογικού πλαισίου για τον εντοπισμό των θαλάσσιων απορριμμάτων και πετρελαιοκηλίδων με υψηλή χωρική και θεματική ανάλυση.

Πιο συγκεκριμένα, στο πρώτο στάδιο παρουσιάζεται η διερεύνηση των δυνατοτήτων των δορυφόρων υψηλής χωρικής ανάλυσης Sentinel-2, Landsat-8 και Planet στον εντοπισμό των θαλάσσιων απορριμμάτων. Εφόσον όταν ξεκίνησε η συγκεκριμένη διατριβή η έρευνα για την ανίχνευση των θαλάσσιων απορριμμάτων από τηλεπισκοπικά δεδομένα ήταν σε αρκετά πρώιμο στάδιο, το πρώτο βήμα ήταν η συλλογή αξιόπιστων δεδομένων αναφοράς με σκοπό να επικαιροποιηθούν οι δορυφορικές παρατηρήσεις. Επομένως, η έρευνα αρχικά επικεντρώθηκε στον Κόλπο της Ονδούρας (Καραϊβική Θάλασσα), μια περιοχή που είναι από τις πιο επιβαρυμένες από την πλαστική ρύπανση παγκοσμίως. Η συγκεκριμένη περιοχή αποδείχθηκε ιδανική για τη μελέτη της πλαστικής ρύπανσης και την εξαγωγή σημαντικών αποτελεσμάτων σχετικά με τα απορρίμματα, τη φασματική τους συμπεριφορά, τις πηγές τους (ποτάμια) αλλά και την πορεία που ακολουθούν. Επομένως, η παρούσα διατριβή συνεισφέρει σημαντικά σε κρίσιμα ερωτήματα που έχει θέσει η επιστημονική κοινότητα σχετικά με τη μελέτη βασικών χαρακτηριστικών των απορριμμάτων με τη χρήση δορυφορικών δεδομένων.

Η δεύτερη ενότητα αφορά στη μεθοδολογία αξιολόγησης (benchmarking) του εντοπισμού των θαλάσσιων απορριμμάτων με τη χρήση των πολυφασματικών δεδομένων υψηλής χωρικής ανάλυσης Sentinel-2. Πιο συγκεκριμένα, δημιουργήθηκε το σύνολο των δεδομένων MARIDA (*Marine Debris* Archive) με παρατηρήσεις πλαστικών απορριμμάτων αλλά και άλλων αντικειμένων/ στοιχείων όπως *Sargassum* μακροφύκος, πλοία, κ.α. και αναπτύχθηκαν και αξιολογήθηκαν οι αλγόριθμοι μηχανικής μάθησης Random Forest και U-Net. Το συγκεκριμένο σύνολο δεδομένων είναι το πρώτο που διαχωρίζει τα πλαστικά απορρίμματα από άλλα στοιχεία που συνυπάρχουν στην επιφάνεια της θάλασσας σε πολυφασματικά δορυφορικά δεδομένα. Επίσης παρουσιάζονται διεξοδικά η φασματική και στατιστική ανάλυση του συνόλου δεδομένων μέσω διαδραστικής ιστοσελίδας, διευκολύνοντας τους χρήστες να εξετάσουν τη φασματική συμπεριφορά των απορριμμάτων καθώς και των υπόλοιπων κατηγοριών. Τα ανοιχτά δεδομένα MARIDA είναι FAIR εφόσον είναι εντοπίσιμα, προσβάσιμα, διαλειτουργικά και επαναχρησιμοποιήσιμα.

Η τρίτη ενότητα της παρούσας διδακτορικής διατριβής είναι η ανάπτυξη μιας μεθοδολογίας για τον αυτόματο εντοπισμό της θαλάσσιας ρύπανσης και άλλων στοιχείων στην επιφάνεια της θάλασσας με σκοπό την υποστήριξη μελλοντικών επιχειρησιακών συστημάτων παρακολούθησης της θαλάσσιας ρύπανσης. Προς αυτή την κατεύθυνση, δημιουργήθηκε μία βάση δεδομένων με παρατηρήσεις πλαστικών απορριμμάτων και πετρελαιοκηλίδων, καθώς και άλλων στοιχείων στην επιφάνεια της θάλασσας σε παγκόσμια κλίμακα για 15

κατηγορίες. Η βάση αυτή αποτελείται από 1,5 εκατομμύρια εικονοστοιχεία και 174 διαφορετικές εικόνες Sentinel-2 που καλύπτουν τη χρονική περίοδο 2015-2022. Επί προσθέτως, υλοποιήθηκε και αξιολογήθηκε ένα καινοτόμο μοντέλο βαθιάς μηχανικής μάθησης, το MariNeXt, το οποίο οδήγησε σε υψηλότερές τιμές από τα υπόλοιπα μοντέλα αναφοράς (baselines) σε όλες τις μετρικές αξιολόγησης. Τέλος, μέσω ποιοτικής και ποσοτικής αξιολόγησης γίνεται λεπτομερής ανάλυση και παρουσίαση των αποτελεσμάτων σχετικά με τη δυνατότητα διάκρισης των απορριμμάτων και πετρελαιοκηλίδων από τα υπόλοιπα στοιχεία/ αντικείμενα στα δορυφορικά δεδομένα.

Η παρούσα διατριβή αναπτύσσεται σε 6 κεφάλαια: Στο Κεφάλαιο 1, περιγράφεται το αντικείμενο της επιστημονικής μελέτης και αναπτύσσονται οι βασικές πτυχές και προκλήσεις στον εντοπισμό της θαλάσσιας ρύπανσης με τη χρήση των διαθέσιμων δορυφορικών δεδομένων. Παράλληλα, προσδιορίζεται το κίνητρο, οι στόχοι της διατριβής καθώς και η συνεισφορά της στην επιστημονική κοινότητα. Ακόμη, παρατίθενται οι δημοσιεύσεις που προέκυψαν από τη συγκεκριμένη διατριβή και τέλος παρουσιάζεται η δομή του τεύχους.

Στο Κεφάλαιο 2, αρχικά, αναπτύσσεται η διαθέσιμη βιβλιογραφία σχετικά με το πρόβλημα της πλαστικής ρύπανσης καθώς και τις μεθόδους που έχουν χρησιμοποιηθεί τα τελευταία χρόνια για την παρακολούθηση της. Επιπλέον, γίνεται ανασκόπηση των μεθοδολογιών που έχουν προταθεί από τη σύγχρονη βιβλιογραφία για τον εντοπισμό της θαλάσσιας ρύπανσης και άλλων στοιχείων στην επιφάνεια της θάλασσας με τη χρήση πολυφασματικών εικόνων υψηλής χωρικής ανάλυσης. Ακόμη, γίνεται σύντομη περιγραφή της διαθεσιμότητας και των χαρακτηριστικών των δορυφορικών εικόνων υψηλής χωρικής ανάλυσης Sentinel-2, Landsat-8 και Planet. Τέλος, παρουσιάζονται οι φασματικές υπογραφές όλων των εξεταζόμενων στοιχείων στην επιφάνεια της θάλασσας καθώς και αναλύονται περιπτώσεις, με οπτικά παραδείγματα, στις οποίες ο διαχωρισμός τους βασει μόνο της φασματικής πληροφορίας είναι αρκετά δύσκολος.

Στο Κεφάλαιο 3 αναπτύσσεται η μεθοδολογία για την παρακολούθηση των θαλάσσιων απορριμμάτων και τον εντοπισμό της πηγής τους (ποτάμια) και της πορείας τους με τη χρήση των πολυφασματικών δεδομένων από τους δορυφόρους Landsat-8, Planet & Sentinel-2. Πιο συγκεκριμένα, αναλύεται διεξοδικά η συλλογή των δεδομένων αναφοράς και των ωκεανογραφικών δεδομένων, η προ-επεξεργασία των εικόνων και η διαδικασία αναγνώρισης των πλαστικών απορριμμάτων και του μακροφύκους *Sargassum* στα δορυφορικά δεδομένα στην περιοχή του κόλπου της Ονδούρας (Καραϊβική Θάλασσα). Έπειτα, παρουσιάζονται αναλυτικά οι φασματικές υπογραφές των εξεταζόμενων στοιχείων καθώς και τα βασικά χαρακτηριστικά (πηγές, τροχιά, ταχύτητα, βάρος) των επιπλεόντων πλαστικών απορριμμάτων. Τέλος, παρατίθενται η σύγκριση των αποτελεσμάτων με τα αποτελέσματα που

προέκυψαν από μοντέλο τελευταίας τεχνολογίας καθώς και με την υπάρχουσα βιβλιογραφία.

Στο Κεφάλαιο 4 αναλύεται η μεθοδολογία αξιολόγησης εντοπισμού των θαλάσσιων απορριμμάτων (*Marine Debris* benchmarking) στα πολυφασματικά δορυφορικά δεδομένα Sentinel-2. Αρχικά γίνεται αναλυτική περιγραφή της συλλογής των δεδομένων αναφοράς (καταγραφές από πολίτες-επιστήμονες, βιβλιογραφία, μέσα ενημέρωσης) που αφορούν σε παράκτιες περιοχές και ποτάμια που είναι επιβαρυμένα από την πλαστική ρύπανση. Στη συνέχεια γίνεται διεξοδική παρουσίαση της διαδικασίας αναγνώρισης των απορριμμάτων στις δορυφορικές εικόνες Sentinel-2 και του πρωτόκολλου ψηφιοποίησης που ακολουθήθηκε. Επιπλέον, παρατίθενται διεξοδική περιγραφή και στατιστική ανάλυση του συνόλου δεδομένων (MARIDA) που προέκυψε, καθώς και η οπτικοποίηση των δορυφορικών παρατηρήσεων βάσει της φασματικής τους πληροφορίας. Έπειτα, αναπτύσσονται και αξιολογούνται ποιοτικά και ποσοτικά οι μέθοδοι μηχανικής μάθησης Random Forest και U-Net και τα αποτελέσματα που εξήχθησαν συγκρίνονται με την πρόσφατη βιβλιογραφία. Τέλος, με βάση το σύνολο δεδομένων MARIDA, εξετάζονται περαιτέρω φασματικοί δείκτες και χωρικά στοιχεία, αναλύεται η συμβολή τους στην επιτευχθείσα ακρίβεια του μοντέλου, και προτείνονται συγκεκριμένοι συνδυασμοί για τον διαχωρισμό των υπό μελέτη επιπλεόντων στοιχείων στα πολυφασματικά δορυφορικά δεδομένα.

Στο Κεφάλαιο 5, παρουσιάζεται η μεθοδολογία εντοπισμού της θαλάσσιας ρύπανσης (θαλάσσια απορρίμματα και πετρελαιοκηλίδες) στα πολυφασματικά δορυφορικά δεδομένα Sentinel-2. Αρχικά, γίνεται διεξοδική περιγραφή της νέας βάσης δεδομένων (MADOS) καθώς και παρουσίαση των θεματικών κατηγοριών που συμπεριλαμβάνονται με οπτικά παραδείγματα. Ακόμη, παρουσιάζεται αντιπροσωπευτική περιοχή μελέτης (Αν . Μεσόγειος) του συνόλου δεδομένων MADOS στην οποία έγινε καταγραφή και μελέτη της παρουσίας πετρελαιοκηλίδας με τη χρήση πολυφασματικών εικόνων Landsat-8, Planet & Sentinel-2. Επιπλέον, αναλύεται το καινοτόμο μεθοδολογικό πλαίσιο για τον εντοπισμό της θαλάσσιας ρύπανσης, το MariNeXt, που αναπτύχθηκε βάσει μοντέλου τελευταίας τεχνολογίας το οποίο στη συνέχεια αξιολογήθηκε ποσοτικά και ποιοτικά. Τέλος, ακολουθεί εκτενής ερμηνεία της συμπεριφοράς του MariNeXt καθώς και αναλυτική περιγραφή, με οπτικά παραδείγματα, των περιπτώσεων που ο εντοπισμός των θαλάσσιων απορριμμάτων και των πετρελαιοκηλίδων είναι εφικτός στα πολυφασματικά δεδομένα.

Τέλος το Κεφάλαιο 6 παρουσιάζει τα συμπεράσματα της διατριβής σχολιάζοντας τα βασικά μεθοδολογικά στοιχεία και τα πειραματικά ευρήματα. Ακόμη, παρέχεται συζήτηση και σκέψεις για βελτιώσεις, και μελλοντικές επεκτάσεις του προτεινόμενου μεθοδολογικού πλαισίου. Μετά το Κεφάλαιο 6 ακολουθεί η Βιβλιογραφία.

Οι βασικές ιδέες και τα αποτελέσματα της διατριβής έχουν δημοσιευτεί σε επιστημονικά περιοδικά και έχουν παρουσιαστεί σε διεθνή συνέδρια με σημαντικό αριθμό αναφορών. Η ανάλυση αυτή συνδυαστικά με τη διεξοδική αξιολόγηση της προτεινόμενης μεθοδολογίας μπορεί να συνεισφέρει σημαντικά στη δημιουργία των μελλοντικών επιχειρησιακών συστημάτων παρακολούθησης της θαλάσσιας ρύπανσης σε υψηλή χωρική και θεματική ανάλυση.

TABLE OF CONTENTS

List of Abbreviations

AWEI	Automated Water Extraction Index
BSI	Bare Soil Index
CloudS	Cloud Shadows
CNN	Convolutional Neural Networks
DenS	Dense *Sargassum*
DL	Deep Learning
EO	Earth Observation
FAI	Floating Algae Index
FDI	Floating Debris Index
GLCM	GrayLevel Co-occurrence Matrix
L8	Landsat-8
MD	Marine Debris
MixWater	Mixed Water
ML	Machine Learning
MNDWI	Modified Normalised Water Index
MWater	Marine Water
NatM	Natural Organic Material
NDVI	Normalized Difference Vegetation Index
NDMI	Normalised Moisture Index
NDSII	Normalised Difference Snow and Ice Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
NN	Neural Networks
NRD	Near Infrared-Red Difference
Oil	Oil Spill
OilPlat	Oil Platform
PI	Plastic Index
RF	Random Forest
S2	Sentinel-2
SI	Shadow Index
SLWater	Sediment-Laden Water
SpFA	Sparse Floating Algae
SpS	Sparse *Sargassum*
SVM	Support Vector Machine
SWater	Shallow Water
SWIR	Short-Wave Infrared
TOA	Top-of-Atmosphere
TWater	Turbid Water
Waves	Waves & Wakes

1 Introduction

This doctoral dissertation aims to study, design, implement and validate cutting-edge methods for efficiently exploiting high-resolution multispectral satellite data towards detecting marine pollutants on the sea surface. This chapter briefly presents the current challenges in marine pollution detection from spaceborne data, the overall motivation, the objectives and contributions of this PhD thesis, as well as the scientific publications derived from this research.

1.1 Challenges & Motivation

Monitoring the marine environment is of fundamental importance as it supports the sustainable use of ocean resources, enables environmental change assessment, and protects human health from water pollution (Danovaro et al., 2016). Current monitoring approaches rely mainly on in situ campaigns via conventional methods, e.g., ships, nets, trawls, laboratory analysis etc. (Cadiou et al., 2020; Eriksen et al., 2018; Parinos et al., 2019) and Earth Observation (EO) data due to several and significant benefits. Indeed, EO data are beneficial to improve understanding of the marine environment, providing cost-effective measurements of the sea surface at several spatiotemporal scales with global coverage.

EO data can also support the UN Sustainable Development Goals (SDGs) by providing reliable information on various targets related to environmental and socio-economic aspects. For instance, EO data can help monitor the marine environment by providing information on marine and coastal ecosystems' health, ocean acidification, coastal erosion, and sustainable management of fisheries (Canonico et al., 2019; UNDP, 2015), contributing to Goal 14, i.e., to conserve the oceans, seas and marine resources for sustainable development. Among the targets outlined in Goal 14, reducing marine pollution by 2025 is considered significantly crucial. This is particularly important because marine pollutants such as *Oil Spills* and *Marine Debris* are of major concern due to their adverse impacts on marine life, human health, and the economy (Barboza et al., 2018; Beaumont et al., 2019; Shahidul Islam and Tanaka, 2004). Additionally, the EU Mission 'Restore our Ocean and Waters' has set its sights on protecting and restoring the health of the oceans aiming at zero pollution by the year 2030.

Hence, effective marine pollution monitoring systems are crucial to understand its sources and develop effective strategies to prevent and mitigate its impact on the environment. Detecting and monitoring marine pollution forms or other sea surface features from EO imagery with Machine Learning (ML) has already concentrated a vast amount of research and development efforts. However, most remote sensing and ML methods have been designed to detect a single pollutant/sea feature, or they target a small number of categories without taking into consideration other

competing classes, usually resulting in several false positives (Heiselberg, 2020; Sannigrahi et al., 2022; Solé Gómez et al., 2022). Marine monitoring solutions should be designed to identify any sea surface feature or water-related class accurately while being false-alarm resistant under complex weather and sea state conditions. Additionally, in some cases, marine classes[1] present similar spectral and spatial patterns, posing additional challenges on classification tasks in optical satellite images. For instance, Sea snot has been recently regarded to significantly influence *Marine Debris* detection (Hu et al., 2022). At the same time, the discrimination of floating matters in radar data is considered significantly more difficult (Qi et al., 2022).

To explore the unique characteristics of marine pollutants and other floating materials, spectral analysis has been performed with optical imagery, also considering the competing thematic classes (Hu, 2022; Qi et al., 2020; Qi and Hu, 2021). Apart from the spectral methods, only a few studies have considered advanced Machine Learning (ML) techniques that exploit other patterns (e.g., spatial) and try to address the detection of challenging cases such as small and texture-complex floating objects (Mifdal et al., 2021; Solé Gómez et al., 2022). Moreover, very few optical multispectral datasets are openly available to support spectral analysis studies or research towards marine monitoring solutions (Liu et al., 2017; Mifdal et al., 2021). This is because identifying and annotating sea surface features in optical satellite data is quite challenging due to the properties of current sensors (e.g., S2 different band resolutions) (Hu, 2022, 2021) and the complex optical properties of specific features (Dierssen, 2019; Garaba and Dierssen, 2018).

To this end, the motivation for this dissertation was to exploit high-resolution globally distributed satellite data for marine pollutants monitoring, considering various sea features that affect their detection, and propose a methodology based on spectral analysis and advanced ML techniques to contribute to future monitoring solutions development. We mainly focused on *Marine Debris* pollution, a challenging and continuously growing issue, by collecting reports in coastal areas, investigating monitoring from multi-source satellite data, and benchmarking. Oil slicks from different sources (i.e., ship accidents, offshore refineries, discharges due to natural disasters and seafloor natural seepages) were also considered and mapped in detail. Additionally, particular focus is given to floating *Sargassum* macroalgae, other materials co-existing with marine pollutants on the sea surface, and water-related classes targeting an open-access database and a mapping framework with a thematic analysis of 15 representative marine classes.

[1] Throughout this dissertation, the term 'marine classes' refers to all classes considered in the analysis, including marine pollutants, floating objects, and water-related classes.

1.2 Objectives and contributions

The objectives of this dissertation were the following:

- ✓ Study the state-of-the-art in marine pollutant detection with optical remote sensing
- ✓ Exploit multisource and multitemporal high-resolution data to detect and identify plastic debris source, transport, distribution and accumulation patterns, as well as the natural processes (e.g., precipitation, currents, wind) that determine their occurrence.
- ✓ Design and conduct an open-access satellite database including major marine pollutants, i.e., *Marine Debris* and *Oil Spill*, and other competing marine classes.
- ✓ Design, develop and evaluate a mapping framework for detecting marine pollution forms and other features on the sea surface.
- ✓ Identify and discuss the potential in detecting and discriminating marine pollutants in high-resolution satellite data.

The main contributions of this PhD Thesis are:

- ✓ The assessment of the capabilities of the current Planet, Sentinel-2 and Landsat-8 sensors in monitoring *Marine Debris* in the frame of sources, pathways, tracking and spatial distributions, contributing to recent open issues highlighted by the research community.
- ✓ Benchmarking of *Marine Debris* detection from S2 images by openly providing a novel *Marine Debris* dataset (MARIDA) and ML baselines for weakly supervised semantic segmentation.
- ✓ The development of a state-of-the-art and efficient classification framework (MariNeXt) for marine pollutants (*Marine Debris* and *Oil Spill*) and other sea surface features detection from multispectral data.
- ✓ The creation of an open-access and well-curated dataset (MADOS) with 15 marine classes consisting of 1481155 S2 pixels and 174 S2 images.
- ✓ The extraction and analysis of important findings regarding the discrimination potentials of marine pollutants and considered marine classes based on extensive quantitative and qualitative evaluation.

Additional contributions:

- ✓ The systematic recording and statistical analysis of spectral signatures of pollutants and other marine classes for the spectral patterns mapping.
- ✓ The investigation of satellite discrimination of *Marine Debris* and *Sargassum* macroalgae in Honduras Gulf (Caribbean Sea).
- ✓ The mapping of the extent of the plastic pollution over Honduras Gulf, triggering further research, clean-up activities (4ocean) and prevention solutions (Ocean Cleanup) in the region.

- ✓ The collection of in situ observations/reports at the global scale for *Marine Debris* verification in satellite data with a focus on the Caribbean Sea.
- ✓ The validation of state-of-the-art global plastic debris modelling procedures.
- ✓ The investigation of the effect of the earthquake activity and ocean conditions on the oil slick spreading in Zakynthos Island (E. Mediterranean).
- ✓ The quantitative and qualitative evaluation and comparison of shallow and deep learning methods for the considered marine classes detection.
- ✓ The application of state-of-the-art SegNeXt model in remote sensing data.
- ✓ The assessment of the contribution of the spectral and spatial information to the classification process.
- ✓ The investigation of spectral indices and spatial features at multiple scales and suggestion of specific combinations for the enhancement of marine classes discrimination.

1.3 Scientific publications, Software and Dissemination of research

Scientific journals

Kikaki, A., Karantzalos, K., Power, C. A., & Raitsos, D. E. (2020). Remotely sensing the source and transport of marine plastic debris in Bay Islands of Honduras (Caribbean Sea). *Remote Sensing*, *12*(11), 1727.

Kikaki, K., Kakogeorgiou, I., Mikeli, P., Raitsos, D. E., & Karantzalos, K. (2022). MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data. *PloS one*, *17*(1), e0262247.

Kikaki, K., Kakogeorgiou, I., Hoteit, I. & Karantzalos, K. (2023). Detecting Marine Pollutants and Sea Surface Features with Deep Learning in Sentinel-2 Imagery.

Conferences

Mikeli, P., **Kikaki, K.,** Kakogeorgiou, I., & Karantzalos, K. (2022). How Challenging is the Discrimination of Floating Materials on the sea surface using High Resolution Multispectral Satellite data?. *XXIV ISPRS*, *43*, 151-157.

Kikaki, A., Kakogeorgiou, I., Mikeli, P., Raitsos, D. E., and Karantzalos, K. (2021). Detecting and Classifying Marine Plastic Debris from high-resolution multispectral satellite data, *EGU General Assembly 2021*, online, 19–30 Apr 2021, EGU21-15243.

Kikaki A., Karegeorgis A., Rousakis G., Karantzalos, K. (2019) *Oil Spill* detection by satellite remote sensing in Zakynthos Island, Ionian Sea, E. Mediterranean. *12th International Conference of the Hellenic Geographical Society*, Athens, November.

Kikaki A., Karantzalos K., Kapsimalis V. (2018). Monitoring of Marine Pollution using Multitemporal high-resolution remote sensing data. *11th International Conference of the Hellenic Geographical Society,* Lavrion, 12-15 April.

Kikaki, K., Kakogeorgiou, I., Karantzalos, K. (2018). Detecting *Sargassum* blooms and Plastic *Marine Debris* from High-resolution Multispectral Satellite data. *XXIV Ocean Optics*, Dubrovnik, 7-12 October.

Invited Talks

Kikaki, K., Kakogeorgiou, I., Mikeli, P., Raitsos, D. E., & Karantzalos, K. (2022). MARIDA: A Sentinel-2 dataset for *Marine Debris* detection. *2nd Workshop of the IOCCG Task Force on Remote Sensing of Marine Litter and Debris*, 8 March.

Interviews

ESA Success story: https://sentinel.esa.int/web/success-stories/-/copernicus-sentinel-2-benchmark-dataset-for-detection-of-dense-marine-plastic-accumulations
Planet Explorer: https://www.planet.com/pulse/all-eyes-on-marine-plastic/
Hakai magazine: https://hakaimagazine.com/news/scientists-can-spy-shrimp-eggs-from-space/

Research Community contributions

Member of International Ocean Colour Coordinating Group (IOCCG) Task Force on Remote Sensing of Marine Litter and Debris (Algorithms & Applications topic)

Scientific Software & Interactive Online Material

Marine Debris Archive (MARIDA) Interactive Online Material with a detailed overview and analysis of our Dataset: https://marine-debris.github.io/.
Marine Debris Archive (MARIDA) Quick Start Guide providing openly code and pre-trained models: https://github.com/marine-debris/marine-debris.github.io
Marine Debris & Oil Spill (MADOS) implementation code and ML models will be provided soon.

1.4 Thesis Roadmap

Below is a brief summary of the following chapters in this dissertation (Figure 1.1): Chapter 2 consists of the work related to marine pollution detection from multispectral satellite data and the current challenges based on a detailed spectral analysis. Chapter 3 presents the investigation of floating debris monitoring with satellite data. Chapter 4 introduces a novel benchmark for *Marine Debris* detection. Chapter 5 presents the methodology for automated detection of marine pollutants and other features. Finally, Chapter 6 summarizes the conclusions and suggests methods and ideas for future work.

Figure 1.1 Thesis Roadmap.

2 Marine Pollution Detection: Current Challenges and Related Work

This chapter provides an overall introduction and background related to detecting marine pollutants and sea surface features studied in this dissertation. Special focus is given to *Marine Debris* detection methods, a more recent research field. The first section 2.1 presents the background related to plastics' distribution dynamics and monitoring. In section 2.2, we provide the related work regarding *Marine Debris* and *Oil Spill* detection from S2 remote sensing data. Section 2.3 presents a brief description of satellite data specifications and existing datasets for marine pollutants detection. Furthermore, section 2.4 constitutes the related work about other sea features detection. Finally, in section 2.5, we demonstrate the spectral analysis of all considered features and identify the discrimination challenges.

2.1 Plastic pollution: Distribution Dynamics & Monitoring

With evidence of a significant cumulative increase over the last 60 years, plastic debris in marine environments is currently considered one of the most topical issues in marine pollution due to its potential impact on human and ecosystem health (Maximenko et al., 2019; Ostle et al., 2019). Plastics infiltrate the marine environment through diverse land-based and sea-based activities. Additionally, plastics' fate in the marine environment is not predetermined but is influenced by a variety of factors such as the characteristics of the plastics themselves (size, shape, density), climatic conditions (precipitation, air intensity, temperature, solar radiation), sea currents and waves, and biological interference (Galgani et al., 2015; Sebille et al., 2020).

The global amount of microplastic debris was estimated between 93 and 236 thousand metric tonnes for 2014 (Sebille et al., 2015), while there is also sufficient evidence that the largest plastics accumulator (Great Pacific Garbage Patch) is constantly growing (Lebreton et al., 2018). Plastic debris has been found in fish (Rochman et al., 2015) in different trophic levels and habitats. In fact, plastic debris was found in >25% of individual animals and in >50% of the species collected in fish markets in USA and Indonesia, revealing that plastic debris poses a threat to marine animals and human health. Plastic debris events have been reported even in deep-sea organisms (such as Cnidaria, Echinodermata and Arthropoda) in the equatorial mid-Atlantic and the SW Indian ocean, mainly due to their exposure to human waste and ingestion of plastic microfibres (Taylor et al., 2016). Microplastics from various sources have been also found on seafloor sediments (Martin et al., 2017), in sea surface microlayer in estuarine systems (Anderson et al., 2018) and in arctic waters (Cózar et al., 2017), indicating that *Marine Debris* is abundant across all marine systems.

Runoff (including riverine outflow) and beach users are important sources of *Marine Debris* in coastal regions, while offshore transport by ocean currents has been shown to be the most important driving force to the open waters (Willis et al., 2017). Between 1.15 and 2.41 million tonnes of plastic waste enters the ocean every year from rivers, with over 74% of emissions occurring between May and October (Lebreton et al., 2017) (Figure 2.1). Mani et al. (Mani et al., 2015) indicated almost 900 thousand plastic particles per km^{-2} along the Rhine River, reflecting the significant role of rivers as debris pathways to the open seas. The large mass of mismanaged plastic waste (Jambeck et al., 2015) directly affects the amount of plastics in marine ecosystems, while the constant increase in plastic production (Law et al., 2010) can potentially have a significant impact on the amount of expected marine plastic debris discharges.

Nevertheless, understanding plastic debris distribution dynamics is challenging as it is influenced by its own physical characteristics (density, size) as well as environmental features (such as winds, waves, thermohaline gradients) (Zhang, 2017). In particular, time series of plastic concentrations in the western North Atlantic Ocean (Law et al., 2010) (1986 to 2008) showed that the highest concentrations were associated with sea surface currents transport, indicating that floating plastic debris is a passive tracer of ocean dynamics. Lusher et al. (Lusher et al., 2015) found microplastics in surface and subsurface Arctic waters that entered the sea via coastal sewage and wastewater and were transported by large-scale currents to the open ocean. They also showed that sea surface temperature is an important predictor of microplastics abundance, influencing water mass circulation (i.e., cyclonic eddies, gyres and seafronts). Kataoka et al. (Kataoka et al., 2018) used high-resolution aerial imagery in order to estimate debris distribution around Vancouver Island (Canada) and relate its accumulation to Ekman flow and prevailed winds.

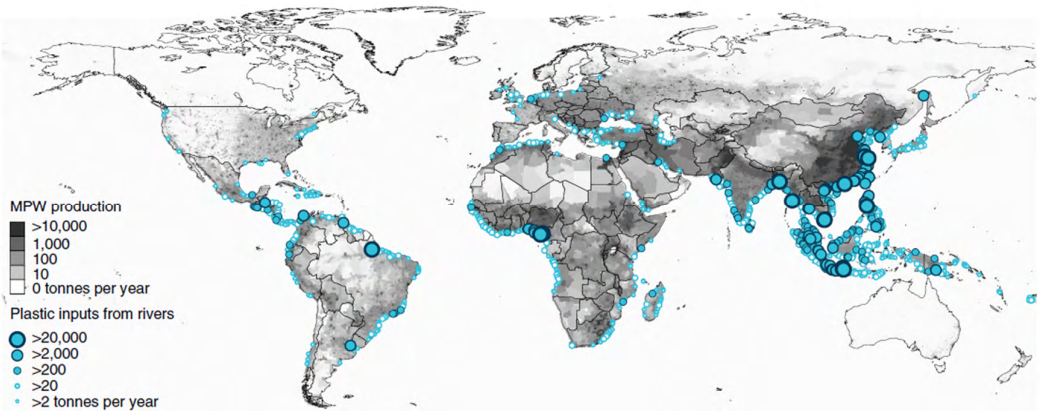


Figure 2.1 Mass of river plastic flowing into oceans in tonnes per year (Lebreton et al., 2017).

The majority of litter-detection studies have been primarily conducted using conventional methods (e.g., trawl, glass plate, net). These methodological approaches are effective; however, they have weaknesses including the relatively high-cost, the need for personnel, the deployment proximity to the land, and the low spatiotemporal coverage of the acquired data. Alternatively, satellite remote sensing could provide adequate information in detecting plastic debris at the surface of the oceans at high spatial and temporal scales. The next section describes in detail the remote sensing methods utilized so far for plastic debris detection and monitoring.

2.2 Detecting *Marine Debris* & *Oil Spill* from remote sensing data

Marine Debris

To tackle the *Marine Debris* issue, several solutions for detecting (Martínez-Vicente et al., 2019; Maximenko et al., 2019), cleaning (Zielinski et al., 2019) and preventing (Schmaltz et al., 2020) have been developed and validated. Among those, detecting and monitoring floating litter has recently gained the attention of most research and development efforts (Bellou et al., 2021). For instance, Light Detection And Ranging (LIDAR) imagery (Ge et al., 2016) and high-resolution commercial satellite data (Acuña-Ruz et al., 2018) have been used with machine learning techniques for beach debris identification and classification. Worldview-2, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Synthetic Aperture Radar (SAR) satellite datasets have been efficiently employed in monitoring marine plastic debris events, for instance, after the great east Japan earthquake on March 2011, when a remarkable amount of >1.5 million tonnes of debris was generated (Arii et al., 2014; Matthews et al., 2017). Satellite data coupled with ocean models have also been employed to investigate the temporal variability of *Marine Debris*, its anthropogenic sources (Agustin et al., 2015; Pichel et al., 2012) and its accumulation patterns (Atwood et al., 2019). Except for satellite imagery, remote sensing data from manned aircraft (Garcia-Garin et al., 2021), unmanned aerial vehicles (UAVs) (Bao et al., 2018; Jakovljevic et al., 2020; Martin et al., 2018; Papakonstantinou et al., 2021; Wolf et al., 2020), bridge-mounted (Lieshout et al., 2020) and underwater-cameras (Politikos et al., 2021) have also been employed for floating *Marine Debris* detection.

Among current satellite sensors, S2 MSI is the most frequently used for monitoring *Marine Debris* on the sea surface in coastal waters (Mukonza and Chiang, 2022). Since identifying and experimenting with floating marine litter in the open sea is difficult, many efforts have been concentrated on artificial targets, coupled with various methods, i.e., image fusion (Kremezi et al., 2022), spectral indices (Themistocleous et al., 2020), OBIA, Spectral Unmixing (SU), Matched Filtering (MF) (Topouzelis et al., 2020), and ML (e.g., semi-supervised fuzzy c-means, K-means,

Random Forest, SVM etc.) (Basu et al., 2021; Jamali and Mahdianpari, 2021). In a different direction, other studies have focused on detecting floating *Marine Debris* in real cases using S2 images (Table 2.1). Regarding preprocessing, most studies were based on Surface Reflectance values. To understand *Marine Debris* behavior and enhance the efforts in *Marine Debris* detection, different spectral analysis methods have been considered (Arias et al., 2021; Biermann et al., 2020; Ciappa, 2022, 2021; Gonzaga et al., 2021; Sannigrahi et al., 2022), as well as simulation experiments have been performed by Hu (2021).

Table 2.1: Publications for Marine Debris (MD) mapping using Sentinel-2 data during the period 2018-2022.

General info					Pre-processing		Classification/ Detection	
Publication	Main purpose	Study Area	Time span	# of S2 scenes	Atm. Corr.	Reflectance	Method	# of classes
Biermann et al., 2020	MD detection	Global	2018-2019	12	Yes	Surface Refl.	Naïve Bayes	5
Mifdal et al., 2021	Float. Objects detection	Global	2018-2020	15	Yes	Surface Refl., TOA	SVM, RF, Naïve Bayes, U-Net	2
Gonzaga et al., 2021	MD detection	Manila Bay, Philippines	2019-2020	6	Yes	Surface Refl.	Naïve Bayes, MTMF, FDI, NDVI, PI	5
Arias et al., 2021	Windrows detection	Greece, Honduras	N/A	40	Yes	Surface Refl., TOA	Wasp Spectral Index	1
Ciappa, 2021	MD detection	Hawaii, N. Adriatic, S.Domingo	2020	6	Yes	Surface Refl.	Red edge bands, NDVI, PI, FDI	5
Olyaei et al., 2022	MD detection	Global	2015-2021	63	Yes	Rayleigh	Deep Knockoff, SVM, RF	3
Sannigrahi et al., 2022	MD detection	Mediterranean Sea	2018-2019 & 2021	27	Yes	Surface Refl.	SVM, RF	3
Sole Gomez et al., 2022	River debris detection	Drina, Los Angeles & Yangtzee	2018-2020	30	Yes	Surface Refl.	U-Net, U-Net3DE, DeepLab V3+	3
Ciappa, 2022	MD detection	N.Adriatic, Australia	2020	6	Yes	Surface Refl.	Spectral bands	4
Russwurm et al., 2023	MD detection	Global	2015-2021	89	Yes	Surface Refl.	U-Net, Label Refinement Module	1
Booth et al., 2023	MD detection & density mapping	Global	2015-2021	63	Yes	Surface Refl.	U-Net, MAP-Mapper	1
Duarte & Azevedo, 2023	MD detection	Global	2016-2021	25	Yes	Surface Refl.	GANs, SI, XGBoost	7

Abbreviations: SVM= Support Vector Machine, RF= Random Forest, MTMF= Mixture Tuned Matched Filtering, FDI= Floating Debris Index, NDVI= Normalized difference vegetation index, PI= Plastic Index, GANs= Generative Adversarial Networks, SI= Spectral Indices, XGBoost= Extreme Gradient Boosting, N/A= Not Mentioned.

Machine learning methods have recently gained recognition for plastics identification (Politikos et al., 2023). Using S2 data, well-established supervised

learning techniques, i.e., naïve Bayes, SVM, Random Forest (RF), U-Net, DeeplabV3+ (Basu et al., 2021; Biermann et al., 2020; Booth et al., 2022; Gonzaga et al., 2021; Jamali and Mahdianpari, 2021; Mifdal et al., 2021; Rußwurm et al., 2023; Sannigrahi et al., 2022; Solé Gómez et al., 2022), as well as Generative Adversarial Networks (GANs) (Duarte and Azevedo, 2023) have also been applied for various *Marine Debris* detection tasks. Concerning the number of classes, Biermann et al. (2020); Ciappa (2021); Gonzaga et al. (2021) considered five classes.

Oil Spill

Oil Spill detection from satellite imagery has been extensively studied over the past few decades (Al-Ruzouq et al., 2020). Specifically, radar sensors dominate due to their all-weather and all-day capabilities (Dong et al., 2022). Nevertheless, exploiting optical data is crucial to separate oil from look-alikes (i.e., biogenic slicks, low wind speed, internal waves, ship wakes, algae, rain) and extract information about oil types and thickness (Garcia-Pineda et al., 2020; Zhao et al., 2014). (Hu et al., 2009; Pisano et al., 2015; Sun et al., 2015; Zhao et al., 2014) explored the optical properties of oil slicks in multispectral data to determine sun glint effects on oil-water contrast and investigated the differentiation from other look-alikes using interpretation, spectral indices, and statistical methods. Landsat-8 images have been also utilized for *Oil Spill* detection and concentration estimation.

Table 2.2: Publications for Oil Spill (Oil) mapping using Sentinel-2 data during the period 2018-2022.

General Info					Pre-processing		Classification/ Detection	
Publication	Main purpose	Study Area	Time span	# of S2 scenes	Atm. Corr.	Reflectance	Method	# of classes
Kolokoussis & Karathanassi, 2018	Oil detection	Zakynthos Isl. & Saronikos Gulf (Greece)	2017	4	No	TOA	OBIA, SI	1
Sun et al., 2018	Oil detection	Gulf of Mexico	2004-2016	24	Yes	Rayleigh	Statistical analysis	1
Althawadi & Hashim, 2019	Oil detection & Calibration	Arabian Gulf	2016-2017	7	Yes	Surface Refl.	Region-based segmentation, Regression	1
Rajendran et al., 2021	Oil detection	Ambarnaya river, Russia	2020	14	No	TOA	SI, Dynamical permafrost model, DSM	5
Rajendran et al., 2021	Oil detection	Mauritius	2020	6	Yes	Surface Refl.	PSA, Band ratios, DSM	3
Rajendran et al., 2021	Oil detection	Mauritius	2020	8	No	TOA	Band ratios, DSM	1
Argamosa et al., 2021	Oil detection	Manila Bay, Philippines	2020	2	No	TOA	PCA, Gradient boosting algor.	2
Samra & Ali, 2022	Oil detection	Nile Delta	2019-2021	N/A	Yes	Rayleigh	Band ratios, DSM	1
Tysiac et al., 2022	Oil detection	Ambarnaya river, Russia	2020	6	No	TOA	SVM, SI	6

Abbreviations: OBIA= Object-Based Image Analysis, SI= Spectral Indices, DSM= Decorrelation Stretch Method, PSA= Parallelepiped Supervised Algorithm, PCA= Principal Component Analysis, N/A= Not Mentioned, SVM= Support Vector Machine.

Recent studies for *Oil Spill* detection based on S2 optical imagery are summarized in Table 2.2. Starting from the pre-processing steps, most studies were based on Top-of-Atmosphere (TOA) reflectance products. Moreover, several methods have been focused on spectral indices (Kolokoussis and Karathanassi, 2018; Rajendran et al., 2021a; Tysiąc et al., 2022), spectral bands ratios (Abou Samra and Ali, 2022; Rajendran et al., 2021c, 2021b) and Object-Based Image Analysis (OBIA) (Kolokoussis and Karathanassi, 2018) (Table 2.2).

Support Vector Machine (SVM), Gradient Boosting Algorithm (Argamosa et al., 2022; Tysiąc et al., 2022) and Parallelepiped Supervised Algorithm (PSA) (Rajendran et al., 2021c) have also been applied for *Oil Spill* detection. Additionally, statistical methods have been implemented to detect *Oil Spills* and study oil slick distributions in S2 images (Althawadi and Hashim, 2019; Argamosa et al., 2022; Sun et al., 2018a). *Oil Spill* thickness was also determined through the Decorrelation Stretch Method (Abou Samra and Ali, 2022; Rajendran et al., 2021b). Most studies were designed to detect one class (i.e., *Oil Spill*), while only a few (Rajendran et al., 2021a; Tysiąc et al., 2022) considered five (i.e., *Oil Spill*, *Water*, *Snow*, *Vegetation*, *Wetlands)* or six (i.e., *Oil Spill*, *Water*, *Ground*, *Vegetation*, *Buildings*, *Snow*) classes.

2.2.1 Satellite data and benchmark datasets

Satellite sensors with high spatial, spectral and temporal resolution are required to develop automated detection methods and future operational solutions for monitoring the marine environment. Two Copernicus Sentinel-2 satellites (Sentinel-2A and 2B) were launched successfully on June 23, 2015, and March 7, 2017, with Multi-Spectral Instrument (MSI) sensor. The twin satellites cover global coastal waters at high spatial resolutions (10 m, 20m and 60 m). Additionally, they have frequent revisit time of approximately five days (depending on the latitude), delivering a large amount of free geospatial data. They also acquire data in 13 spectral bands ranging from the visible to the Short-Wave Infrared (SWIR). In our study, we collected and utilized the S2 Level 1C products (Top-of-Atmosphere reflectance) through EarthExplorer (USGS) and Coperinicus Open Access Hub .

Landsat-8 was launched on February 11, 2013 by National Aeronautics and Space Administration (NASA) and United States Geological Survey (USGS), carrying two instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The OLI measures in the visible, near-infrared, and shortwave infrared parts of the spectrum at a spatial resolution of 30 m with a revisit time of 16 days. L8 OLI also consists of a panchromatic band (15 m resolution), enabling pan-sharpening methods. L8 level1T data were acquired from EarthExplorer (USGS). PlanetScope provides imagery at 3 m to 5 m spatial resolution with almost daily

worldwide coverage. It consists of four bands (i.e., Blue, Green, Red, Near Infrared), while the newest generation of Planet satellites carry additional bands (i.e., Coastal Blue, Green I, Yellow, and Red Edge). Surface reflectance data were obtained from Planet Labs, USA.

In general, most of the currently available marine remote sensing datasets focus on detecting specific objects such as vessels ("Airbus Ship Detection Challenge," 2021; Heiselberg and Heiselberg, 2017; Liu et al., 2017; Tang et al., 2015). Datasets for cloud detection over the ocean (Kristollari and Karathanassi, 2020) and *Sargassum* macroalgae extraction (Ody et al., 2019; Wang and Hu, 2021a) have also been developed with a limited number of classes. Regarding *Marine Debris*, hyperspectral measurements have already been conducted, exploring sensors' capabilities in distinguishing plastics from other features such as vegetation, natural material, and water types (Garaba et al., 2020; Garaba and Dierssen, 2020; Knaeps et al., 2021; Tasseron et al., 2021). The aforementioned studies openly provided their data.

Despite the challenging and continuously growing issue of marine pollution, the available multispectral datasets are relatively limited in number and do not usually employ open-access high-resolution satellite images over geographically extended areas. These facts prohibit satellite data exploitation from ML frameworks and operational solutions. In particular, for *Marine Debris,* (Arias et al., 2021; Booth et al., 2022; Duarte and Azevedo, 2023; Mifdal et al., 2021; Rußwurm et al., 2023; Sannigrahi et al., 2022; Solé Gómez et al., 2022) employed more than 15 S2 scenes, while for *Oil Spill* only Sun et al. (2018) utilized a similar amount of S2 scenes. Additionally, most studies' annotated/ reference data were sparse and restricted to one or a few small geographical regions. Regarding *Oil Spill,* only Kolokoussis and Karathanassi (2018) considered two different areas, while for *Marine Debris* (Biermann et al., 2020; Booth et al., 2022; Duarte and Azevedo, 2023; Mifdal et al., 2021) examined globally distributed events. Finally, most studies were limited to a single up to a three-year period, while Sun et al. (2018) collected *Oil Spill* data for more than six years.

We highlight that most works in Tables 2.1 and 2.2 do not provide publicly available data. Although extensive Synthetic Aperture Radar (SAR) datasets with *Oil Spills* have already been introduced (Dong et al., 2022; Krestenitis et al., 2019), no S2 open-access dataset exists aiming to benchmark *Oil Spill* detection methods. For *Marine Debris*, Mifdal et al. (2021) openly provided annotations of floating objects (one general class) using global S2 imagery. Moreover, Topouzelis et al. (2020) offered satellite observations from the artificial debris targets in Mytilene, Greece.

2.3 Detecting Other Sea Surface features

***Sargassum* macroalgae**

Regarding *Sargassum* detection from optical data, most remote sensing studies have relied on floating algae indexes based on their red-edge spectral property in the NIR. Hu (2009) proposed the Floating Algae Index (FAI) using MODIS 250 m bands to detect large macroalgae blooms, including *Sargassum* macroalgae. An alternative Floating Algae Index (AFAI) based on 1 km MODIS bands was developed by (Wang and Hu, 2016), proposing band combinations for cloud-masking and a new method for *Sargassum* quantification. Additionally, research has been focused on investigating *Sargassum* spectral properties and the potential of discrimination from other sea surface features (e.g., other algae species, Oil) by employing pure spectroscopy and field or laboratory data (Hu et al., 2015; Qi et al., 2020).

Several studies have successfully applied vegetation and algae indices such as FAI, AFAI, NDVI, DVI, MCI on Sentinel-3 and high-resolution data for *Sargassum* detection and monitoring (Gower et al., 2013; Ody et al., 2019). The influence of parameters such as *Sargassum* immersion depth and water turbidity was also investigated to determine the relationship between AFAI (derived from the S2 sensor) and fractional coverage of *Sargassum* (Descloitres et al., 2021).

Regarding detection techniques using ML, the Random Forest algorithm, along with FAI and vegetation indices, were applied to Landsat-8 imagery to monitor spatial and temporal *Sargassum* distribution on Mexican coasts (Chávez et al., 2020; Cuevas et al., 2018). A fully automated workflow has been suggested to extract *Sargassum* from S2 imagery using FAI and Convolutional Neural Networks (CNNs) while removing noise patterns (i.e., cloud artifacts and wave-induced glints) (Wang and Hu, 2021a). Furthermore, RF has been applied to distinguish *Sargassum* from *Ulva* (Xiao et al., 2022), and VGGUnet has been employed for *Sargassum* extraction and biomass quantification in high-resolution optical data (Wang and Hu, 2021b).

Ships & Wakes

For *Ship* identification (i.e., size, type), statistical, threshold-based, and transform-domain methods have been applied on different optical sensors (Kanjir et al., 2018), while DL techniques on S2 data have been assessed (Štepec et al., 2019; Xie et al., 2020). CNNs have also been examined to distinguish ships from icebergs (Heiselberg, 2020), and self-supervised learning was exploited to generate binary ship/no-ship segmentation outputs (Ciocarlan and Stoian, 2021). Other competing bright classes, such as *Waves* and *Wakes,* have been studied based on SVMs (Heiselberg and Heiselberg, 2017; Song et al., 2014), while Liu et al. (2021) proposed a Radon transform method for *Wake* identification, effectively eliminating false alarms such as *Waves* and *Oil Spill*.

Sea snot

Reflectance spectra of *Sea snot/ mucilage*, extracted from multiple optical sensors, were examined by Hu et al. (2022) to understand the potential of discrimination from *Marine Debris.* Moreover, spectral indices (Kavzoglu and Goral, 2022), OBIA (Kavzoğlu et al., 2021) and SVM (Tuzcu Kokal et al., 2022) were adopted for *Sea snot* detection, as well as validation against in situ reference data was performed by Yagci et al. (2022). In addition, the correlation with Chlorophyll-a and Turbidity was explored based on S2 imagery (Sunar et al., 2022).

2.4 Spectral analysis and discrimination challenges

Representative examples of marine pollutants and other features, along with their spectral shapes in S2 images, are selected to be demonstrated in this section. For each case, the spectral signature from a specific pixel/ target is shown. The utilized S2 spectral bands are: 443 nm, 490 nm, 560 nm, 665 nm, 705 nm, 740 nm, 783 nm, 842 nm, 865 nm, 1610 nm and 2190 nm. Additionally, for *Marine Debris*, *Sargassum* and Sea snot, the mean spectra from the 5x5 pixel window (centered at the target/ pixel) are calculated to eliminate the spectral distortions due to mixed band resolutions (Hu, 2022, 2021). The subtraction technique was also performed due to its effectiveness in minimizing the impact of variable subpixel coverage and nearby water type (Hu, 2021; Qi et al., 2020).

Marine Debris

Marine Debris, especially plastics, have complex properties, diversifying in color, chemical composition, size and immersion depth. In fact, floating marine plastic debris can have various spectral signatures, occasionally similar to other features and constituents, such as vegetation, sediments and water (Garaba et al., 2018; Garaba and Dierssen, 2018). Additionally, spectral discrimination from other objects (e.g., ships, foam) is not straightforward, and differentiating from bright features (waves, sunglint, clouds) is currently considered very challenging (Garaba and Dierssen, 2018).

In Figure 2.2, the S2 image shows elongated *Marine Debris* trajectories at Motagua river front as citizen scientist Caroline Power reported. Inspection of the reflectance spectra (Fig. c) shows the typical peak at NIR band which has been already indicated by past studies based on spectra extraction from individual pixels (Sannigrahi et al., 2022)(Sannigrahi et al., 2022) or plastic targets (Biermann et al., 2020; Topouzelis et al., 2020). However, when averaging, a spectral trough at 833 nm appears due to the elimination of the spectral distortions, confirming the cautious notes by Hu, 2022. Further subtraction of nearby water was applied, and its SAM to plastic was estimated using the plastic bags endmember suggested by Hu, 2021.

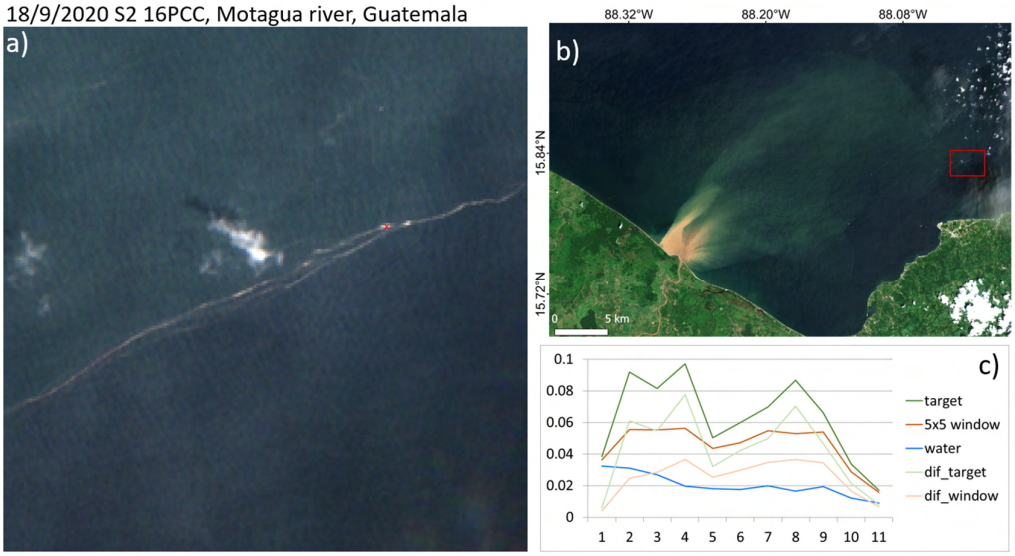


Figure 2.2 a) RGB patch showing plastic debris at Motagua river front in Guatemala. b) The location is outlined by a red square in the L2A Sentinel-2 image. c) The spectral signatures of the target and 5x5 pixels window as well as the reflectance difference between them and the nearby water are demonstrated.

The SAM value of 8,1 revealed a "strong" spectral similarity (Garaba and Dierssen, 2018). The spectral similarity between the plastic pixel and the plastic endmember was calculated using the following S2 bands: 560, 665, 704, 742, 842 and 865 nm.

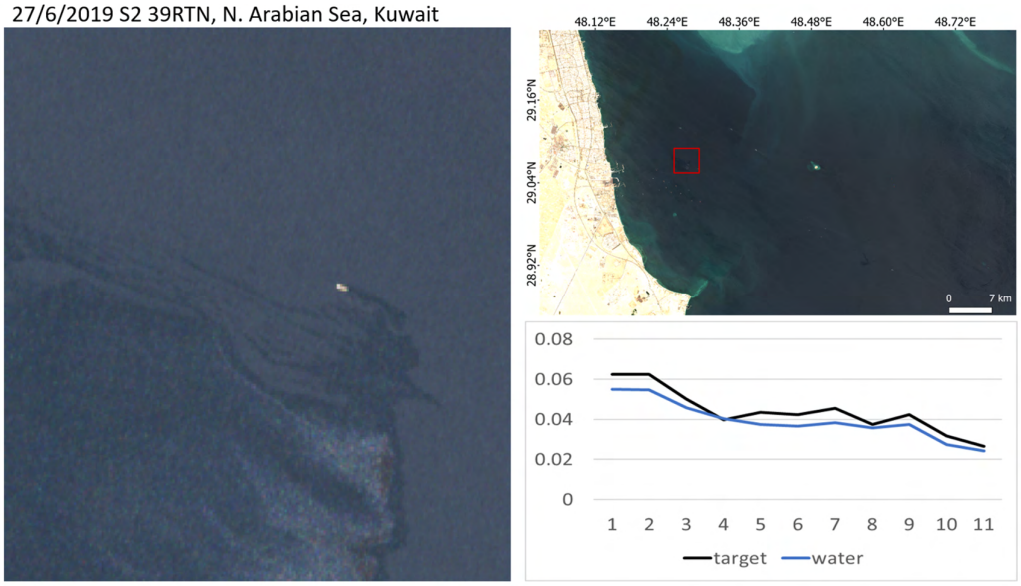


Figure 2.3 a) RGB patch showing Oil Spill in N. Arabian Sea. b) The location is outlined by a red square in the L2A Sentinel-2 image. c) The spectral signatures of the target and the nearby water are demonstrated.

Oil Spill

Oil shows absorption and reflection peaks in the visible and infrared parts, which can generate dark, gray, shiny, opaque, and rainbow colorations due to the variation in the incidence of light (BAOAC, 2007).

Oil can be detected from the point of view of optical remote sensing since its different inherent optical properties (i.e., scattering, absorption) from water contribute to oil-water contrast. Furthermore, oil-water contrast depends on sun glint, observing geometry, sea state (wind speed), oil thickness and state, and spatial/spectral resolutions of remote sensing images (Hu et al., 2021; Sun and Hu, 2019). Additionally, when interpreting optical remote sensing data, factors such as mixed pixels and water types (clear or turbid) need to be considered (Hu et al., 2021). The spectral signatures indicate higher reflectance values for *Oil Spill* than water in all S2 bands (Figure 2.3).

***Sargassum* macroalgae**

Since 2011, massive *Sargassum fluitans/ natans* blooms have inundated Caribbean and Atlantic coasts, posing a threat to tourism, fisheries, marine life and health. In January 2019, large amounts of *Sargassum* were washed ashore on the Roatan coastal zone in Honduras, captured by the Sentinel-2 sensor (Figure 2.4). Field measurements suggest that in comparison with other floating algae species (e.g., *Trichodesmium, Ulva*), *Sargassum* has a distinctive reflectance curvature of 630 nm due to its Chlorophyll pigments, resulting in a unique spectral signature (Hu et al., 2015). The spectral signatures demonstrate the typical red-edge reflectance of the floating vegetation and the local trough at 842 nm band (Figure 2.4). After averaging and subtracting, we note that the reflectance at 560 nm is slightly higher than the red band, indicating the presence of Chlorophyll pigment. For the SAM calculation, we used the same bands described in the *Marine Debris* section and utilized the *Sargassum* endmember recommended by Hu (2021). Its SAM value was 2,6, suggesting a "very strong" spectral similarity (Garaba and Dierssen, 2018).

Sea snot

Sea snot (or mucilage) refers to macroaggregations excreted from microorganisms, including phytoplankton, ranging from few cm to several km (Yagci et al., 2022). Here, we focus on the recent event in the Marmara Sea in the Spring of 2021, which raised awareness among the research community to monitor it using remote sensing data (Hu et al., 2022; Tuzcu Kokal et al., 2022; Yagci et al., 2022). On 24 June 2021, a S2 image demonstrates sea snot trajectories originating from the Sea of Marmara that reached the coasts of Limnos Island in the N. Aegean Sea. The spectral signatures show higher reflectance at red band than at green, possibly indicating a non-living algae's presence (Figure 2.5). Our findings follow the field measurements in the Gulf of Izmit collected by Yagci et al. (2022), who observed

that the highest surface reflectance values were in green, red and NIR bands, as the lowest were in SWIR bands.

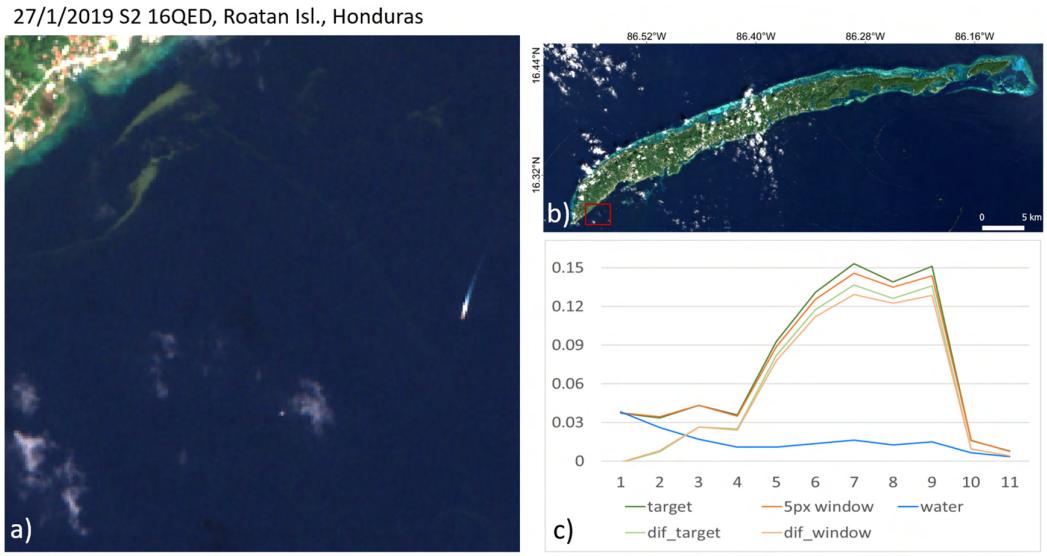


Figure 2.4 a) RGB patch showing dense Sargassum macroalgae trajectories in Roatan Island (Honduras). b) The location is outlined by a red square in the L2A Sentinel-2 image. c) The spectral signatures of the target and 5x5 pixels window and the reflectance difference between them and the nearby water are demonstrated.

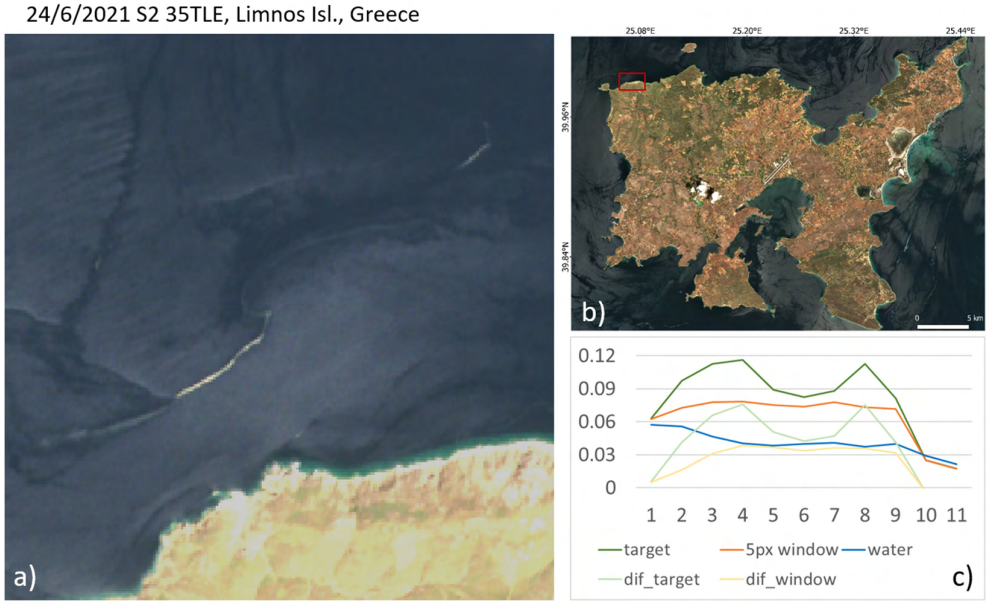


Figure 2.5 a) RGB patch showing Sea snot in Limnos Island (N. Aegean Sea). b) The location is outlined by a red square in the L2A Sentinel-2 image. c) The spectral signatures of the target and 5x5 pixels window and the reflectance difference between them and the nearby water are demonstrated.

Ships & Wakes

Vessels result in high reflectance at Red-Green-Blue (RGB) and Near-Infrared (NIR) bands; and sometimes reflect a different color from white (e.g., red) (Heiselberg, 2016). In general, vessel detection is still challenging, as complex sea state features (i.e., waves, sunglint) or other look-alikes (i.e., clouds, small islands) can be detected as false alarms (Kanjir et al., 2018). Additionally, turbulent wakes and Kelvin waves created from moving vessels have similar spectral characteristics to ships (Kanjir et al., 2018), yet their reflectance at the NIR band is much lower (Heiselberg, 2016) (Figure 2.6).

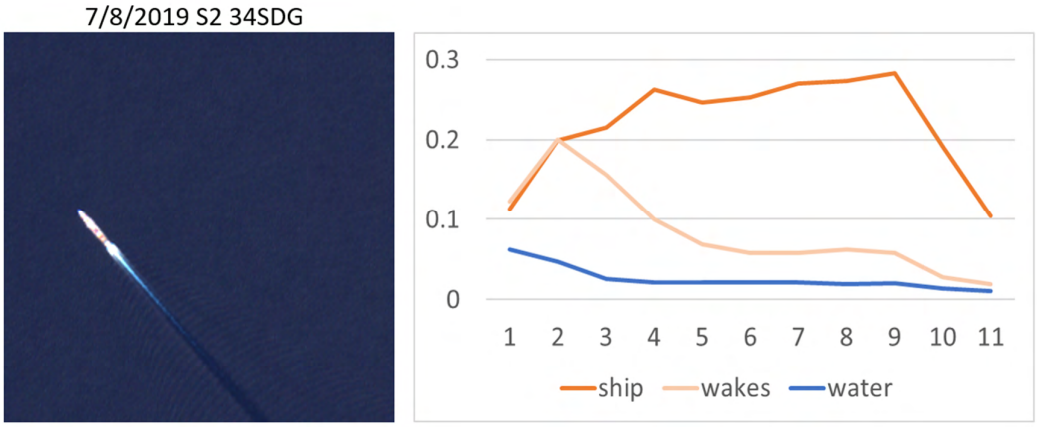


Figure 2.6 Left: RGB patch showing Ship and Wakes in the Ionian Sea. Right: The spectral signatures of Ship, Wakes and the nearby Water are demonstrated.

Waves-Whitecaps

Waves affect pixel values in optical data as wave slopes reflect the different amounts of sun light, and breaking waves (whitecaps) create bright patches (Kanjir et al., 2018). The high variation in the background reflectivity (clutter), proportional to the wind speed and sea state, leads to erroneous detection of objects at the sea surface (e.g., ships). In general, whitecaps appear "white" or spectrally flat in the visible (400 nm-700 nm), as occasionally, their reflectance can be enhanced at NIR band (Garaba and Dierssen, 2020) (Figure 2.7). Differentiating whitecaps from other look-alike features (e.g., ice, plastics) is still challenging for ocean color remote sensing and future sensor development (Garaba and Dierssen, 2020).

Water-related Classes

Clear water efficiently absorbs NIR to SWIR light with high reflectance at shorter wavelengths. However, surface features (e.g., foam, plastics, floating macroalgae), constituents within the water (e.g., sediments, phytoplankton) or shallow seafloor can make the water much brighter than normal (Garaba and Dierssen, 2020). Specifically for complex coastal waters, various spectral response patterns can be recorded due to great variability in optical properties (Ruddick et al., 2006). The

Figure 2.8 demonstrates a representative example of coastal waters ich in sediments below. Sediment-laden water shows a characteristic spectral pattern with two major peaks at 665 nm and 783 nm and a lower at 865 nm due to the increased concentrations of suspended sediments. This part of the spectrum has already been proposed to map turbidity and suspended matter using remote sensing (Dogliotti et al., 2015).

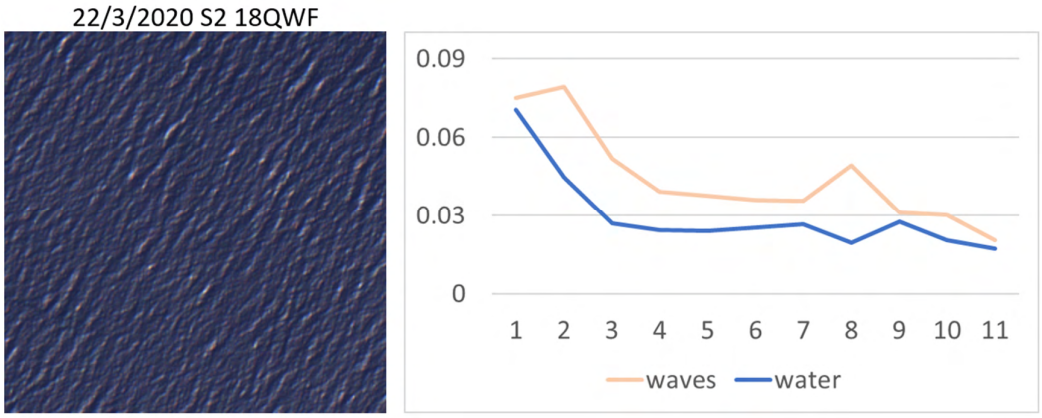


Figure 2.7 Left: RGB patch showing Waves in Caribbean Sea. Right: The spectral signatures of Waves and nearby Water are demonstrated.

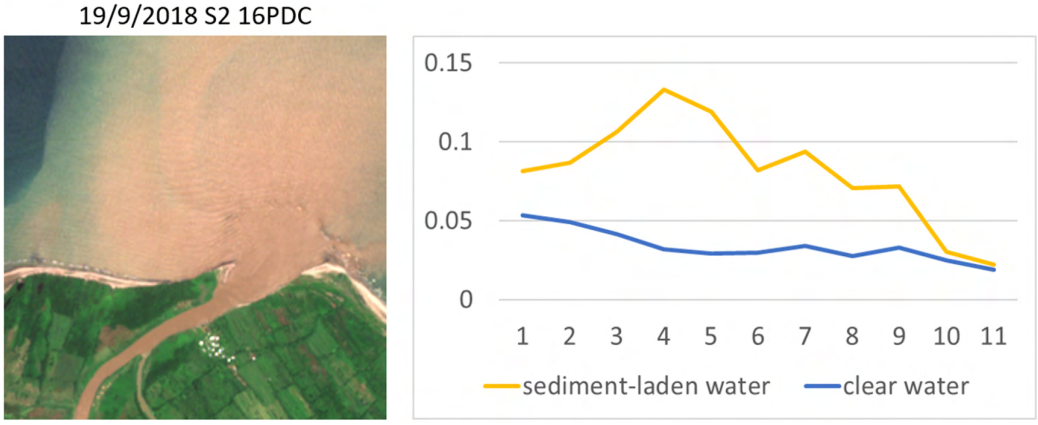


Figure 2.8 Left: RGB patch showing Sediment-Laden Water in Guatemala. Right: The spectral signatures of Sediment-Laden Water and nearby Clear Water are demonstrated.

Jellyfish

The recent increase of jellyfish blooms has been correlated with global changes, including warming temperatures, overfishing, eutrophication, hypoxia etc. (Condon et al., 2013). Based on laboratory measurements, sea jellies reflectance spectra resemble those of transparent plastic bottles due to the lack of pigment absorption (Van der Zande et al., 2014). An event with suspected Jellyfish identified by Qi et al. (2020) is presented below in Figure 2.9.

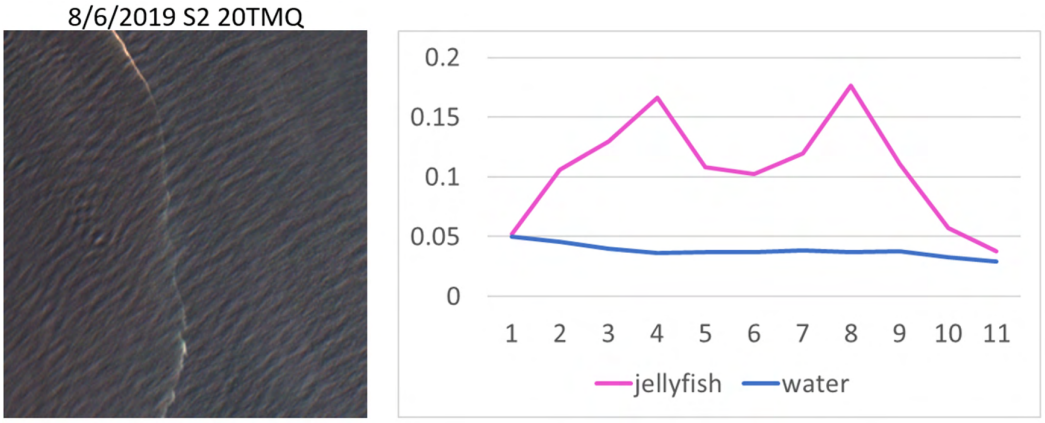


Figure 2.9 Left: RGB patch showing suspected Jellyfish in Nova Scotia. Right: The spectral signatures of Jellyfish and the nearby Water are demonstrated.

2.4.1 Challenging discrimination cases

By collecting numerous observations of marine pollutants and other sea surface features and recording their spectral shapes, we argue that their discrimination is quite challenging (Figure 2.10). In fact, the sea features complex optical properties and variable sea conditions can lead to various spectral signatures. Discrimination between competing classes becomes even more demanding when it is based only on spectral signatures (Figures 2.11-2.13).

More specifically, features such as Foam, Sea snot, Sediment-laden Water, Dense *Sargassum* macroalgae and man-made Ships and Oil Platforms exhibit higher reflectance values than other water-related classes (i.e., Turbid, Shallow and Marine Water), Waves, Wakes, *Oil Spill* and floating matters (i.e., *Marine Debris*, Natural Organic Material, Sparse floating algae and Jellyfish) (Figure 2.10). The largest overlapping of reflectance values is observed at the Blue and SWIR bands. Only man-made Ships and Oil Platforms can be clearly separated from the other features at SWIR wavelengths. Below, three representative challenging discrimination cases in S2 imagery (10 m spatial resolution) are demonstrated (Figures 2.11-2.13).

Figure 2.11 demonstrates Marine Water, Dense & Sparse Debris and a Ship in the region of Haiti. The spectral signatures analysis shows that Dense Debris and Ship are spectrally close. Sparse Debris is spectrally similar to Water, indicating the significant impact of density on spectral behavior. *Marine Debris* spectral signatures can also be close to suspected Jellyfish. Additionally, these features occasionally present similar spatial patterns (Figure 2.12). Another challenging case is the discrimination between *Oil Spill* and Water (Figure 2.13). Figure indicates that their separation based only on spectral shapes is quite demanding (Figure 2.13).

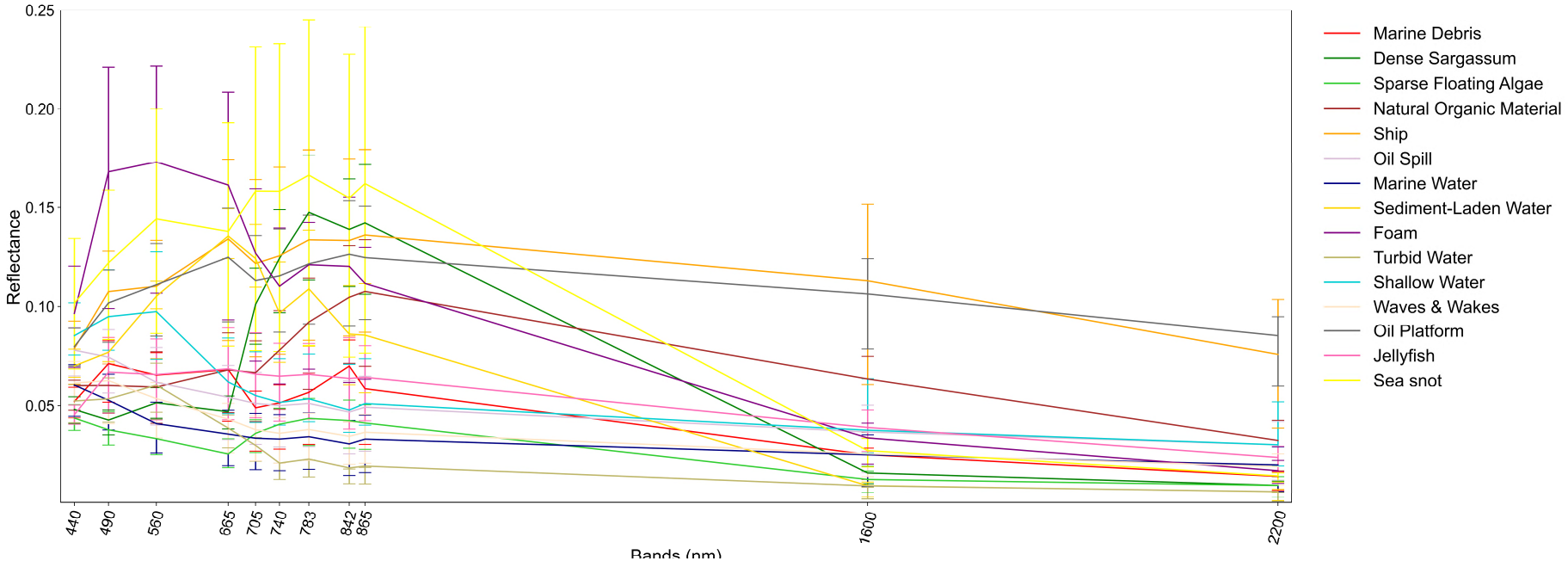


Figure 2.10 The Spectral Signatures of Marine Pollutants and other competing Sea Surface Features. The mean Spectral Signatures are presented with 25-75 percentiles as error bars.

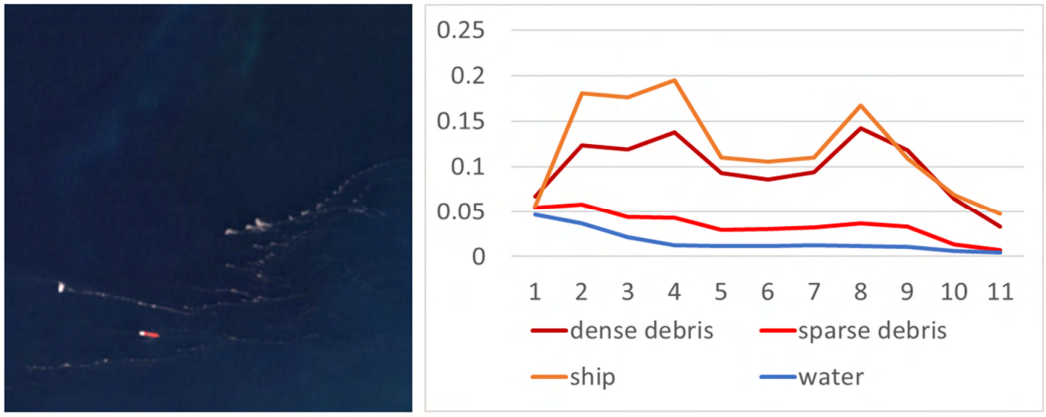


Figure 2.11 Left: RGB patch showing Marine Debris, Ship and Marine water in the region of Haiti. Right: The spectral signatures of Dense & Sparse Debris, Ship and the nearby Water are demonstrated.

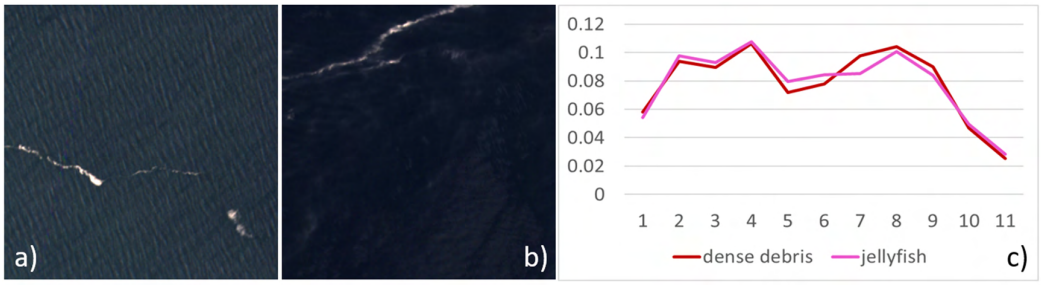


Figure 2.12 a) RGB patch showing Marine Debris in Motagua river. b) RGB patch showing Jellyfish in Nova Scotia. c) The spectral signatures of Dense Debris and Jellyfish are presented.

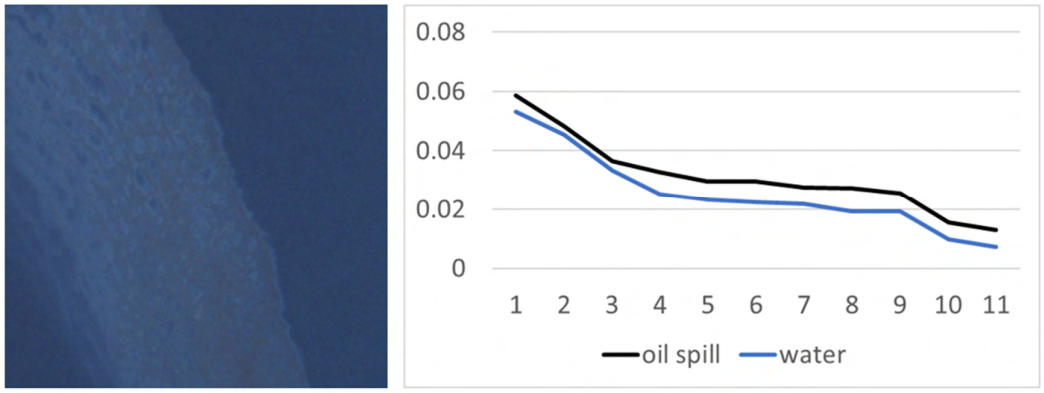


Figure 2.13 Left: RGB patch showing Oil Spill and Marine water. Right: The spectral signatures of Oil Spill and the nearby Water are demonstrated.

3 Detect and Track floating Plastic Debris in Multitemporal Data

In this chapter, we assess the capabilities of current satellite sensors in detecting and tracking *Marine Debris*. Section 3.1 demonstrates the specifications of Landsat-8, Planet, Sentinel-2 and the ancillary data characteristics. Next, section 3.2 presents the methodological approach for collecting in situ reports and identifying *Marine Debris* in multitemporal high-resolution satellite images over Honduras Gulf. Furthermore, experimental results on spatial distribution and major characteristics of *Marine Debris* are described in section 3.3. Section 3.4 presents the discussion regarding marine litter monitoring from satellite sensors. Conclusions from the whole chapter are provided in section 3.5.

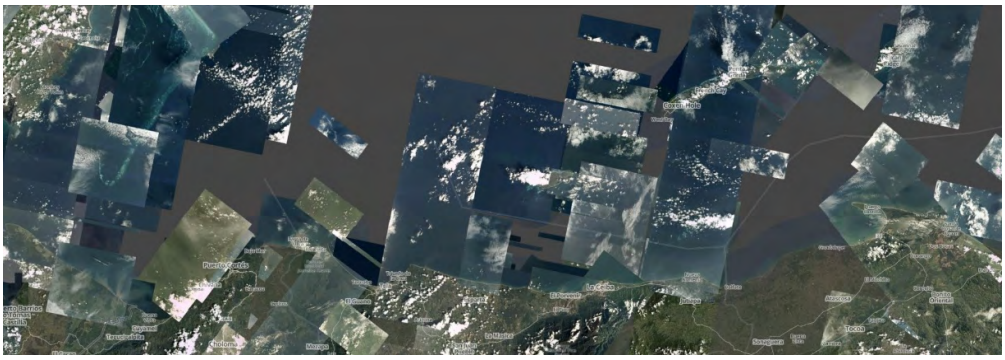
3.1 Multitemporal satellite data and Ground Truth

From June 2014 to December 2018, a total amount of 125 Landsat-8 (L8) OLI level1T data and 340 Sentinel-2 (S2) MSI level1C images were collected. In particular, three L8 path/row (i.e., 19/49, 18/49 and 17/49) and eight S2 tiles (i.e., 16PCC, 16PDC, 16PEC, 16PFC, 16PGC, 16QFD, 16QED and 16QDD) were selected, covering the study area (Figure 3.1). The satellite data with a cloud presence of over 25% were rejected and were not further processed. The rest of the data were atmospherically corrected, and surface reflectance values were extracted using ACOLITE atmospheric processor (Vanhellemont and Ruddick, 2018). Moreover, more than 400 high-resolution Planet images were analyzed for mapping the observed plastic debris and *Sargassum* in the study area.

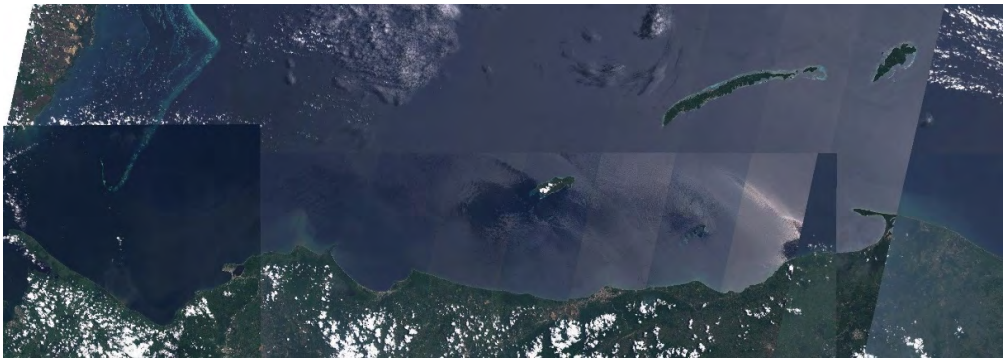
Between 2014 and 2019, in situ data were collected through vessel and diving expeditions around Bay islands (Table 3.1). Additional data were obtained from local stakeholders who were engaged in beach cleanup, professional diving, the "Dive Against Debris" initiative, fishing and media in Honduras, and the Solid Waste Management office of the Municipality of Zacapa in Guatemala. Information about floating plastic debris and *Sargassum* occurrences and trajectories was systematically recorded and used as reference data against the multitemporal observations on the satellite data. For *Sargassum* events, reports from the South Florida Optical Oceanography Laboratory were also employed and analyzed. Meteorological and sea state data (e.g., precipitation, wind, sea surface currents) were also acquired for a six-year period, i.e., 2014-2019. These data were associated and linked with the multitemporal satellite observations and debris events. In particular, daily precipitation data were acquired from the La Ceiba station (15° 44′ 53″N, 86° 50′ 36″W) through the National Oceanographic Data Center of National Oceanic and Atmospheric Administration (NOAA) (https://www.nodc.noaa.gov/). Moreover, data regarding the wind direction and speed were derived from the National Data

Buoy Center (https://www.ndbc.noaa.gov/) and NASA Prediction of Worldwide Energy Resources (https://power.larc.nasa.gov/data-access-viewer). Sea surface currents velocity and direction was estimated using daily ocean physics analysis data from (CMEMS, http://marine.copernicus.eu/).

(a) Planet Data *(September – October 2016)*



(b) Sentinel-2 Data *(October 2016)*



(c) Landsat-8 Data *(October 2016)*

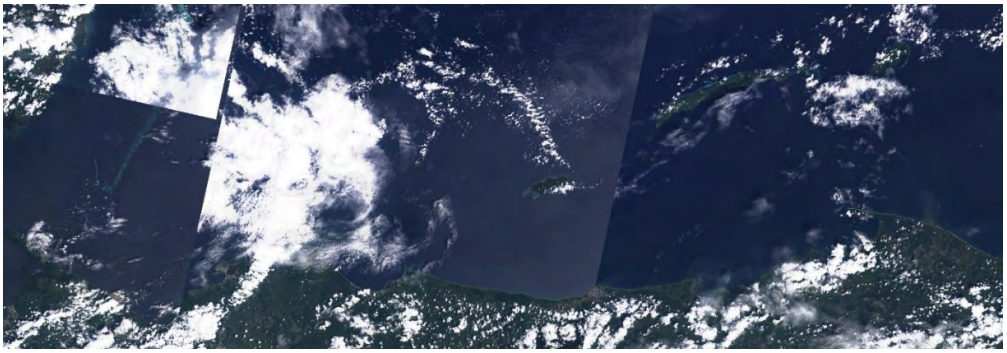


Figure 3.1 Indicative satellite data for late September and October 2016 plastic debris event. (a) Planet satellite images from 3m to 5m spatial resolution. (b) Sentinel-2 images at 10 m spatial resolution. (c) Landsat-8 images at 30m resolution.

Table 3.1. Reported plastic debris events along with corresponding coordinates (WGS'84).

Detected Debris Events	Date of *In Situ* Data Collection	Location of *In Situ*
September 2014	30/9	16.2193° N, 86.5960° W
November 2015	29/11	16.0667° N, 86.3965° W
October 2016	22/10	16.1843° N, 86.4233° W
November 2016	3/11	16.1182° N, 86.4958° W
October 2017	17/10	16.0367° N, 86.5878° W
February 2018	8/2	16.2516° N, 86.6008° W
March 2018	8/3	16.2955° N, 86.5416° W
October 2018	19/10	16.2999° N, 86.3897° W
September 2019	28/8	14.9827° N, 89.5442° W

3.2 Photo-Interpretation & Identification of Floating Marine Debris

Based on the multitemporal reference data from in situ observations, numerous satellite images from different satellite missions were collected and annotated with plastic debris polygons in a Geographic Information System environment (i.e., QGIS). In particular, an intensive and laborious manual on-screen digitization was performed from photo-interpretation experts, annotating numerous images with plastic debris and *Sargassum* slicks. Apart from these two primary categories (i.e., plastic debris and *Sargassum*), plastics were also subdivided into individual polygons according to the observed density of the plastics, i.e., (a) dense, (b) medium and (c) sparse. In total, around 1500 pixels of "dense", 3200 pixels of "medium" and 5000 pixels of "sparse" pixels were digitized. Based on the in situ data and verified debris events around the Roatan and Cayos Cochinos islands, numerous pixels containing clear material observations (i.e., not mixed pixels) were collected towards calculating the spectral signatures of both plastic debris and *Sargassum* from the multitemporal multispectral data. In order to do so, around 1600 reflectance values (Figure 3.2) were extracted from the satellite data and compared with the literature (Acuña-Ruz et al., 2018; Biermann et al., 2020; Hu et al., 2015). Since we collected and processed numerous satellite images, the detection and verification of plastics were performed on data acquired under various weather and ocean conditions.

Along with the events that were directly verified with in situ observations, additional debris events were detected based on the observed spectral signatures, plastic debris characteristics reported in the literature, the detected debris size and pattern, the successful detection of debris on multitemporal data, the successful

correlation of the estimated trajectories and sea surface currents direction, their association with specific river discharges and corresponding rainfall events.

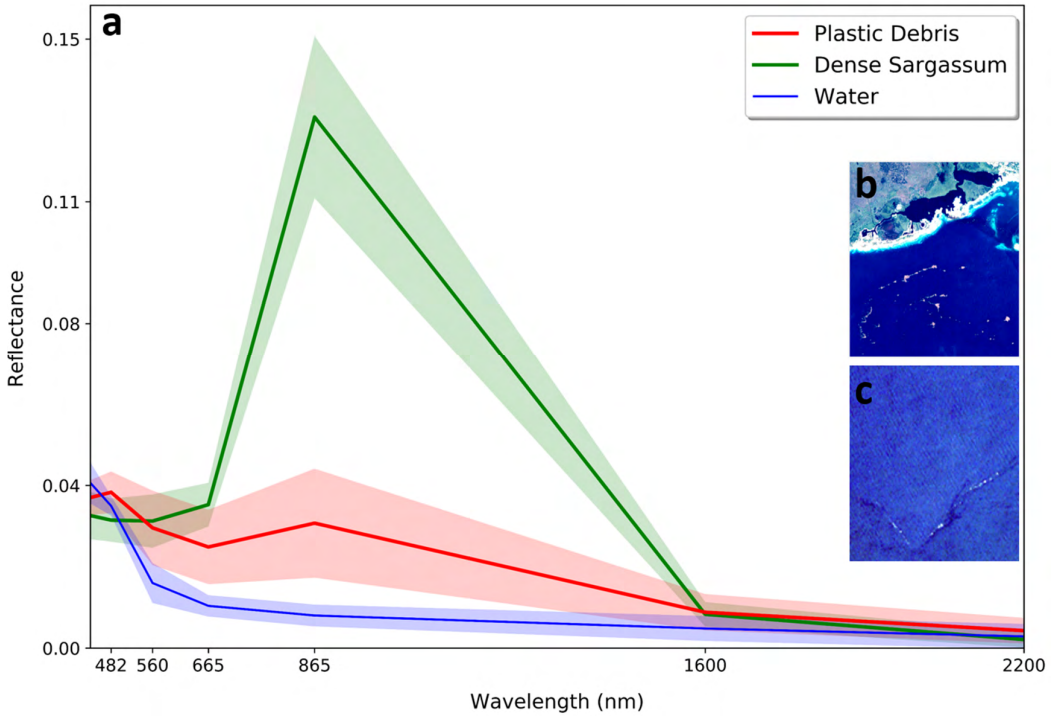


Figure 3.2 The systematically recorded, from the multispectral remote sensing data, spectral signatures of plastic debris, dense Sargassum and seawater along with their standard deviation (n = 1600). (a) Spectral signatures of plastic debris (red), dense Sargassum (green) and water (blue). (b) An indicative Planet image with dense Sargassum. (c) Indicative Planet image with plastic debris.

The detected debris was verified based on the aforementioned criteria and was discriminated from any other sea surface structure that may resemble plastic debris i.e., mainly seawater convergences and fronts. In Table 3.2, all the detected and verified events, the corresponding satellite data and the verification manner are presented.

In order to quantify the detected plastics, the area of every debris slick was calculated based on the number of pixels at every given polygon and the spatial resolution of the satellite image. An adjustment based on material density (i.e., dense, medium and sparse) was applied for every plastic debris pixel. In order to estimate the weight of the detected plastics per polygon, we made the following assumptions based also on the corresponding in situ observations: the average thickness of the observed dense floating plastics was approximately 30 cm, which results in a plastic mass of 5000 tonnes per km2. Although field and laboratory observations are accurate and direct methods for plastics properties examination, prediction models can provide an automated and cost-effective indication of discharged plastic debris weight per river based on waste management, population density and hydrological information. For this reason, the satellite-estimated weight for all events was also

compared with the ones derived from state-of-the-art global prediction models (Lebreton et al., 2017). In order to do so, the required geospatial data were obtained and monthly plastic mass midpoint estimates (Lebreton et al., 2017) were used for each river (Table 3.3).

Table 3.2. The detected 20 plastic debris events for the years 2014 to 2019 based on satellite and in situ data.

Detected Events (date)	**Satellite Data & Acquisition Dates (day/month), *clouds >25%**			**Validation**
	Landsat-8 (L8)	**Sentinel-2 (S2)**	**Planet (P)**	
September 2014	20/9, 29/9	-	-	in situ *(30/9)*
November 2015	3/11, 10/11	29/11	24/11	in situ *(29/11)*
January 2016	22/1	15/1	26/1	photo-interpretation
February 2016	23/2	14/2*	16/2	photo-interpretation
September 2016	25/9*	4/9, 11/9, 24/9	29/9	photo-interpretation
October 2016	4/10, 20/10	14/10	28/9, 8/10	in situ (22/10)
November 2016	5/11*	3/11	9/11*	in situ (3/11)
January 2017	15/1*	12/1, 25/1, 31/1	29/1	photo-interpretation
February 2017	25/2	21/2	14/2, 27/2	photo-interpretation
March 2017	13/3	23/3	6/3, 22/3, 26/3	photo-interpretation
August 2017	20/8*	10/8, 15/8, 30/8	25/8, 26/8	photo-interpretation
October 2017	21/9, 28/9	11/10, 14/10, 9/10	27/9, 7-8/10, 15-17/10	in situ (17/10)
January 2018	11/1	22/1, 24/1	7/1, 8/1	photo-interpretation
February 2018	12/2*	16/2, 26/2	11/2, 24/2, 25/2	in situ (8/2)
March 2018	16/3	8/3, 18/3	13/3, 17/3	in situ (8/3)
September 2018	15/9, 24/9	14,9 19/9	13/9, 19/9	photo-interpretation
October 2018	26/10, 17/10	14/10, 24/10, 29/10	21-22/10	in situ (19/10)
November 2018	27/11	13/11, 23/11, 28/11	24/11	photo-interpretation
December 2018	20/12	13/12, 28/12	14/12, 19-22/12, 25/12	photo-interpretation
September 2019	18/9	4/9, 9/9	3/9, 6/9	in situ (28/8)

In several cases of in situ collections, the floating plastic debris was mixed at a certain extent with macroalgae species or organic material (wood, tree branches etc.). This is expected and in accordance with the literature (Martínez-Vicente et al., 2019). Due to the spatial resolution of the satellite data (i.e., 3 m to 30 m), discrimination of plastics from other floating features was not achievable at the pixel level. Spectral unmixing was also beyond the scope of this work. For this reason, throughout this study, only the clear—inside the range of their standard deviation (Figure 3.2)—plastic pixels were considered and annotated. We estimated the weight with the same manner for all events; however, these estimations are representative for our study area with cases of dense plastic masses (Lechner et al., 2014).

3.3 Marine Debris Sources & Pathways Detection

The next subsections describe (i) the process of detecting and verifying the plastic debris from space (Figure 3.3 and Figure 3.4), (ii) its spatial distribution and extent in the study area (Figure 3.5), its major characteristics (Table 3.3), (iii) the detection of debris source (Figure 3.7) and (iv) the observed dominant transport trajectories of floating debris (Figure 3.8).

3.3.1 Verifying Satellite Observations

Plastic debris pixels identification was achieved through interpretation of satellite-derived data from three satellite missions (Figure 3.1). Using successive satellite images before and after the reported plastic events, we managed to detect their sources at the rivers' mouths as well as their transport trajectories in Bay Islands area. In almost all cases that plastics were detected and tracked, they were accumulated at dense masses following linear trajectories.

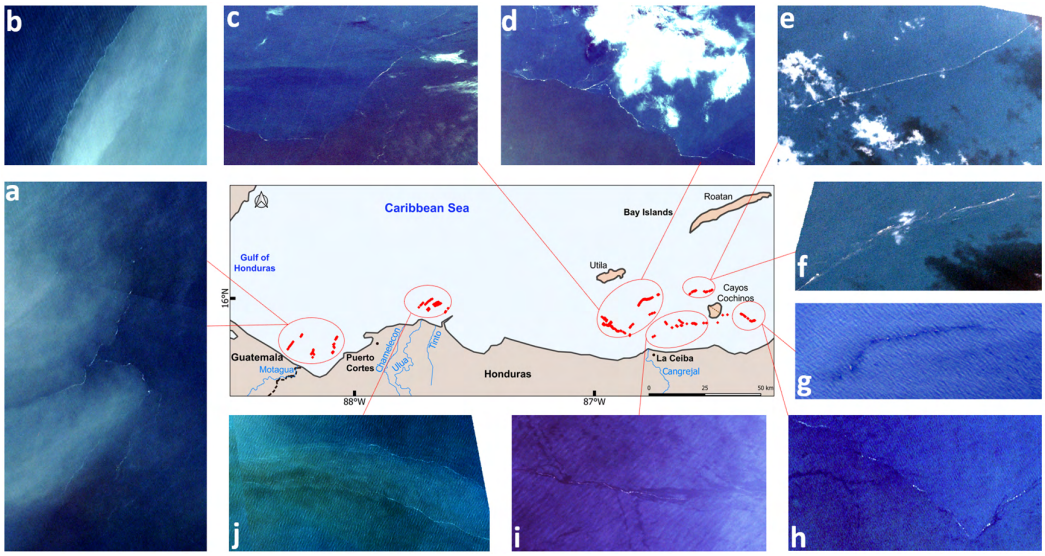


Figure 3.3 The detected plastic debris (red dots in the map) and snapshots of the corresponding satellite images at the Gulf of Honduras and Bay Islands during late September–October 2017. (a,b) Planet images with the detected plastic debris from the Motagua River on 27/9/2017. (c, d) The location of the detected debris on 15/10/2017. (e, f) Two days afterwards (17/10/2017), debris reached Cayos Cochinos island. (g) Debris trails were detected (9/10/2017) on a Sentinel-2 image. (h) Same as (g) but on a Planet image. (i) Debris detected on 7/10/2017 indicating that River Cangrejal also contributes to plastic pollution. (j) Planet data (27/9/2017) with the detected plastics originating from the Chamelecon, Ulua and Tinto rivers.

Plastic debris was occasionally detected in fronts, which were identified in satellite data through blue color differentiation due to discrete water masses (Figure 3.3).

Our results were verified with in situ data that were collected around the Bay Islands i.e., Roatan, Utila and Cayos Cochinos (Figure 3.4). Collected data indicated dense plastic masses of 6 km length (on average), and their widths varied from 1 m to 40 m. Plastic debris in the Bay Islands was dominated by macroplastics and

secondary microplastics, reflecting its proximity to land and plastic pollution sources. In some cases (i.e., October 2017, November 2016 and November 2015), satellite data captured exactly the region of observed plastic debris, and in other cases (i.e., September 2014, October 2016, October 2018,) plastics were detected in satellite data up to 10 km from the in situ collections (Figure 3.4).

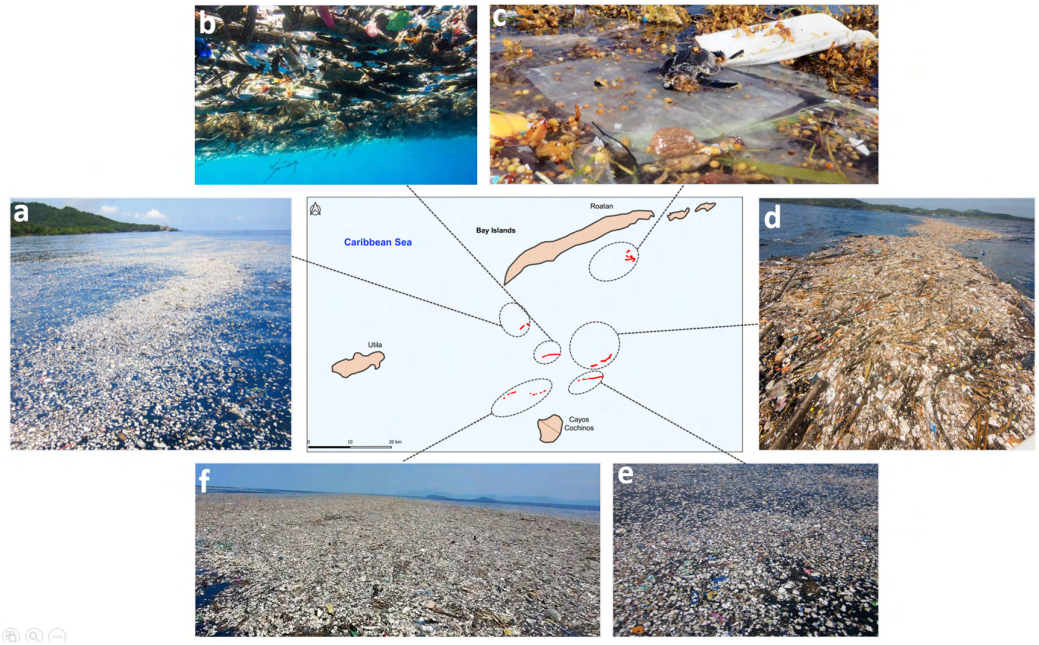


Figure 3.4 Detected plastic debris in satellite data (red dots in the map) and the corresponding in situ verification during the years of 2014–2018 around Bay Islands. (a) Collected in situ data southwest of Roatan island in September 2014. (b) In situ observations and underwater capture in November 2016. (c) Collected plastics in Roatan area. Sargassum macroalgae and a dead juvenile turtle were also recorded (October 2018). (d) Observed plastics during October 2016. Organic material (i.e., wood) was also recorded. (e) Observed plastic masses during November 2015. (f) Large plastic masses recorded in October 2017. Caroline Power acquired the photos of this figure.

3.3.2 Spatial Distribution and Descriptive Information of Plastic Debris

A validation procedure was applied on all reported plastic debris events from in situ collections, leading to plastic debris pixels detection and annotation in the available satellite data from 2014 onwards. In our study area, there is no evidence of a specific region where plastics accumulate. Instead, plastic debris seems to be distributed and travel all over the region, indicating a potential dynamic circulation influencing its transport around the Bay Islands. Our detections highlight that Guatemala's and Honduras' river mouths are the main sources of plastic pollution in the study area.

In particular, the Motagua River (Guatemala) and Ulua River (Honduras) seem to be major sources of frequent plastic debris discharges, followed by Tinto, Cangrejal and Aguan rivers (Honduras). Other relatively large (e.g., Sico,

Chamelecon) and smaller rivers (e.g., Lean, Cuero) in Honduras have been also detected as occasional plastic debris sources (Figure 3.5).

In order to obtain more descriptive information about plastic debris source and transport in the Bay Islands, we examined their characteristics during all recorded events. Table 3.3 summarizes the satellite-detected events and provides information regarding the surface area covered by debris, its estimated weight, the detected trajectory based on multitemporal observations and oceanographic conditions for the corresponding dates, i.e., surface currents and wind speed/ direction. In total, 20 plastic debris events were recorded based on in situ collections, satellite observations and photo-interpretation.

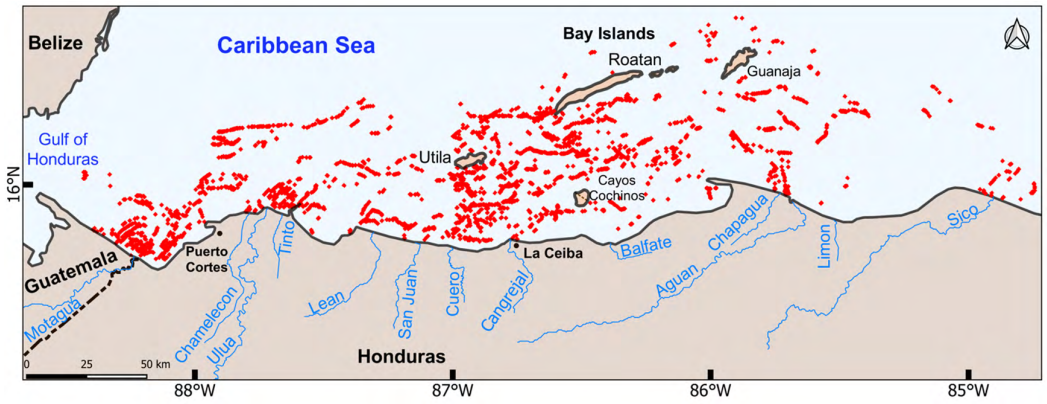


Figure 3.5 Total satellite-detected marine plastic debris from 2014 to 2019 in the southeast Gulf of Honduras and Bay Islands in the Caribbean Sea (red dots in the map). Plastics debris enters Caribbean Sea through river discharges. Plastic debris travels long distances dispersed in the entire study area.

Analyzing the numerous collected data, we found that plastic debris events occurred during late summer to early Spring (August to March). During the specific events, plastics were still accumulated in dense masses, reinforcing the capability of their detection, yet their spread, as well as their progressive degradation in smaller plastic debris, did not allow further detection in available satellite data. Table 3.3 demonstrates that September, October, January and November were the months with a significant presence of floating plastics at sea. Regarding the duration of their presence at sea surface during the study period, in 2016 and 2017, the events lasted for five months, while in 2018 they lasted for more than half a year. Based on the outcomes of debris detection and quantification for all reported events, we found that the total area of detected plastic debris was 0.77 km^2 in 2018, 0.5 km^2 in 2017 and 0.23 km^2 in 2016.

During the recorded events, currents appeared to have an SW–NE direction with velocities ranging from 0.06 to 0.26 m/s (Figure 3.6) and easterly winds prevailed with recorded speeds of 1.3 to 6.3 m/s. Based on the multitemporal satellite data, debris tracking (i.e., detection of the same plastics at different locations/ images on different dates) was achieved for 13 out of the 20 plastic debris events.

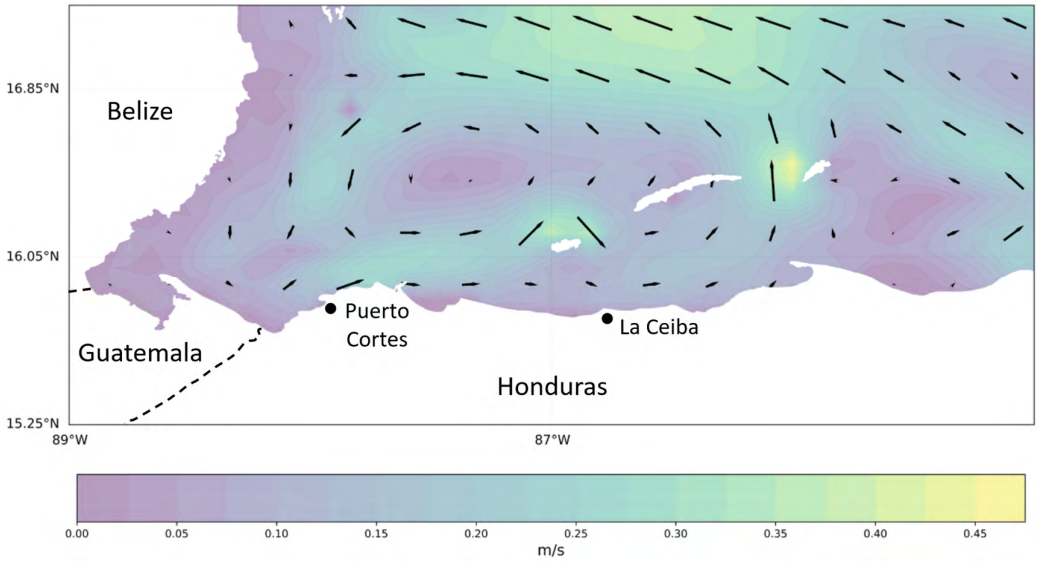


Figure 3.6 Sea surface currents direction and velocity for the 9th of October 2017 as derived from CMEMS (Global ocean 1/12° physics analysis and forecast daily product). Honduras gyre is recorded west of Roatan island.

Combining information from in situ data and remote sensing observations, we calculated the plastic debris speed for those events. According to Table 3.3, plastics' travelling speed ranged from 0.02 to 0.14 m/s; however, in every case their speed was found to be lower than the currents' speed. This indicates that debris travels at a slower pace than the speed of sea surface currents, probably due to its heavy weight and submerged structure. For this reason, we examined if larger, heavier patches could lead to relative lower travelling speeds. However, testing such hypothesis offers no such evidence, since no correlation was found between the estimated debris weight and its speed ($r^2 = 0.02$, $p > 0.05$, $n = 13$), indicating that the travelling pace is mainly controlled by the intensity of ocean currents regardless of its weight. For almost all cases indeed, we found that prevailing currents were strong enough to transport plastic debris. The estimated speed of plastic debris (Table 3.3) was highly correlated with currents' velocity ($r^2 = 0.75$, $p < 0.005$, $n = 13$), yet debris speed did not present any correlation with wind speed ($r^2 = 0.01$, $p > 0.05$, $n = 13$). The latter result could be expected, as prevailing eastern wind direction affected debris movement only for the detected event during November 2015 (Table 3.3).

Table 3.3. The detected 20 plastic debris events in Bay Islands for the years 2014 to 2019. The detected area, weight, sources, travelling trajectories, wind and currents conditions are also presented for each event.

YEAR	MONTH	Detected Source/ River Discharges	Detected Area *Km²*	Estimated Weight (Actual) *tonnes*	Detected Debris Direction	Travel Distance *days/Km*	Estimated Speed *m/s*	Wind Direction	Wind Speed *m/s*	Currents Direction	Currents Velocity *m/s*
2014	September	Ulua, Motagua	0.026	130	SE–NW	1d/7km	0.08	ESE	1.3	SE–NW	0.09
2015	November	Cangrejal, Motagua	0.028	140	NE–SW	5d/11km	0.02	ENE	4.9	SW–NE	0.06
2016	January	Limon, Motagua, Chamelecon, Aguan	0.052	260	SW–NE	4d/50km	0.14	NNW	5.2	SW–NE	0.22
	February	Chapagua,Tinto,Ulua, Aguan	0.019	95	SE–NW	7d/82km	0.14	ENE	5.6	SE–NW	0.26
	September	Motagua, Lean, Chamelecon, Sico	0.047	235	SW–NE	5d/15km	0.04	E	5	SW–NE	0.06
	October	Aguan, Cangrejal, Ulua, Tinto, Balfate	0.055	275	SW–NE	12d/65km	0.06	NE	2.5	SW–NE	0.07
	November	Motagua, Cuero, San Juan	0.061	305	–	-	-	NE	5	SW–NE	0.1
2017	January	Ulua, Tinto, Lean, Motagua, Cuero	0.112	560	–	-	-	E	5.3	E–W	0.12
	February	Motagua	0.026	130	–	-	-	SE	4.4	SW–NE	0.08
	March	Aguan, Sico, Limon	0.085	425	SE–NW	7d/38km	0.06	ENE	5.9	SE–NW	0.12
	August	Ulua, Tinto, Cuero, Motagua	0.131	655	SW–NE	11d/105km	0.11	E	4.1	SW–NE	0.13
	October	Motagua, Ulua, Tinto, Cangrejal, Chamelecon	0.14	700	SW–NE	2d/8km	0.04	ESE	3	SW–NE	0.1
2018	January	Motagua, Ulua, Aguan, Sico	0.133	665	SW–NE	4d/15km	0.04	NNW	3.8	SW–NE	0.1
	February	Motagua, Ulua, Tinto, Aguan	0.035	175	–	-	-	E	6.3	SE–NW	0.15
	March	Lean, Ulua, Cangrejal	0.024	120	–	-	-	E	5.3	SE–NW	0.14
	September	Motagua, Ulua, Tinto, Chamelecon, Cangrejal	0.138	690	SW–NE	5d/21km	0.05	E	5.6	SW–NE	0.06
	October	Motagua, Ulua, Tinto, Chamelecon, Lean, Cuero, Cangrejal	0.154	770	SW–NE	6d/38km	0.07	ENE	2.7	SW–NE	0.08
	November	Aguan, Motagua, Lean	0.135	675	SE–NW	1d/9km	0.1	ENE	3.6	SE–NW	0.12
	December	Cangrejal, Cuero, Motagua, Aguan, Tinto	0.149	745	–	-	-	NNW	5.6	W–E	0.12
2019	September	Motagua, Tinto, Ulua	0.053	265	-	-	-	NE	2.3	S–N	0.08

3.3.3 Mapping the Origin & Travelling Trajectories: Representative Cases

For all of the recorded events (n = 20) we found that the detected floating debris was originating from certain river mouths (Figure 3.5), and the discharges were related to major recorded rainfalls. All satellite sensors captured plastic debris mostly at river plumes, while only the very high-resolution Planet sensor recorded plastics at the rivers' tidal area. Satellite data were also capable of verifying in situ observations of massive plastic debris areas along large rivers in the mainland. For example, in late August/ early September 2019, debris source was verified with in situ observations along the Motagua River around the city of Zacapa (Guatemala), which is 170 km away from the river's estuary. The detected debris floating along the Motagua River (Figure 3.7) was only partially captured on custom-made litter-stopping booms (nets with plastics, Figure 3.7c), while the majority of plastics (i.e., mostly plastic bottles, plates and food wrappers) were discarded through Motagua River mouth. Satellite data evidently detected plastic debris areas during September 2019 (Figure 3.7e,f,g).

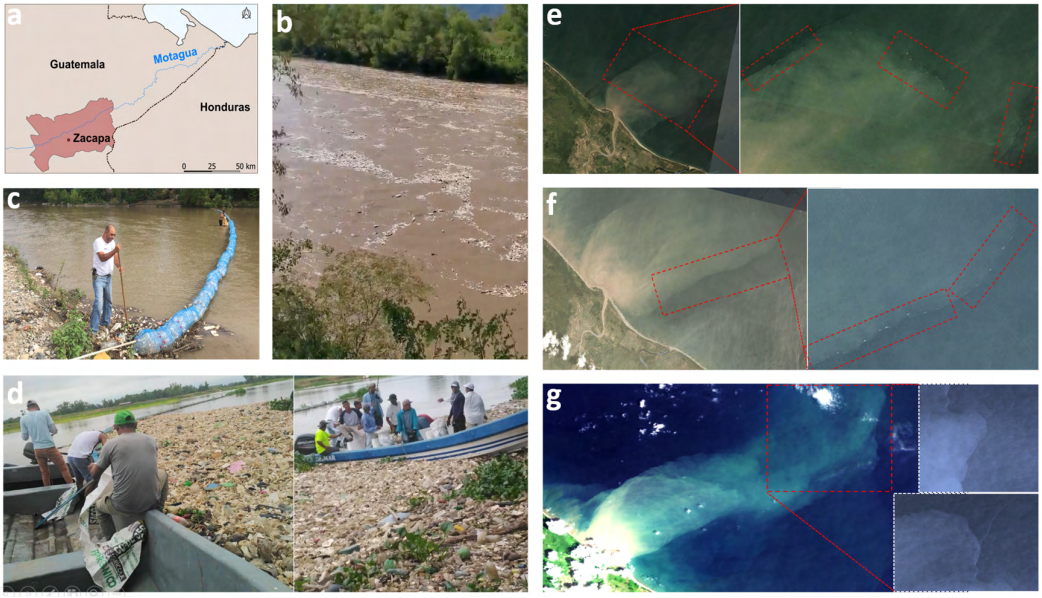


Figure 3.7 Detected floating plastic debris along the Motagua River in the Zacapa province of Guatemala and the corresponding discharges detected in satellite data during late August/ September 2019. (a) Location of Zacapa city. (b) Recorded plastic debris along the Motagua River. (c,d) Plastics partially collected by booms, which managed to collect only a small amount of debris, since they were rapidly overwhelmed and debris overpassed them. (e) Detected plastic discharges at the Motagua River mouth on a Planet image (3/9/2018). (f) Same as (e) but on 6 September 2018. (g) Detected plastic debris in Motagua River mouth on Landsat-8 data (18/9/2018). Julio R Guzman Perdomo acquired the photos (b), (c) and (d) of this figure.

Based on the detection of floating debris at multiple sequential dates, we managed to track plastic debris transport and estimate its travelling trajectories for 13 out of 20 reported events (Figure 3.8). A quite representative case was the

reported event of late September–October 2017, during which remarkable amounts of plastics were discharged following the rainfalls of the 24th–25th of September 2017 and onwards. In particular, on the 7th and 9th of October, plastics from Cangrejal River were tracked on Planet and Sentinel-2 data in a distance of 16 km up to 48 km from the river's mouth.

Tracking of floating debris was achieved in the multitemporal satellite data between the 7th and 9th of October. On the 7th of October, two plastic slicks were detected on a Planet image (1 km and 600 m length, respectively) originating from the Cangrejal River. On the 9th of October (Planet and Sentinel-2 images), the same debris was detected and tracked (1500 m length) travelling towards Cayos Cochinos, indicating that the two previous instances had been accumulated into this larger one. Based on detected river discharges, identified plastic slicks, tracking and surface circulation, we estimated plastic debris travelling trajectories in the Bay Islands area for October 2017. Our findings indicate that sea surface currents transported plastic debris in a SW-NE direction, and due to the local dynamic circulation, debris travelled over a long distance up to 170 km from the source (Figure 3.8).

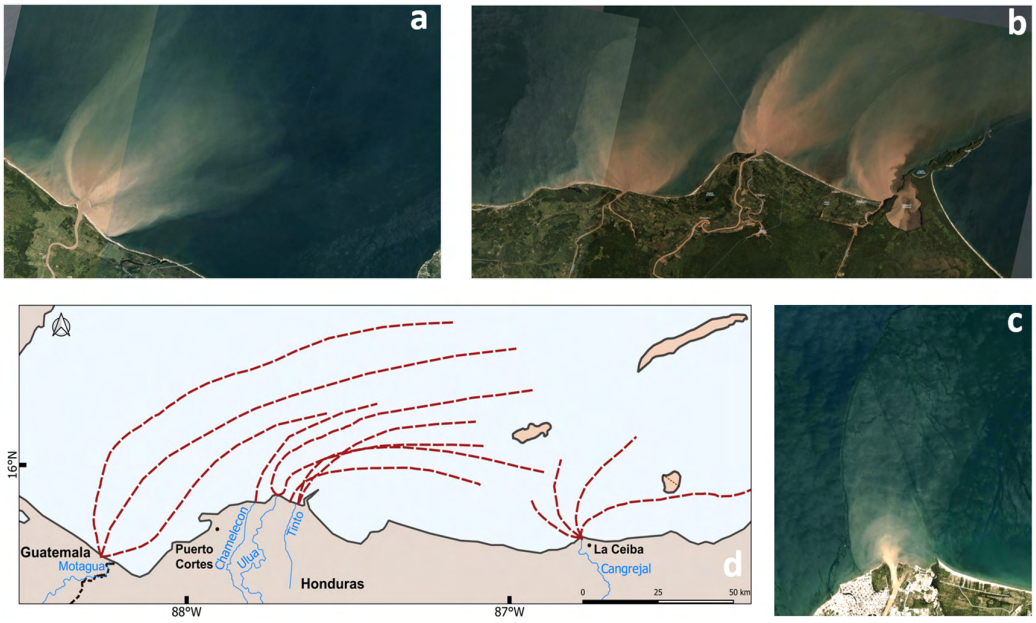


*Figure 3.8 Plastic debris source, corresponding river discharges and estimated debris trajectories based on multitemporal satellite detections in the Bay Islands and Gulf of Honduras. (**a**) Motagua River discharges as recorded during the 27th of September 2017. (**b**) Detected discharges of Chamelecon, Ulua and Tinto rivers on the 27th of September 2017. (**c**) Detected Cangrejal River discharge on the 20th of September 2017. (**d**) Estimated litter travelling trajectories (red dashed lines) based on the multitemporal detections. Plastic pathways followed, in general, a SW–NE direction similar to currents direction acquired from Copernicus Marine Environment Monitoring Service (CMEMS).*

3.3.4 Comparison with state-of-the-art models

Last, to further validate the satellite-derived outcomes, we compared our results with state-of-the-art global plastic debris modeling procedures (Lebreton et al., 2017). Regarding the estimated weight and maximum plastics input per season, our results differ slightly from model predictions. In particular, the global model predictions, for most of the rivers within our study area, suggest that the season with the maximum input of plastics is from October to December. However, evidence from satellite observations shows that the maximum plastic debris season is longer (i.e., from September to January). Regarding weight estimations, model predictions seem to underestimate river plastic inputs during the period of January to March, when river plastic inputs detected from remote-sensing were important. However, in order to obtain more accurate comparison results, model outputs need to be produced during the same period of time. Our weight estimations may be higher than previously reported in other regions (e.g., the Mediterranean Sea) (Suaria et al., 2016); however, plastic patches observed here were remarkably dense. Due to the heterogeneity of polymers that marine plastic debris is composed of, additional in situ and laboratory measurements would increase the accuracy of weight estimation. To this end, future studies should incorporate a more multidisciplinary approach to reach an integrated *Marine Debris* monitoring system (Maximenko et al., 2019), where model outputs could be further improved by synergistically combined satellite-derived observations and in situ collections.

3.4 Discussion

Plastics are rapidly increasing in both coastal and oceanic environments, highlighting the urgent need for continuous tracking and monitoring, as well as for implementing prevention measures. Here, we examined the effectiveness of satellite sensors in monitoring plastic debris, with the Bay Islands province as a case study. Over this region, large dense masses of floating plastics frequently occurred, suggesting that it is an ideal region to investigate plastic debris dynamics. In situ collection revealed several types of plastics, including bags, bottles, plates, caps, forks, spoons and food wrappers.

Additionally, the frequently observed floating macroalgae *Sargassum* slicks in the region, originating from the east (Wang and Hu, 2017), shows that this province is ideal for satellite discrimination of plastic debris from floating macroalgae species. Our results indicated that satellite sensors are capable of detecting plastic debris, as we managed to identify all of the reported events. Moreover, our findings are contributing to recent open issues and critical questions (Martínez-Vicente et al., 2019) regarding plastic debris source, transport, distribution and natural processes that influence them. In particular, based on in situ collection (vessel expeditions and citizen science reports) and remote sensing, we attempted to respond to several of

the suggested questions. Each identified event was described in detail, regarding the plastic debris source (i.e., specific rivers) and amount (i.e., detected floating debris size and estimated weight). Using temporarily dense satellite observations, we also estimated (for certain events) plastic debris velocities and trajectories, leading to an overall quantitative and qualitative description of plastic debris characteristics in the Bay Islands of the Caribbean Sea.

Driven from numerous satellite observations during the reported events (2014–2019), plastic debris covered a total area of 1.6 km2 with a total estimated weight of 8000 tonnes (Table 3.3). Our findings are in accordance with the local stakeholders, which indicate that floating plastic debris has been increasing from 2016 onwards. In particular, currently, shores have to be cleaned up from local authorities more frequently than in 2016 and the years before i.e., from 1–2 times per year, steadily to minimum 3–4 times, depending on the year. Moreover, our results indicated that plastic debris was distributed all over the region, highlighting that there is an urgent need for local authorities to take prevention measures at the national level. Additional in situ data need to be collected, also within the water column, in order to examine the biological and chemical mechanisms that control plastics degradation, movement and accumulation in the study area.

Furthermore, based on verified multiple observations and cross-validation between the different satellite sensors and in situ data, the main source of the plastic debris was found to be the river discharges from the corresponding catchment basins of Honduras and Guatemala. Plastics enter the Caribbean Sea through the river estuaries, travelling over long distances up to 200 km from the source, due to dynamic surface circulation. Plastic debris was systematically recorded after major rainfall events between the months of August and March, which includes the main rainfall season (i.e., October to February) (Schmitt and Gischler, 2016). Large rivers (e.g., Motagua, Ulua, and Aguan) were the major plastic pathways towards Bay Islands, confirming recent studies reporting on the significant impact of river discharges on marine plastic pollution (Kataoka et al., 2019; Simon-Sánchez et al., 2019). Apart from these major contributors, smaller rivers that pass through cities (e.g., La Ceiba City—Cangrejal River) and plastic waste sourcing from major harbors (e.g., Puerto Cortes harbor) were also detected to contribute to plastic pollution in the study area.

Based on multitemporal satellite, meteorological (e.g., precipitation) and oceanographic datasets (e.g., currents), plastic floating debris was tracked and monitored, allowing us to record and estimate its travel distances and trajectories. The detected and verified debris trajectories indicated that sea surface currents significantly affect plastic debris transportation and distribution in the Bay Islands region. This is in accordance with the literature (Law et al., 2010) since floating debris has been proven to be strongly affected by ocean circulation. Floating debris was detected occasionally at fronts, which is also in accordance with the literature (Pichel

et al., 2012). Plastic debris followed, in general, an eastward direction. Regarding the main oceanographic features, in most cases, the Honduras gyre (Carrillo et al., 2015; Chérubin et al., 2008) was observed (~100 km) north of our study area. In cases that the Honduras gyre was located closer to our study area (Figure 3.6), we did not observe any relevant floating debris patterns, indicating the gyre as a transporting driver. However, further investigation is needed to explore how smaller features (such as the cyclonic eddy between Cayos Cochinos and Roatan islands) possibly contribute to plastic debris transportation in Bay Islands area.

Regarding oceanic conditions that were recorded during reported debris events over the Bay Islands, we found that the main current direction was SW–NE and the mean current velocity was 0.11 m/s. Easterly winds prevailed, and the mean wind speed was found to be 4.4 m/s. The multitemporal tracking of plastic debris indicated a mean debris velocity of 0.07 m/s. Our findings about the mean current velocity and wind conditions are consistent with previous studies (Mehrtens et al., 2001; Potemra, 2012). According to the literature, the travelling speed of the detected *Marine Debris* varies from 0.05 m/s when it follows the geostrophic speed of currents, like in North Pacific (Lebreton et al., 2012), to 0.35 m/s after tsunami (Matthews et al., 2017), and 0.5 m/s when debris velocity is constrained by boundary currents like the Gulf Stream.

Satellite-derived monitoring of plastic debris certainly has acknowledged weaknesses, and for this reason, we managed to track its distribution in 13 out of 20 recorded events. Continuous tracking of debris from the source (i.e., river mouth) towards the open ocean was not always possible. This fact is attributed to the lack of daily satellite observations, the relative high cloud coverage (cases with >25% coverage is common during rainfalls), and the spatial resolution of open-access data (i.e., S2 and L8), which cannot capture debris spread, especially during stormy conditions (indeed, in most of the cases that debris tracking failed, the wind speed was greater than 5 m/s and currents velocity was ≥0.1 m/s, while rainfall ranged from 49 mm to 292 mm). Comprehensive *Marine Debris* monitoring is certainly interdisciplinary by nature. Satellite-derived approaches allow the detection and tracking of floating debris at the surface, rather than providing information on the submerged litter (Maximenko et al., 2019). More field campaigns accompanied with auxiliary observations have to be carried out in order to examine in more detail how the precipitation amount, wind speed and vertical mixing influence plastics abundance and trajectories on sea surface (Goldstein et al., 2013). For instance, it has been suggested that future efforts should involve a thorough investigation of physical processes that affect plastic debris transport across multiple scales (e.g., Ekman, waves, Langmuir circulation), to successfully report their vertical and horizontal distribution in the water column (Sebille et al., 2020).

The identified gradual increase in plastic marine pollution must be further studied in relation also with the rapid urbanization and Municipal Solid Waste

Management (MSWM) in the Caribbean Sea. In particular, Honduras' urban population has been increased from 22.7% to 55.3% during the period of 1960–2016, while in Guatemala, the corresponding percentage increased from 31.1% to 52%. For both countries, MSWM is very challenging (Aparcana, 2017), as open-air waste burning, uncontrolled open-air dumps and disposal in water bodies are common waste activities (Hettiarachchi et al., 2018). Lack of recycling and insufficient waste treatment technology is also an essential issue leading to high plastic amounts in landfills and dumps (Hettiarachchi et al., 2018). According to an earlier study (Lu, 2013), there is no policy for waste management in most urban areas in Guatemala, while in Honduras, only urban areas have access to municipal waste collection services. It is worth mentioning that in the largest cities of Honduras, only 28% of the domestic waste is collected and only 3.7% of the total waste ends up in a sanitary or controlled landfill.

3.5 Conclusions

The high-resolution multispectral satellite data can contribute significantly to the efficient monitoring of marine litter dynamics, the detection of its origin and the dominant marine plastic pathways. Ocean circulation plays an important role in *Marine Debris* transportation, as in most examined cases, surface currents lead to the dispersion of plastic debris in the whole region. Satellite and in situ observations should be further analyzed synergistically to monitor the plastic debris dynamics in major accumulation hotspots such as the North Atlantic gyre, which is the most essential "attractor" in the surrounding region, as well as the possibility of the Honduras gyre being the local accumulation center of debris. Limitations regarding mainly the continuous tracking of plastic debris (e.g., on a daily basis) highlight the need for automated machine learning algorithms capable of plastic pollution detection using any additional optical or radar earth observation data. The successive continuation of plastic debris detection efforts should involve an automated approach (such as machine learning) encompassing both high-resolution satellite and in situ observations. New technology and specific satellite sensors capable of direct tracking of plastic debris could alter our understanding of marine plastic dynamics remarkably. A new waste management plan must be set up for Guatemala and Honduras, including both citizens and industries. Recycling and integrated waste management systems, should be implemented everywhere within a country (including smaller towns and villages in the mainland) and not only in the coastal large cities. A well-designed management strategy is required to protect marine life and public health at a global scale. Except for national waste-management laws and local stakeholders' involvement, management strategies should facilitate a collaboration between *Marine Debris* networks. Eventually, international conventions

can play a key role in reducing the produced plastic waste quantity and plastic hazards through their lifecycle.

4 A Novel Marine Litter Detection Benchmark

In this Chapter, we design, develop and evaluate a novel Marine litter detection benchmark. In particular, section 4.1 describes the benchmark MARIDA dataset properties, collected data (i.e., satellite images and in stu reports) characteristics and the annotation process. Next, a thorough statistical and spectral analysis of MARIDA dataset is presented in section 4.2. Section 4.3 presents the experimental design and the developed Machine Learning baselines are evaluated qualitatively and quantitatively. Section 4.4 details the classification outputs and identifies the future challenges of MARIDA. In section 4.5, we provide further Machine Learning experiments by exploiting both spatial and spectral information and gain significant insights regarding the challenging task of *Marine Debris* detection and discrimination from other competing classes. Conclusions from the whole chapter are provided in section 4.6.

4.1 MARIDA: A novel Marine Litter Dataset

Here, we introduce MARIDA, an open-source dataset consisting of annotated georeferenced polygons/pixels on S2 satellite imagery. MARIDA was designed to be temporally and geographically well-distributed; thus, we used open-access data from the S2 satellite sensor which coverage includes global coastal waters. S2 is capable of detecting and continuous monitoring large floating debris, as it provides multispectral data at a spatial resolution of 10 m and 20 m with a frequent revisit time of 2–5 days.

Regarding *Marine Debris* ground-truth data, reported events were collected from citizen scientists and social media over coastal areas and river mouths. After identifying these cases in S2 satellite data, the events were verified with very high-resolution satellite data (whenever possible due to availability), and the corresponding *Marine Debris* pixels were annotated. Additionally, sea surface features that co-occurred on satellite images were annotated: *Ships*, *Sargassum macroalgae*, *Foam*, *Waves* and *Natural Organic Material* (i.e., vegetation and woody), water types (i.e., *Clear*, *Turbid Water* and *Sediment-Laden Water*), *Shallow Coastal Waters* including benthic habitats, *Clouds* and *Cloud Shadows*. Regarding the annotation procedure, three image-interpretation experts annotated the satellite images by assessing the spectral and spatial patterns of all features, considering the limitations of the S2 sensor (i.e., different band resolutions and limited signal-to-noise ratio) (Hu, 2021). Finally, an inter-annotator agreement protocol was established to merge the annotated data and aggregate the confidence levels derived from the three experts (see the Annotation process and protocol section).

The current benchmark dataset aims to support real-world scientific issues that could eventually not only facilitate research efforts in *Marine Debris*, but also offer

operational monitoring solutions. Thus, MARIDA consists of realistic, non-iconic and non-ideal (e.g., with term ideal, we refer to cloud-free data during calm sea state conditions) satellite observations. MARIDA's annotations are also sparse to reduce the potentially noisy labels due to the complexity of sea surface features. The annotated polygons with real cases on S2 images (10 m resolution) do not correspond to thematic class endmembers or pure/clear pixels (in some cases, we annotated sparse *Marine Debris* pixels or floating materials pixels under very thin clouds).

For constructing MARIDA, a specific process was designed and followed, including three major steps (Figure 4.1): i) collection of reports (ground-truth data and literature) regarding floating *Marine Debris* events in coastal areas, ii) satellite data acquisition and processing, auxiliary weather data collection, spectral indices calculation, image interpretation and annotation, statistical analysis, and iii) MARIDA dataset generation and ML benchmarking.

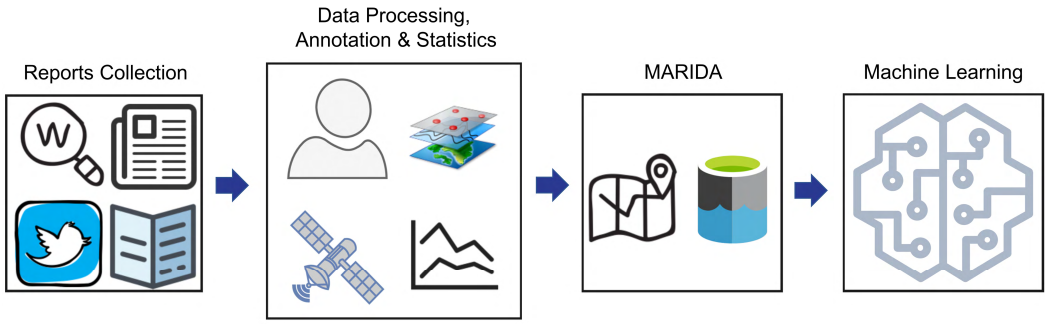


Figure 4.1 Schematic diagram representing the different steps for the construction of Marine Debris Archive-MARIDA.

4.1.1 *Marine Debris* Ground Truth and Reports Collection

For a seven-year period (2015–2021), we gathered reports on marine litter and plastic pollution across coastal areas and river mouths in several countries (Table 4.1). The reports included observations gathered by photographers and citizen scientists, and information extracted from media, social media, and ocean clean-up activities. The URLs of the reports used are included in the Table 4.2.

In addition to ground-truth data collection, the MARIDA dataset also included published satellite-derived data on *Marine Debris* detection (Biermann et al., 2020; Kikaki et al., 2020), and observations from rivers that have been reported in the literature as major polluters (Cordova and Nurhati, 2019; Jang et al., 2014; Lebreton et al., 2017; Schmidt et al., 2017; van Emmerik et al., 2020; Zhao et al., 2019). Table 4.1 demonstrates the source of the reported data (i.e., ground-truth and indicated by literature), as well as the corresponding date and location, when available. For each area, corresponding S2 tiles are also included (Table 4.1).

Table 4.1: Collected Marine Debris reports across different countries and continents for the period 2015-2021. The table shows the regions along with the reported events information (source, date and exact location).

Continent/ Country	S2 Tile	Source	Date	Location (WGS'84)
C. America/ Guatemala	16PCC	Citizen Scientist	18/9/20	15.836206° N, 88.022087° W
C. America/ Guatemala	16PCC	Photographer	16/6/18	15.827222° N, 88.047500° W
C. America/ Guatemala	16PCC	Kikaki et al. (2020)	4/9/19	14.9827° N, 89.5442° W
C. America/ Honduras	16PDC	Citizen Scientist	18/9/20	16.1490° N, 87.6282° W
C. America/ Honduras	16PEC, 16QED	Citizen Scientist	23/9/20	16.042194° N, 86.432081° W
C. America/ Honduras	16PEC	Kikaki et al. (2020)	29/11/15	16.0667° N, 86.3965° W
N. America/ S. Domingo	19QDA	Media	13/7/18	18.467723° N, 69.886808° W
N. America/ Haiti	18QWF/ QYF/ QYG	4ocean Clean-Ups	20/3/20	-
N. America/ Haiti	18QWF/ QYF/ QYG	4ocean Clean-Ups	5/1/21	-
N. America/ Haiti	18QWF/ QYF/ QYG	4ocean Clean-Ups	9/12/20	-
N. America/ Haiti	18QWF/ QYF/ QYG	4ocean Clean-Ups	15/12/20	-
Asia/ Indonesia	50LLR	Social Media	4/3/18	8.715828° S, 115.446799° E
Asia/ Vietnam	48PZC	Social Media	23/11/19	15.994762° N, 108.27417° E
Asia/ Philippines	51PTS	Social Media	18/5/19	-
Asia/ Philippines	51PTS	Social Media	16/7/16	-
Europe/ Scotland	30VWH	Biermann et al. (2020)	20/4/18	-
Africa/ South Africa	36JUN	Biermann et al. (2020)	24/4/19	-
Asia/ South Korea	52SDD	Jang et al. (2014)	-	-
Asia/ Indonesia	48MXU/ MYU	Cordova & Nurhati (2019)	-	-
Asia/ China	51RVQ	Zhao et al. (2019)	-	-

Table 4.2. Source of Marine Debris reports with available links. All links were last accessed on 24 July 2021.

Site	Source	URL Link
Santo Domingo	Media	https://www.youtube.com/watch?v=sFy1Vmm42zQ
Indonesia/Bali	Social Media	https://www.youtube.com/watch?app=desktop&v=31CdhLMV7Es
Vietnam/ Danang	Social Media	https://www.tripadvisor.com/ShowUserReviews-g298085-d456220-r728403711-Non_Nuoc_Beach-Da_Nang.html
Haiti/ La Gonave Gulf	4ocean Clean-Ups	https://www.facebook.com/4oceanBracelets/
Philippines/ Manila Bay	Social Media	https://www.dreamstime.com/manila-philippines-may-ocean-plastic-pollution-manila-bay-shore-manila-philippines-may-ocean-plastic-pollution-manila-bay-image149765103
Philippines/ Manila Bay	Social Media	https://ichkaufnix.com/2016/07/16/der-plastikstrand-oder-wie-wir-uns-selbst-vergiften/

4.1.2 Satellite Data Processing and Annotation

Satellite Data

Based on the ground-truth events, the corresponding S2 level1C images were acquired for the exact reported dates and locations using a mean time window of 10 days. Additionally, for the regions that are significantly affected by plastic pollution (such as river discharges), the seasonality and the periods of maximum plastic presence were examined. We also extended our research for the entire 2015 to 2021 period, focusing on the major recorded rainfalls (https://power.larc.nasa.gov/data-access-viewer/).

At an early stage for selecting images with potential *Marine Debris*, we visually inspected S2 Red-Green-Blue (RGB) composites along with very high-resolution Planet (https://www.planet.com) and Google Earth imagery, when available. The S2 data in which the visual inspection indicated *Marine Debris* occurrence were further processed. Rayleigh reflectance values were extracted at 10 m resolution for 11 bands using ACOLITE atmospheric processor, excluding Vapour (Band 9) and Cirrus (Band 10). To improve the accuracy of the following annotation step, FDI (Biermann et al., 2020) and FAI (Hu, 2009) spectral indices were calculated.

Annotation Protocol

During the annotation step, three image-interpretation experts had access to the gathered data, including reports, S2, Planet satellite imagery, and computed spectral indices. The annotators digitized *Marine Debris* based on ground-truth events, considering S2 sensor limitations, and employing domain knowledge about its spectral behaviour (Biermann et al., 2020; Hu, 2021; Qi et al., 2020; Topouzelis et al., 2019) and its accumulation patterns (i.e., fronts, marine litter windrows) (Cózar et al., 2021). A laborious and intensive image interpretation and manual assessment of each pixel were performed for all selected images leading to *Marine Debris* annotations at

pixel level. In addition, diverse floating objects, sea state features, water types and clouds were annotated based on image interpretation and established spectral patterns (Dierssen, 2019; Hu et al., 2015; Kanjir et al., 2018; Kubryakov et al., 2021; Liu et al., 2021). Wind data were also utilized (https://power.larc.nasa.gov/data-access-viewer/) to examine the possibility of whitecaps, which may appear similar to plastics in human eye (Dierssen, 2019).

Expert annotators recorded the thematic class and their confidence level for each digitized polygon. In particular, all annotated polygons were labelled with three confidence levels (i.e., #1 for high confidence, #2 for moderate and #3 for low confidence level). After the annotation step, an inter-annotators agreement protocol was established, which is described below:

i. For *Marine Debris*, *Natural Organic Material* and *Sparse Sargassum*, which occasionally can have similar spectral behaviour (Hu, 2021), the intersection per two annotators extracted (i.e., an agreement between at least two annotators regarding the class label). If so, the lowest confidence level that was originally assigned was kept for these cases.
ii. For the other features, the union of the annotated data was calculated. If at least two contradictory annotated classes existed for the same digitized area, the annotation was excluded. For the rest of the cases, where the three experts agreed regarding polygon labeling, the lowest confidence score was kept.

For each annotation, *Marine Debris* report existence was also recorded (i.e., #1 when exact date and locations were identified and matched to the available reports, #2 when patches were identified at a distance of either up to 20km or up to 6 days apart from the reported locations and dates; and #3 for no recorded reports close to the detected debris). Additionally, the cases that debris was detected based on previous studies reporting river discharges, were labelled under category 3.

Regarding our annotation strategy, clouds/thin clouds were determined by their intensity, texture and shape, as well as by their adjacent shadows. Regarding thin clouds, following the same manner as for the other features/ classes, spectral signatures were examined, and confidence levels were assigned. After the inter-agreement protocol, the union of the annotations was estimated, while in cases with different confidence level labelling, the lowest confidence level was kept (we note that #1 is for high confidence level, #2 for moderate and #3 for low confidence level).

In cases that floating materials were below thin clouds or semi-transparent features such as aerosols or haze, the experts proceeded with the annotation of the considered floating feature only when its spectral profile/ shape/ pattern was not significantly distorted (assigning also confidence levels in all cases).

Refining Data

In order to improve the quality of our annotated data, the structure of the recorded high-dimensional observations (i.e., 11 multispectral bands) was visualized and explored. Specifically, to examine the pairwise distances between the high-dimensional annotated pixels, we utilized t-distributed Stochastic Neighborhood Embedding (t-SNE) algorithm proposed by Maaten and Hinton (2008), using Spectral Angle Mapping (SAM) (Garaba and Dierssen, 2018; Kruse et al., 1993) as a distance metric. By representing our data in a 2D space, spectral patterns of thematic classes were mapped and outliers were identified and further explored (revisit the data to determine if they had been erroneously annotated).

The annotation procedure resulted in a vector dataset of the digitized polygons, in a shapefile format. The dataset was converted into a raster structure, which was finally cropped into non-overlapping 256x256 pixel-sized patches. After the cropping, each patch was available for extra visual inspection.

4.2 The MARIDA Benchmark Dataset

MARIDA contains 1381 patches, consisting of 837,357 annotated pixels, based on 63 S2 scenes acquired from 2015 to 2021. MARIDA provides patches (Figure 4.2) with corresponding masks of pixel-wise annotated classes and confidence levels in the format of GeoTiff. For each patch, the assigned multi-labels are given in a JSON file. In addition, MARIDA includes shapefiles data in WGS'84/ UTM projection, with file naming convention following the below scheme: s2_dd-mm-yy_ttt, where s2 denotes the S2 sensor, dd denotes the day, mm the month, yy the year and ttt denotes the S2 tile. Shapefiles data include the class of each annotation, along with the confidence score and the report description. The produced dataset is composed of geodata, covering different sites around the globe (Figure 4.3). The selected study sites are distributed over eleven countries (i.e., Honduras, Guatemala, Haiti, Santo Domingo, Vietnam, South Africa, Scotland, Indonesia, Philippines, South Korea and China).

4.2.1 Thematic class distribution

To demonstrate the descriptive overview of MARIDA, the class and pixel distributions are presented in Tables 4.3 and 4.4 and their spectral and statistical analysis are illustrated in Figures 4.4 and 4.5. More specifically, the 15 different classes of MARIDA are shown in Table 4.3, which includes the class description, the corresponding number of provided image patches, and all acronyms of the annotated classes. Regarding the class distribution, the *MWater* class has been digitized in 870 patches due to its implicit abundance in satellite data and straightforward annotation.

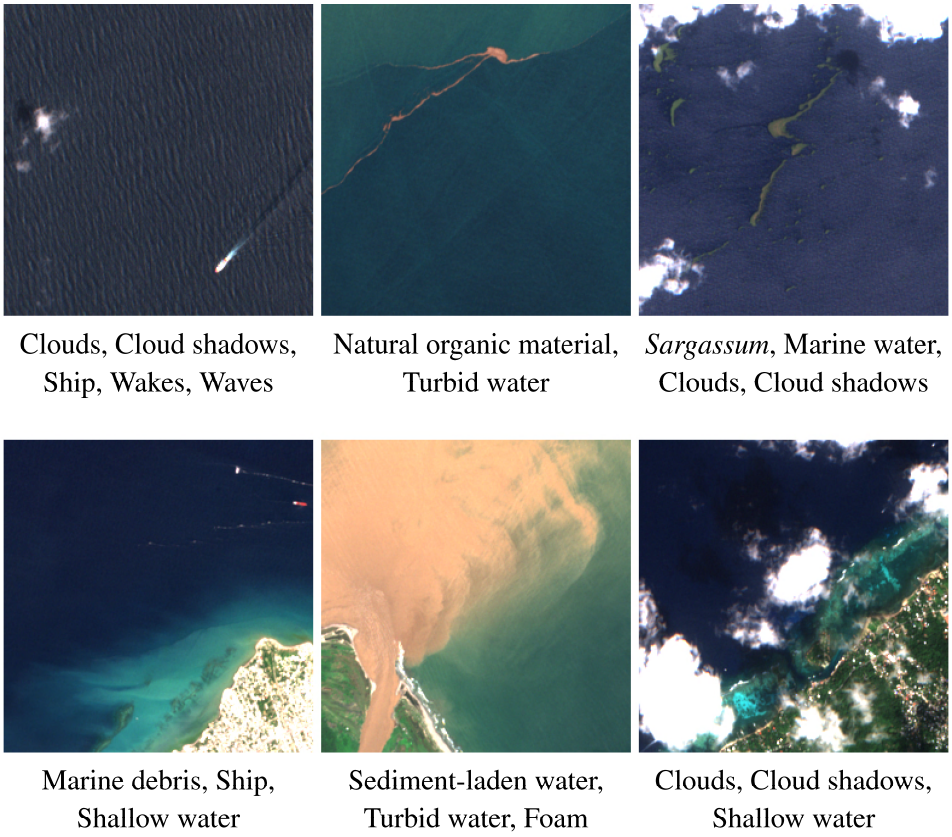


Figure 4.2 Example of S2 patches provided in MARIDA.

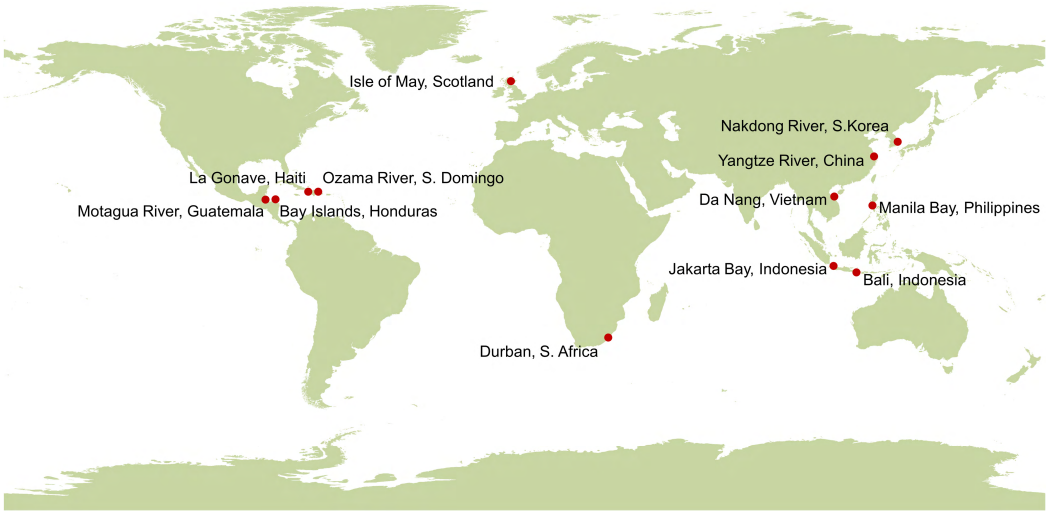


Figure 4.3 The sites (red dots in the map) where Marine Debris events were reported, and corresponding Sentinel-2 satellite images were acquired and processed. Marine Debris and other features that co-existed were annotated in considered satellite data. The corresponding map is acquired from Natural Earth (http://www.naturalearthdata.com/).

As proposed by Hu (2021), we have included additional *MWater* pixels that were close to *Marine Debris* pixels, in order not only to facilitate further experiments with SAM, but also run experiments on pixel windows (3x3 or 5x5) and reflectance difference. The second-highest number of 373 patches were labelled as *Marine Debris*, indicating the high variety of annotations in different patches. *Cloud*, *Ship* and *Turbid Water* were annotated in a sufficient number of patches (~200), as they are plenty in the natural environment and easily identified by annotators.

The rest of the categories were digitized in fewer patches (appr. 50–100). Some of the considered categories, such as *SLWater*, *Sargassum* blooms, *CloudS*, *SWater* were easily digitized with compact, not extended polygons, while *Foam*, *NatM*, *Wakes* and *Waves* required a laborious and intensive manual assessment. Considering that MARIDA is a *Marine Debris*-oriented dataset, we provide only a certain number of indicative cases with the classes mentioned above. The artifact due to the dissimilar S2 band resolutions led to a specific spectral signature primarily recorded on surrounding water pixels of *Marine Debris*, *SpS* and *Ship*. This class was labelled as *MixWater*, as it corresponds to water, and digitized around annotated *Marine Debris* pixels. For more details about patches and class co-occurrence, readers are referred to the online material (https://marine-debris.github.io/).

Table 4.3 The thematic classes of MARIDA. Name, description and corresponding number of patches are presented for each class. All acronyms are stated here.

Class Name	Acronym	Description	Number of Patches
Marine Debris	MD	Floating plastics or other polymers, mixed anthropogenic debris	373
Dense *Sargassum*	DenS	Dense floating *Sargassum* macroalgae	49
Sparse *Sargassum*	SpS	Sparse floating *Sargassum* macroalgae	106
Natural Organic Material	NatM	Vegetation & Wood	71
Ship	Ship	Sailing & Anchored Vessels	182
Clouds	Cloud	Clouds including thin Clouds	181
Marine Water	MWater	Clear Water	870
Sediment-Laden Water	SLWater	High-Sediment river discharges with brown colour	51
Foam	Foam	Foam recorded at river fronts or coastal wave breaking area	59
Turbid Water	TWater	Turbid waters close to coastal areas	220
Shallow Water	SWater	Coastal waters, including coral reefs and submerged vegetation	64
Waves	Waves	Waves	54
Cloud Shadows	CloudS	Cloud Shadows	71
Wakes	Wakes	Wakes & Waves from a sailing vessel	106
Mixed Water	MixWater	Water near floating materials	140
Total			**1381**

Apart from the per-patch analysis, we also discuss the pixel-level distribution of MARIDA classes. Table 4.4 summarizes the class distribution in pixel level for each S2 tile, indicating that MARIDA provides numerous pixels annotated in 17 S2 tiles.

Table 4.4. MARIDA's class distribution at pixel-level. For Sentinel-2 tiles description, the reader is referred to Table 1. For classes acronyms, the reader is referred to Table 4.3.

S2 Tile	MD	DenS	SpS	NatM	Ship	Cloud	MWater	SLWater	Foam	TWater	SWater	Waves	CloudS	Wakes	MixWater	# of pixels	# of S2 scenes
16PCC	1496	2048	574	78	3322	62082	60169	285886	712	99501	3960	3417	3585	5929	191	532950	19
16PDC	143	49	226	78	96	13507	15258	85449	334	24923	2251	0	883	253	75	143525	6
16PEC	129	222	645	193	485	11678	19341	11	86	27080	3782	108	1733	1115	51	66659	6
16QED	0	474	691	0	90	4098	1719	0	0	0	5910	0	1841	221	0	15044	2
18QWF	0	0	0	0	0	0	324	0	0	0	0	1461	0	0	0	1785	1
18QYF	1112	4	200	154	408	7977	1360	0	0	0	1038	0	314	48	58	12673	13
18QYG	90	0	0	7	0	373	222	0	0	831	277	0	106	0	15	1921	1
19QDA	0	0	21	3	11	0	110	0	0	5	40	0	0	0	0	190	1
30VWH	27	0	0	0	36	3505	24393	0	0	0	0	0	1975	0	0	29936	1
36JUN	46	0	0	0	625	3500	600	0	0	300	0	0	0	18	0	5089	1
48MXU	208	0	0	0	71	5807	194	0	0	382	45	0	489	15	12	7223	2
48MYU	24	0	0	0	223	0	291	0	0	10	48	0	0	611	0	1207	2
48PZC	24	0	0	0	298	4108	2079	0	48	4129	0	0	765	171	1	11623	3
50LLR	41	0	0	3	27	402	485	0	41	0	18	841	0	72	5	1935	1
51PTS	38	0	0	20	17	0	35	0	0	0	0	0	0	0	0	110	2
51RVQ	17	0	0	0	0	363	163	0	0	0	0	0	37	0	0	580	1
52SDD	4	0	0	328	94	0	2416	1591	4	451	0	0	0	37	2	4927	1
Total pixels	3399	2797	2357	864	5803	117400	129159	372937	1225	157612	17369	5827	11728	8490	410	837377	63
Perc. %	0,41	0,33	0,28	0,1	0,69	14,02	15,42	44,54	0,15	18,82	2,07	0,70	1,40	1,01	0,05	100	

Overall, most given pixels correspond to Honduras Gulf, a known plastic polluted region where a thorough remote-sensing study has been previously conducted by Kikaki et al. (2020), based on ground-truth data. It should be noted that, although we avoided digitizing extended regions with water or clouds, the produced dataset cannot be balanced at pixel-level due to the implicit different size and characteristics of considered sea features. Indeed, our goal was to create a Marine Debris-oriented dataset. To this end, we provide a significant number of 3339

Marine Debris pixels in total. The 1625 pixels were digitized and annotated with high confidence, based on reports and domain knowledge. Additionally, 1235 pixels were labelled with moderate and 539 pixels with low confidence (Table 4.5). For scenes with large garbage trajectories and high confidence annotations, the readers are referred to 18 September 2020 (tile 16PCC) and 14 March 2020 (tile 18QYF), where ground-truth events were available. An indicative case with dense marine litter patches at Motagua river mouth was also evident on 4 September 2016 (tile 16PCC). For other scenes with high-confidence *Marine Debris* annotated data, the reader can consider the online material (https://marine-debris.github.io/).

Table 4.5. The distribution of the confidence scores in pixel level for the classes of Marine Debris, Natural Organic Material and Sparse Sargassum.

Confidence Level	***Marine Debris***	**Natural Organic Material**	**Sparse *Sargassum***
High	1625 (47.81%)	556 (64.35%)	2052 (87.06%)
Moderate	1235 (36.33%)	201 (23.26%)	290 (13.3%)
Low	539 (15.86%)	107 (12.38%)	15 (0.64%)
Total	**3399**	**864**	**2357**

4.2.2 Spectral signatures & statistical analysis

To study the spectral behavior of *Marine Debris* annotated data, we extracted the mean spectral signatures for each scene, leading to a detailed analysis presented thoroughly in the online material. The mean spectral reflectance of annotated pixels with high confidence in MARIDA is depicted in Figure 4.4. The mean spectral signatures are presented along with 25–75 percentiles as error bars to demonstrate the variation along with the skewness of their distribution. Atmospheric correction process, diverse proportions of floating *Marine Debris* within pixels, differences resulting from colours and immersion, and mixed conditions in the natural environment led to high variability of recorded *Marine Debris* spectral signatures.

However, the recorded *Marine Debris* mean spectral reflectance is very similar with the corresponding simulated signature proposed recently by (Hu, 2021). Slightly higher values in our data indicate different debris proportions within pixels. In comparison with previous studies (Biermann et al., 2020; Kikaki et al., 2020; Topouzelis et al., 2019), which exploited S2 imagery, higher reflectance at Green and Red bands was observed, possibly due to the denser patches that we recorded. Additionally, the mean spectral signature of high-confidence *NatM* was considered for comparison, as in some cases with low subpixel proportions, their spectral discrimination was not straightforward. Regarding *Marine Debris* and *NatM* comparison, it was found that their discrimination might be possible in 865 nm and SWIR bands.

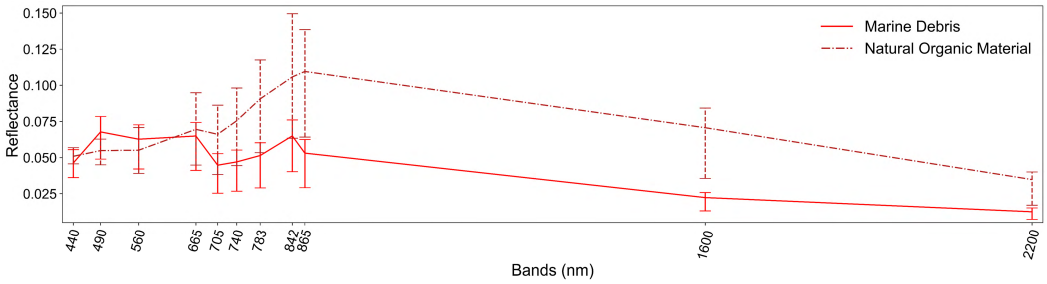


Figure 4.4 The spectral signatures of the Marine Debris and Natural Organic Material classes derived from the annotations with the high confidence levels. The mean spectral signatures are presented with 25-75 percentiles as error bars.

By applying t-SNE algorithm along with spectral signatures analysis described above important insights were gained about spectral behaviour of floating *Marine Debris* and the potential of spectral discrimination from other features with similar patterns such as *SpS*, *Ship*, *Waves* and *NatM*.

Figure 4.5 presents t-SNE results for the considered features, indicating the different confidence level for each annotation with a different symbol. Based on the recorded data, a well-shaped *Marine Debris* cluster was developed, which is discrete from other clusters. Very sparse recorded *Marine Debris* (e.g., 20 April 2018 in Scotland) led to a smaller separate cluster between *Waves* and *Marine Debris*. A well-shaped *Ship* cluster was also mapped, yet some annotated *Ship* pixels were depicted in *Marine Debris* cluster due to the similar polymer types. Respectively, some dense *Marine Debris* pixels were mapped in the *Ship* cluster. Some *Ship* pixels were also depicted close to *Waves* pixels; this is evident in cases with moving vessels, where discrimination of boundary *Ship* pixels from water-related classes (i.e., *Wakes*) was challenging for a human expert.

Occasionally, NatM cannot be spectrally separated from *Marine Debris* (e.g., 18 September 2020 at Motagua river mouth). Mixed conditions at the river mouth, low coverage at pixel-level and potentially colored marine litter (e.g., green or brown) led to uncertainties represented with low confidence *Marine Debris* and *NatM* annotations. However, dense Natural woody debris has a discrete spectral signature (e.g., 7 October 2018 at Nakdong river mouth). This fact was also confirmed by a smaller (but well-shaped) *NatM* cluster depicted in brown color (Figure 4.5). A discrete *SpS* cluster was also formed, including *NatM* (i.e., vegetation). In some cases, the *SpS* annotated pixels have been mapped in the *Marine Debris* and Waves clusters, though, the majority of these cases corresponded to sparse floating materials that were detected at a lower subpixel level. This fact confirms that sparse floating vegetation pixels in some cases cannot be spectrally discriminated from sparse marine litter pixels (e.g., 4 March 2018 in Bali) (Hu, 2021).

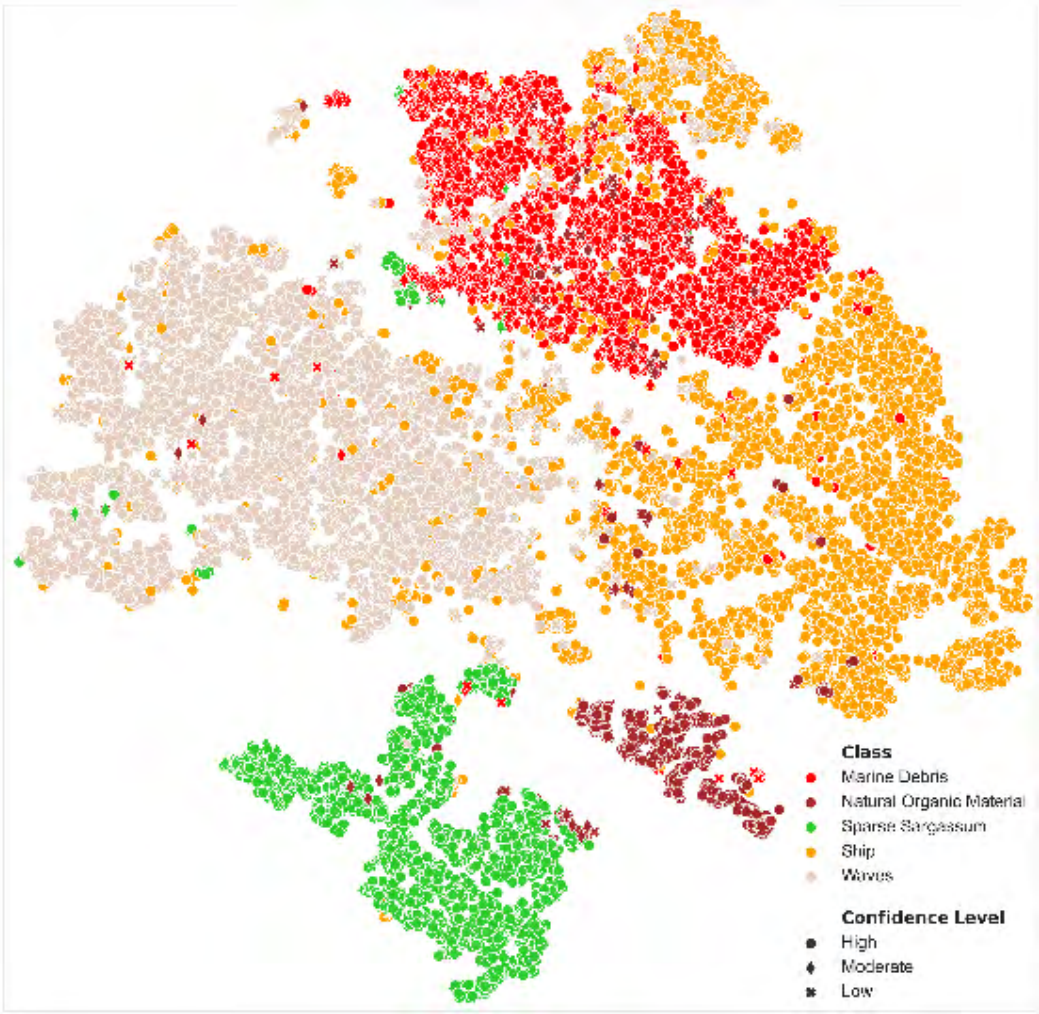


Figure 4.5 A 2D embedding using T-SNE algorithm with SAM metric for the classes: Marine Debris, Ships, Sparse Sargassum, Natural Organic Material and Waves. Each class is represented with a different color. Different symbols demonstrate the confidence level of annotations.

4.3 Detection with Machine Learning

In order to trigger more research efforts towards *Marine Debris* detection methods and solutions, we provide software baselines for *weakly supervised* pixel-level semantic segmentation tasks, by employing a Random Forest model (RF) (Breiman, 2001) and an U-Net architecture (Ronneberger et al., 2015).

Random Forest

In particular, RF is a well-established supervised model, which has been widely used in remote sensing and computer vision community. A RF classifier consists of many decision trees and uses averaging to improve the predictive performance and control over-fitting. For our RF model, we extracted features similar to the first place team of Track 2 of the 2020 IEEE GRSS Data Fusion Contest (Robinson et al., 2021).

We trained three different RF models: i) one based on spectral signatures of each pixel (RF_{SS}), ii) one based on spectral signatures and calculated spectral indices (RF_{SS+SI}), and iii) one with spectral signatures, spectral indices, and extracted Gray-Level Co-occurrence Matrix (GLCM) (Haralick et al., 1973) textural features ($RF_{SS+SI+GLCM}$) in order to incorporate the spatial information. The extracted spectral indices were NDVI, NDWI, FAI, FDI, Shadow Index (SI), Normalized Difference Moisture Index (NDMI), Bare Soil Index (BSI) and NRD (Hu, 2021), which are broadly used in remote sensing studies. To compute the GLCM features, Rayleigh corrected RGB composites were converted to grayscale images which consequently were quantized in 16 bins-level. The selected GLCM features were Contrast (CON), Dissimilarity (DIS), Homogeneity (HOMO), Energy (ENER), Correlation (COR) and Angular Second Moment (Haralick et al., 1973). For those features extraction, a window of size 13 x 13 was used.

U-Net

The U-Net is a well-established deep learning model for semantic segmentation. Its architecture consists of two parts, the down-sampling and the up-sampling part. The first part encodes the input image yielding a low dimensional representation using successive blocks of 3 x 3 convolutions for features extraction and max-pooling layers for down-sampling. The feature maps/ produced channels are doubled in each block, while the spatial dimensions are reduced by half. The second part decodes the internal representation using successive up-convolution layers to create the final segmentation output.

For our task, the first input layer of U-Net was modified to adapt to the 11 Rayleigh reflectance S2 bands, and the final classification layer was changed to output the MARIDA classes. We also used 4 down-sampling and up-sampling blocks, as well as 16 hidden channels produced by the initial down-sampling block.

Evaluation Metrics

To assess pixel-level semantic segmentation performance, we relied on three metrics. Our main evaluation metric was the Jaccard Index or Intersection-over-Union (IoU) (Everingham et al., 2015). In addition, the average for each class F_1 score (Macro-F_1/ mF_1) and the Pixel Accuracy (PA) for the per-class assessment were employed.

To evaluate the performance of RF model and U-Net architecture, we relied on the Intersection-over-Union (IoU):

$$IoU = \frac{TP}{TP + FP + FN}, \tag{1}$$

where TP is the number of true positives, FP the number of false positives and FN the number of false negatives. Intuitively, IoU is equal to the ratio of intersection between predicted and ground truth area to the union between predicted and ground truth area. Considering that we deal with a multi-class task, the main evaluation metric for the overall assessment is the mIoU which is the average IoU over all classes c:

$$mIoU = \frac{1}{n}\sum_{c}^{n} IoU_c \tag{2}$$

The second metric is the average for each class F_1 score (Macro-F_1/ mF_1), which is the harmonic mean between precision $P = TP/(TP + FP)$ and recall $R = TP/(TP + FN)$:

$$mF_1 = \frac{1}{n}\sum_{c}^{n} F_{1_c} = \frac{1}{n}\sum_{c}^{n} \frac{2P_cR_c}{P_c + R_c} \tag{3}$$

The third metric is the Pixel Accuracy (PA) for the per-class assessment and the corresponding mean PA (mPA) for the overall assessment. For each class, PA is the ratio of the correctly predicted pixels to the total number of pixels. For the multi-class pixel-level classification, this metric is equivalent to recall R.

MARIDA is designed to be beneficial for several remote sensing applications and tasks which are described in detail in the following section (Discussion). However, it primarily aims to benchmark weakly supervised pixel-level semantic segmentation learning methods. In particular, the produced dataset falls into incomplete-supervision due to sparsely annotated data, inexact-supervision due to sensor limitations (i.e., 10 m resolution, different bands resolution), and inaccurate supervision derived from potential slightly noisy annotations (i.e., sensor noise, human error).

Experimental Design

MARIDA was split into train, validation and test disjoint sets. The data were not split randomly; instead, each data split was produced as a representative subset of the whole dataset. For instance, the dataset was divided into subsets which were ensured to have balanced class distribution (Table 4.6). It should be noted that the data of each scene/unique date were retained in the same set. The split was selected

to be ~50/25/25%. More specifically, the split contains 694 training (429,412 px), 328 validation (213,102 px) and 359 test (194,843 px) patches.

Table 4.6. Class distribution (%) for each split in MARIDA. All acronyms are stated in Table 4.3.

Class	Train	Validation	Test	Total
MD	27.38	25	28.13	27.01
DenS	3.46	3.66	3.62	3.55
SpS	8.07	7.93	6.69	7.68
NatM	5.04	6.1	4.46	5.14
Ship	12.82	12.8	14.21	13.18
Cloud	13.4	13.11	12.53	13.11
MWater	62.25	64.63	62.95	63
SLWater	3.46	4.27	3.62	3.69
Foam	4.32	4.27	4.18	4.27
TWater	16.57	15.24	15.32	15.93
SWater	5.48	3.35	4.18	4.63
Waves	3.31	5.49	3.62	3.91
CloudS	5.62	5.79	3.62	5.14
Wakes	7.2	8.23	8.08	7.68
MixWater	9.94	10.67	10.03	10.14

Due to the moderate size of MARIDA and aiming at a *Marine Debris*-oriented dataset, the initial 15 classes were aggregated to 11 classes. The categories of *Wakes*, *CloudS*, *Waves* and *MixWater* were grouped with *MWater* and formed a water super-class, as they semantically belong to the same class as well as present similar spectral profiles (see online material).

Regarding RF training, all models (RF_{SS}, RF_{SS+SI}, $RF_{SS+SI+GLCM}$) were composed of 125 trees, each with a maximum depth of 20 nodes. Due to pixel-level class distribution, which is by nature imbalanced (e.g., *Marine Water* px contrary to *Marine Debris* px), we used class weighting inversely proportional to class frequencies in the training set. Additionally, the annotators' confidence score was utilized such that low confidence samples contribute less to the training process. Specifically, the weights for high, moderate and low confidence samples were 1, 2/3 and 1/3, respectively. The final selection of RF hyperparameters described above was based on grid search in the validation set.

During the U-Net training process, the Adam algorithm was employed to minimize the Cross-Entropy loss with an initial learning rate of $2x10^{-4}$. Moreover, we utilized early stopping based on the loss of the validation set and trained for 44 epochs. After the 40^{th} epoch, the learning rate was reduced to $2x10^{-5}$. The selected batch size was 5 samples. We also employed random rotations of the input images by -90°, 0°, 90°, or 180° and horizontal flips in order to augment the dataset. The selection of the hyperparameters above and training set-up was based on grid search in the validation set. It should be noted that the U-Net model was trained from

scratch. To address the unbalanced data limitation, the proposed weighting scheme was used on the Cross-Entropy loss for each class (Paszke et al., 2016). Finally, it should be mentioned that in our U-Net baseline, in contrast to RF, we did not experiment with the annotators' confidence levels.

4.3.1 Quantitative & Qualitative evaluation

This subsection describes the quantitative and qualitative assessment of our ML baseline outcomes in MARIDA. To evaluate our results quantitatively, we demonstrate the scores for all metrics per class on the test set (Table 4.7). Overall, our results indicate that RFSS+SI+GLCM leads to the highest average scores for all metrics, followed by RFSS+SI and RFSS, which provide almost equivalent average scores. Regarding scores per class, for *SWater*, U-Net provides the highest scores, while for *Ship*, *Clouds*, *MWater* and *Foam*, RFSS+SI+GLCM performs best. For *DenS*, RFSS+SI leads to the highest scores, as for *SpS*, RFSS+SI+GLCM leads to higher scores for IoU and F1. For *TWater*, both RFSS+SI+GLCM and U-Net achieve similarly high scores. It is noteworthy to highlight that for *SLWater*, all RF models and U-Net achieve for all metrics the highest scores (i.e., 1).

Regarding *Marine Debris*, RF_{SS+SI} performs the highest scores, while adding spatial information does not improve the classification performance results (i.e., IoU and F_1 decreased slightly). Future experiments with different window sizes for the extraction of GLCM textural features may lead to higher scores. We have to note that, for the *NatM* class, all models lead to low scores. *NatM* presents similar spectral behavior to *Marine Debris*, while both follow the same spatial patterns (e.g., linear trajectories). In this case, adding spectral indices or textural information leads to lower scores than the initial. Especially, U-Net predicts only few annotated *NatM* pixels on the test set.

Table 4.7. Evaluation scores obtained by RF_{SS}, RF_{SS+SI}, $RF_{SS+SI+GLCM}$ and U-Net for each class on Marine Debris Archive. The highest scores are highlighted. All acronyms are stated in Table 4.3.

	RF_{SS}			**RF_{SS+SI}**			**$RF_{SS+SI+GLCM}$**			**U-Net**		
Class	**IoU**	**PA**	**F_1**	**IoU**	**PA**	**F_1**	**IoU**	**PA**	**F_1**	**IoU**	**PA**	**F_1**
MD	0.55	0.91	0.71	**0.67**	**0.92**	**0.8**	0.65	**0.92**	0.79	0.33	0.7	0.5
DenS	0.87	0.92	**0.93**	**0.88**	**0.93**	**0.93**	0.87	**0.93**	**0.93**	0.6	0.6	0.75
SpS	0.53	0.91	0.69	0.69	**0.92**	0.82	**0.83**	0.9	**0.91**	0.66	0.89	0.79
NatM	**0.31**	**0.47**	**0.47**	0.17	0.27	0.29	0.18	0.31	0.31	0.02	0.02	0.04
Ship	0.54	0.72	0.7	0.47	0.7	0.64	**0.67**	**0.82**	**0.8**	0.62	0.76	0.76
Clouds	0.75	0.85	0.86	0.74	0.82	0.85	**0.84**	**0.86**	**0.91**	0.62	0.62	0.76
MWater	0.66	0.82	0.79	0.65	0.83	0.79	**0.75**	**0.93**	**0.86**	0.61	0.88	0.76
SLWater	**1**	**1**	**1**	0.99	**1**	**1**	0.99	**1**	**1**	0.99	0.99	**1**
Foam	0.23	0.29	0.37	0.31	0.48	0.47	**0.6**	**0.74**	**0.75**	0.55	0.55	0.71
TWater	0.74	0.78	0.85	0.8	0.83	0.89	**0.88**	0.92	**0.94**	0.84	**0.95**	0.91
SWater	0.08	0.25	0.16	0.13	0.33	0.23	0.3	0.37	0.46	**0.45**	**0.67**	**0.62**
Average	0.57	0.72	0.69	0.59	0.73	0.7	**0.69**	**0.79**	**0.79**	0.57	0.69	0.69

Except for the quantitative evaluation described above, a qualitative (visual) assessment of our baseline results on the test set was also performed (Figure 4.6). As it is easily noticed, the two models, RFSS+SI+GLCM and U-Net, provide similar results. Nevertheless, U-Net seems more robust to S2 noise and single pixels with sharp spectral differences than RF. U-Net is capable of modeling the shapes and spatial patterns of sea features, and appeared to be no sensitive in isolated pixels/ spikes, potentially due to the inherent multiple-scale information (successive convolutional layers). On the other hand, RFSS+SI+GLCM is more prone to S2 noise and mixed bands resolutions artifact. In particular, in RFSS+SI+GLCM results, some pixels around *Marine Debris* and *SpS* are classified as *Cloud* (Figure 4.6B,D).

In both models, small vessels are classified as *Marine Debris* (Figure 4.6C), which is expected due to similar polymer types that are composed and possibly similar floating material proportion within pixel. Regarding *Cloud*, RFSS+SI+GLCM predicts more accurately the considered class than U-Net (Figure 4.6C,D), while U-Net predicts better the *SWater* habitats (Figure 4.6C). The latter fact can be also seen in the highest scores in all U-Net metrics (Table 4.7). In the coastal zone, both models lead to similar results. However, in U-Net classification images, some *Foam* pixels are predicted as *Marine Debris*, while in RF results, some *TWater* pixels are classified as *MWater* (Figure 4.6A).

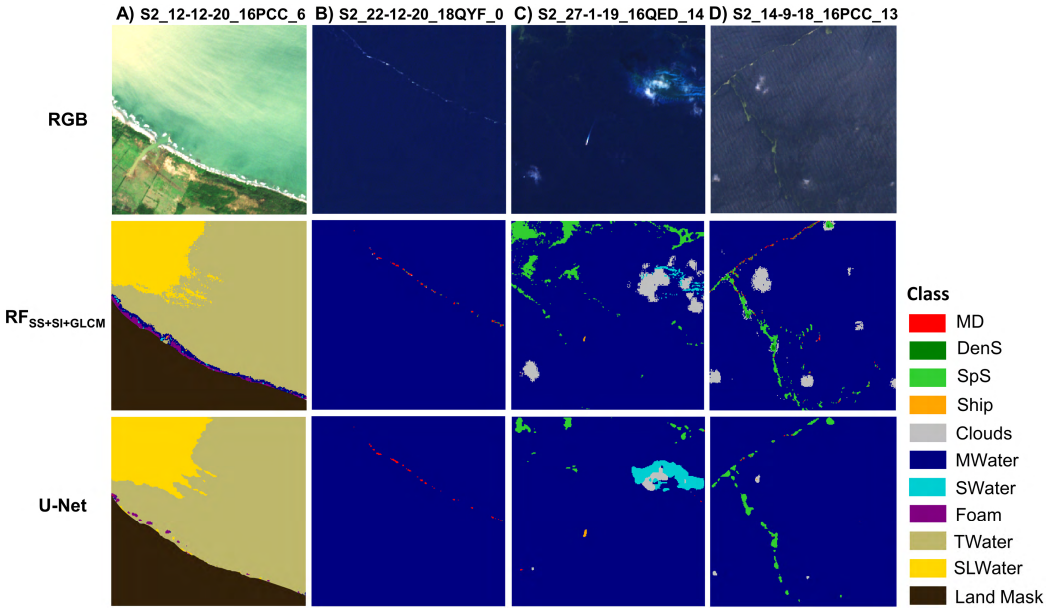


Figure 4.6 Classification results extracted by the baseline RFSS+SI+GLCM and U-Net models. Selected indicative cases demonstrate (A) S2_12-12-20_16PCC_6, (B) S2_22-12-20_18QYF_0, (C) S2_27-1-19_16QED_14 and (D) S2_14-9-18_16PCC_13 patches on test set. RGB patches are derived from Sentinel-2 data which were freely downloaded from https://earthexplorer.usgs.gov/. All acronyms are stated in Table 4.3.

By assessing our baseline experiments quantitatively and qualitatively, we observe that there is a consistency between metric scores and classification outputs in

general. Yet, in some cases, the classification is still challenging. For instance, although both models achieve high scores (in comparison with other classes) for *SpS* (Table 4.7), in some cases with very sparse conditions, *SpS* pixels are classified as *Marine Debris* (Figure 4.6D).

4.4 Discussion

In this Chapter, a new dataset (MARIDA) is introduced towards triggering the research community at improving and developing new methods for detecting *Marine Debris* and discriminating from other sea surface features that co-exist. Based on the collected ground-truth, literature review and intensive image interpretation, MARIDA provides 3399 *Marine Debris* pixels, labelled in different S2 tiles across various countries, different seasons, years and sea state conditions. Thus, MARIDA is an important geodata source for evaluating existing detection methods and developing new techniques based on available S2 data.

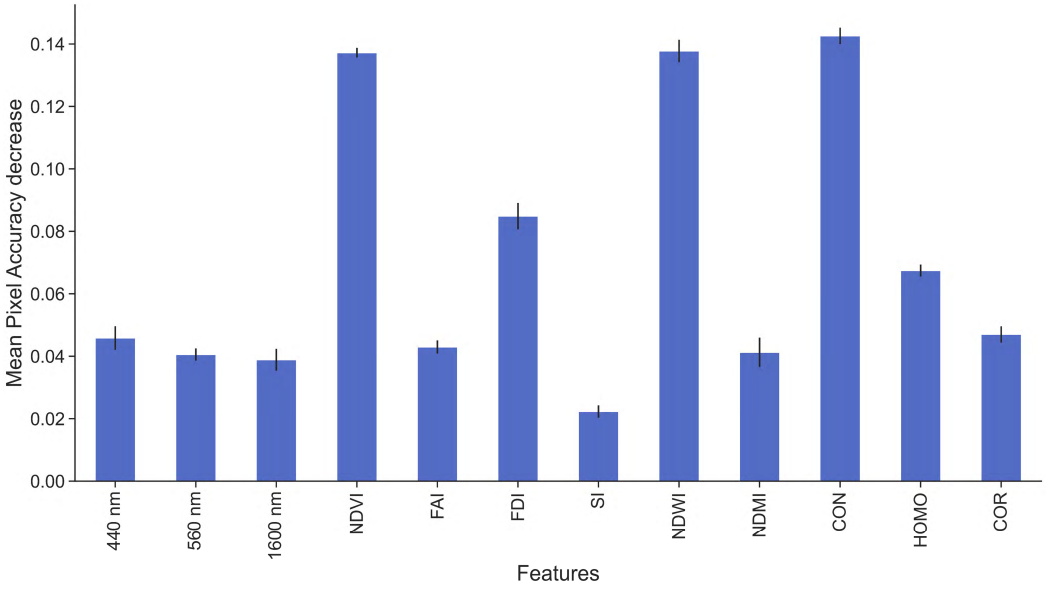


Figure 4.7 In order to inspect our developed model with the best performance, and subsequently understand the importance of each input feature, we applied the permutation feature importance. Specifically, for each feature, we performed random shuffling between different samples on test set to obtain the amount of decrease in the model's accuracy. For each feature, we performed this procedure 20 repeated times to estimate the mean decrease. However, due to the multicollinearity of the input features, permuting one feature does not affect model's accuracy as long as similar information is still included. Thus, we calculated the Spearman Correlation among input features on train set, in order to form highly correlated groups and keep only a single feature from each group . we selected a cut-off threshold to form the groups with the most correlated features. Followingly, we trained the $RF_{SS+SI+GLCM}$ *from scratch only on these features and obtained almost identical results (mIoU = 0.67, mPA = 0.79,* $mF_1 = 0.78$*), compared to initial approach. This fact reveals that the selected subset of features preserves the same amount of input information. We observe that the largest accuracy decrease occurs by permuting CON, NDWI, NDVI and FDI.*

After training four different models, the results showed that the developed RFSS+SI+GLCM achieved the highest scores for all metrics; yet it seems more prone to S2 noise and different bands resolutions than the deep U-Net architecture. Further experimentation with RFSS+SI+GLCM indicated that the most distinctive feature is the spatial feature CON (i.e, a measure of the intensity difference between a pixel and its neighbour), followed by NDWI, NDVI and FDI (Figure 4.7). This fact is also in line with Tasseron et al. (2021) who recommended that the combination of FDI and NDVI can be efficient in the separation of vegetation and *Marine Debris*.

Low-confidence annotations were also included in our dataset, revealing challenging cases where no ground-truth events existed, and thus, human-experts attempted to identify the floating materials/ features based on domain knowledge, image interpretation and statistical analysis. Indicative cases include the sparse floating materials detected at fronts (e.g., 1 December 2019 in Jakarta Bay), very turbid conditions (e.g., 12 January 2017 in Honduras), and windrows (29 August 2017 at Yangtze river mouth) where human-experts could not easily define if they were dominated by dense foam or plastic concentrations. In addition, the spectral discrimination of *Marine Debris* from *NatM* was not straightforward in some cases (e.g., 18 September 2020 PCC). This issue was also observed in a previous study by Moshtaghi et al. (2021), demonstrating that the considered floating materials (e.g., brown *Marine Debris* and woody debris) can have similar spectral patterns.

Regarding MARIDA limitations, it should be noted that the dataset is not optimally balanced geographically due to the lack of open-access in situ data reporting marine litter cases worldwide. MARIDA dataset can be augmented in future works with other datasets (e.g., clouds), other recorded features such as macroalgae species (e.g., *Ulva*, *Noctiluca*), jellyfish blooms (Qi et al., 2020) and future collections of additional verified *Marine Debris* events.

Due to S2 spatial resolution, the annotation procedure was occasionally not straightforward. For example, the discrimination between boundary *Ship* pixels and *Wakes* in moving ships was challenging for all experts. Thus, these cases potentially induced slight noise to the dataset. Certain S2 images with erroneous atmospheric corrections, such as the S2 image acquired on 23 September 2020 (Bay Islands, Honduras), were excluded, even though a major *Marine Debris* event was reported in the region during this date. Furthermore, high cloud coverage did not allow marine litter detection in all available S2 images in Santo Domingo, where a significant event was reported (July 2018).

The ACOLITE Dark Spectrum Fitting (DSF) algorithm was selected in this work after the recommendation from several studies (Biermann et al., 2020; Hu, 2021; Kikaki et al., 2020; Topouzelis et al., 2019), reporting that ACOLITE performed well in detecting marine litter. However, ACOLITE performs simple pixel replication and no interpolation (such as bilinear or cubic) or other more sophisticated methods such as pan-sharpening to resample the S2 20 m and 60 m bands to 10 m.

Despite the limitations mentioned above, MARIDA is designed to be a multi-task dataset with various future aspects. Firstly, the RF model used here can be further enhanced by using spatial information at multiple scales (e.g., GLCM features at different windows size). Further feature-engineering and selection of the most distinctive bands, might improve the RF performance as well. Also, the experimentation with the denoising of the prediction masks (as a meta-classifier) can create more accurate classification outputs.

Regarding U-Net, experimentation with different loss functions and different weighting schemes can potentially address the class imbalance. For instance, the Focal Loss (Lin et al., 2017) may help the model focus on classes that have not been trained well. Furthermore, the exploitation of annotators' confidence level information should be incorporated into the learning process. Another arising challenge is the combination of the predictions from multiple models (ensemble methods), potentially leading to more promising results. Experimentation with other improved or more sophisticated architectures can also be examined. The integration of advanced pre-processing techniques (i.e., cloud masking, denoising algorithms) should improve *Marine Debris* detection and sea features classification outcomes (Wang and Hu, 2021a).

Other tasks derived by MARIDA that could be further explored are the unsupervised classification methods and/or clustering analysis, for better understanding the spectral patterns of sea features. In addition, the produced dataset can be used to evaluate existing spectral indices such as FDI, FAI and optimal thresholds tuning, as well as the development of new spectral indices. Last but not least, by providing annotated water pixels close to *Marine Debris*, we encourage the readers to further experiment with subtracting nearby water pixels (i.e., reflectance difference), windows-size and x subpixel proportion (Hu, 2021).

4.5 Further Experimentation on MARIDA

4.5.1 Exploiting Spectral & Spatial Information

Kikaki et al. (2022) proposed a set of spectral indices used in MARIDA classification baselines. To further investigate the degree of sea surface feature separation, we used a more extensive set of spectral indices. These indices were chosen to enhance the spectral differences of competing classes.

More specifically, we used the additional spectral indices: i) Enhanced Vegetation Index (EVI) for vegetation and *Sargassum* macroalgae mapping ii) Modified Normalised Water Index (MNDWI) and Automated Water Extraction Index (AWEI) for water-features extraction, and iii) Normalised Difference Snow and Ice Index (NDSII) to highlight bright or other objects.

Highlighting spectral properties via spectral indices is essential for detecting and classifying materials and features on the sea surface. However, in challenging cases, this information alone seems to be insufficient. For this reason, following MARIDA analysis, we propose the combination of spectral information with spatial indices that utilize information about the texture (i.e., GLCM, LBP), as well as edges, corners, and flat image regions (i.e., Gaussian, Sobel, Hessian Eigenvalues). It should be noted that this combination has not been widely investigated in the literature for this challenging task.

Kikaki et al. (2022) proposed a classification baseline in which GLCM features are used to provide texture information. We experimented with the Contrast (CON), the Correlation (COR), the Homogeneity (HOMO), the DIS (Dissimilarity), the ASM (Angular Second Moment) and ENER (Energy). These features are computed on a quantized version of a single-band image. In this work, we used grayscale images derived by RGB composites (Robinson et al., 2021) and quantized them in 16 bins level. To extract a GLCM feature for a specific region (window around a pixel), the associated GLCM matrix has to be computed first. In our case, a GLCM is a 16 x 16 (defined by the number of bins) matrix containing the probability of each pixel value i co-occurring with a pixel value j, for defined distance offsets inside the selected window (we used a sliding 13 x 13 window). Finally, the GLCM is multiplied by a weight factor which depends on the selected texture feature. Although GLCMs provide useful spatial information for distinguishing sea surface features, they are computationally expensive, especially when calculated for multiple scales. To overcome this obstacle and include the scale information into the process while ensuring low computational costs: we utilized the Gaussian of the grayscale images derived by RGB composites, the Sobel of the Gaussian image, and the Eigenvalues $\lambda$of the Hessian Matrix of the Gaussian image at different scale levels (for standard deviation σ = 1, 2, 4, 8, 16).

In order to enhance the included texture information, we also utilized the Local Binary Patterns features (LBP, LBP UNI), which inform about the uniformity of local texture. Intuitively, LBP examines the neighbors of a center pixel and determines if the neighbor pixel values are more or less than the center pixel value.

4.5.2 Feature Selection using Machine Learning

This section presents a quantitative and qualitative evaluation based on the extracted results after exploiting the 51 spectral and spatial features (S2 bands values, GLCM, Local Binary patterns, Gaussian, Sobel of Gaussian Hessian Eigenvalues of Gaussian) and applying the Random Forest classifier.

Firstly, we quantitatively assess the performance of the applied Random Forest classifier (i.e., RF+). To evaluate our results and compare them with the corresponding outcomes from MARIDA, we relied on three metrics, i.e., IoU, Recall

and F1. The Table 4.8 demonstrates the scores for all metrics per class obtained by our RF+, as well as by MARIDA baseline models (RF∗ and U-NET∗). Overall, we observe that the proposed spectral and spatial features improve the classification performance, as it is indicated by the higher average scores that our model achieves for all metrics.

Regarding scores per class, Sediment-Laden Water still has the highest IoU, Recall and F1 scores (i.e., 0.99 - 1.00). For *Dense Sargassum*, *Sparse Sargassum*, *Ship*, *Clouds*, RF+ achieves an improvement of > +2% for IoU, > +1% for Recall and +2% for F1 compared with the RF∗ and U-NET∗. Interestingly, for the Foam class, RF+ improves IoU by +23%, Recall by +11% and F1 by +16%. Additionally, a significant improvement of +21% for IoU, +18% for Recall and +25% for F1 can also be seen for the *Natural Organic Material* class. However, both RF models provide the same results for the *Marine Debris* class. As far as the class *Shallow Water* is concerned, the U-Net model still achieves the highest scores for all metrics. To qualitatively evaluate the performance of our Random Forest model (RF+), we visually inspected the produced prediction maps and compared them to the respective classification results extracted by MARIDA baselines (Figure 4.8). Although RF+ and RF∗ provide the same scores for *Marine Debris* (Table 4.8), it seems that RF+ predicts better the specific class (Figure 4.8). Especially, the prediction of *Sargassum* (Figure 4.8) by RF+ is significantly improved; this fact is also consistent with the higher scores that our model achieves (Table 4.8). Finally, the performance of RF+ appears to be better than the U-Net∗ and RF∗ models over the coastal region (Figure 4.8), where *Foam*, *Shallow Water*, and *Turbid Water* co-exist.

Table 4.8. Quantitative evaluation of the proposed RF⁺, compared with MARIDA (Kikaki et al., 2022) RF⁎ and U-Net⁎ models.

Method	RF⁎			U-Net⁎			RF⁺		
	IoU	Recall	F_1	IoU	Recall	F_1	IoU	Recall	F_1
Marine Debris	**65**	**92**	**79**	33	70	50	**65**	**92**	**79**
Dense *Sargassum*	87	93	93	60	60	75	**90**	**94**	**95**
Sparse *Sargassum*	83	90	91	66	89	79	**86**	**91**	**93**
Natural Organic Material	18	31	31	2	2	4	**39**	**49**	**56**
Ship	67	82	80	62	76	76	**69**	**83**	**82**
Clouds	84	86	91	62	62	76	**87**	**88**	**93**
Marine Water	75	93	86	61	88	76	**80**	**95**	**89**
Sediment-Laden Water	**99**	**100**	**100**	99	99	**100**	99	**100**	99
Foam	60	74	75	55	55	71	**83**	**83**	**91**
Turbid Water	88	92	94	84	**95**	91	**90**	92	**95**
Shallow Water	30	37	46	**45**	**67**	**62**	31	45	47
Average	69	79	79	57	69	69	**74**	**83**	**83**

Overall, it seems that using the spatial information at multiple scales leads to better classification results with less noise (e.g., less isolated pixels classified as *Marine Debris*) and improved shape of predicted features (e.g., *Clouds*).

For further experimentation, we calculated Spearman Correlation to form highly correlated groups among the included spectral and spatial features and keep only one from each group. The selected features, each of which represents a different group, are demonstrated in Figure 4.9. By applying the permutation feature importance as described in (Kikaki et al., 2022) on the non-highly correlated features, we managed to identify the most important features, i.e., NDWI, H EIG 2 S16, CON, NDVI, FDI and SOBEL S16. Instead, individual bands (e.g., green) and the LBP group do not contribute to the classification process. Interestingly, H_EIG_2 seems to be more important than H_EIG_ 1.

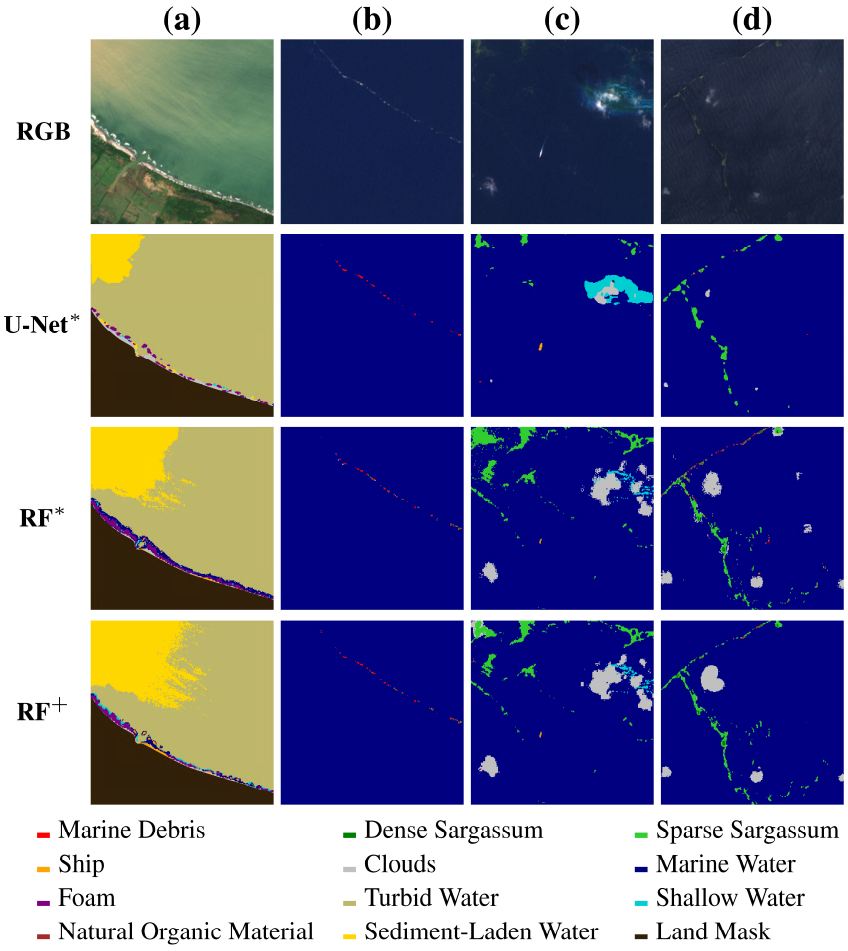


Figure 4.8 Classification results obtained by our RF+ and MARIDA baseline RF and U-Net* models. (a) S2 12-12-20 16PCC 6, (b) S2 22-12-20 18QYF 0, (c) S2 27-1-19 16QED 14, (d) S2 14-9-18 16PCC 13.*

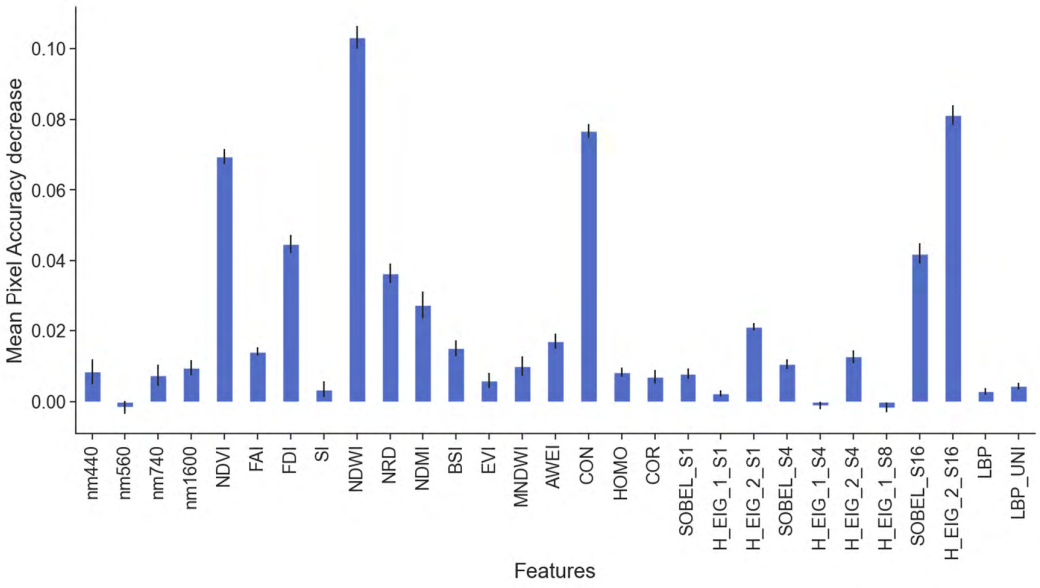


Figure 4.9 Mean Accuracy Pixel Decrease of 29 non high correlated features based on Random Forest Permutation method. SOBEL_σ sort for Sobel filter of the Gaussian of standard deviation σ. H_EIG_λ_σ sort of Hessian Eigenvalue λ (1st or 2nd) of the Gaussian of standard deviation σ.

4.5.3 Proposing new Combinations of Spectral & Spatial patterns

Finally, in order to further examine the ability of spectral and spatial features in competing classes discrimination, we visualize them per couple, and discuss the extracted insights based on the scattergrams analysis (Figure 4.10). we focused on the major competing cases for *Marine Debris* detection on the sea surface, as indicated by Kikaki et al. (2022). For this reason, we investigate the discrimination between *Marine Debris*, *Sargassum* macroalgae, *Natural Organic Material*, *Ship* and *Foam*.

Marine Debris vs *Sparse Sargassum* (Figure 4.10a,b): Regarding the spectral patterns of the considered floating materials, *Marine Debris* presents slightly lower peak at NIR and higher values at SWIR. Well-established vegetation index NDVI tends to have positive values for *Sparse Sargassum* and negative values for *Marine Debris* (4.10a). Nevertheless, there is an overlapping area where NDVI values are close to zero, possibly reflecting the cases with low subpixel proportions (i.e., sparse conditions). Concerning FDI, it significantly contributes to the classification process, as it is the fourth most important feature (Figure 4.9). For this index there is a strong theoretical justification (Tasseron et al., 2021) as well, yet we confirm that FDI alone does not adequately separate the specific materials based on S2 data (Biermann et al., 2020). This fact probably corresponds to the level of *Marine Debris* submersion, which is higher at the sea surface than in a controlled environment. On the other hand, the concurrent use of NDVI seems to highlight *Sparse Sargassum* and *Marine Debris* spectral patterns differences. The combination NDVI & EVI could potentially separate these features as well (Figure 4.10b).

Marine Debris vs *Ship* (Figure 4.10c): *Marine Debris* and *Ship* present similar spectral properties due to the same polymer composition. Additionally, they can be distributed in similar spatial patterns. For instance, small vessels and *Marine Debris* pixels can be depicted as individual pixels, leading to a challenging discrimination task. This is can also be observed in Figure 4.10c, as none of the considered models is able to predict the small ship correctly. Due to these challenges, the well-established NDVI and FDI fail to distinguish *Marine Debris* from *Ship*. On the contrary, the GLCM CON texture feature appears to be promising in this case. The GLCM CON presents the high *Ship* contrast to is background; thus, the variety of its value might be linked to the ship size. Moreover, improved results are achieved by combining CON with the NDMI index. *Marine Debris* has higher NDMI value than Ship, based on the fact that in the case of floating and partially submerged *Marine Debris* the moisture is higher.

Marine Debris vs *Natural Organic Material* (Figure 4.10d): The *Natural Organic Material* class consists of woody and vegetation debris which tends to accumulate on the sea surface in very similar patterns to *Marine Debris*. As shown in Figure 4.10d, NDVI seems to contribute to the discrimination of the considered materials, as it captures the reflectance values difference at the red and NIR bands. The additional use of MNDWI enhances their separation. The specific spectral index, by using green and swir bands, extracts water information and removes background noise (i.e., built-up area, vegetation) (Xu, 2006). *Marine Debris* has higher values in MNDWI and lower (mostly negative) values in NDVI than *Natural Organic Material*.

Marine Debris vs *Foam* (Figure 4.10e): *Foam*, compared to *Marine Debris*, has higher reflectance values across spectral spectrum, presenting a peak at green and a local minimum at 740nm as well (Kikaki et al., 2022). Furthermore, *Foam* accumulation patterns in the wave breaking zone are different to *Marine Debris*. The concurrent use of NDWI and Hessian Eigenvalue 2 of the Gaussian of standard deviation $\sigma = 16$ seems to enhance the discrimination of these two classes. The NDWI has lower values for *Marine Debris* than for *Foam*; this is propably due to the fact that *Foam* is a water-related class. Additionally, the Hessian Eigenvalue 2 of the Gaussian of standard deviation$\sigma = 16$ is higher for *Marine Debris* than for *Foam*.

Sparse Sargassum vs *Natural Organic Material* (Figure 4.10f): *Sparse Sargassum* and *Natural Organic Material* are floating organic debris which follow similar linear trajectories. Except for the common spatial patterns, the specific floating materials can present similar spectral signatures, as natural organic debris may contain vegetation such as leaves or plants. The EVI, as expected, highlights the vegetation, yet there is an overlapping area where vegetation debris and *Sparse Sargassum* macroalgae cannot be discriminated. Instead, BSI decreases the overlapping area, further enhancing the separation between these two classes. Specifically, the BSI index receives higher values for *Natural Organic Material*, which properties appear to be closer to the soil, than *Sparse Sargassum* that mostly exhibits negative values.

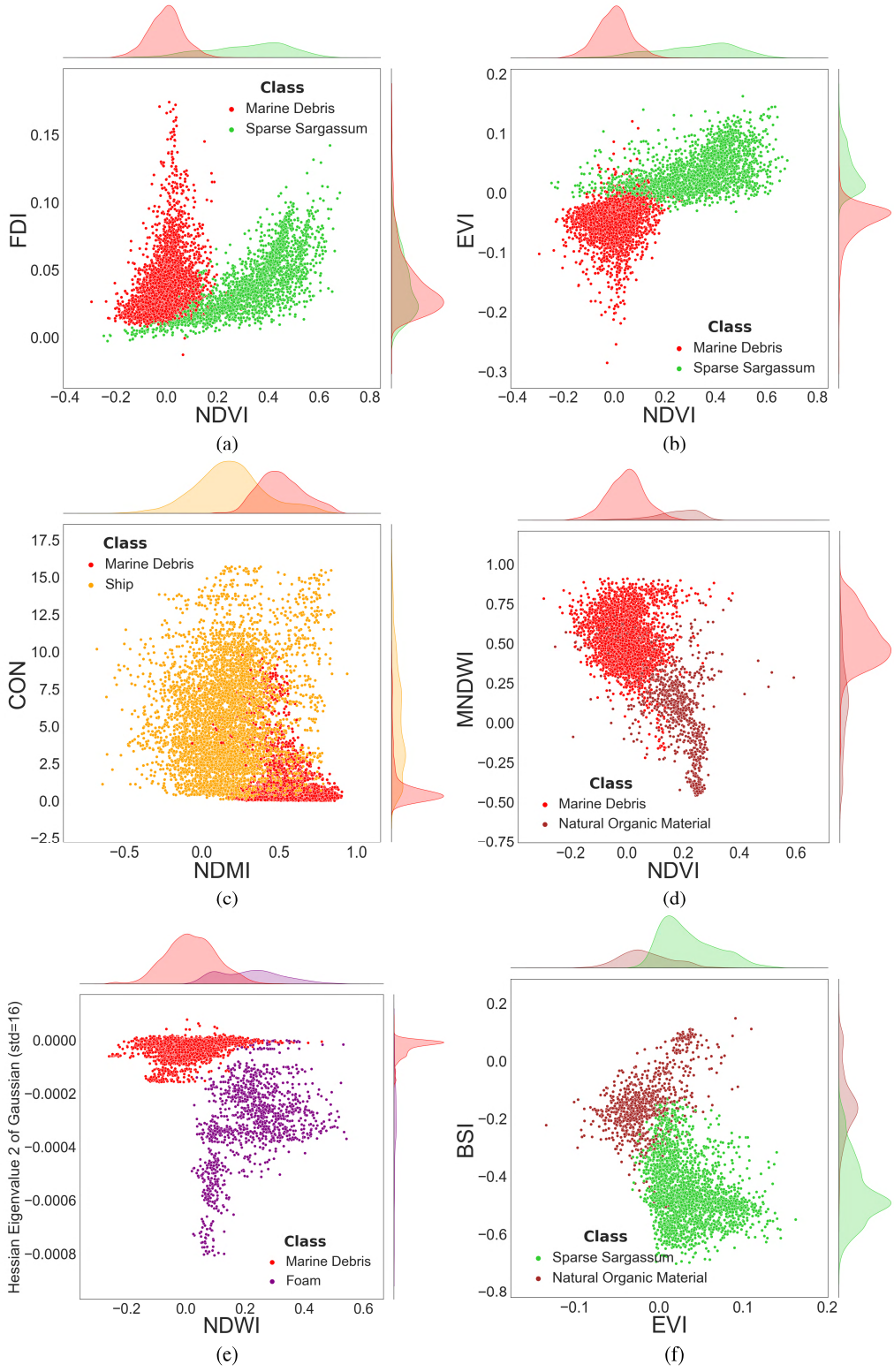


Figure 4.10 Scattergrams for the qualitative evaluation of challenging cases: (a) NDVI and FDI for Marine Debris and Sparse Sargassum, (b) NDVI and EVI for Marine Debris and Sparse Sargassum, (c) NDMI and CON for Marine Debris and Ship, (d) NDVI and MNDWI for Marine Debris and Natural Organic Material, (e) NDWI and Hessian Eigenvalue 2 of Gaussian of σ = 16 for Marine Debris and Foam (f) EVI and BSI for Natural Organic Material and Sparse Sargassum.

4.6 Conclusions

In this Chapter, we present MARIDA, a benchmark dataset for the detection of *Marine Debris* on S2 multispectral satellite data. MARIDA challenges the research community by: i) offering annotations of *Marine Debris* and various sea features that co-occur in realistic cases, ii) providing a detailed overview of MARIDA as well as spectral signatures analysis of annotated data, iii) evaluating ML algorithms, and iv) identifying application cases and open issues. Considering that marine litter research is increasing significantly and plastic debris monitoring using remote sensing is still challenging, we provide a *Marine Debris* dataset appropriate for future detection experiments, ML classification tasks and water quality monitoring. Except for spectral and spatial features examined in this study, temporal features can be also investigated to assess their contribution to specific marine classes detection (e.g., *Turbid Water*, *Shallow Water*). We envisage the continuous expansion of this dataset, including additional cases from the global oceans.

5 Marine Pollution detection against other Competing Sea Surface Features

This Chapter introduces a *Marine Debris* and *Oil Spill* dataset (MADOS) and a mapping framework for marine pollution detection. Section 5.1 describes MADOS itself and the included thematic classes. A representative study area of MADOS (Ionian Sea) is also demonstrated. In section 5.2, we present all ML baselines and our novel DL model, named MariNeXt, for *weakly-supervised semantic segmentation* tasks. In section 5.3, we evaluate our framework and perform a thorough ablation study. Next, in section 5.4, we discuss MariNeXt's predictive capabilities and the potential in detecting and discriminating marine pollutants in satellite data. The conclusions of the whole chapter are presented in section 5.5.

5.1 Introducing the MADOS Benchmark Dataset

5.1.1 MADOS Data Processing & Curation

Satellite sensors with high spatial, spectral and temporal resolution are required to develop operational solutions for monitoring the marine environment. Currently, the openly available Copernicus S2 multispectral data cover global coastal waters at high spatial resolutions (10 m, 20m and 60 m). Additionally, S2 satellite has a frequent revisit time of approximately five days (depending on the latitude), delivering a large amount of free geospatial data. It also acquires data in 13 spectral bands, which range from the visible to the Short-Wave Infrared (SWIR). Therefore, in our study, we utilized the S2 multispectral satellite images. We collected S2 Level 1C products (Top-of-Atmosphere reflectance) through EarthExplorer (USGS).

The annotation was performed by three image-interpretation experts who had access to all collected data. After the annotation step, an inter-annotators agreement protocol was established and data refining was achieved using t-distributed Stochastic Neighborhood Embedding (t-SNE) algorithm following MARIDA methods and protocols (Kikaki et al., 2022). Similar to previous studies on floating materials detection (Qi et al., 2020; Sun et al., 2018a; Zhao et al., 2014), we applied Acolite to retrieve the Rayleigh reflectance. Based on the annotated data, we cropped the S2 scenes into non-overlapped image patches. In the produced dataset, the initial spatial resolution of each spectral band was retained without any upsampling technique. Each provided image corresponds to 2.4x2.4 km^2; thus, the 10m spectral bands form image patches of 240x240 pixels, the 20m patches of 120x120 pixels and the 60m patches of 40x40 pixels.

5.1.2 MADOS thematic classes and annotations

We have constructed MADOS, which contains annotated S2 data based on 174 scenes and 47 tiles captured between 2015-2022. Numerous annotations were created consisting of 1481155 pixels based on studies, reports and data on marine pollution as well as intensive photo-interpretation (Table 5.1). Finally, our globally distributed dataset (Figure 5.1) contains 15 different thematic classes (Table 5.2).

Table 5.1: The sites where publications and reports were collected for MADOS construction.

Site	Publications	Reports
	Marine debris	
Global	(Kikaki et al., 2022) (MARIDA)	
Guatemala & Honduras	(Kikaki et al., 2020)	Caroline Power (Citizen scientist)
Haiti	(Kikaki et al., 2022)	4ocean Clean-ups
Manila, Philippines	(Gonzaga et al., 2021)	
Calabria, Italy	(Sannigrahi et al., 2022)	
Beirut, Lebanon	(Sannigrahi et al., 2022)	
	Oil spills	
Zakynthos Isl., Greece	(Kolokoussis and Karathanassi, 2018)	
Saronikos Gulf, Greece	(Parinos et al., 2019)	
Mauritius	(Rajendran et al., 2021)	https://www.youtube.com/watch?v=kDOobamVtAA&ab_channel=GuardianNews
Solomon Isl.		https://www.solomontimes.com/news/growing-concerns-over-bulker-oil-spill/8889
East China Sea	(Sun et al., 2018)	
Red Sea	(Vankayalapati et al., 2023)	
Peru		(Mega, 2022)
Arabian Gulf	(Zhao et al., 2014), (Sardi et al., 2020)	KAUST & NTUA reports and analysis
	Jellyfish	
Nova Scotia	(Qi et al., 2020)	
	Sea snot	
Marmara Sea	(Hu et al., 2022)	

We focus on two major marine pollutants, i.e., *Oil Spill* and *Marine Debris*. First, oil slicks from different sources (i.e., ship accidents, offshore platforms, discharges due to natural disasters and seafloor natural seepages) were recorded. Our aim was to develop a model capable of detecting one *Oil Spill* thematic class and not different oil types. However, oil forms of different thicknesses were considered and annotated in MADOS, including all categories established by NOAA (2016): a) Dark, b) Mousse (emulsified), c) Transitional dark, d) Metallic, e) Rainbow and f) Sheen (Figure 5.2). Regarding *Marine Debris*, we used the open-access MARIDA dataset, which distinguishes litter from various sea features. We also increased the number of *Marine Debris* pixels by adding new observations based on recent reports and publications.

Following previous studies (Biermann et al., 2020; Hu et al., 2022; Kikaki et al., 2020; Martínez-Vicente et al., 2019), we additionally considered floating materials

that have been suggested to be spectrally similar to plastics, or tend to be aggregated with floating macroplastics on the ocean surface. Thus, *Foam*, *Sargassum*, *Natural Organic Material* (including driftwood), *Ships* and *Sea Snot* trajectories (Hu et al., 2022) were annotated (Figure 5.2). Suspected *Jellyfish* investigated by Qi et al., (2020) were also taken into account to evaluate the discrimination from *Marine Debris* and *Sea Snot,* as previous studies have indicated the potential of spectral similarities (Hu et al., 2022; Van der Zande et al., 2014). Classes sometimes observed concurrently with oil slicks, such as *Oil Platform* and *Sparse Floating Algae,* were also considered and added. Moreover, sea state look-alikes were annotated and aggregated with the *Marine Water* class. Other water-related classes recorded on our study regions (i.e., *Sediment-Laden Water, Turbid Water, Shallow Water, Waves and Ship Wakes)* have also been annotated. Also, sparse *Sargassum* was grouped with other sparse floating algae species, forming a *Sparse Floating Algae (SpFa)* super-class. We note that MADOS consists of realistic observations recorded under diverse weather conditions aiming to support real-world applications (Figure 5.2).

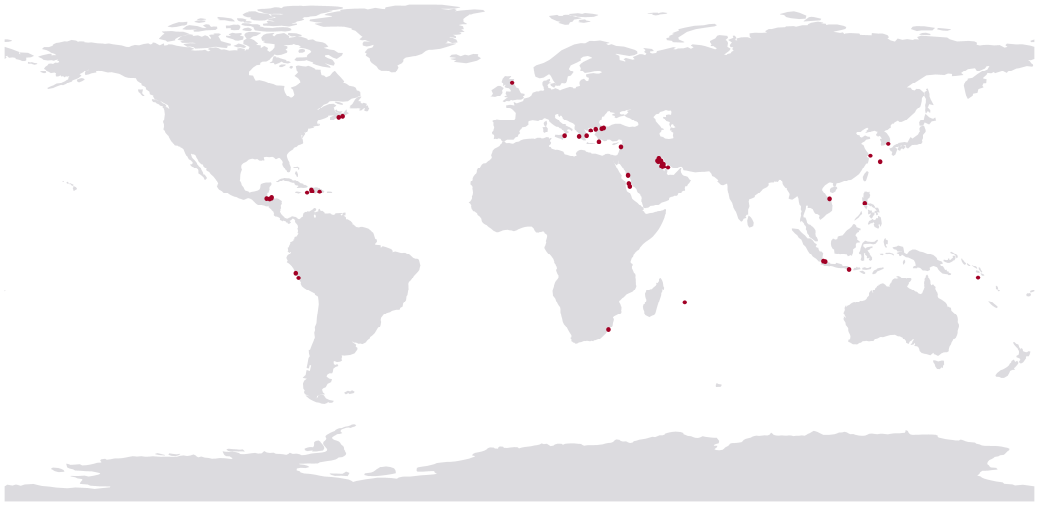


Figure 5.1 The globally distributed Sentinel-2 tiles from which MADOS data were sampled.

We note that clouds and cloud shadows were not included in our analysis. We consider this a different, quite challenging, task that the remote sensing scientific community is working on. Therefore, one could integrate any of the standard software tools (such as MAJA, Fmask, Sen2Cor) into the processing pipeline to address cloud and shadow detection.

From this point in the paper, we will use the acronyms presented in Table 5.2. The Table 5.2 defines these acronyms to facilitate clarity throughout the text. Overall, MADOS consists of 1481155 annotated pixels composed of 4696 *MD* pixels and 234568 *Oil* pixels, corresponding to 469600 m^2 and 23456800 m^2 areas, respectively. Additionally, 2803 image patches are offered, each one covering 2.4x2.4 km^2.

Table 5.2: MADOS Thematic Classes. Per-class description, acronym, number of image patches and corresponding annotated pixels.

#	Class Name	Acronym	Description	# of patches (240x240)	# of pixels
1	*Marine Debris*	MD	Floating plastics, mixed anthropogenic debris	472	4696
2	*Oil Spill*	Oil	Oil types including crude oil, emulsified, oil on water, metallic, sheen & rainbow	361	234568
3	Dense *Sargassum*	DenS	Dense floating *Sargassum* macroalgae	77	4308
4	Sparse Floating Algae	SpFA	Sparse *Sargassum & other* macroalgae	157	4972
5	Natural Organic Material	NatM	Vegetation & Wood	92	1454
6	Ship	Ship	Sailing & Anchored Vessels	298	13222
7	Marine Water	MWater	Clear Water & sea state Look-alikes	1360	496702
8	Sediment-Laden Water	SLWater	High-Sediment river discharges with brown colour	57	376061
9	Foam	Foam	Foam recorded at river fronts or coastal wave breaking	58	1225
10	Turbid Water	TWater	Turbid waters close to coastal areas	232	170306
11	Shallow Water	SWater	Coastal waters, including coral reefs and submerged vegetation	140	122753
12	Waves & Wakes	Waves	Waves & Ship Wakes	224	20374
13	Oil Platform	OilPlat	Offshore Oil platform	348	15701
14	Jellyfish	Jellyfish	Jellyfish	12	2541
15	Sea snot	Sea snot	Sea snot/ mucilage	63	12272
	Total			2803	1481155

In the Table 5.2, we observe that in MADOS, the distribution of samples per class at the pixel- or patch-level is not uniform. As expected, water-related classes can be found and digitized to a significantly larger extent than the usually small floating objects like floating plastics. Specifically, the class *MWater* has been annotated in 1360 patches, while the second-highest number of 472 patches was *MD* and the third-highest number of 361 patches was *Oil*. The considerable amount of annotated images indicates the high diversity of annotated *Oil and MD* pollutants. On the other hand, *Jellyfish* and *NatM* are the least annotated classes as they are not frequent in the marine environment and were not our primary focus.

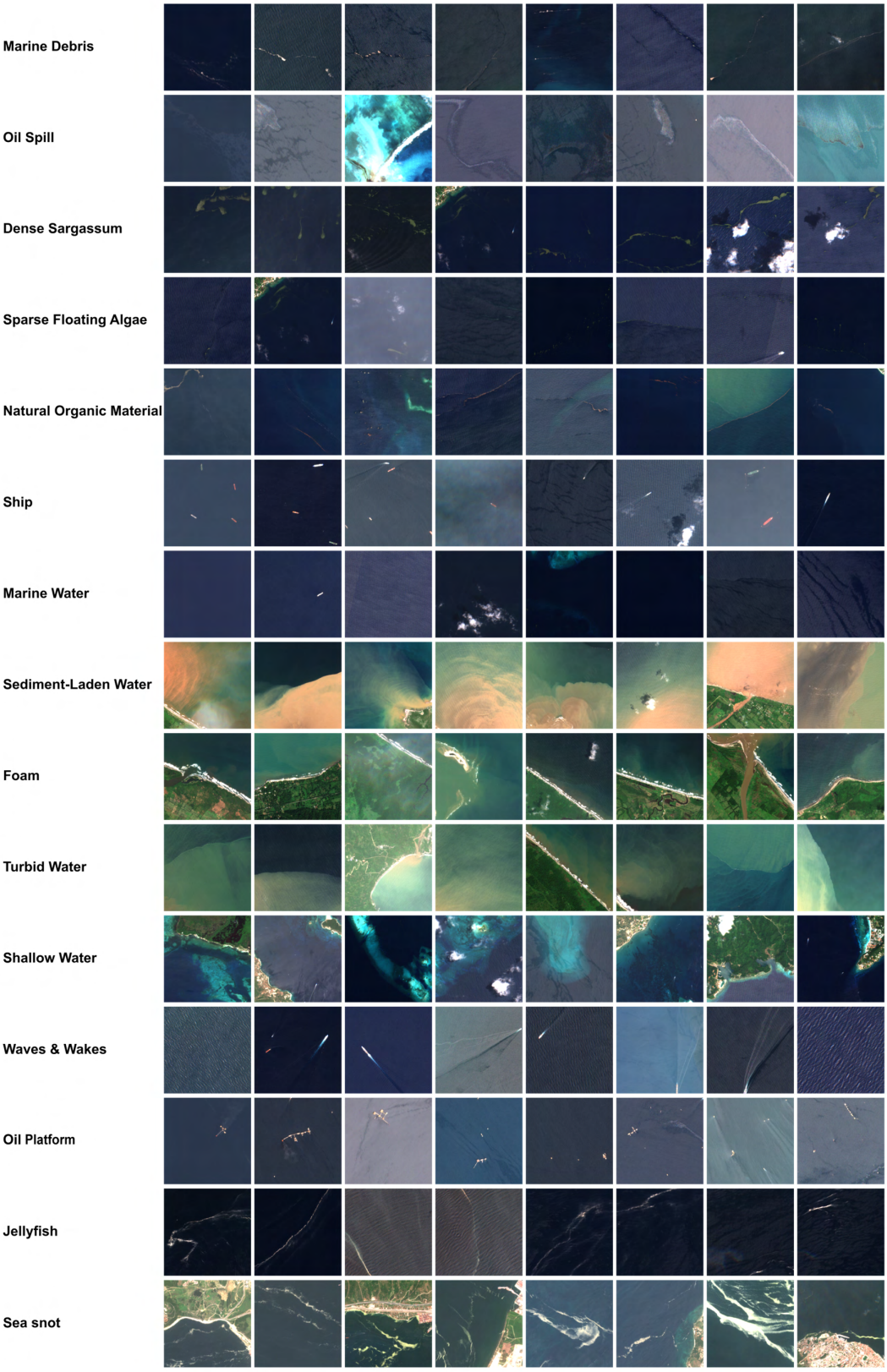


Figure 5.2 An overview of MADOS provided patches. Marine pollutants and Sea surface features were annotated under various weather and sea state conditions.

5.1.3 Oil Spill Detection in a MADOS Study Site: Zakynthos Isl., E.Mediterranean

High-resolution satellite data from different sensors (i.e., Sentinel-2, Landsat-8 and Planet) were used in order to monitor and identify the origin, structure and spatial extent of naturally occurring sea surface oil slicks in the coastal zone of Zakynthos Island (Ionian Sea, E. Mediterranean). We also examined the potential effect of earthquake activity on the oil slick behavior during 2018-2019, a period including a strong earthquake measuring magnitude M_w 6.8 occurred in the Ionian Sea (26 October 2018). Based on successive satellite observations along with ocean conditions description, we also explored *Oil Spill* spreading.

Zakynthos Island: A Representative MADOS Study Area

The island of Zakynthos is one of the most seismically active regions in the Mediterranean Sea as it is located very close to Cephalonia fault and to the convergent boundary between the Eurasian and African plates (Papazachos and Papazachou, 1997). The study site includes two natural oil springs: a) the "Herodotus springs" of Keri Lake, located in the southern part of island (Figure 5.3b) expanding over an area of 3 km^2 at 1 m elevation (Avramidis et al., 2017); and b) offshore sea surface oil slicks, commonly visible in the study area (Figure 5.3c). In the latter area, continuous natural seepage of oil has been recorded at depths of ~150 m. Earlier studies on marine sediments conducted by HCMR, highlighted the presence of petroleum hydrocarbons with peculiar molecular profiles, not characteristic of other marine sediments in the Hellenic region. Previous studies of HCMR (2008) by side scan sonar revealed the presence of several depressions on the seafloor covering an overall area of 100x30 m, where natural hydrocarbon seepage occurs (Figure 5.3d).

Satellite & Auxilliary Data

S2 multispectral data from April 2018 to September 2019 were collected as well as L8 data were additionally obtained for earthquake events description. In total, two L8 path rows, i.e., 184/34 and 185/34, and one S2 tile, i.e., 34SDG, were used for this study. Planet satellite images (resolution=3 m) were also used for validation. Satellite data with a cloud presence of over 25% were rejected and not further processed. The rest of the data were atmospherically corrected and surface reflectance values were extracted based on the ACOLITE atmospheric processor (Vanhellemont and Ruddick, 2016). After atmospheric correction proposed indices (i.e. B2/B11 and StdDev(B2)*B2/B11 by Kolokoussis & Karathanassi, 2018) were applied in all S2 data for *Oil Spill* detection. CMEMS (http://marine.copernicus.eu) data were used for the study of surface currents velocity and wind data from National Observatory of Athens for wind conditions description.

Regarding the earthquake activity period, multispectral (Sentinel-2 and Landsat-8) and radar data (Sentinel-1) before and after the specific seismic events were obtained from Copernicus and USGS official websites. During July 2019, successive satellite images were used in order to estimate *Oil Spill* velocity and describe its evolution along with ocean conditions.

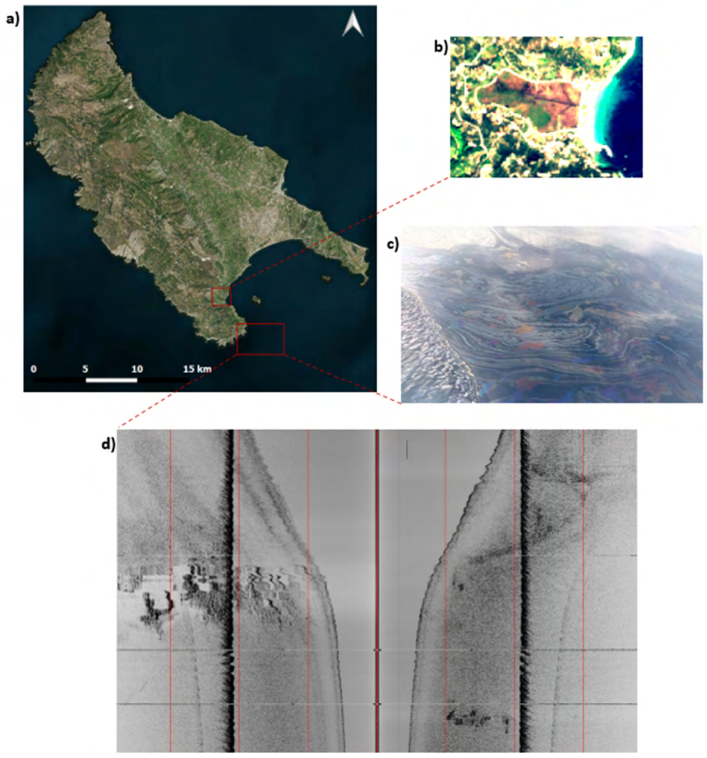


Figure 5.3 (a) The island of Zakynthos, (b) Herodotus springs of Keri Lake (Sentinel-2 image), (c) Observed sea surface oil slick (http://patrastimes.gr/), (d) Side scan sonar profile illustrating seafloor depressions associated with oil seepage on the left (by Dr. G. Rousakis HCMR).

Oil Spill Monitoring

Collected multispectral satellite data during October 2018-September 2019, confirmed that *Oil Spill* is a permanent feature of study area, due to the natural oil seep located underwater. It has to be emphasized that largest oil slicks were tracked after seismic events. In almost all cases one oil slick of mean 6 km length and width was ranged from 100 m to 2 km. Landsat-8 image acquired on 2 July 2019 demonstrates indicative *Oil Spill* position at sea surface (Figure 5.4). Earthquake events recorded in Zakynthos area during the period late October 2018-Early September 2019, are given in Table 5.3. The greatest earthquake occurred on 26 October 2018 (M_w 6.8), which led to the largest *Oil Spill* in study area for the corresponding period. On 26 October 2018 four *Oil Spill* features were tracked in a satellite image captured by Sentinel-2, 9 hours after the earthquake. Figure 5.5

presents the *Oil Spill* areas in RGB composite. The length of these areas was 2-9 km, as the width was 25 m-2 km. Bands ratio B2/B4 was also applied in the specific image in order to outline the *Oil Spill* features (Figure 5.6).

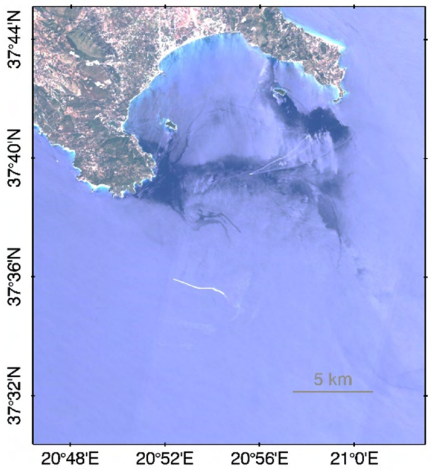


Figure 5.4 Top-of-Atmosphere RGB composite of Landsat-8 acquired on 2 July 2019.

Table 5.3. Recorded seismic events and satellite data acquisition dates.

Earthquake activity	**Earthquake magnitude M_w**	**Landsat-8**	**Sentinel-2**
26 October 2018	6.8		26/10
11 January 2019	3.7	16/1	
27 February 2019	3.8		28/2
17 March 2019	4.5	21/3	25/3
23 April 2019	3.8	29/4	29/4
20 May 2019	3.9	24/5	
20 June 2019	3.8	25/6, 2/7	23/6
28, 30 July 2019	3.9, 4.2	27/7, 3/8	28/7
14 August 2019	3.5	19/8	17/8
6 September 2019	3.7	13/9	6/9

In late July 2019 two successive Landsat-8 and Sentinel-2 images were used to study oil slick spreading, i.e., 27 July (Landsat-8) – 28 July (Sentinel-2). The observed oil slick covered a distance of 4.5 km in 24 hours. It moved in a SE-NW direction and its recorded speed was 0.52 m/s (Figure 5.8). Surface currents model outputs (CMEMS) indicated that mean currents velocity was 0.32 m/s with a NE-SW direction (Figure 5.7). Southern winds prevailed and the average winds speed was 2.6 m/s. The specific event confirms that ocean conditions (i.e., winds and currents)

contribute to *Oil Spill* spreading (De Dominicis et al., 2016; Zodiatis et al., 2018). Information from forecasting models about *Oil Spill* weathering and three-dimensional evolution need additionally to be regarded in a future study.

Experimental results showed that between 2018-2019 large oil slicks were recorded after seismic events as the largest was outlined after strong earthquake measuring magnitude M_w 6.8 in October 2018. Successive satellite images during late July 2019 confirmed that *Oil Spill* spreading is affected by ocean parameters (i.e., winds and currents). A detailed comparison between remote sensing and *in situ* measurements as well as results from *Oil Spill* forecasting models need to be included in a future study.

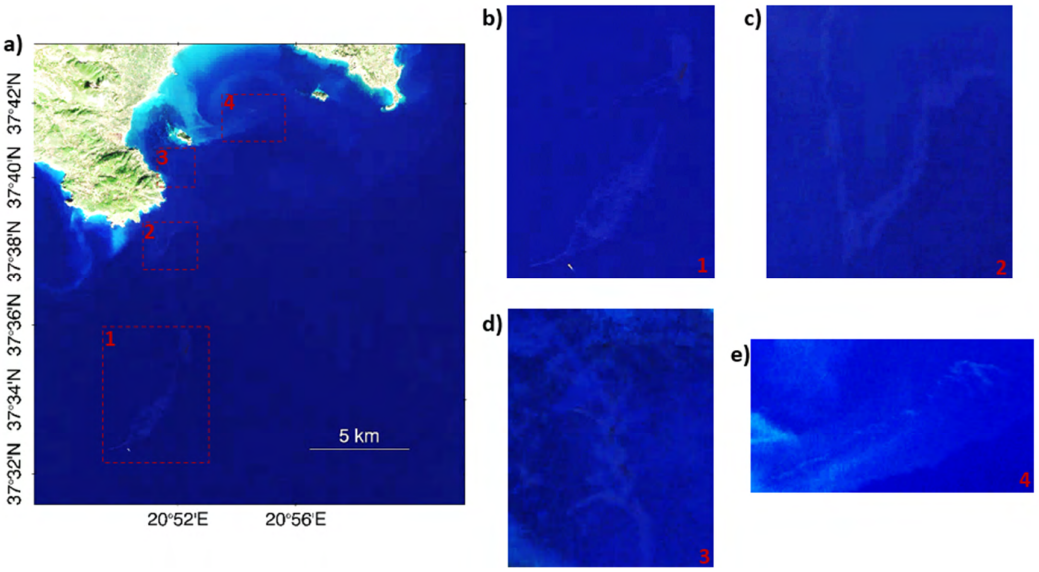


Figure 5.5 (a) Atmospherically corrected Sentinel-2 RGB composite on 26 October 2018. (b)-(e) Oil Spill detection. The RGB images are stretched for illustration purposes.

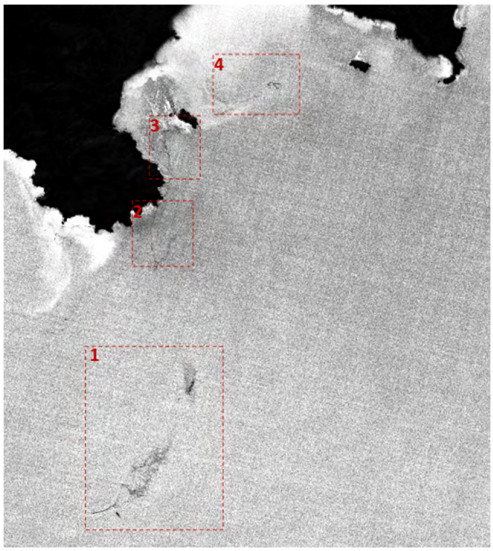


Figure 5.6 Band ratio B2/B4 calculated for S2 image acquired on 26 October 2018.

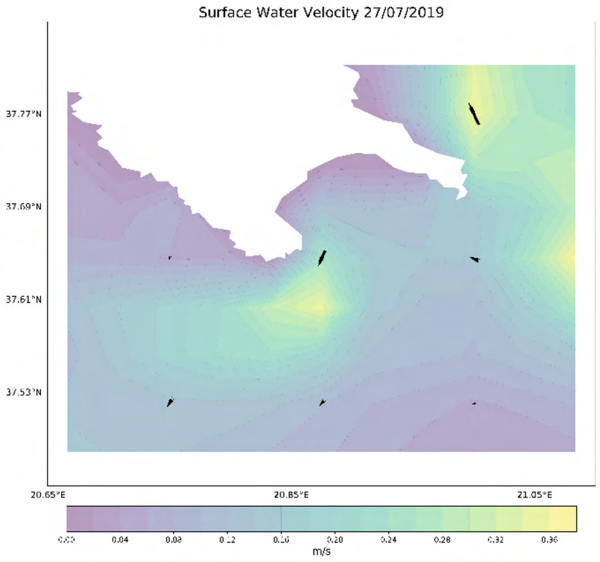


Figure 5.7 Ocean currents velocity and direction for 27 July 2019 (CMEMS after proc.).

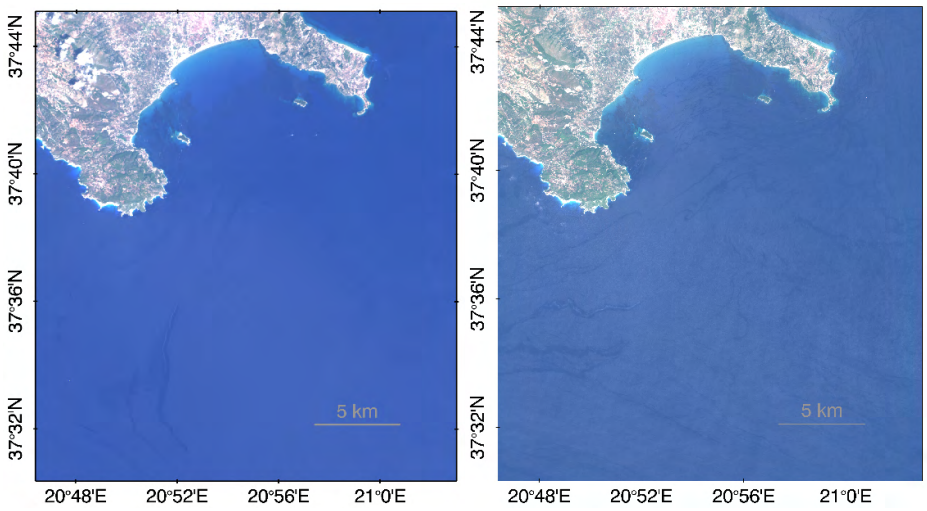


Figure 5.8 Top-of-Atmosphere RGB composites. L8 Image acquired on 27 July 2019 (on the left). S2 Image acquired on 28 July 2019 (on the right).

5.2 Machine learning classification frameworks

We present and discuss well-established and state-of-the-art models employed as baselines in our experiments. Then, we describe in detail our MariNeXT framework. Finally, we show our experimental design, including the evaluation protocol and implementation details.

5.2.1 Baselines

Random Forest & U-Net

We examined two different baseline models using RF: i) one based on just the spectral signatures and ii) one with spectral signatures, spectral indices and other spatial features. In particular, the employed spectral indices were: NDVI, NDWI, FAI, FDI, Shadow Index (SI), Normalized Difference Moisture Index (NDMI), Bare Soil Index (BSI) and NRD. Following Mikeli et al. (2022), for the spatial features, we utilized the grayscale images derived by RGB composites and computed: GrayLevel Co-occurrence Matrix (GLCM), the Gaussian at multiple scales, the Sobel of the Gaussians, and the Eigenvalues of the Hessian Matrix of the Gaussians. For RF and U-Net the reader is referred to the section 4.1.4.

SegNeXt

We also employed a state-of-the-art convolution-based, encoder-decoder architecture named SegNeXt (Guo et al., 2022). The encoder consists of a sequential four-stage pyramid structure named Multi-Scale Convolutional Attention Network (MSCAN). Each stage contains a down-sampling step followed by multiple blocks. The multiple blocks are inspired by a ViT-like (Vision Transformer) structure (Dosovitskiy et al., 2020), with the main difference being that the Multi-Scale Convolutional Attention (MSCA) replaces the self-attention module. The MSCA

module consists of a depth-wise 5 x 5 convolution layer followed by a multi-branch depth-wise strip convolution layer that efficiently approximates the 7 x 7, 11 x 11 and 21 x 21 large-kernel convolutions, extracting information at multiple scales. The down-sampling steps consist of 3 x 3 convolution and batch normalization layers. The spatial dimensions of the first stage features are H/4 x W/4 and are further reduced by half for each following stage, where W and H are the width and height of the input image, respectively. Moreover, the decoder concatenates the output feature maps extracted from the last three stages, discarding the features from the first stage. Guo et al. (2022) highlight that stage-one features contain too much low-level information that does not perform well in their evaluation setup with RGB images of everyday objects, humans and actions. In order to capture global context, the Hamburger (Geng et al., 2022) module is adopted, which factorizes the input features using low-rank matrix decomposition, discarding the irregular noise and redundant features while keeping the clean signal subspace.

5.2.2 The MariNeXt Framework

In this work, we have designed a semantic segmentation framework based on recent convolutional attention mechanisms (Guo et al., 2022) that are able to encode contextual information effectively depicted in multispectral satellite imagery in the marine environment. In particular, we have adopted the aforementioned SegNeXt architecture and proposed improvements in three directions, i.e., modeling, training and testing (Figure 5.9). These improvements were tailored for marine mapping applications and significantly boosted the performance of *weakly-supervised* marine pollution detection. The three directions are described in the following paragraphs.

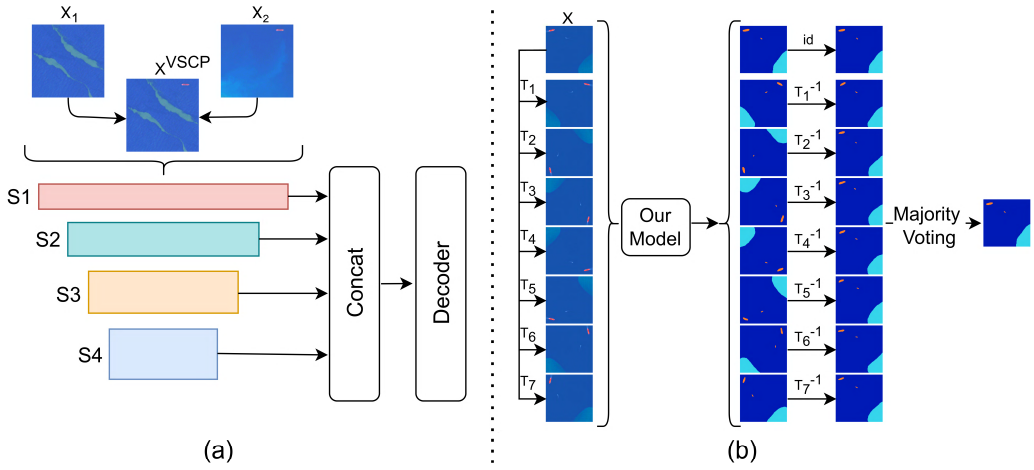


Figure 5.9 Simplified overview of MariNeXt, which consists of two main parts: a) During the training process, the input batch of images is augmented using the VSCP module combing the samples, which are then passed as new input. Then MariNeXt exploits the S1 (1st-Stage Higher Resolution Features) to produce fine-grained predictions. b) During testing, the TTA strategy is employed to produce several augmented views of the input image and aggregate the predictions by major voting, producing refined predictions.

1st-Stage Higher Resolution Features

Regarding the modeling part, we reexamine the design of SegNeXt that discards the first-stage features. This implementation choice leads SegNeXT to output H/8 x W/8 prediction maps, which are nearest-neighbor upsampled to match the input H x W image. As shown by Guo et al. (2022), the high-resolution features might contain too much low-level information for the RGB images of everyday objects. However, in our case, various floating materials such as *Marine Debris* might be presented only on a few pixels turning high-resolution features essential for effective modeling. Additionally, this approach enables the model to output H/4 x W/4 prediction maps. Thus, we show that exploiting additional high-resolution features from the encoder improves the performance in the considered marine environment mapping task.

Very Simple Copy-Paste (VSCP)

As discussed in previous sections, the collection of ground truth data is challenging, involves high-cost laborious work, while collecting high-quality and confident annotations is not straight forward. For such cases, and unbalanced datasets as MADOS, considering data augmentation techniques can guide the model to learn better feature representations. Towards this direction, we have integrated a robust data augmentation strategy that fits perfectly to the particularities of marine classes.

More specifically, Copy-Paste augmentation approaches are established techniques for object detection and instance segmentation tasks (Dvornik et al., 2018; Dwibedi et al., 2017; Fang et al., 2019). These methods are related to MixUp (Zhang et al., 2022) and CutMix (Yun et al., 2019) augmentation strategies. However, compared to MixUp and CutMix, Copy-Paste methods do not transfer an object's whole bounding box but only the exact annotated pixels. Furthermore, as everyday objects depicted in RGB images are placed in specific and oriented positions, most previous works on Copy-Paste utilized the surrounding context to place the copied object in the appropriate environment. Recently, Simple Copy-Paste (Ghiasi et al., 2021) proposed placing the objects in random positions and demonstrated promising results in instance segmentation.

In our case, MADOS consists of S2 satellite images captured at 786 km mean orbital altitude in the aquatic environment. This fact means that the surrounding context and the relative scale between the objects are always the same. Note also that every floating material can co-occur with any other over the same background environment (i.e., *Water*). For these reasons, in our case, there are not significantly different surrounding contexts or scale issues. Therefore, inspired by Ghiasi et al. (2021), we propose a straightforward and Very Simple Copy-Paste (VSCP) training augmentation approach that fits excellently in our *weakly-supervised semantic segmentation* task (Figure 5.10).

In particular, given a batch $X \epsilon \mathbb{R}^{(N_{batch},C,W,H)}$ and $Y \epsilon \mathbb{Z}^{(N_{batch},W,H)}$, where X are the input images, Y the groundtruth masks, N_{batch} is the batch size, C is the spectral dimension, and W and H are the width and height, respectively, our approach augments the current training batch size by $N_{VSCP} = \left\lfloor \frac{N_{batch}}{2} \right\rfloor$, where $\lfloor \cdot \rfloor$ outputs the greatest integer less than or equal to the input. Specifically, given the i and the $i + N_{VSCP}$ samples in the current training batch, we construct a new augmented image by adding the annotated regions and replacing the pixels to the image i from the image $i + N_{VSCP}$. The new synthesized image contains the annotations from both i and $i + N_{VSCP}$ using the background (non-annotated regions) of the i image (Figure 5.10).

More formally, each pixel element $x_{n,c,w,h}^{VSCP}$ and $y_{n,w,h}^{VSCP}$ of the N_{VSCP} augmented images $X^{VSCP} \epsilon \mathbb{R}^{(N_{VSCP},C,W,H)}$ with the corresponding ground truth masks $Y^{VSCP} \epsilon \mathbb{Z}^{(N_{VSCP},W,H)}$, are:

$$x_{n,c,w,h}^{VSCP} = \begin{cases} x_{n+N_{VSCP},c,w,h}, & \text{if } y_{n+N_{VSCP},w,h} \neq -1 \\ x_{n,c,w,h}, & \text{otherwise} \end{cases}$$

$$y_{n,w,h}^{VSCP} = \begin{cases} y_{n+N_{VSCP},w,h}, & \text{if } y_{n+N_{VSCP},w,h} \neq -1 \\ y_{n,w,h}, & \text{otherwise} \end{cases}$$

for $n = 1, \dots, N_{VSCP}$, $c = 1, \dots, C, w = 1, \dots, W$ and $h = 1, \dots, H$. Note that the non-annotated pixels have $y_{n,w,h} =$ -1.

Test-Time Augmentation (TTA)

In a similar manner with (Gonzalo-Martín et al., 2021; Moshkov et al., 2020; Wang et al., 2019b, 2019a), during the training procedure, we have also considered augmentations during the inference (test-time) execution in order to refine MariNeXt predictions by averaging the outputs across the different augmented views. To do so, we have explored a simple and effective Test-Time Augmentation (TTA) strategy.

In particular, for each input image $X \in \mathbb{R}^{C \times W \times H}$, a trained model $f: \mathbb{R}^{C \times W \times H} \rightarrow \mathbb{N}^{W \times H}$ produces an output $Y \epsilon N^{W \times H}$ forming the assigned integers that correspond to the predicted class for each pixel. Using N_T different transformations $\{\mathcal{T}_i\}_{i=1}^{N_T}$ we obtain a set of predictions $\{Y^{\mathcal{T}_i}\}_{i=1}^{N_T}$:

$$Y^{\mathcal{T}_i} = \mathcal{T}_i^{-1} f(\mathcal{T}_i(X)), \text{ for } i = 1, \dots, N_T$$

Then for each pixel we aggregate the results using the majority voting which corresponds to the mode of the sampling distribution:

$$y_{w,h}^{TTA} = MODE(\{Y_{w,h}^{\mathcal{T}_i}\}_{i=1}^{N_T}), \text{ for } w = 1, \dots, W \text{ and } h = 1, \dots, H$$

The final predicted image is the $Y^{TTA} \epsilon N^{W \times H}$ with elements $y_{w,h}^{TTA}$.

In order to keep the initial spectral bands and object scales as in the raw S2 images, we do not apply any color or scale transformation. Specifically, we combine rotations of 0°, 90°, 180°, 270° with the horizontal flipping, resulting in $N_T = 8$ different transformations applied in test-time execution.

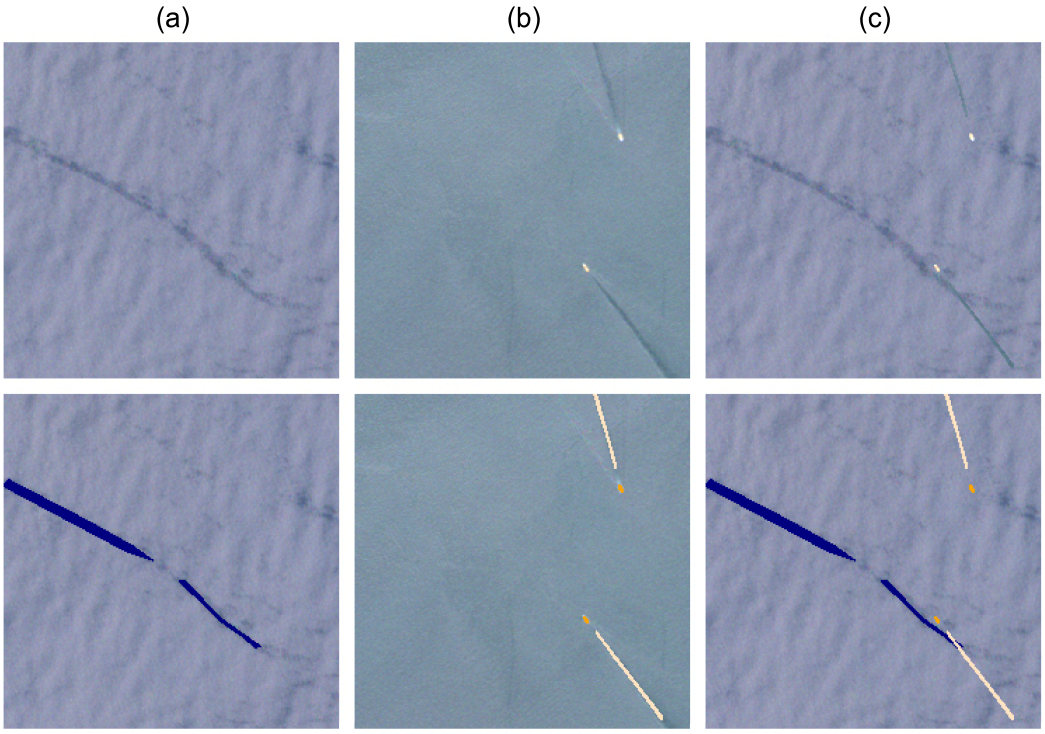


Figure 5.10 MariNeXt integrates the Very Simple Copy-Paste (VSCP) augmentation technique for constructing new images, enriching MADOS. Given pairs of input images and annotations, VSCP combines the input and annotated pixels of two images to create a new synthetic input pair increasing the class co-occurrence of input data. The two input images show a) Sea state look-alikes and Marine Water and b) Ships, Wakes and Marine Water. We also demonstrate the c) synthesized image. RGB images (top) and our corresponding annotations (down) are presented.

5.2.3 Experimental design

Evaluation on MADOS

In order to train and assess MariNeXt and the other baseline models, we split MADOS into three disjoint train, validation and test sets. The train set is employed to train the models explicitly, the validation set to monitor the learning process and the test set to perform our evaluation. Each of these sets retains the class distribution of the whole MADOS dataset. Additionally, the cropped patches from each scene/ unique date are kept on the same set. The splits contain 1433 training (761887 pixels), 642 validation (358121 pixels) and 728 test (361147 pixels) patches. The number of patches and pixels corresponds to approximately 50%/25%/25% of the initial MADOS dataset. We note that each data split was created as a representative subset of the entire MADOS dataset in terms of pixels and patches statistics (Table 5.4).

Table 5.4 Patches and pixels class distribution (%) for the train, validation and test sets in MADOS.

Set	MD	Dens	SpFA	NatM	Ship	Oil	MWater	SLWater	Foam	TWater	SWater	Waves	OilPlat	Jellyfish	Sea snot
	Patches														
Train	17.45	3.07	6.21	3.63	9.56	14.24	48.92	2.02	2.16	8.72	5.02	6.84	10.4	0.42	2.16
Validation	15.26	1.87	4.83	3.43	11.68	11.99	47.82	2.34	2.02	7.32	4.52	9.19	14.33	0.62	2.49
Test	17.03	2.88	5.08	2.47	11.81	10.99	48.35	1.79	1.92	8.24	5.36	9.2	14.7	0.27	2.2
Total	16.84	2.75	5.6	3.28	10.63	12.88	48.52	2.03	2.07	8.28	4.99	7.99	12.42	0.43	2.25
	Pixels														
Train	0.34	0.24	0.34	0.14	0.78	18.45	34.77	20.64	0.06	11.69	9.19	1.31	0.92	0.18	0.96
Validation	0.32	0.33	0.18	0.03	1.04	11.86	25.77	35.06	0.1	10.91	11.17	1.12	1.1	0.26	0.75
Test	0.27	0.36	0.49	0.08	0.99	14.27	38.62	25.82	0.11	11.68	3.53	1.77	1.33	0.08	0.62
Total	0.32	0.29	0.34	0.1	0.89	15.84	33.53	25.39	0.08	11.5	8.29	1.38	1.06	0.17	0.83

To assess the models' performance quantitatively, we relied on three evaluation metrics: i) Jaccard Index or Intersection-over-Union (IoU), ii) F_1 score, iii) Overall Accuracy (OA), and iv) per class Recall. The evaluation metrics were calculated as follows:

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c} \tag{1}$$

where TP_c are true positives, FP_c are false positives and FN_c are false negatives of each c class.

Since we deal with a multi-class task, the evaluation metric for the overall assessment is the mIoU which is the average IoU over all classes c:

$$mIoU = \frac{1}{n}\sum_c^n IoU_c \tag{2}$$

The second metric is the average for each class F_1 score (Macro-F_1/ mF_1), which is the harmonic mean between Precision $P = TP/(TP + FP)$ and Recall $R = TP/(TP + FN)$:

$$mF_1 = \frac{1}{n}\sum_c^n F_{1c} = \frac{1}{n}\sum_c^n \frac{2P_cR_c}{P_c + R_c} \tag{3}$$

We also considered the Overall Accuracy:

$$OA = \frac{\sum_c^n TP_c}{\sum_c^n TP_c + FP_c} \tag{4}$$

Implementation details

In all experiments, we upsampled the 20m and 60m bands to 10m using the nearest-neighbor resampling approach. For both RF models, we sticked to the training setup as proposed by (Kikaki et al., 2022; Mikeli et al., 2022). In a more

detailed way, we employed 125 trees, each with a maximum depth of 20 nodes and used a class weighting scheme to handle the class imbalance while the annotators' confidence levels were also exploited. Regarding the spatial features, the grayscale images derived by RGB composites were utilized in order to extract the GLCM features using a 13 x 13 sliding window. We computed the Gaussian, the Sobel of the Gaussians, and the Eigenvalues of the Hessian Matrix of the Gaussians at the different scale levels using the standard deviation of $\sigma = 1, 2, 4, 8, 16$.

For the DL models, we based our implementation on MARIDA (Kikaki et al., 2022), SegNext (Guo et al., 2022), timm (Wightman, 2019) and mmsegmentation (MMSegmentation Contributors, 2020) libraries. For all models, we adapted the first input layer to handle the 11 S2 bands and the final layer to output the MADOS 15 classes. Regarding SegNeXt, we used the smallest *Tiny* (SegNeXt-T) version. The same was adopted in MariNeXt. All DL models were trained from scratch with Adam algorithm for 80 epochs with an initial learning rate of $2x10^{-4}$ minimizing the Cross-Entropy loss. After epochs 45 and 65 the learning rate was set to $2x10^{-5}$ and $2x10^{-6}$, respectively. As detailed in Table 5.5, we set the batch size for UNet to 15 and for SegNeXt and MariNeXt to 5. For data augmentations, we used 0°, 90°, 180° or 270° random rotations as well as horizontal and vertical flips. Following MARIDA, we used the same weighting scheme to alleviate the effects of the imbalanced class distribution. During training, we maintained the *exponential moving average* (EMA) of the trained parameters θ by according to $\theta' \leftarrow \alpha\theta' + (1 - \alpha)\theta$, setting $\alpha = 0.999$. We found this process extremely beneficial to produce smooth loss curves as well as robust and stable models.

Table 5.5 F1 scores of SegNeXt, U-Net and MariNeXt models over the different Batch sizes. The scores are measured on MADOS test set and averaged over five trials.

Method	Batch Size			
	5	10	15	20
U-Net	53.6	59.7	**63.8**	61.2
SegNeXt	**60.6**	58.5	54.9	52.2
MariNeXt	**76.0**	74.4	69.1	62.0

5.3 MariNeXt Performance Evaluation

In this section, we compare our method with the other employed models and evaluate its performance both quantitively and qualitatively. In addition, an ablation study is performed to assess the main components of the developed MariNeXt framework.

5.3.1 Comparing MariNeXt with ML baselines

Firstly, the performance of all applied models is quantitatively evaluated. Overall, we observe that MariNeXt achieved the best classification performance, as indicated by the highest average scores for all evaluation metrics, i.e., 76.0% for F_1, 64.3% for mIoU and 89.1% for OA (Table 5.6). Thus, MariNeXt outperformed all baselines by at least 11.6% in F1 and 11.9% in mIoU metrics.

The second highest scores were obtained by RF^{++} (i.e., 64.4% for F_1 and 52.4% for mIoU), followed by U-Net, which led to 63.8% for F_1 and 51.0% for mIoU. Regarding the rest two classifiers, the default RF reached the lowest scores for all metrics while SegNeXt achieved an improvement of +4.1% for F_1, +5.3% for mIoU and +19.5% for OA. We note that the default SegNeXt achieved the second-highest OA score (i.e., 86.6%) (Table 5.6).

Table 5.6 Evaluation of MariNeXt against ML baselines on MADOS test set using F_1, mean Intersection of Union (mIoU) and Overall Accuracy (OA) metrics, averaged over five trials. The highest scores are highlighted with bold fonts.

Method	F_1 (%)	mIoU (%)	OA (%)
RF	56.6	43.9	67.1
RF^{++} (Mikeli et al., 2022)	64.4	52.4	83.8
U-Net (Ronneberger et al., 2015)	63.8	51.0	82.9
SegNext (Guo et al., 2022)	60.6	49.2	86.6
MariNeXt	**76.0**	**64.3**	**89.1**

5.3.2 Quantitative and Qualitative evaluation

To further assess the performance of developed MariNeXt and study the agreement between the annotated data and the predictions, we present and discuss the corresponding confusion matrix (Table 5.7). The scores are demonstrated for all metrics per class on the test set. Our analysis and discussion are concentrated on the achieved F_1 scores, but similar conclusions can also be derived from the other metrics. Specifically, the highest scores (i.e., >80%) for F_1 were achieved for classes with typical optical characteristics and spectra, such as *SLWater*, water-related classes (i.e., *MWater*, *TWater*), *DenS*, *SpFA*, classes with distinct textural patterns such as *OilPlat* and *Foam*, as well as floating organisms like *Jellyfish* and *Sea snot*.

On the other hand, *Waves*, *MD* and *NatM* showed the least F_1 scores (i.e., <58%). For *Oil*, *Ship* and *SWater*, our method led to F_1 values >62%. Moreover, one can observe that *MD*, a quite challenging case, presented relatively low rates for F_1 (i.e., 51.3%) and IoU (i.e., 34.6%). Nevertheless, 70.4% of its testing pixels were correctly classified, indicating that there is a high possibility that our method detects *MD* accurately. Also, 11.1% were classified as *Ship* and 12.7% as *MWater*. This can be

attributed to the fact that plastics are spectrally similar to ships, as well as they can be probably spectrally close to *MWater* under certain circumstances (e.g., low density, submersion, polymer type, environmental conditions).

The *Oil* class, consisting of several oil types, presented high scores for both F_1 (i.e., 78.3%) and IoU (i.e., 64.4%). In addition, 70.6% of its testing data were correctly predicted as *Oil*. However, a proportion of 28.3% for *Oil* testing labeled pixels were classified as *MWater*, indicating similar spectral behavior.

Regarding artificial materials and man-made objects, *OilPlat,* characterized by distinct spatial structures, achieved a high F_1 rate (i.e., 84.3%) and a very high Recall value (i.e., 94.6%). On the contrary, *Ship* class composed of varying sizes and cruising speeds reached a 62.8% F_1 score. Almost half of the testing annotated data were correctly classified as *Ship*, while 39.3% were wrongly predicted as *OilPlat*. Apparently, these two classes share common spectral characteristics due to the similar materials they consist of; thus, their discrimination is considered challenging.

Water-related classes, such as *SLWater*, *MWater* and *TWater* held high F_1 scores of over 90%. *Foam* also achieved high F_1 score of over 80% due to its distinct spatial pattern that presents in coastal areas (wave breaking). The least accurate water-classes were *SWater* and *Waves*, with 63.7% and 37.7% F_1 scores, respectively. Recall values were 61.2% for *SWater* and 28.4% for *Waves*. This was mainly affected by the confusions between *SWater* with *MWater* or *TWater*; and *Waves* with *MWater* or *Oil*.

Floating macroalgae species presented high F_1 rates over 80%. However, as Table 5.7 indicates, there is some misclassification cases between *DenS* and *SpFA*. Specifically, 28.2% of *DenS* testing annotated data were predicted as *SpFA*. In fact, this was expected since both classes belong to the same super-class of floating macroalgae which contains mainly *Sargassum*.

Table 5.7 The resulting Confusion Matrix after applying MariNeXt on MADOS test set. The metrics of Recall, IoU, Precision and F_1 are presented. Columns represent the predicted labels and rows represent the reference labels (%). All acronyms are stated in Table 5.2.

Predicted / Ground Truth	MD	DenS	SpFA	NatM	Ship	Oil	MWater	SLWater	Foam	TWater	SWater	Waves	OilPlat	Jellyfish	Sea snot	# of pixels (test set)	Recall (%)
MD	**70.4**	0.0	2.1	0.0	11.1	0.1	12.7	0.0	0.0	0.0	0.0	3.1	0.4	0.0	0.0	973	**70.4**
DenS	0.5	**70.0**	28.2	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1303	**70.0**
SpFA	1.2	1.3	**90.4**	0.2	0.0	0.0	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1786	**90.4**
NatM	22.7	0.0	5.3	46.1	17.6	0.1	5.7	0.0	0.0	0.2	0.0	2.3	0.0	0.0	0.0	278	46.1
Ship	3.1	0.0	0.0	0.0	**52.7**	2.1	0.3	0.2	0.0	0.4	0.0	1.6	39.3	0.3	0.0	3574	**52.7**
Oil	0.0	0.0	0.0	0.0	0.0	**70.6**	28.3	0.0	0.0	0.1	0.2	0.7	0.0	0.0	0.0	51528	**70.6**
MWater	0.1	0.0	0.0	0.0	0.0	2.6	**95.7**	0.0	0.0	1.1	0.2	0.4	0.0	0.0	0.0	139465	**95.7**
SLWater	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**99.3**	0.0	0.7	0.0	0.0	0.0	0.0	0.0	93254	**99.3**
Foam	0.5	0.0	0.0	0.0	4.9	0.1	0.3	0.0	**80.6**	1.1	0.0	7.4	0.0	0.0	5.2	387	**80.6**
TWater	0.0	0.0	0.0	0.0	0.1	0.0	1.5	0.1	0.0	**89.8**	7.6	1.0	0.0	0.0	0.0	42168	**89.8**
SWater	0.0	0.0	0.0	0.0	0.0	0.0	26.1	0.0	0.0	12.4	**61.2**	0.3	0.0	0.0	0.0	12732	**61.2**
Waves	2.9	0.0	0.0	0.0	2.2	17.3	41.7	0.0	0.1	1.2	6.1	28.4	0.1	0.0	0.0	6401	28.4
OilPlat	0.1	0.0	0.0	0.0	3.5	1.1	0.3	0.0	0.0	0.0	0.0	0.5	**94.6**	0.0	0.0	4790	**94.6**
Jellyfish	1.1	0.0	0.0	0.0	0.0	1.5	0.0	0.0	0.0	0.0	0.0	13.1	0.0	**84.4**	0.0	274	**84.4**
Sea snot	22.4	0.0	0.0	0.0	0.0	7.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	**69.6**	2234	**69.6**
IoU (%)	34.6	68.6	73.0	42.7	45.9	64.4	82.9	99.2	77.3	82.1	47.1	23.2	72.9	81.8	68.9	**mIoU:**	**64.3**
Precision (%)	40.6	97.7	79.2	83.8	77.9	88.8	86.2	99.9	95.0	90.8	67.9	57.1	76.1	96.2	98.7	**OA :**	**89.1**
F1-score (%)	**51.3**	**81.2**	**84.3**	**57.2**	**62.8**	**78.3**	**90.7**	**99.6**	**87.1**	**90.1**	**63.7**	37.7	**84.3**	**89.9**	**81.4**	**F1-macro:**	**76.0**

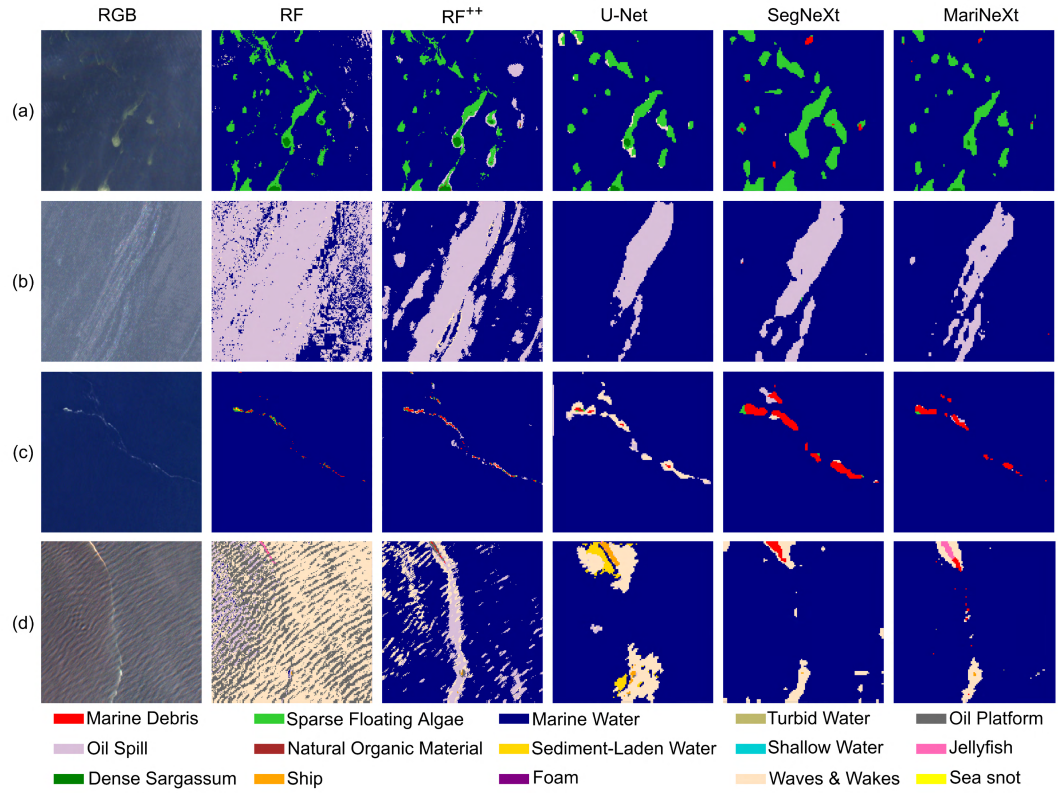


Figure 5.11 Prediction maps extracted by the applied baselines and MariNeXt. The RGB images show: a) Dense & sparse Sargassum, Marine Water and haze/thin clouds, b) Oil Spill and Marine Water, c) Marine Debris and Marine Water, and d) Jellyfish and rough sea state. The RGB images were stretched for illustration purposes.

Despite the number of testing pixels being the lowest compared to the other categories, *Jellyfish* achieved a F_1 score of almost 90%. As Table 5.7 shows, 13.1% of the testing data were classified as *Waves*, a class of relatively high-intensity values in the optical spectrum. The *Sea snot* class also reached a high F_1 value of 81.4%, while 22.4% of its testing labeled data were classified as *MD*. Moreover, the quite challenging *NatM* class, which contains vegetation and wood materials, resulted in a F_1 score of 57.2%. Again, we note that 22.4% of its testing annotations were classified as *MD*.

Apart from the quantitative assessment, a qualitative evaluation of the applied models on the test set was also performed. To describe the predictive power of MariNeXt and employed baselines, we demonstrate representative classification results in Figure 5.11. By inspecting all resulting maps, it is clean that MariNeXt achieved the highest predictive performance. In particular, it led to high-resolution results, being the most robust to isolated pixels/ spikes and weather conditions (Figure 5.11). Especially in the case of *Oil*, our model could accurately discriminate *Oil* from *MWater* in fine detail (Figure 5.11b). Although MariNeXt resulted in coarser outputs than RF models due to its convolutional nature, it also managed to predict small objects such as *Jellyfish* (Figure 5.11d).

We generally observe an agreement between our MariNeXt prediction results and metric scores. For instance, a small part of *Sargassum* pixels of low biomass densities was misclassified as *MD* (Figure 5.11a), some *MWater* pixels around debris were predicted as *Oil* (Figure 5.11c), and sparse *Jellyfish* pixels were classified as *MD* (Figure 5.11d). These misclassification cases were also observed in the quantitative results in Table 5.7.

Regarding the other ML baselines, it is observed that SegNeXt led to similar classification results with MariNeXt but with relatively coarser outputs. This also supports our framework design choices by including first-stage features and exploiting their high-resolution feature representations. U-Net also yielded satisfactory results for *Sargassum* and *Oil* while it misclassified several *MD* pixels as *Waves* and almost all *Jellyfish* as *Ship*. Concerning RF models, the default RF, which reached the lowest scores for all metrics (Table 5.7), failed to discriminate *Oil* from *MWater* and *Waves* from *MWater*, presenting the highest sensitivity to S2 sensor noise and isolated pixels (Figure 5.11b,d). However, adding spectral (indices) and texture/ spatial information substantially improved the classification performance results. In fact, RF^{++} reached the second highest metric scores (Table 5.7), providing high-resolution outputs for *Sargassum* and *Oil* classes (Figure 5.11a,b). It also achieved to predict some pixels correctly as *Waves*. On the other hand, in the case of small-width objects (i.e., *MD* and *Jellyfish*), it led to misclassification results as part of *MD* pixels was identified as *SpFA*, *Waves* and *Oil*, and *Jellyfish* was classified as *OilPlat*, *MD* and *Ship* (Figure 5.11c,d).

By evaluating the baselines qualitatively and quantitatively, we observe that although RF^{++} led to the second-highest rates (Table 5.7), its extracted maps presented enough scattered, isolated predictions. This points out that RF^{++} is prone to artifacts (e.g., S2 mixed bands, low-density pixels) and weather conditions. Besides, the qualitative interpretation showed that SegNeXt could capture the semantic context of depicted classes. Note that SegNext produces H/8 x W/8 prediction maps, which are simply nearest-neighbor upsampled. Therefore, it produces coarse-grained prediction maps leading to relatively low F_1 and mIoU scores (Table 5.7).

5.3.3 Ablation study for MariNeXt

In order to demonstrate the impact of the design choices regarding the proposed MariNeXt framework, we performed a detailed ablation by adding/ removing different components. We report the corresponding F_1 and mIoU values for each experiment in Table 5.8. After a close look at all experiments, we observe that TTA is an essential feature, as it improved the model's performance in all experiments, i.e., either adding 1^{st}-Stage Features or not, either with VSCP or not. In particular, by integrating this augmentation strategy, the model produced more robust predictions based on eight different augmentations of the same image. Without adding features

of the 1st stage, TTA resulted in an increase of 2.8% for F_1 and 2.4% for mIoU, as with the additional features the corresponding values were 1% and 2%, respectively.

Additionally, our VSCP augmentation strategy was critical for better classification results (Table 5.8). This was expected since increasing the training samples improves the model's accuracy and generalization. In the case of default SegNeXt, an improvement of 5.3% for F_1 and 3.8% for mIoU was obtained, as for the extra high-resolution features, an enhancement of 6.4% for F_1 and 6% for mIoU was performed. Table 6 also shows the importance of utilizing higher-resolution features of the 1st stage. The additional features led to an improvement of 5.7% for F_1 and 4.4% for mIoU compared to the default model.

Table 5.8 Ablation study on the proposed approach. SegNeXt is considered as a baseline for the ablation.

1st-Stage Features	VSCP	Increased BS	TTA	F_1	mIoU
✗	✗	✗	✗	60.6	49.2
✗	✗	✗	✓	63.4	51.6
✗	✓	✗	✗	65.9	53.0
✓	✗	✗	✗	66.3	53.6
✓	✗	✓	✗	66.1	53.6
✓	✗	✗	✓	67.3	55.6
✗	✓	✗	✓	69.0	56.2
✓	✓	✗	✗	72.7	59.6
✓	✓	✗	✓	**76.0***	**64.3***

VSCP: Very Simple Copy-Paste, BS: Batch Size, TTA: Test-Time Augmentation, *MariNeXt

Furthermore, we wanted to investigate if the improved performance with the additional mixed samples was due to the increased batch size. Thus, increasing the batch size with unmixed samples was also examined, leading to lower metric scores (Table 5.8). As evident in Table 5.8, the best combination is 1st-Stage Features, with VSCP and TTA, which shows the effectiveness of our method over the baseline SegNeXt. Also, we highlight that each component is orthogonal to the others, as each one benefits the performance independently. Last but not least, we experimented with four different BSs (5, 10, 15, 20), indicating that the batch size selection was crucial for the performance of the applied models (Table 5.9).

Table 5.9 F1 scores of SegNeXt, U-Net and MariNeXt models over the different Batch sizes. The scores are measured on MADOS test set and averaged over five trials.

Method	Batch Size			
	5	10	15	20
U-Net	53.6	59.7	**63.8**	61.2
SegNeXt	**60.6**	58.5	54.9	52.2
MariNeXt	**76.0**	74.4	69.1	62.0

To further understand the effectiveness of MariNeXt's components, we systematically enabled/ disabled them from our framework inspecting the class-wise F_1 scores (Table 5.10). By integrating high-resolution 1st stage features, F_1 scores were significantly increased by over 12% for the *MD*, *DenS*, *SpFA* and *Foam* classes. High-resolution features were beneficial for our task as we focused on pixel-based predictions with objects presented at significantly different scales (e.g., *Oil* vs. MD) Furthermore, by enabling TTA and VSCP, we observe an increase of over 20% for *NatM*, *Jellyfish* and *Foam* (Table 5.10), which are the most under-representative classes in the MADOS dataset. By integrating TTA, MariNeXt delivers more robust predictions, improving the classification performance. At the same time, with VSCP, MariNeXt synthesizes additional images on the fly, tackles unbalanced data and strengthens its performance on our *semantic segmentation* task.

Table 5.10 Per class evaluation of our employed strategies on MADOS test set using F1 score, averaged over five trials.

Method	F_1-Score														
	MD	DenS	SpFA	NatM	Ship	Oil	MWater	SLWater	Foam	TWater	SWater	Waves	OilPlat	Jellyfish	Sea snot
MariNeXt (Without 1st Stage)	39.5	63.5	71.8	55.3	54	77.2	87.6	99.9	61.1	87	64.9	35.6	82.9	81.6	72.8
MariNeXt (Without TTA+VSCP)	45.6	72.0	76.6	30.6	61.3	80.2	90.0	98.4	66.8	86.4	56.3	43.2	83.4	25.9	77.2
MariNeXt	51.3	81.2	84.3	57.2	62.8	78.3	90.7	99.6	87.1	90.1	63.7	37.7	84.3	90.0	81.4

5.4 Discussion

5.4.1 Potentials in detecting and discriminating marine pollutants in S2 data

Based on our experimental results, the developed MariNeXt framework can distinguish the two primary marine pollution forms, i.e., *MD* and *Oil*, on optical S2 images. The quantitative assessment presented in Table 5.7 indicated that only 0.1% of *MD* was classified as *Oil*, while 0.02% of *Oil* was predicted as *MD*. Furthermore, the calculated T-SNE plot (Figure 5.12) indicates a clear separation potential in the 2D space between these two pollutants. Based on our experiments in this study, we conclude that it seems feasible to develop operational marine monitoring solutions that consider both classes concurrently without major misclassification issues between them.

Oil Spill

Based on both quantitative (Table 5.7) and qualitative results (Figure 5.11), we reveal the most challenging cases for *Oil*. Most of these cases correspond to sparse *Oil* regions without distinct spatial patterns classified as *MWater* (Figure 5.11). Moreover, *Waves* are frequently misclassified as *Oil* (Figure 5.13a,b), (Table 5.7). The main source of these misclassifications is that *Oil*, *MWater* (including sea-state look-alikes) and *Waves* can be spectrally similar, leading to a significant overlap in T-SNE plot (Figure 5.14). We argue that these observations are due to the high spectral variations derived from different oil concentrations and weathering processes. These outcomes are in accordance with the literature since *Oil* has been proven to present spectral patterns similar to *Water* and *Waves* (Fingas and Brown, 2018; Sun et al., 2015; Trujillo-Acatitla et al., 2022). However, MariNeXt managed to tackle these similarities achieving a relatively high F_1 78.3% for *Oil*.

In order to better understand MariNeXt's efficiency, we present the learned feature space against the initial one (based on Spectral Angle Mapping) (Figure 5.15). Specifically, we extract the output features of the MariNeXt encoder, as they are given to the decoder and visualize them (Figure 5.15b). During the training process, MariNeXt managed to learn rich representations of the input data and alleviate the challenging issue of *Oil*, *Waves* and *MWater*, by separating them to a certain extent in the feature space (Figure 5.15b).

Marine Debris

For *MD*, the most competing classes were *Ship*, *Sea snot* and *NatM* (Table 5.7, Figure 5.13b,c,d). As can be observed in the corresponding T-SNE plot (Figure 5.16), despite the overall distinct cluster representations, there are several low-density *NatM and Sea snot pixels* that overlap with the *MD* cluster. Dense *MD* pixels are also depicted in the *Ship* cluster, which is in accordance with the literature (Kikaki et al., 2022). Interestingly, the main *MD* and *Sea snot* clusters are distinguished (Figure 5.16), suggesting that the considered classes can be differentiated under certain circumstances (e.g., Chlorophyll pigments presence), as was also reported by Papageorgiou et al. (2022). Although *MD* detection from current sensors is quite challenging (Martinez-Vicente, 2022), we show that state-of-the-art deep networks like MariNeXt can contribute to addressing certain overlaps, e.g., *MD* against *Ship* and *Sea snot* (Figure 5.17). Still, the most competing feature for *MD* is *NatM* due to the usually similar spatial patterns and spectral behavior (Figure 5.17b).

Water-related Classes

In general, *MWater*, *SLWater* and *TWater* are performing well with a relatively small number of false positives. However, in some cases, *SWater* is misclassified as *TWater*, and *Waves* and *SWater* are predicted as *MWater* (Table 5.7, Figure 5.11d). The corresponding T-SNE plot (Figure 5.18a) also confirms that almost all water-related

classes exhibit similar spectral characteristics. The most well-separated cluster was formed by *SLWater* (Figure 5.18a), which presents distinct spectral behavior due to the increased concentrations of suspended sediments (highest Recall in Table 5.7). Recent studies have also reported this outcome (Kikaki et al., 2022; Mikeli et al., 2022). On the other hand, the *Waves* class (with relatively high-intensity values) presents the greatest spectral variance derived from the different sea state conditions. Although MariNeXt contributes to the separation of the water-related classes, we highlight that the *Waves* class is the most challenging, and its accurate detection is crucial for the overall performance of any operational marine mapping system.

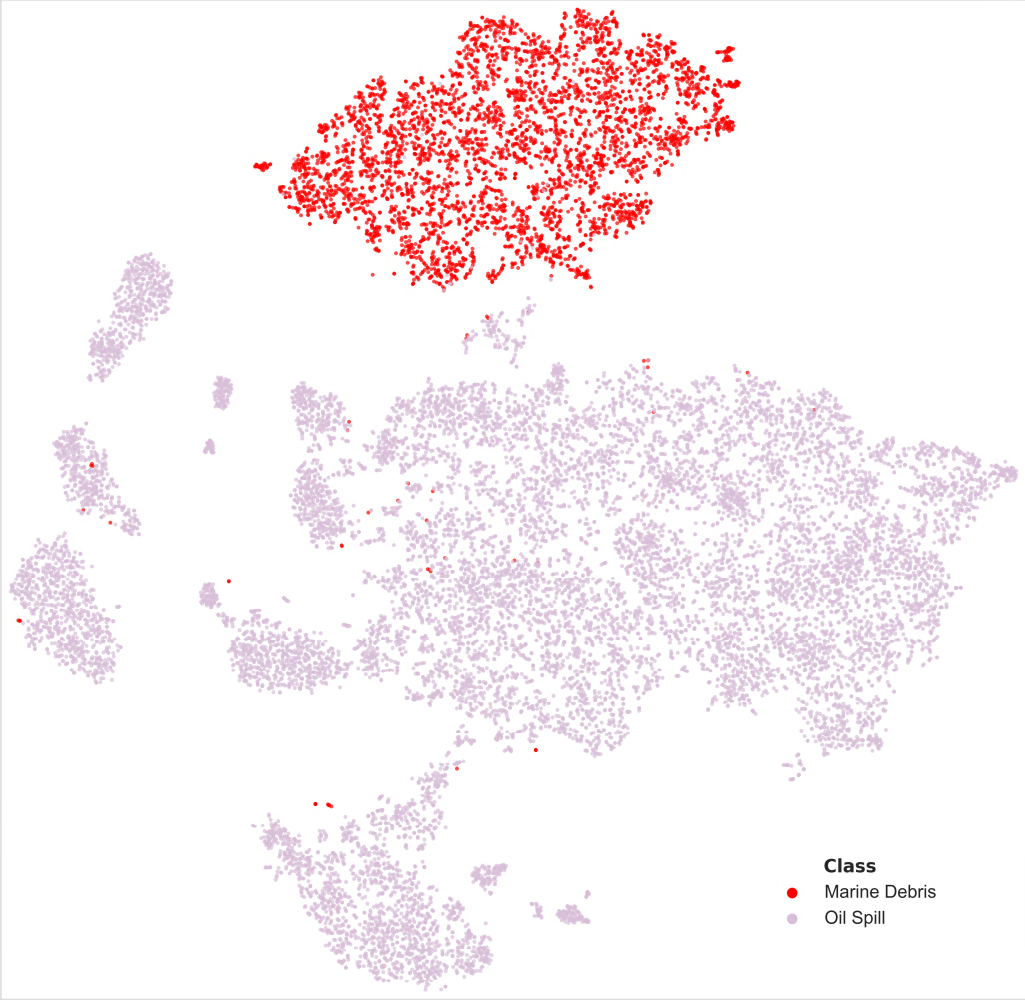


Figure 5.12 T-SNE Analysis on MADOS dataset using Spectral Angle Mapping for Marine Debris and Oil Spill.

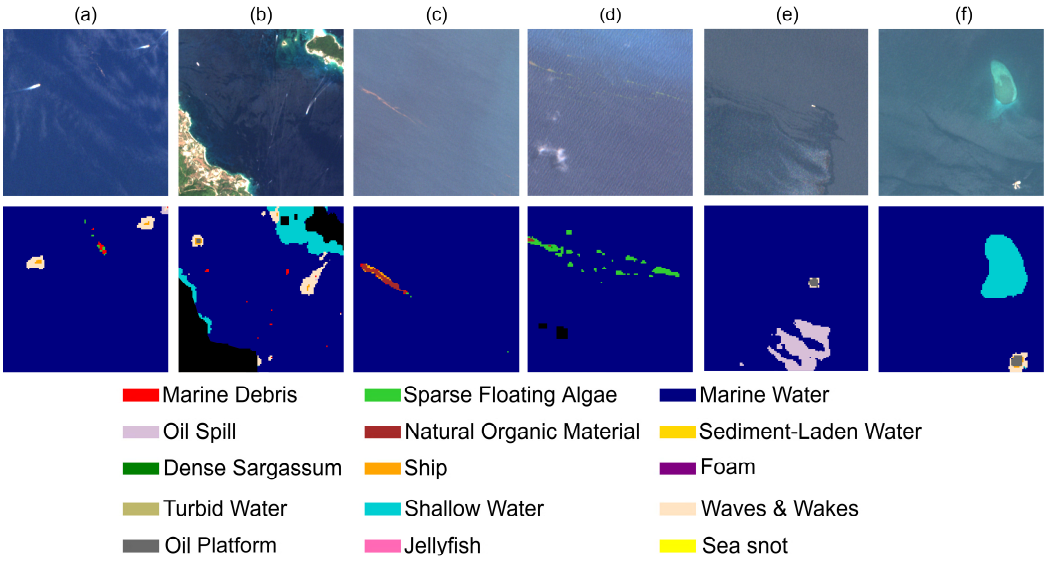


Figure 5.13 Classification results obtained by our MariNeXt framework. We demonstrate representative cases of: a) Ships, Wakes, Natural Organic Material and Marine Water, b) Ships, Wakes, Shallow & Marine Water, c) Natural Organic Material, Marine Water and haze, d) Dense & sparse Sargassum, Marine Water and haze, e) Oil Spill, Oil Platform and Marine Water, and f) Shallow & Marine Water, Oil Platform. The RGB images were stretched for illustration purposes. Land, Clouds and thin clouds are masked with black color.

Other classes

Moreover, *DenS* and *SpFA* are well-distinguishable, with relatively high F_1 scores of over 80% (Table 5.7, Figure 5.11). This is in accordance with the literature that has indicated that *Sargassum* has unique spectral behavior that differs from other floating materials and algae species (Hu et al., 2015; Qi and Hu, 2021). Nevertheless, based on our experiments and the MADOS dataset, *spFA* can be confused with *DenS* or other floating vegetation (i.e., *NatM* when vegetation is prominent) (Table 5.7, Figure 5.11 and Figure 5.13d,e). Moreover, man-made classes like *OilPlat* and *Ship*, composed of similar materials, are expectedly grouped together in the T-SNE diagram (Figure 5.19a). Interestingly the *Jellyfish* class, which does not form a distinct cluster in the T-SNE plot, is mainly dispersed in categories of relatively high-intensity values like *OilPlat* and *Ship*. Previous studies (Van der Zande et al., 2014) have indicated through laboratory observations that *Jellyfish* reflectance spectra resemble those of transparent plastic bottles. However, we demonstrate that dense *Jellyfish* trajectories can be spectrally close to dense polymer and/or artificial classes. MariNeXt managed to capture the different spatial patterns of *OilPlat*, *Ship* and *Jellyfish*, resulting in a relatively small number of misclassification cases (Figure 5.19b, Figure 5.13a,b,f,g,h).

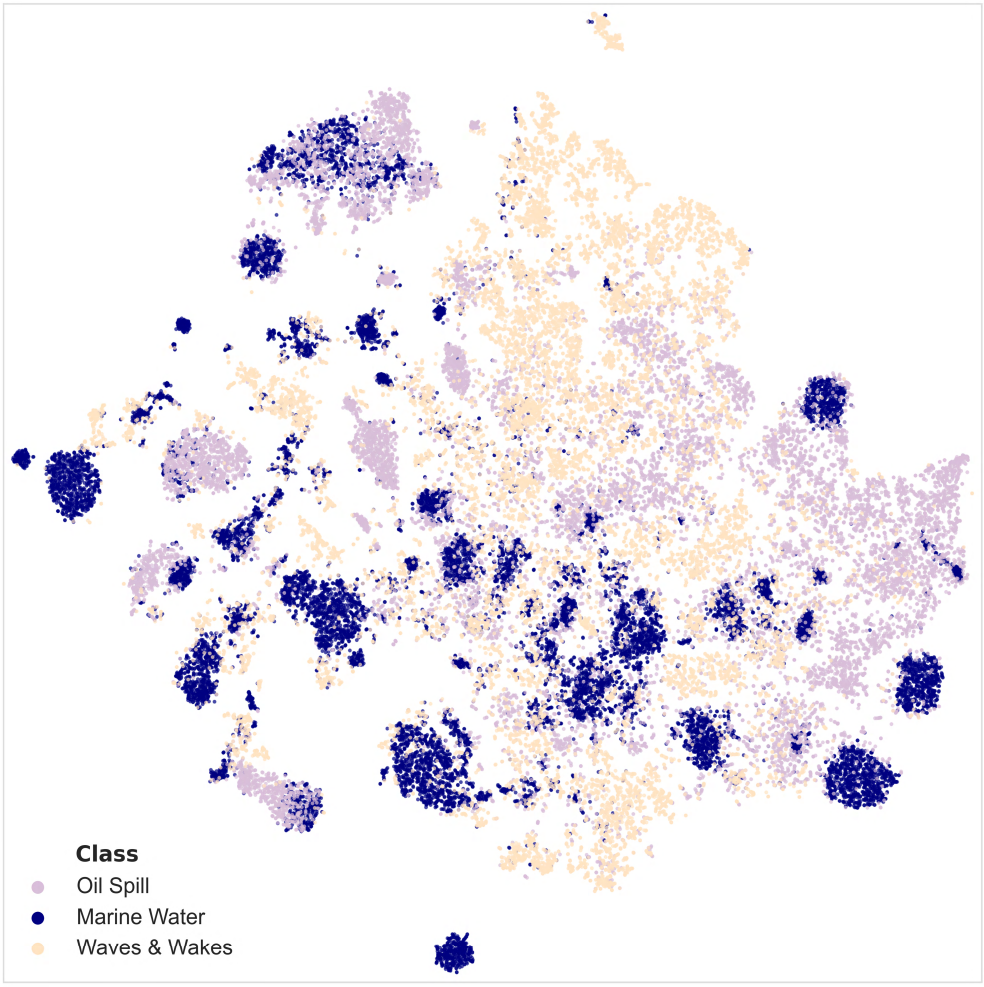


Figure 5.14 T-SNE Analysis on MADOS dataset using Spectral Angle Mapping for Oil Spill, Marine Water and Waves & Wakes.

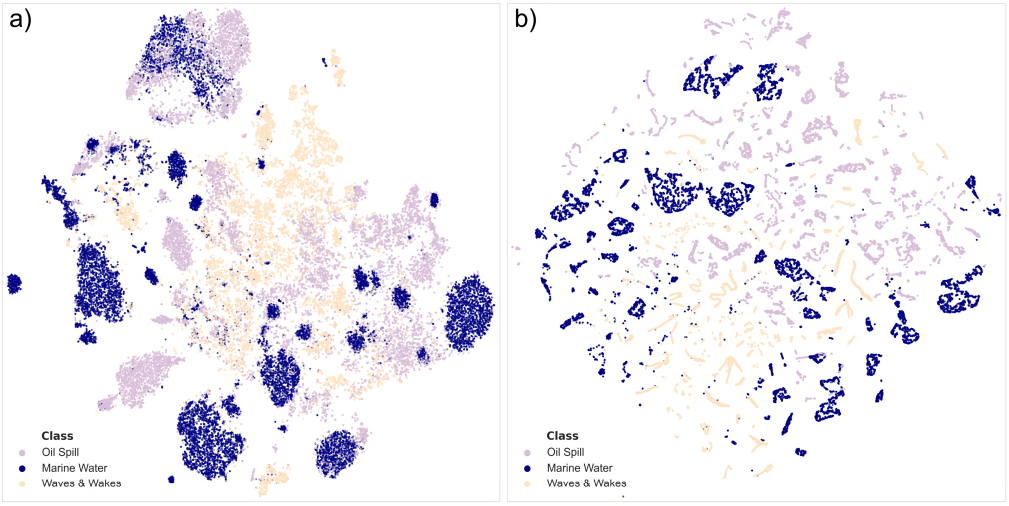


Figure 5.15 T-SNE Analysis using (a) SAM and (b) MariNeXt features for Oil Spill Waves & Wakes and Marine Water. For fair comparison, we use only validation and testing samples.

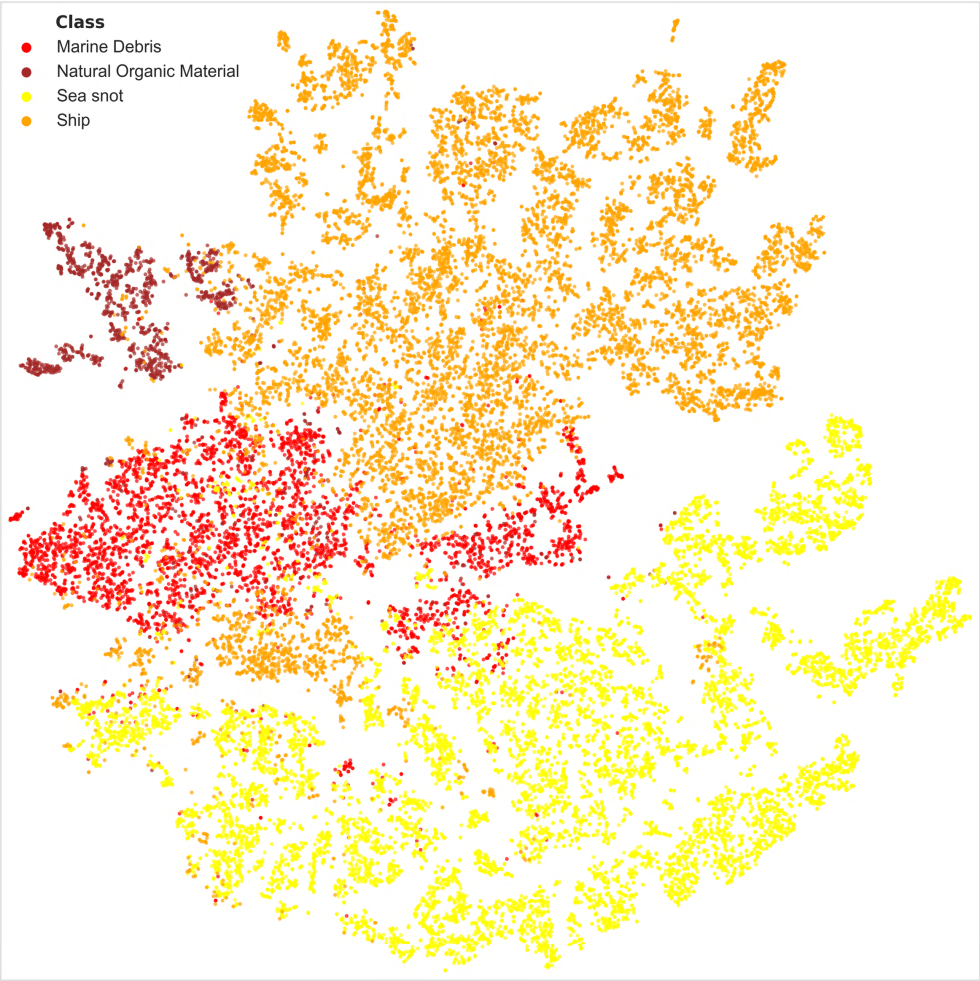


Figure 5.16 T-SNE Analysis on MADOS dataset using Spectral Angle Mapping for Marine Debris, Ship, Natural Organic Material and Sea snot.

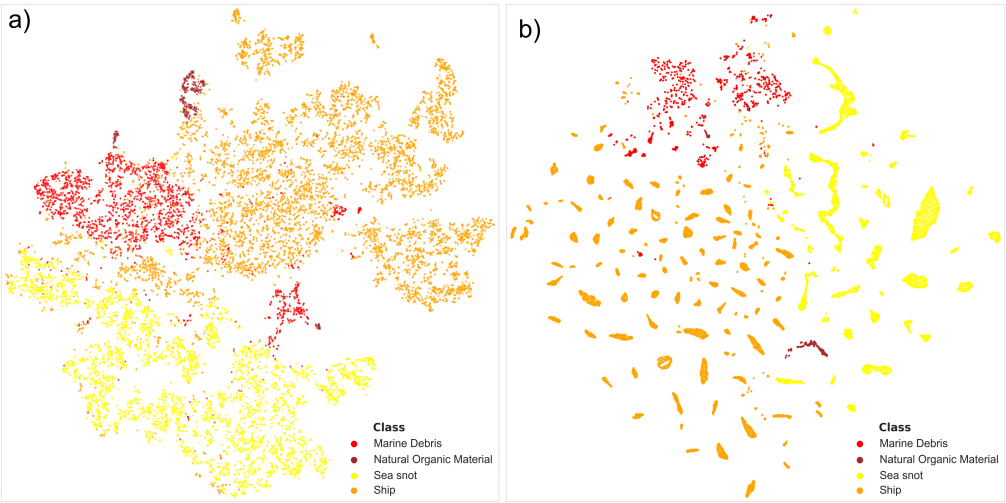


Figure 5.17 T-SNE Analysis using (a) SAM and (b) MariNeXt features for Marine Debris, Ship, Natural Organic Material and Sea snot. For fair comparison, we use only validation and testing samples.

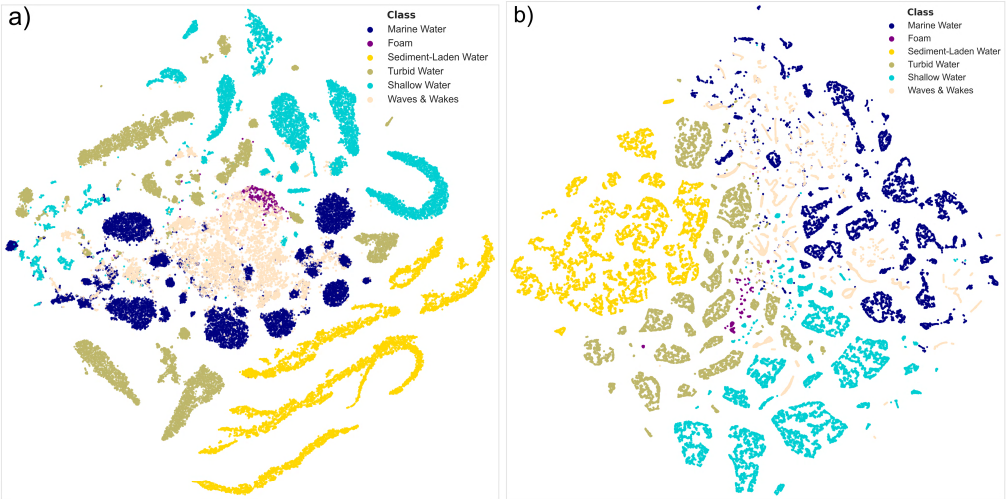


Figure 5.18 T-SNE Analysis on MADOS dataset using SAM for Marine Water, Foam, Turbid Water, Shallow Water, Waves & Wakes and Sediment-Laden Water.

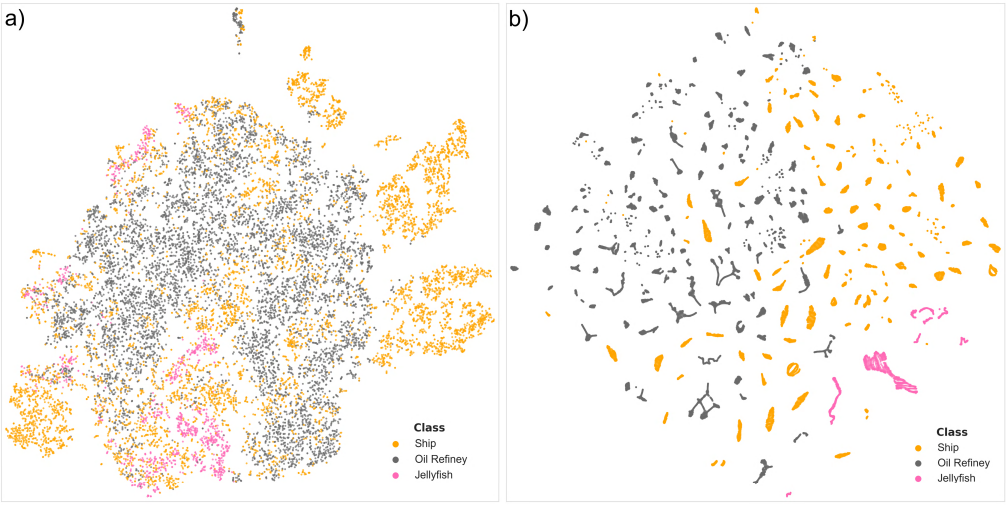


Figure 5.19 T-SNE Analysis using (a) SAM and (b) MariNeXt features for Ship, Oil Platform and Jellyfish. For fair comparison, we use only validation and testing samples.

5.4.2 Limitations & Challenges

MADOS fills a current gap in benchmark data availability for marine pollution detection, and the MariNeXt framework seems capable of classifying competing marine classes. That being said, some limitations need to be considered. First, MADOS is, by nature, an unbalanced dataset. Classes like MWater and Oil, which are abundant in the dataset, tend to be predicted more frequently than under-representative classes like Jellyfish and NatM. MADOS can be expanded in the future with other observations to enhance the predictive capabilities of MariNeXt. This will boost the detection accuracies on these less representative classes,

particularly improving the model's performance when applied to large-scale real-world applications. Furthermore, S2 data limitations (Hu, 2022), combined with diverse marine environmental conditions, and not always consistent atmospheric correction processes, led to high variability in MADOS reflectance values. These challenges need to be addressed in any operational remote sensing system. Also, during our experimental validation, we observed some systematic false positives such as Waves around OilPlat, OilPlat mixed with Ship and MD spikes (Figure 5.13). Post-processing strategies can tackle these systematic spurious correlations and keep only the main floating object.

Putting the limitations aside, MADOS is a valuable dataset that benchmarks ML algorithms for marine pollution detection from open S2 data, supporting the development of future operational marine monitoring solutions. In addition to the semantic segmentation, MADOS can contribute to other ML and remote sensing tasks, including semi-, self- or unsupervised and transfer learning. We also encourage the community to examine pan-sharpening methods or other upsampling techniques beyond Nearest Neighbor resampling that has been mostly employed.

5.5 Conclusions

This study introduces MADOS, a globally-distributed benchmark dataset for Marine Debris and Oil Spill detection from Sentinel-2 satellite imagery. MADOS consists of competing marine classes and almost 1.5 million annotated pixels, being to the best of our knowledge, the largest available Sentinel-2 dataset for marine pollutants. We have also introduced MariNeXt, a novel DL framework capable of recognizing specific spatial patterns and distinguishing competing classes that cannot be spectrally separated. Through our experiments, MariNeXt led to promising prediction maps and outperformed the other ML baselines in all quantitative metrics.

We also demonstrate that Oil and MD tend to be separable and detectable. Most false positives and negatives are related to low-density floating matters, objects composed of similar materials or presenting high-intensity values, and some water-related classes (i.e., Waves). By considering, without exceptions, the competing sea surface features and water-related classes, reliable operational monitoring solutions can be developed using the openly available multispectral data. Still, several challenges exist, including the scarcity of in situ data, the need for more balanced datasets, and the requirement for more sophisticated machine learning models capable of accommodating multi-temporal data to facilitate informed decision-making.

6 Conclusions and Future Perspectives

6.1 Conclusions

In this dissertation, we have thoroughly studied various aspects of marine pollutants detection and discrimination from other sea surface features using multitemporal multispectral satellites. The main contribution is the development of a novel framework that exploits high-resolution satellite data for marine pollution mapping at a global scale.

Based on multitemporal satellite data from three current sensors, we identified the sources, pathways and spatial distribution of marine litter in Honduras, a highly polluted area of the second-largest coral reef in the world. This information is essential and can be used to inform policies and actions to reduce marine litter and protect the marine environment. Using Sentinel-2 satellite data, two highly reliable datasets, i.e., MARIDA & MADOS, were created based on existing in situ reports and literature on marine pollution, ocean data and photo-interpretation. Machine Learning baselines were developed, and a novel mapping framework (i.e., MariNeXt) was proposed for marine pollution detection, formulating a *weakly-supervised semantic segmentation* task. Under the proposed framework, spectral and spatial features were successfully utilized, while their contribution to the classification process has been evaluated. The proposed MariNeXt framework resulted into an overall F_1 of 76% for the 15 thematic classes, achieving an improvement of 12% compared with the applied RF model that had the best performance. Furthermore, aiming to contribute to future operational solutions, MariNeXt framework capabilities and findings regarding the discrimination of marine pollutants from various sea surface features were discussed in detail.

To the best of our knowledge, there is no other dataset and framework for marine pollution detection at 10 m spatial resolution with such a thematic analysis and quite promising outcomes. Additionally, we highly regard that we openly provided a quick start code, an interactive online material with our detailed statistical and spectral analysis and the created databases to trigger the research community for further experiments.

6.2 Future work

In this section, we present some suggestions for further experiments inspired by the work of this dissertation:

- ✓ Improve MariNext predictive capabilities by adding more observations from marine pollutants and other features that are less common on the sea surface or their identification is more challenging (e.g., Jellyfish).

- ✓ Increase the thematic resolution by additionally including types of classes (e.g., add Oil types and different macroalgae species).
- ✓ Apply our proposed methodology to satellite data with higher spatial and/or temporal resolution (e.g., Planet satellite imagery) or data from new sensors.
- ✓ Evaluate existing cloud masking techniques and develop new methods for accurate cloud and cloud shadow detection.
- ✓ Apply our methodology to multitemporal satellite data to assess marine pollution levels in our study areas (e.g., Haiti) for the last seven years. Evaluate our framework in different regions globally.
- ✓ Utilize other state-of-the-art computer vision approaches in order to exploit the vast majority of the non-annotated remote sensing data
- ✓ Exploit temporal features (e.g., time series) to enhance marine pollutants detection and monitoring potentially.
- ✓ Evaluate the synergy between the multispectral Sentinel-2 satellite images and radar data (e.g., Sentinel-1) for marine pollution monitoring.
- ✓ Investigate advanced pan-sharpening techniques to increase the capabilities of the Sentinel-2 sensor.
- ✓ Organize fieldwork and perform in situ measurements to validate further satellite results and examine spectra of different polymers, oil types and algae species.

7 BIBLIOGRAPHY

Abou Samra, R.M., Ali, R.R., 2022. Monitoring of oil spill in the offshore zone of the Nile Delta using Sentinel data. Marine Pollution Bulletin 179, 113718. https://doi.org/10.1016/j.marpolbul.2022.113718

Acuña-Ruz, T., Uribe, D., Taylor, R., Amézquita, L., Guzmán, M.C., Merrill, J., Martínez, P., Voisin, L., Mattar B., C., 2018. Anthropogenic marine debris over beaches: Spectral characterization for remote sensing applications. Remote Sensing of Environment 217, 309–322. https://doi.org/10.1016/j.rse.2018.08.008

Afgatiani, P.M., Putri, F.A., Suhadha, A.G., Ibrahim, A., 2020. Determination of Sentinel-2 spectral reflectance to detect oil spill on the sea surface. Sustinere: Journal of Environment and Sustainability 4, 144–154. https://doi.org/10.22515/sustinere.jes.v4i3.115

Agustin, A.E., Merrifield, M.A., Potemra, J.T., Morishige, C., 2015. Temporal variability of marine debris deposition at Tern Island in the Northwestern Hawaiian Islands. Marine Pollution Bulletin 101, 200–207. https://doi.org/10.1016/j.marpolbul.2015.10.076

Airbus, 2021, 2021. Airbus Ship Detection Challenge [WWW Document]. URL https://kaggle.com/c/airbus-ship-detection (accessed 6.16.21).

Al-Ruzouq, R., Gibril, M.B.A., Shanableh, A., Kais, A., Hamed, O., Al-Mansoori, S., Khalil, M.A., 2020. Sensors, Features, and Machine Learning for Oil Spill Detection and Monitoring: A Review. Remote Sensing 12, 3338. https://doi.org/10.3390/rs12203338

Althawadi, J.J.A., Hashim, M., 2019. AN APPROACH OF VICARIOUS CALIBRATION OF SENTINEL-2 SATELLITE MULTISPECTRAL IMAGE BASED ON SPECTRAL LIBRARY FOR MAPPING OIL SPILLS, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the ISPRS WG IV/I
6th International Conference on Geomatics and Geospatial Technology (GGT 2019) (Volume XLII-4/W16) - 1–3 October 2019, Kuala Lumpur, Malaysia, Copernicus GmbH, pp. 117–121. https://doi.org/10.5194/isprs-archives-XLII-4-W16-117-2019

Anderson, Z.T., Cundy, A.B., Croudace, I.W., Warwick, P.E., Celis-Hernandez, O., Stead, J.L., 2018. A rapid method for assessing the accumulation of microplastics in the sea surface microlayer (SML) of estuarine systems. Sci Rep 8, 9428. https://doi.org/10.1038/s41598-018-27612-w

Aparcana, S., 2017. Approaches to formalization of the informal waste sector into municipal solid waste management systems in low- and middle-income countries: Review of barriers and success factors. Waste Management 61, 593–607. https://doi.org/10.1016/j.wasman.2016.12.028

Argamosa, R.J.L., Blanco, A.C., Reyes, R.B., 2022. Mapping Oil Spills on sea surface from Sentinel-2 images using Proncipal Components and categorical boosting, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the ISPRS Joint International Conference Geospatial Asia-Europe 2021 and GeoAdvances 2021;6 October 2021, online, Copernicus GmbH, pp. 33–38. https://doi.org/10.5194/isprs-archives-XLVI-4-W3-2021-33-2022

Arias, M., Sumerot, R., Delaney, J., Coulibaly, F., Cozar, A., Aliani, S., Suaria, G., Papadopoulou, T., Corradi, P., 2021. Advances on Remote Sensing of Windrows as Proxies for Marine Litter Based on Sentinel-2/MSI Datasets, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the

2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 1126–1129. https://doi.org/10.1109/IGARSS47720.2021.9555139

Arii, M., Koiwa, M., Aoki, Y., 2014. Applicability of SAR to Marine Debris Surveillance After the Great East Japan Earthquake. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 1729–1744. https://doi.org/10.1109/JSTARS.2014.2308550

Atwood, E.C., Falcieri, F.M., Piehl, S., Bochow, M., Matthies, M., Franke, J., Carniel, S., Sclavo, M., Laforsch, C., Siegert, F., 2019. Coastal accumulation of microplastic particles emitted from the Po River, Northern Italy: Comparing remote sensing and hydrodynamic modelling with in situ sample collections. Marine Pollution Bulletin 138, 561–574. https://doi.org/10.1016/j.marpolbul.2018.11.045

Bao, Z., Sha, J., Li, X., Hanchiso, T., Shifaw, E., 2018. Monitoring of beach litter by automatic interpretation of unmanned aerial vehicle images using the segmentation threshold method. Marine Pollution Bulletin 137, 388–398. https://doi.org/10.1016/j.marpolbul.2018.08.009

BAOAC, 2007, n.d. Bonn Agreement | Working Together For Cleaner Seas [WWW Document]. Bonn Agreement. URL https://www.bonnagreement.org/ (accessed 1.4.23).

Barboza, L.G.A., Dick Vethaak, A., Lavorante, B.R.B.O., Lundebye, A.-K., Guilhermino, L., 2018. Marine microplastic debris: An emerging issue for food security, food safety and human health. Marine Pollution Bulletin 133, 336–348. https://doi.org/10.1016/j.marpolbul.2018.05.047

Basu, B., Sannigrahi, S., Sarkar Basu, A., Pilla, F., 2021. Development of Novel Classification Algorithms for Detection of Floating Plastic Debris in Coastal Waterbodies Using Multispectral Sentinel-2 Remote Sensing Imagery. Remote Sensing 13, 1598. https://doi.org/10.3390/rs13081598

Beaumont, N.J., Aanesen, M., Austen, M.C., Börger, T., Clark, J.R., Cole, M., Hooper, T., Lindeque, P.K., Pascoe, C., Wyles, K.J., 2019. Global ecological, social and economic impacts of marine plastic. Marine Pollution Bulletin 142, 189–195. https://doi.org/10.1016/j.marpolbul.2019.03.022

Bellou, N., Gambardella, C., Karantzalos, K., Monteiro, J.G., Canning-Clode, J., Kemna, S., Arrieta-Giron, C.A., Lemmen, C., 2021. Global assessment of innovative solutions to tackle marine litter. Nat Sustain 4, 516–524. https://doi.org/10.1038/s41893-021-00726-2

Biermann, L., Clewley, D., Martinez-Vicente, V., Topouzelis, K., 2020. Finding Plastic Patches in Coastal Waters using Optical Satellite Data. Sci Rep 10, 5364. https://doi.org/10.1038/s41598-020-62298-z

Booth, H., Ma, W., Karakus, O., 2022. High-precision Density Mapping of Marine Debris and Floating Plastics via Satellite Imagery. https://doi.org/10.48550/arXiv.2210.05468

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cadiou, J.-F., Gerigny, O., Koren, Š., Zeri, C., Kaberi, H., Alomar, C., Panti, C., Fossi, M.C., Adamopoulou, A., Digka, N., Deudero, S., Concato, M., Carbonell, A., Baini, M., Galli, M., Galgani, F., 2020. Lessons learned from an intercalibration exercise on the quantification and characterisation of microplastic particles in sediment and water samples. Marine Pollution Bulletin 154, 111097. https://doi.org/10.1016/j.marpolbul.2020.111097

Canonico, G., Buttigieg, P.L., Montes, E., Muller-Karger, F.E., Stepien, C., Wright, D., Benson, A., Helmuth, B., Costello, M., Sousa-Pinto, I., Saeedi, H., Newton, J., Appeltans, W., Bednaršek, N., Bodrossy, L., Best, B.D., Brandt, A., Goodwin, K.D., Iken, K., Marques, A.C., Miloslavich, P., Ostrowski, M., Turner, W., Achterberg, E.P., Barry, T., Defeo, O.,

Bigatti, G., Henry, L.-A., Ramiro-Sánchez, B., Durán, P., Morato, T., Roberts, J.M., García-Alegre, A., Cuadrado, M.S., Murton, B., 2019. Global Observational Needs and Resources for Marine Biodiversity. Frontiers in Marine Science 6.

Carrillo, L., Johns, E.M., Smith, R.H., Lamkin, J.T., Largier, J.L., 2015. Pathways and Hydrography in the Mesoamerican Barrier Reef System Part 1: Circulation. Continental Shelf Research 109, 164–176. https://doi.org/10.1016/j.csr.2015.09.014

Chávez, V., Uribe-Martínez, A., Cuevas, E., Rodríguez-Martínez, R.E., van Tussenbroek, B.I., Francisco, V., Estévez, M., Celis, L.B., Monroy-Velázquez, L.V., Leal-Bautista, R., Álvarez-Filip, L., García-Sánchez, M., Masia, L., Silva, R., 2020. Massive Influx of Pelagic Sargassum spp. on the Coasts of the Mexican Caribbean 2014–2020: Challenges and Opportunities. Water 12, 2908. https://doi.org/10.3390/w12102908

Chérubin, L.M., Kuchinke, C.P., Paris, C.B., 2008. Ocean circulation and terrestrial runoff dynamics in the Mesoamerican region from spectral optimization of SeaWiFS data and a high resolution simulation. Coral Reefs 27, 503–519. https://doi.org/10.1007/s00338-007-0348-1

Ciappa, A.C., 2022. Marine Litter Detection by Sentinel-2: A Case Study in North Adriatic (Summer 2020). Remote Sensing 14, 2409. https://doi.org/10.3390/rs14102409

Ciappa, A.C., 2021. Marine plastic litter detection offshore Hawai'i by Sentinel-2. Marine Pollution Bulletin 168, 112457. https://doi.org/10.1016/j.marpolbul.2021.112457

Ciocarlan, A., Stoian, A., 2021. Ship Detection in Sentinel 2 Multi-Spectral Images with Self-Supervised Learning. Remote Sensing 13, 4255. https://doi.org/10.3390/rs13214255

Condon, R.H., Duarte, C.M., Pitt, K.A., Robinson, K.L., Lucas, C.H., Sutherland, K.R., Mianzan, H.W., Bogeberg, M., Purcell, J.E., Decker, M.B., Uye, S., Madin, L.P., Brodeur, R.D., Haddock, S.H.D., Malej, A., Parry, G.D., Eriksen, E., Quiñones, J., Acha, M., Harvey, M., Arthur, J.M., Graham, W.M., 2013. Recurrent jellyfish blooms are a consequence of global oscillations. Proceedings of the National Academy of Sciences 110, 1000–1005. https://doi.org/10.1073/pnas.1210920110

Cordova, M.R., Nurhati, I.S., 2019. Major sources and monthly variations in the release of land-derived marine debris from the Greater Jakarta area, Indonesia. Sci Rep 9, 18730. https://doi.org/10.1038/s41598-019-55065-2

Cózar, A., Aliani, S., Basurko, O.C., Arias, M., Isobe, A., Topouzelis, K., Rubio, A., Morales-Caselles, C., 2021. Marine Litter Windrows: A Strategic Target to Understand and Manage the Ocean Plastic Pollution. Front. Mar. Sci. 8. https://doi.org/10.3389/fmars.2021.571796

Cózar, A., Martí, E., Duarte, C.M., García-de-Lomas, J., van Sebille, E., Ballatore, T.J., Eguíluz, V.M., González-Gordillo, J.I., Pedrotti, M.L., Echevarría, F., Troublè, R., Irigoien, X., 2017. The Arctic Ocean as a dead end for floating plastics in the North Atlantic branch of the Thermohaline Circulation. Science Advances 3, e1600582. https://doi.org/10.1126/sciadv.1600582

Cuevas, E., Uribe-Martínez, A., Liceaga-Correa, M. de los Á., 2018. A satellite remote-sensing multi-index approach to discriminate pelagic Sargassum in the waters of the Yucatan Peninsula, Mexico. International Journal of Remote Sensing 39, 3608–3627. https://doi.org/10.1080/01431161.2018.1447162

Danovaro, R., Carugati, L., Berzano, M., Cahill, A.E., Carvalho, S., Chenuil, A., Corinaldesi, C., Cristina, S., David, R., Dell'Anno, A., Dzhembekova, N., Garcés, E., Gasol, J.M., Goela, P., Féral, J.-P., Ferrera, I., Forster, R.M., Kurekin, A.A., Rastelli, E., Marinova, V., Miller, P.I., Moncheva, S., Newton, A., Pearman, J.K., Pitois, S.G., Reñé, A., Rodríguez-Ezpeleta, N., Saggiomo, V., Simis, S.G.H., Stefanova, K., Wilson, C., Lo Martire, M., Greco, S., Cochrane, S.K.J., Mangoni, O., Borja, A., 2016. Implementing and Innovating Marine Monitoring Approaches for Assessing Marine Environmental Status. Frontiers in Marine Science 3.

De Dominicis, M., Bruciaferri, D., Gerin, R., Pinardi, N., Poulain, P.M., Garreau, P., Zodiatis, G., Perivoliotis, L., Fazioli, L., Sorgente, R., Manganiello, C., 2016. A multi-model assessment of the impact of currents, waves and wind in modelling surface drifters and oil spill. Deep Sea Research Part II: Topical Studies in Oceanography, Physical, chemical and biological observations and modelling of oil spills in the Mediterranean Sea 133, 21–38. https://doi.org/10.1016/j.dsr2.2016.04.002

Descloitres, J., Minghelli, A., Steinmetz, F., Chevalier, C., Chami, M., Berline, L., 2021. Revisited Estimation of Moderate Resolution Sargassum Fractional Coverage Using Decametric Satellite Data (S2-MSI). Remote Sensing 13, 5106. https://doi.org/10.3390/rs13245106

Dierssen, H.M., 2019. Hyperspectral Measurements, Parameterizations, and Atmospheric Correction of Whitecaps and Foam From Visible to Shortwave Infrared for Ocean Color Remote Sensing. Front. Earth Sci. 7. https://doi.org/10.3389/feart.2019.00014

Dogliotti, A.I., Ruddick, K.G., Nechad, B., Doxaran, D., Knaeps, E., 2015. A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters. Remote Sensing of Environment 156, 157–168. https://doi.org/10.1016/j.rse.2014.09.020

Dong, Y., Liu, Y., Hu, C., MacDonald, I.R., Lu, Y., 2022. Chronic oiling in global oceans. Science 376, 1300–1304. https://doi.org/10.1126/science.abm5940

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. Presented at the International Conference on Learning Representations.

Duarte, M.M., Azevedo, L., 2023. Automatic Detection and Identification of Floating Marine Debris Using Multi-Spectral Satellite Imagery. IEEE Transactions on Geoscience and Remote Sensing 1–1. https://doi.org/10.1109/TGRS.2023.3283607

Dvornik, N., Mairal, J., Schmid, C., 2018. Modeling Visual Context Is Key to Augmenting Object Detection Datasets, in: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), Computer Vision – ECCV 2018, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 375–391. https://doi.org/10.1007/978-3-030-01258-8_23

Dwibedi, D., Misra, I., Hebert, M., 2017. Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection, in: 2017 IEEE International Conference on Computer Vision (ICCV). Presented at the 2017 IEEE International Conference on Computer Vision (ICCV), pp. 1310–1319. https://doi.org/10.1109/ICCV.2017.146

Eriksen, M., Liboiron, M., Kiessling, T., Charron, L., Alling, A., Lebreton, L., Richards, H., Roth, B., Ory, N.C., Hidalgo-Ruz, V., Meerhoff, E., Box, C., Cummins, A., Thiel, M., 2018. Microplastic sampling with the AVANI trawl compared to two neuston trawls in the Bay of Bengal and South Pacific. Environmental Pollution 232, 430–439. https://doi.org/10.1016/j.envpol.2017.09.058

Everingham, M., Eslami, S.M.A., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A., 2015. The Pascal Visual Object Classes Challenge: A Retrospective. Int J Comput Vis 111, 98–136. https://doi.org/10.1007/s11263-014-0733-5

Fang, H.-S., Sun, J., Wang, R., Gou, M., Li, Y.-L., Lu, C., 2019. InstaBoost: Boosting Instance Segmentation via Probability Map Guided Copy-Pasting, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Presented at the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 682–691. https://doi.org/10.1109/ICCV.2019.00077

Fingas, M., Brown, C.E., 2018. A Review of Oil Spill Remote Sensing. Sensors 18, 91. https://doi.org/10.3390/s18010091

Galgani, F., Hanke, G., Maes, T., 2015. Global Distribution, Composition and Abundance of Marine Litter, in: Bergmann, M., Gutow, L., Klages, M. (Eds.), Marine Anthropogenic Litter. Springer International Publishing, Cham, pp. 29–56. https://doi.org/10.1007/978-3-319-16510-3_2

Garaba, S.P., Acuña-Ruz, T., Mattar, C.B., 2020. Hyperspectral longwave infrared reflectance spectra of naturally dried algae, anthropogenic plastics, sands and shells. Earth System Science Data 12, 2665–2678. https://doi.org/10.5194/essd-12-2665-2020

Garaba, S.P., Aitken, J., Slat, B., Dierssen, H.M., Lebreton, L., Zielinski, O., Reisser, J., 2018. Sensing Ocean Plastics with an Airborne Hyperspectral Shortwave Infrared Imager. Environ. Sci. Technol. 52, 11699–11707. https://doi.org/10.1021/acs.est.8b02855

Garaba, S.P., Dierssen, H.M., 2020. Hyperspectral ultraviolet to shortwave infrared characteristics of marine-harvested, washed-ashore and virgin plastics. Earth System Science Data 12, 77–86. https://doi.org/10.5194/essd-12-77-2020

Garaba, S.P., Dierssen, H.M., 2018. An airborne remote sensing case study of synthetic hydrocarbon detection using short wave infrared absorption features identified from marine-harvested macro- and microplastics. Remote Sensing of Environment 205, 224–235. https://doi.org/10.1016/j.rse.2017.11.023

Garcia-Garin, O., Monleón-Getino, T., López-Brosa, P., Borrell, A., Aguilar, A., Borja-Robalino, R., Cardona, L., Vighi, M., 2021. Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a web application in R. Environmental Pollution 273, 116490. https://doi.org/10.1016/j.envpol.2021.116490

Garcia-Pineda, O., Staples, G., Jones, C.E., Hu, C., Holt, B., Kourafalou, V., Graettinger, G., DiPinto, L., Ramirez, E., Streett, D., Cho, J., Swayze, G.A., Sun, S., Garcia, D., Haces-Garcia, F., 2020. Classification of oil spill by thicknesses using multiple remote sensors. Remote Sensing of Environment 236, 111421. https://doi.org/10.1016/j.rse.2019.111421

Ge, Z., Shi, H., Mei, X., Dai, Z., Li, D., 2016. Semi-automatic recognition of marine debris on beaches. Sci Rep 6, 25759. https://doi.org/10.1038/srep25759

Geng, Z., Guo, M.-H., Chen, H., Li, X., Wei, K., Lin, Z., 2022. Is Attention Better Than Matrix Decomposition? Presented at the International Conference on Learning Representations.

Ghiasi, G., Cui, Y., Srinivas, A., Qian, R., Lin, T.-Y., Cubuk, E.D., Le, Q.V., Zoph, B., 2021. Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2917–2927. https://doi.org/10.1109/CVPR46437.2021.00294

Goldstein, M.C., Titmus, A.J., Ford, M., 2013. Scales of Spatial Heterogeneity of Plastic Marine Debris in the Northeast Pacific Ocean. PLOS ONE 8, e80020. https://doi.org/10.1371/journal.pone.0080020

Gonzaga, M.L.R., Wong, M.T.S., Blanco, A.C., Principe, J.A., 2021. Utilization of SENTINEL-2 Imagery in the Estimation of Plastics among Floating Debris Along the Coast of Manila Bay. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 46W6, 177–184. https://doi.org/10.5194/isprs-archives-XLVI-4-W6-2021-177-2021

Gonzalo-Martín, C., García-Pedrero, A., Lillo-Saavedra, M., 2021. Improving deep learning sorghum head detection through test time augmentation. Computers and Electronics in Agriculture 186, 106179. https://doi.org/10.1016/j.compag.2021.106179

Gower, J., Young, E., King, S., 2013. Satellite images suggest a new Sargassum source region in 2011. Remote Sensing Letters 4, 764–773. https://doi.org/10.1080/2150704X.2013.796433

Guo, M.-H., Lu, C.-Z., Hou, Q., Liu, Z.-N., Cheng, M.-M., Hu, S., 2022. SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation. Presented at the Advances in Neural Information Processing Systems.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Heiselberg, H., 2020. Ship-Iceberg Classification in SAR and Multispectral Satellite Images with Neural Networks. Remote Sensing 12, 2353. https://doi.org/10.3390/rs12152353

Heiselberg, H., 2016. A Direct and Fast Methodology for Ship Recognition in Sentinel-2 Multispectral Imagery. Remote Sensing 8, 1033. https://doi.org/10.3390/rs8121033

Heiselberg, P., Heiselberg, H., 2017. Ship-Iceberg Discrimination in Sentinel-2 Multispectral Imagery by Supervised Classification. Remote Sensing 9, 1156. https://doi.org/10.3390/rs9111156

Hettiarachchi, H., Ryu, S., Caucci, S., Silva, R., 2018. Municipal Solid Waste Management in Latin America and the Caribbean: Issues and Potential Solutions from the Governance Perspective. Recycling 3, 19. https://doi.org/10.3390/recycling3020019

Hu, C., 2022. Remote detection of marine debris using Sentinel-2 imagery: A cautious note on spectral interpretations. Marine Pollution Bulletin 183, 114082. https://doi.org/10.1016/j.marpolbul.2022.114082

Hu, C., 2021. Remote detection of marine debris using satellite observations in the visible and near infrared spectral range: Challenges and potentials. Remote Sensing of Environment 259, 112414. https://doi.org/10.1016/j.rse.2021.112414

Hu, C., 2009. A novel ocean color index to detect floating algae in the global oceans. Remote Sensing of Environment 113, 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012

Hu, C., Feng, L., Hardy, R.F., Hochberg, E.J., 2015. Spectral and spatial requirements of remote measurements of pelagic Sargassum macroalgae. Remote Sensing of Environment, Special Issue on the Hyperspectral Infrared Imager (HyspIRI) 167, 229–246. https://doi.org/10.1016/j.rse.2015.05.022

Hu, C., Li, X., Pichel, W.G., Muller-Karger, F.E., 2009. Detection of natural oil slicks in the NW Gulf of Mexico using MODIS imagery. Geophysical Research Letters 36. https://doi.org/10.1029/2008GL036119

Hu, C., Lu, Y., Sun, S., Liu, Y., 2021. Optical Remote Sensing of Oil Spills in the Ocean: What Is Really Possible? Journal of Remote Sensing 2021. https://doi.org/10.34133/2021/9141902

Hu, C., Qi, L., Xie, Y., Zhang, S., Barnes, B.B., 2022. Spectral characteristics of sea snot reflectance observed from satellites: Implications for remote sensing of marine debris. Remote Sensing of Environment 269, 112842. https://doi.org/10.1016/j.rse.2021.112842

Ivanov, A.Yu., Gerivani, H., 2021. Oil leaking and seeping site in the Persian Gulf detected and studied by satellite observations. Marine Georesources & Geotechnology 39, 1481–1496. https://doi.org/10.1080/1064119X.2020.1853286

Jakovljevic, G., Govedarica, M., Alvarez-Taboada, F., 2020. A Deep Learning Model for Automatic Plastic Mapping Using Unmanned Aerial Vehicle (UAV) Data. Remote Sensing 12, 1515. https://doi.org/10.3390/rs12091515

Jamali, A., Mahdianpari, M., 2021. A Cloud-Based Framework for Large-Scale Monitoring of Ocean Plastics Using Multi-Spectral Satellite Imagery and Generative Adversarial Network. Water 13, 2553. https://doi.org/10.3390/w13182553

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L., 2015. Plastic waste inputs from land into the ocean. Science 347, 768–771. https://doi.org/10.1126/science.1260352

Jang, Y.C., Lee, J., Hong, S., Mok, J.Y., Kim, K.S., Lee, Y.J., Choi, H.-W., Kang, H., Lee, S., 2014. Estimation of the annual flow and stock of marine debris in South Korea for management purposes. Marine Pollution Bulletin 86, 505–511. https://doi.org/10.1016/j.marpolbul.2014.06.021

Kanjir, U., Greidanus, H., Oštir, K., 2018. Vessel detection and classification from spaceborne optical images: A literature survey. Remote Sensing of Environment 207, 1–26. https://doi.org/10.1016/j.rse.2017.12.033

Kataoka, T., Murray, C.C., Isobe, A., 2018. Quantification of marine macro-debris abundance around Vancouver Island, Canada, based on archived aerial photographs processed by projective transformation. Marine Pollution Bulletin, SI: Japanese Tsunami Debris 132, 44–51. https://doi.org/10.1016/j.marpolbul.2017.08.060

Kataoka, T., Nihei, Y., Kudou, K., Hinata, H., 2019. Assessment of the sources and inflow processes of microplastics in the river environments of Japan. Environmental Pollution 244, 958–965. https://doi.org/10.1016/j.envpol.2018.10.111

Kavzoglu, T., Goral, M., 2022. Google Earth Engine for Monitoring Marine Mucilage: Izmit Bay in Spring 2021. Hydrology 9, 135. https://doi.org/10.3390/hydrology9080135

Kavzoğlu, T., Tonbul, H., Çölkesen, İ., Sefercik, U.G., 2021. The Use of Object-Based Image Analysis for Monitoring 2021 Marine Mucilage Bloom in the Sea of Marmara. International Journal of Environment and Geoinformatics 8, 529–536. https://doi.org/10.30897/ijegeo.990875

Kikaki, A., Karantzalos, K., Power, C.A., Raitsos, D.E., 2020. Remotely Sensing the Source and Transport of Marine Plastic Debris in Bay Islands of Honduras (Caribbean Sea). Remote Sensing 12, 1727. https://doi.org/10.3390/rs12111727

Kikaki, K., Kakogeorgiou, I., Mikeli, P., Raitsos, D.E., Karantzalos, K., 2022. MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data. PLOS ONE 17, e0262247. https://doi.org/10.1371/journal.pone.0262247

Knaeps, E., Sterckx, S., Strackx, G., Mijnendonckx, J., Moshtaghi, M., Garaba, S.P., Meire, D., 2021. Hyperspectral-reflectance dataset of dry, wet and submerged marine litter. Earth System Science Data 13, 713–730. https://doi.org/10.5194/essd-13-713-2021

Kolokoussis, P., Karathanassi, V., 2018. Oil Spill Detection and Mapping Using Sentinel 2 Imagery. Journal of Marine Science and Engineering 6, 4. https://doi.org/10.3390/jmse6010004

Kremezi, M., Kristollari, V., Karathanassi, V., Topouzelis, K., Kolokoussis, P., Taggio, N., Aiello, A., Ceriola, G., Barbone, E., Corradi, P., 2022. Increasing the Sentinel-2 potential for marine plastic litter monitoring through image fusion techniques. Marine Pollution Bulletin 182, 113974. https://doi.org/10.1016/j.marpolbul.2022.113974

Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., Kompatsiaris, I., 2019. Oil Spill Identification from Satellite Images Using Deep Neural Networks. Remote Sensing 11, 1762. https://doi.org/10.3390/rs11151762

Kristollari, V., Karathanassi, V., 2020. Artificial neural networks for cloud masking of Sentinel-2 ocean images with noise and sunglint. International Journal of Remote Sensing 41, 4102–4135. https://doi.org/10.1080/01431161.2020.1714776

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment, Airbone Imaging Spectrometry 44, 145–163. https://doi.org/10.1016/0034-4257(93)90013-N

Kubryakov, A.A., Kudryavtsev, V.N., Stanichny, S.V., 2021. Application of Landsat imagery for the investigation of wave breaking. Remote Sensing of Environment 253, 112144. https://doi.org/10.1016/j.rse.2020.112144

Law, K.L., Morét-Ferguson, S., Maximenko, N.A., Proskurowski, G., Peacock, E.E., Hafner, J., Reddy, C.M., 2010. Plastic Accumulation in the North Atlantic Subtropical Gyre. Science 329, 1185–1188. https://doi.org/10.1126/science.1192321

Lebreton, L., Slat, B., Ferrari, F., Sainte-Rose, B., Aitken, J., Marthouse, R., Hajbane, S., Cunsolo, S., Schwarz, A., Levivier, A., Noble, K., Debeljak, P., Maral, H., Schoeneich-Argent, R., Brambini, R., Reisser, J., 2018. Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. Sci Rep 8, 4666. https://doi.org/10.1038/s41598-018-22939-w

Lebreton, L.C.-M., Greer, S.D., Borrero, J.C., 2012. Numerical modelling of floating debris in the world's oceans. Marine Pollution Bulletin 64, 653–661. https://doi.org/10.1016/j.marpolbul.2011.10.027

Lebreton, L.C.M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., Reisser, J., 2017. River plastic emissions to the world's oceans. Nat Commun 8, 15611. https://doi.org/10.1038/ncomms15611

Lechner, A., Keckeis, H., Lumesberger-Loisl, F., Zens, B., Krusch, R., Tritthart, M., Glas, M., Schludermann, E., 2014. The Danube so colourful: A potpourri of plastic litter outnumbers fish larvae in Europe's second largest river. Environmental Pollution 188, 177–181. https://doi.org/10.1016/j.envpol.2014.02.006

Lieshout, C. van, Oeveren, K. van, Emmerik, T. van, Postma, E., 2020. Automated River Plastic Monitoring Using Deep Learning and Cameras. Earth and Space Science 7, e2019EA000960. https://doi.org/10.1029/2019EA000960

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal Loss for Dense Object Detection, in: 2017 IEEE International Conference on Computer Vision (ICCV). Presented at the 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2999–3007. https://doi.org/10.1109/ICCV.2017.324

Liu, Y., Zhao, J., Qin, Y., 2021. A novel technique for ship wake detection from optical images. Remote Sensing of Environment 258, 112375. https://doi.org/10.1016/j.rse.2021.112375

Liu, Z., Yuan, L., Weng, L., Yang, Y., 2017. A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines. Presented at the 6th International Conference on Pattern Recognition Applications and Methods, pp. 324–331.

Lu, M., 2013. Ocean Waste in the Gulf of Honduras: Where it goes and what to do about it | ELAW [WWW Document]. URL https://elaw.org/ocean-waste-gulf-honduras-where-it-goes-and-what-do-about-it (accessed 6.25.23).

Lusher, A.L., Tirelli, V., O'Connor, I., Officer, R., 2015. Microplastics in Arctic polar waters: the first reported values of particles in surface and sub-surface samples. Sci Rep 5, 14947. https://doi.org/10.1038/srep14947

Maaten, L. van der, Hinton, G., 2008. Visualizing Data using t-SNE. Journal of Machine Learning Research 9, 2579–2605.

Majidi Nezhad, M., Groppi, D., Laneve, G., Marzialetti, P., Piras, G., 2018. Oil Spill Detection Analyzing "Sentinel 2" Satellite Images: A Persian Gulf Case Study. https://doi.org/10.11159/awspt18.134

Mani, T., Hauk, A., Walter, U., Burkhardt-Holm, P., 2015. Microplastics profile along the Rhine River. Sci Rep 5, 17988. https://doi.org/10.1038/srep17988

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M.F., Duarte, C.M., 2018. Use of unmanned aerial vehicles for efficient beach litter monitoring. Marine Pollution Bulletin 131, 662–673. https://doi.org/10.1016/j.marpolbul.2018.04.045

Martin, J., Lusher, A., Thompson, R.C., Morley, A., 2017. The Deposition and Accumulation of Microplastics in Marine Sediments and Bottom Water from the Irish Continental Shelf. Sci Rep 7, 10772. https://doi.org/10.1038/s41598-017-11079-2

Martinez-Vicente, V., 2022. The need for a dedicated marine plastic litter satellite mission. Nat Rev Earth Environ 3, 728–729. https://doi.org/10.1038/s43017-022-00360-2

Martínez-Vicente, V., Clark, J.R., Corradi, P., Aliani, S., Arias, M., Bochow, M., Bonnery, G., Cole, M., Cózar, A., Donnelly, R., Echevarría, F., Galgani, F., Garaba, S.P., Goddijn-Murphy, L., Lebreton, L., Leslie, H.A., Lindeque, P.K., Maximenko, N., Martin-Lauzer, F.-R., Moller, D., Murphy, P., Palombi, L., Raimondi, V., Reisser, J., Romero, L., Simis, S.G.H., Sterckx, S., Thompson, R.C., Topouzelis, K.N., van Sebille, E., Veiga, J.M., Vethaak, A.D., 2019. Measuring Marine Plastic Debris from Space: Initial Assessment of Observation Requirements. Remote Sensing 11, 2443. https://doi.org/10.3390/rs11202443

Matthews, J.P., Ostrovsky, L., Yoshikawa, Y., Komori, S., Tamura, H., 2017. Dynamics and early post-tsunami evolution of floating marine debris near Fukushima Daiichi. Nature Geosci 10, 598–603. https://doi.org/10.1038/ngeo2975

Maximenko, N., Corradi, P., Law, K.L., Van Sebille, E., Garaba, S.P., Lampitt, R.S., Galgani, F., Martinez-Vicente, V., Goddijn-Murphy, L., Veiga, J.M., Thompson, R.C., Maes, C., Moller, D., Löscher, C.R., Addamo, A.M., Lamson, M.R., Centurioni, L.R., Posth, N.R., Lumpkin, R., Vinci, M., Martins, A.M., Pieper, C.D., Isobe, A., Hanke, G., Edwards, M., Chubarenko, I.P., Rodriguez, E., Aliani, S., Arias, M., Asner, G.P., Brosich, A., Carlton, J.T., Chao, Y., Cook, A.-M., Cundy, A.B., Galloway, T.S., Giorgetti, A., Goni, G.J., Guichoux, Y., Haram, L.E., Hardesty, B.D., Holdsworth, N., Lebreton, L., Leslie, H.A., Macadam-Somer, I., Mace, T., Manuel, M., Marsh, R., Martinez, E., Mayor, D.J., Le Moigne, M., Molina Jack, M.E., Mowlem, M.C., Obbard, R.W., Pabortsava, K., Robberson, B., Rotaru, A.-E., Ruiz, G.M., Spedicato, M.T., Thiel, M., Turra, A., Wilcox, C., 2019. Toward the Integrated Marine Debris Observing System. Front. Mar. Sci. 6. https://doi.org/10.3389/fmars.2019.00447

Mega, E.R., 2022. Unprecedented oil spill catches researchers in Peru off guard. Nature. https://doi.org/10.1038/d41586-022-00333-x

Mehrtens, C.J., Rosenheim, B., Modley, M., Young, R.S., 2001. Reef morphology and sediment attributes, Roatan, Bay Islands, Honduras. Carbonates Evaporites 16, 131–140. https://doi.org/10.1007/BF03175831

Mifdal, J., Longépé, N., Rußwurm, M., 2021. Towards Detecting Floating Objects on a global scale with learned Spatial Features Using Sentinel- 2, in: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the XXIV ISPRS Congress, Commission III - 2021 edition, 9 July 2021, Copernicus GmbH, pp. 285–293. https://doi.org/10.5194/isprs-annals-V-3-2021-285-2021

Mikeli, P., Kikaki, K., Kakogeorgiou, I., Karantzalos, K., 2022. How Challenging is the Discrimination of Floating Materials on the sea surface using High Resolution Multispectral Satellite Data?, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the XXIV ISPRS Congress, Commission III - 2022 edition, 11 June 2022, Nice, France, Copernicus GmbH, pp. 151–157. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-151-2022

MMSegmentation Contributors, 2020. OpenMMLab Semantic Segmentation Toolbox and Benchmark.

Moshkov, N., Mathe, B., Kertesz-Farkas, A., Hollandi, R., Horvath, P., 2020. Test-time augmentation for deep learning-based cell segmentation on microscopy images. Sci Rep 10, 5068. https://doi.org/10.1038/s41598-020-61808-3

Moshtaghi, M., Knaeps, E., Sterckx, S., Garaba, S., Meire, D., 2021. Spectral reflectance of marine macroplastics in the VNIR and SWIR measured in a controlled environment. Sci Rep 11, 5436. https://doi.org/10.1038/s41598-021-84867-6

Mukonza, S.S., Chiang, J.-L., 2022. Satellite sensors as an emerging technique for monitoring macro- and microplastics in aquatic ecosystems. Water Emerging Contaminants & Nanoplastics 1, 17. https://doi.org/10.20517/wecn.2022.12

NOAA, 2016. Open Water Oil Identification Job Aid for Aerial Observation | response.restoration.noaa.gov [WWW Document]. URL https://response.restoration.noaa.gov/oil-and-chemical-spills/oil-spills/resources/open-water-oil-identification-job-aid.html (accessed 2.1.23).

Ody, A., Thibaut, T., Berline, L., Changeux, T., André, J.-M., Chevalier, C., Blanfuné, A., Blanchot, J., Ruitton, S., Stiger-Pouvreau, V., Connan, S., Grelet, J., Aurelle, D., Guéné, M., Bataille, H., Bachelier, C., Guillemain, D., Schmidt, N., Fauvelle, V., Guasco, S., Ménard, F., 2019. From In Situ to satellite observations of pelagic Sargassum distribution and aggregation in the Tropical North Atlantic Ocean. PLOS ONE 14, e0222584. https://doi.org/10.1371/journal.pone.0222584

Ostle, C., Thompson, R.C., Broughton, D., Gregory, L., Wootton, M., Johns, D.G., 2019. The rise in ocean plastics evidenced from a 60-year time series. Nat Commun 10, 1622. https://doi.org/10.1038/s41467-019-09506-1

Papageorgiou, D., Topouzelis, K., Suaria, G., Aliani, S., Corradi, P., 2022. Sentinel-2 Detection of Floating Marine Litter Targets with Partial Spectral Unmixing and Spectral Comparison with Other Floating Materials (Plastic Litter Project 2021). Remote Sensing 14, 5997. https://doi.org/10.3390/rs14235997

Papakonstantinou, A., Batsaris, M., Spondylidis, S., Topouzelis, K., 2021. A Citizen Science Unmanned Aerial System Data Acquisition Protocol and Deep Learning Techniques for the Automatic Detection and Mapping of Marine Litter Concentrations in the Coastal Zone. Drones 5, 6. https://doi.org/10.3390/drones5010006

Parinos, C., Hatzianestis, I., Chourdaki, S., Plakidi, E., Gogou, A., 2019. Imprint and short-term fate of the Agia Zoni II tanker oil spill on the marine ecosystem of Saronikos Gulf. Science of The Total Environment 693, 133568. https://doi.org/10.1016/j.scitotenv.2019.07.374

Paszke, A., Chaurasia, A., Kim, S., Culurciello, E., 2016. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. arXiv:1606.02147 [cs].

Pichel, W.G., Veenstra, T.S., Churnside, J.H., Arabini, E., Friedman, K.S., Foley, D.G., Brainard, R.E., Kiefer, D., Ogle, S., Clemente-Colón, P., Li, X., 2012. GhostNet marine debris survey in the Gulf of Alaska – Satellite guidance and aircraft observations. Marine Pollution Bulletin, At-sea Detection of Derelict Fishing Gear 65, 28–41. https://doi.org/10.1016/j.marpolbul.2011.10.009

Pisano, A., Bignami, F., Santoleri, R., 2015. Oil Spill Detection in Glint-Contaminated Near-Infrared MODIS Imagery. Remote Sensing 7, 1112–1134. https://doi.org/10.3390/rs70101112

Politikos, D.V., Adamopoulou, A., Petasis, G., Galgani, F., 2023. Using artificial intelligence to support marine macrolitter research: A content analysis and an online database. Ocean & Coastal Management 233, 106466. https://doi.org/10.1016/j.ocecoaman.2022.106466

Politikos, D.V., Fakiris, E., Davvetas, A., Klampanos, I.A., Papatheodorou, G., 2021. Automatic detection of seafloor marine litter using towed camera images and deep learning. Marine Pollution Bulletin 164, 111974. https://doi.org/10.1016/j.marpolbul.2021.111974

Potemra, J.T., 2012. Numerical modeling with application to tracking marine debris. Marine Pollution Bulletin, At-sea Detection of Derelict Fishing Gear 65, 42–50. https://doi.org/10.1016/j.marpolbul.2011.06.026

Qi, L., Hu, C., 2021. To what extent can Ulva and Sargassum be detected and separated in satellite imagery? Harmful Algae 103, 102001. https://doi.org/10.1016/j.hal.2021.102001

Qi, L., Hu, C., Mikelsons, K., Wang, M., Lance, V., Sun, S., Barnes, B.B., Zhao, J., Van der Zande, D., 2020. In search of floating algae and other organisms in global oceans and lakes. Remote Sensing of Environment 239, 111659. https://doi.org/10.1016/j.rse.2020.111659

Qi, L., Wang, M., Hu, C., Holt, B., 2022. On the capacity of Sentinel-1 synthetic aperture radar in detecting floating macroalgae and other floating matters. Remote Sensing of Environment 280, 113188. https://doi.org/10.1016/j.rse.2022.113188

Rajendran, S., Sadooni, F.N., Al-Kuwari, H.A.-S., Oleg, A., Govil, H., Nasir, S., Vethamony, P., 2021a. Monitoring oil spill in Norilsk, Russia using satellite data. Sci Rep 11, 3817. https://doi.org/10.1038/s41598-021-83260-7

Rajendran, S., Vethamony, P., Sadooni, F.N., Al-Kuwari, H.A.-S., Al-Khayat, J.A., Govil, H., Nasir, S., 2021b. Sentinel-2 image transformation methods for mapping oil spill – A case study with Wakashio oil spill in the Indian Ocean, off Mauritius. MethodsX 8, 101327. https://doi.org/10.1016/j.mex.2021.101327

Rajendran, S., Vethamony, P., Sadooni, F.N., Al-Kuwari, H.A.-S., Al-Khayat, J.A., Seegobin, V.O., Govil, H., Nasir, S., 2021c. Detection of Wakashio oil spill off Mauritius using Sentinel-1 and 2 data: Capability of sensors, image transformation methods and mapping. Environmental Pollution 274, 116618. https://doi.org/10.1016/j.envpol.2021.116618

Robinson, C., Malkin, K., Jojic, N., Chen, H., Qin, R., Xiao, C., Schmitt, M., Ghamisi, P., Hansch, R., Yokoya, N., 2021. Global Land Cover Mapping with Weak Supervision: Outcome of the 2020 IEEE GRSS Data Fusion Contest. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1–1. https://doi.org/10.1109/JSTARS.2021.3063849

Rochman, C.M., Tahir, A., Williams, S.L., Baxa, D.V., Lam, R., Miller, J.T., Teh, F.-C., Werorilangi, S., Teh, S.J., 2015. Anthropogenic debris in seafood: Plastic debris and fibers from textiles in fish and bivalves sold for human consumption. Sci Rep 5, 14340. https://doi.org/10.1038/srep14340

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Ruddick, K.G., De Cauwer, V., Park, Y.-J., Moore, G., 2006. Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters. Limnology and Oceanography 51, 1167–1179. https://doi.org/10.4319/lo.2006.51.2.1167

Rußwurm, M., Gul, D., Tuia, D., 2023. Improved marine debris detection in satellite imagery with an automatic refinement of coarse hand annotations, in: ICLR Workshop on Machine Learning for Remote Sensing.

Sannigrahi, S., Basu, B., Basu, A.S., Pilla, F., 2022. Development of automated marine floating plastic detection system using Sentinel-2 imagery and machine learning models. Marine Pollution Bulletin 178, 113527. https://doi.org/10.1016/j.marpolbul.2022.113527

Sardi, S.S., Qurban, M.A., Li, W., Kadinjappalli, K.P., Manikandan, P.K., Hariri, M.M., Tawabini, B.S., Khalil, A.B., El-Askary, H., 2020. Assessment of areas environmentally sensitive to oil spills in the western Arabian Gulf, Saudi Arabia, for planning and undertaking an effective response. Marine Pollution Bulletin 150, 110588. https://doi.org/10.1016/j.marpolbul.2019.110588

Schmaltz, E., Melvin, E.C., Diana, Z., Gunady, E.F., Rittschof, D., Somarelli, J.A., Virdin, J., Dunphy-Daly, M.M., 2020. Plastic pollution solutions: emerging technologies to prevent and collectmarineplastic pollution. Environment International 144, 106067. https://doi.org/10.1016/j.envint.2020.106067

Schmidt, C., Krauth, T., Wagner, S., 2017. Export of Plastic Debris by Rivers into the Sea. Environ. Sci. Technol. 51, 12246–12253. https://doi.org/10.1021/acs.est.7b02368

Schmitt, D., Gischler, E., 2016. Recent sedimentary facies of Roatan (Bay Islands, Honduras), a Caribbean oceanic barrier reef system. Facies 63, 5. https://doi.org/10.1007/s10347-016-0485-5

Sebille, E. van, Aliani, S., Law, K.L., Maximenko, N., Alsina, J.M., Bagaev, A., Bergmann, M., Chapron, B., Chubarenko, I., Cózar, A., Delandmeter, P., Egger, M., Fox-Kemper, B., Garaba, S.P., Goddijn-Murphy, L., Hardesty, B.D., Hoffman, M.J., Isobe, A., Jongedijk, C.E., Kaandorp, M.L.A., Khatmullina, L., Koelmans, A.A., Kukulka, T., Laufkötter, C., Lebreton, L., Lobelle, D., Maes, C., Martinez-Vicente, V., Maqueda, M.A.M., Poulain-Zarcos, M., Rodríguez, E., Ryan, P.G., Shanks, A.L., Shim, W.J., Suaria, G., Thiel, M., Bremer, T.S. van den, Wichmann, D., 2020. The physical oceanography of the transport of floating marine debris. Environ. Res. Lett. 15, 023003. https://doi.org/10.1088/1748-9326/ab6d7d

Sebille, E. van, Wilcox, C., Lebreton, L., Maximenko, N., Hardesty, B.D., Franeker, J.A. van, Eriksen, M., Siegel, D., Galgani, F., Law, K.L., 2015. A global inventory of small floating plastic debris. Environ. Res. Lett. 10, 124006. https://doi.org/10.1088/1748-9326/10/12/124006

Setiani, P., Ramdani, F., 2018. Oil spill mapping using multi-sensor Sentinel data in Balikpapan Bay, Indonesia, in: 2018 4th International Symposium on Geoinformatics (ISyG). Presented at the 2018 4th International Symposium on Geoinformatics (ISyG), pp. 1–4. https://doi.org/10.1109/ISYG.2018.8612057

Shahidul Islam, Md., Tanaka, M., 2004. Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. Marine Pollution Bulletin 48, 624–649. https://doi.org/10.1016/j.marpolbul.2003.12.004

Simon-Sánchez, L., Grelaud, M., Garcia-Orellana, J., Ziveri, P., 2019. River Deltas as hotspots of microplastic accumulation: The case study of the Ebro River (NW Mediterranean). Science of The Total Environment 687, 1186–1196. https://doi.org/10.1016/j.scitotenv.2019.06.168

Solé Gómez, À., Scandolo, L., Eisemann, E., 2022. A learning approach for river debris detection. International Journal of Applied Earth Observation and Geoinformation 107, 102682. https://doi.org/10.1016/j.jag.2022.102682

Song, Z., Sui, H., Wang, Y., 2014. Automatic ship detection for optical satellite images based on visual attention model and LBP, in: 2014 IEEE Workshop on Electronics, Computer and Applications. Presented at the 2014 IEEE Workshop on Electronics, Computer and Applications, pp. 722–725. https://doi.org/10.1109/IWECA.2014.6845723

Štepec, D., Martinčič, T., Skočaj, D., 2019. Automated System for Ship Detection from Medium Resolution Satellite Optical Imagery, in: OCEANS 2019 MTS/IEEE SEATTLE. Presented at the OCEANS 2019 MTS/IEEE SEATTLE, pp. 1–10. https://doi.org/10.23919/OCEANS40490.2019.8962707

Suaria, G., Avio, C.G., Mineo, A., Lattin, G.L., Magaldi, M.G., Belmonte, G., Moore, C.J., Regoli, F., Aliani, S., 2016. The Mediterranean Plastic Soup: synthetic polymers in Mediterranean surface waters. Sci Rep 6, 37551. https://doi.org/10.1038/srep37551

Sun, S., Hu, C., 2019. The Challenges of Interpreting Oil–Water Spatial and Spectral Contrasts for the Estimation of Oil Thickness: Examples From Satellite and Airborne Measurements of the Deepwater Horizon Oil Spill. IEEE Transactions on Geoscience and Remote Sensing 57, 2643–2658. https://doi.org/10.1109/TGRS.2018.2876091

Sun, S., Hu, C., Garcia-Pineda, O., Kourafalou, V., Le Hénaff, M., Androulidakis, Y., 2018a. Remote sensing assessment of oil spills near a damaged platform in the Gulf of Mexico. Marine Pollution Bulletin 136, 141–151. https://doi.org/10.1016/j.marpolbul.2018.09.004

Sun, S., Hu, C., Tunnell, J.W., 2015. Surface oil footprint and trajectory of the Ixtoc-I oil spill determined from Landsat/MSS and CZCS observations. Marine Pollution Bulletin 101, 632–641. https://doi.org/10.1016/j.marpolbul.2015.10.036

Sun, S., Lu, Y., Liu, Y., Wang, M., Hu, C., 2018b. Tracking an Oil Tanker Collision and Spilled Oils in the East China Sea Using Multisensor Day and Night Satellite Imagery. Geophysical Research Letters 45, 3212–3220. https://doi.org/10.1002/2018GL077433

Sunar, F., Dervisoglu, A., Yagmur, N., Colak, E., Kuzyaka, E., Mutlu, S., 2022. How Efficient can Sentinel-2 data help Spatial Mapping of Mucilage event in the Marmara Sea?, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the XXIV ISPRS Congress, Commission III - 2022 edition,11 June 2022, Nice, France, Copernicus GmbH, pp. 181–186. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-181-2022

Tang, J., Deng, C., Huang, G.-B., Zhao, B., 2015. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. IEEE Trans Geosci Remote Sens 53, 1174–1185. https://doi.org/10.1109/TGRS.2014.2335751

Tasseron, P., van Emmerik, T., Peller, J., Schreyers, L., Biermann, L., 2021. Advancing Floating Macroplastic Detection from Space Using Experimental Hyperspectral Imagery. Remote Sensing 13, 2335. https://doi.org/10.3390/rs13122335

Taylor, M.L., Gwinnett, C., Robinson, L.F., Woodall, L.C., 2016. Plastic microfibre ingestion by deep-sea organisms. Sci Rep 6, 33997. https://doi.org/10.1038/srep33997

Themistocleous, K., Papoutsa, C., Michaelides, S., Hadjimitsis, D., 2020. Investigating Detection of Floating Plastic Litter from Space Using Sentinel-2 Imagery. Remote Sensing 12, 2648. https://doi.org/10.3390/rs12162648

Topouzelis, K., Papageorgiou, D., Karagaitanakis, A., Papakonstantinou, A., Arias Ballesteros, M., 2020. Remote Sensing of Sea Surface Artificial Floating Plastic Targets with Sentinel-2 and Unmanned Aerial Systems (Plastic Litter Project 2019). Remote Sensing 12, 2013. https://doi.org/10.3390/rs12122013

Topouzelis, K., Papakonstantinou, A., Garaba, S.P., 2019. Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018). International Journal of Applied Earth Observation and Geoinformation 79, 175–183. https://doi.org/10.1016/j.jag.2019.03.011

Trujillo-Acatitla, R., Tuxpan-Vargas, J., Ovando-Vázquez, C., 2022. Oil spills: Detection and concentration estimation in satellite imagery, a machine learning approach. Marine Pollution Bulletin 184, 114132. https://doi.org/10.1016/j.marpolbul.2022.114132

Tuzcu Kokal, A., Olgun, N., Musaoğlu, N., 2022. Detection of mucilage phenomenon in the Sea of Marmara by using multi-scale satellite data. Environ Monit Assess 194, 585. https://doi.org/10.1007/s10661-022-10267-6

Tysiąc, P., Strelets, T., Tuszyńska, W., 2022. The Application of Satellite Image Analysis in Oil Spill Detection. Applied Sciences 12, 4016. https://doi.org/10.3390/app12084016

UNDP, 2015. Goal 14: Life Below Water | Sustainable Development Goals | United Nations Development Programme [WWW Document]. UNDP. URL https://www.undp.org/sustainable-development-goals/below-water (accessed 3.5.23).

Van der Zande, D., Green, J., Ruddick, K., 2014. JellySpec: feasibility study for determining the spectral characteristics of jellyfish from Belgian waters. Ocean Optics XXII, Portland, USA.

van Emmerik, T., van Klaveren, J., Meijer, L.J.J., Krooshof, J.W., Palmos, D.A.A., Tanchuling, M.A., 2020. Manila River Mouths Act as Temporary Sinks for Macroplastic Pollution. Front. Mar. Sci. 7. https://doi.org/10.3389/fmars.2020.545812

Vanhellemont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. Remote Sensing of Environment 216, 586–597. https://doi.org/10.1016/j.rse.2018.07.015

Vankayalapati, K., Dasari, H.P., Langodan, S., El Mohtar, S., Sanikommu, S., Asfahani, K., Desamsetti, S., Hoteit, I., 2023. Multi-Mission Satellite Detection and Tracking of October 2019 Sabiti Oil Spill in the Red Sea. Remote Sensing 15, 38. https://doi.org/10.3390/rs15010038

Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T., 2019a. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing 338, 34–45. https://doi.org/10.1016/j.neucom.2019.01.103

Wang, G., Li, W., Ourselin, S., Vercauteren, T., 2019b. Automatic Brain Tumor Segmentation Using Convolutional Neural Networks with Test-Time Augmentation, in: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 61–72. https://doi.org/10.1007/978-3-030-11726-9_6

Wang, M., Hu, C., 2021a. Automatic Extraction of Sargassum Features From Sentinel-2 MSI Images. IEEE Trans Geosci Remote Sens 59, 2579–2597. https://doi.org/10.1109/TGRS.2020.3002929

Wang, M., Hu, C., 2021b. Satellite remote sensing of pelagic Sargassum macroalgae: The power of high resolution and deep learning. Remote Sensing of Environment 264, 112631. https://doi.org/10.1016/j.rse.2021.112631

Wang, M., Hu, C., 2017. Predicting Sargassum blooms in the Caribbean Sea from MODIS observations. Geophysical Research Letters 44, 3265–3273. https://doi.org/10.1002/2017GL072932

Wang, M., Hu, C., 2016. Mapping and quantifying Sargassum distribution and coverage in the Central West Atlantic using MODIS observations. Remote Sensing of Environment 183, 350–367. https://doi.org/10.1016/j.rse.2016.04.019

Wightman, R., 2019. PyTorch Image Models.

Willis, K., Denise Hardesty, B., Kriwoken, L., Wilcox, C., 2017. Differentiating littering, urban runoff and marine transport as sources of marine debris in coastal and estuarine environments. Sci Rep 7, 44479. https://doi.org/10.1038/srep44479

Wolf, M., Berg, K. van den, Garaba, S.P., Gnann, N., Sattler, K., Stahl, F., Zielinski, O., 2020. Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q). Environ. Res. Lett. 15, 114042. https://doi.org/10.1088/1748-9326/abbd01

Xiao, Y., Liu, R., Kim, K., Zhang, J., Cui, T., 2022. A Random Forest-Based Algorithm to Distinguish Ulva prolifera and Sargassum From Multispectral Satellite Images. IEEE

Transactions on Geoscience and Remote Sensing 60, 1–15. https://doi.org/10.1109/TGRS.2021.3071154

Xie, X., Li, B., Wei, X., 2020. Ship Detection in Multispectral Satellite Images Under Complex Environment. Remote Sensing 12, 792. https://doi.org/10.3390/rs12050792

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, 3025–3033. https://doi.org/10.1080/01431160600589179

Yagci, A.L., Colkesen, I., Kavzoglu, T., Sefercik, U.G., 2022. Daily monitoring of marine mucilage using the MODIS products: a case study of 2021 mucilage bloom in the Sea of Marmara, Turkey. Environ Monit Assess 194, 170. https://doi.org/10.1007/s10661-022-09831-x

Yun, S., Han, D., Oh, S.J., Chun, S., Choe, J., Yoo, Y., 2019. Cutmix: Regularization strategy to train strong classifiers with localizable features, in: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6023–6032.

Zhang, H., 2017. Transport of microplastics in coastal seas. Estuarine, Coastal and Shelf Science 199, 74–86. https://doi.org/10.1016/j.ecss.2017.09.032

Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D., 2022. mixup: Beyond Empirical Risk Minimization. Presented at the International Conference on Learning Representations.

Zhao, J., Temimi, M., Ghedira, H., Hu, C., 2014. Exploring the potential of optical remote sensing for oil spill detection in shallow coastal waters-a case study in the Arabian Gulf. Opt. Express, OE 22, 13755–13772. https://doi.org/10.1364/OE.22.013755

Zhao, S., Wang, T., Zhu, L., Xu, P., Wang, X., Gao, L., Li, D., 2019. Analysis of suspended microplastics in the Changjiang Estuary: Implications for riverine plastic load to the ocean. Water Research 161, 560–569. https://doi.org/10.1016/j.watres.2019.06.019

Zielinski, S., Botero, C.M., Yanes, A., 2019. To clean or not to clean? A critical review of beach cleaning methods and impacts. Marine Pollution Bulletin 139, 390–401. https://doi.org/10.1016/j.marpolbul.2018.12.027

Zodiatis, G., Coppini, G., Perivoliotis, L., Lardner, R., Alves, T., Pinardi, N., Liubartseva, S., De Dominicis, M., Bourma, E., Sepp Neves, A.A., 2018. Numerical Modeling of Oil Pollution in the Eastern Mediterranean Sea, in: Carpenter, A., Kostianoy, A.G. (Eds.), Oil Pollution in the Mediterranean Sea: Part I: The International Context, The Handbook of Environmental Chemistry. Springer International Publishing, Cham, pp. 215–254. https://doi.org/10.1007/698_2017_131

8 LIST OF FIGURES