Original Article

Remote sensing of coral reef habitats in Madagascar using Sentinel-2 satellite images

Aina L. Nomenisoa[1,2], Gildas Todinanahary[1], Hubert Z. Edwin[1], Toky Razakarisoa[1], John Bunyan Israel[1], Saverio Raseta[1], Henitsoa Jaonalison[1], Jamal Mahafina[1], Igor Eeckhaut[2]

[1] Institut Halieutique et des Sciences Marines (IH.SM), University of Toliara, PO Box 141, Toliara, Madagascar

[2] Laboratoire de Biologie des Organismes Marins et Biomimétisme, University of Mons, PO Box 7000, Mons, Belgium

Western Indian Ocean JOURNAL OF Marine Science

Open access

Citation:
Nomenisoa ALD, Todinanahary G, Edwin HZ, Razakarisoa T, Israel JB, Raseta S, Jaonalison H, Mahafina J, Eeckhaut I (2024) Remote sensing of coral reef habitats in Madagascar using Sentinel-2 satellite images. Western Indian Ocean Journal of Marine Science 23(2): 41-56 [doi: 10.4314/wiojms.v23i2.4]

Received:
September 2, 2023

Accepted:
June 26, 2024

Published:
October 18, 2024

*** Corresponding author:**
ainaledon294@gmail.com

Abstract
Publicly available Sentinel-2 satellite imagery was used to map the coral reef systems of Toliara in Madagascar, to standardize methods for monitoring reef health and guiding management decisions. Fieldwork conducted between March and December 2021 used georeferenced photoquadrats to assess benthic structure. The satellite image classification was based on the Object-Based Image Analysis (OBIA) and machine learning algorithms, with k-NN achieving the highest overall accuracy at 83 %, followed by the Bayes classifier (79 %), DT (68 %), RT (67 %) and SVM (42 %). The analysis identified distinct surface areas occupied by seagrass (21 km^2), sand (73 km^2), rubble (21 km^2), coral (10 km^2) and algae (6 km^2). Comparative assessment with the Allen Coral Atlas underscored the importance of aligning satellite image analysis with *in-situ* data. The study emphasized the role of selecting appropriate classifier algorithms for precise mapping and stressed the importance of local data collection for accurate habitat mapping. It also showcased the successful application of OBIA with satellite imagery and field data for coral reef mapping, providing insights into habitat health and spatial changes essential for effective conservation.

Keywords: remote sensing, coral reef, Madagascar, Sentinel-2, OBIA, Machine learning algorithms

Introduction
Coral reefs protect shorelines against storms, serve as fish nurseries, and provide socio-economic benefits when associated with tourism and recreation, shoreline protection, fisheries, and biodiversity services (Eakin *et al.*, 2010). However, these ecosystems face significant threats on both global and local scales, primarily due to climate change and anthropogenic pressures (Xu and Zhao, 2014). Nearly half of the world coral reefs have been destroyed or badly damaged in the last 30 years (Wilkinson, 2008). Current trends suggest that between 70 % to 90 % of global coral reefs are at risk of extinction within the foreseeable future (Foo and Asner, 2019), a fate that extends to the reefs of Madagascar as well (van Hooidonk *et al.*, 2016). Effective management and conservation efforts requires comprehensive monitoring strategies that encompass both spatial and temporal dimensions, focusing on the distribution of species on the benthos as well as their associated substrates (Nurlidiasari and Budiman, 2010). Such measures require a reliable method that can efficiently process continuous data into manageable spatial units (Kennedy *et al.*, 2021). The European Space Agency (ESA) has made Sentinel-2 images freely available since 2015 (ESA's Sentinel-2 team, 2015) offering a 10 m spatial resolution (pixel size), which significantly enhances the utility of these images for coral reef mapping. Remote sensed mapping of coral reefs is particularly important for developing countries, where 80 % of the world's coral

reefs occur (UNEP-WCMC, WorldFish Centre, WRI, TNC, 2021). Therefore, it is critical to determine the efficacy of mapping using freely available satellite images from Sentinel-2 (Hedley *et al.*, 2012, 2018; Wouthuyzen *et al.*, 2019; Yunus *et al.*, 2019).

To achieve comprehensive mapping of coral reefs, it is crucial to derive maps of geomorphic zones and benthic communities at various scales (Phinn, *et al.*, 2012). Numerous initiatives have been undertaken globally to map coral reefs and understand their distribution, including coral reefs from Madagascar. However, existing datasets, such as those from the United Nations Environment Program-World Conservation Monitoring Centre (UNEP-WCMC) and the Allen Coral Reef Atlas, have limitations such as irregular updates, and unstable accuracy, hindering their applicability in dynamic monitoring. Although efforts to study coral reefs from Madagascar have provided valuable insights, further studies are necessary to fully comprehend these ecosystems and their changing habitats.

Efficient and cost-effective methods using remote sensing data are needed to delineate comprehensive reef coverage, geomorphic zoning, and benthic composition in Madagascar. Despite numerous studies on mapping coral reefs, variations in processing schemes, data characterization, and classification methods pose challenges, emphasizing the necessity for a tailored methodology. Therefore, there is a need for a comprehensive methodology tailored to coral reefs from Madagascar, encompassing standardized fieldwork data collection and image processing workflows. This study focuses on the barrier reef in Toliara, commonly known as the 'Grand Récif de Toliara' (GRT), and the reefs of Ankilibe and Sarodrano, where local coral reef geomorphology has been conducted

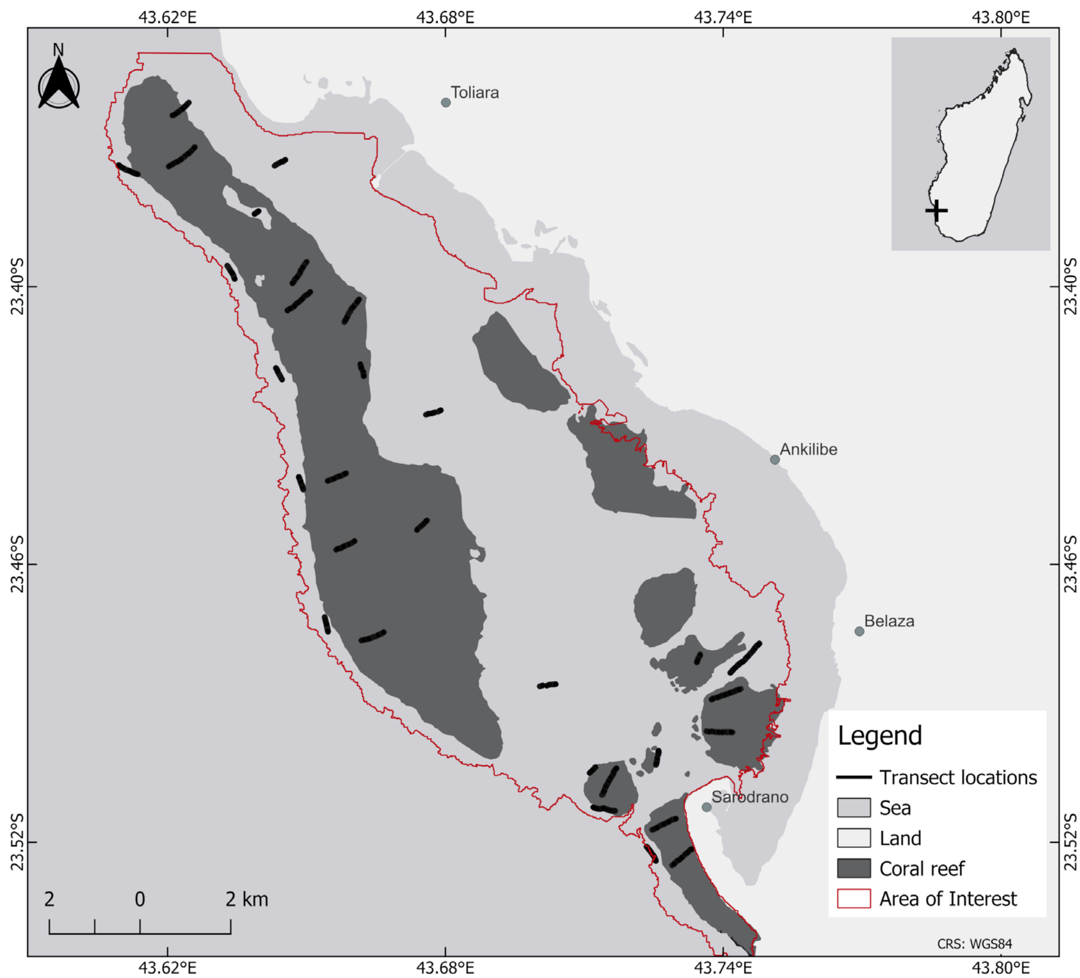

Figure 1. Study area with the placement of the transect lines.

(Pichon, 1972; Battistini *et al.*, 1975; Andréfouët *et al.*, 2013), and several studies focusing on the bio-ecology of coral reefs have been realized (Todinanahary, 2016; Razakandrainy, 2018; Botosoamananto *et al.*, 2021). Undertaking a comprehensive study at the scale of the Toliara region is essential for coral reefs of Madagascar, necessitating standardized methodologies for future data collection and processing. The primary aim of this research was to develop a methodology for coral reef mapping in Madagascar, using freely accessible satellite imagery and advanced remote sensing techniques. Specifically, the study aimed to assess geomorphic zonation and benthic coverage along the barrier reef and fringing reefs of Toliara, using Object-Based Image Analysis (OBIA) applied to Sentinel-2 satellite data, alongside fieldwork data acquisition. This study seeks to establish a foundational understanding of nearshore benthic communities in Madagascar, serving as a vital resource for marine scientists to effectively track ecosystem changes. Furthermore, it aims to furnish policymakers with essential data for monitoring the health of these reefs and formulating sustainable management strategies over the long term.

Materials and methods

Study area, field data collection and processing

The study was carried out within the Bay of Toliara, and was focused on the fringing reef of Sarodrano, the fringing reef of Ankilibe, and the barrier reef of Toliara (Fig. 1). The selection of these sites was based on several factors: (1) accessibility, (2) ecological variability, encompassing both degraded and healthy areas to facilitate the discernment of local stressors, and (3) ecological complexity, given the co-existence of a barrier reef, a fringing reef, and patch reefs within the same coral reef system. Fieldwork was conducted between March and December 2021, during low spring tides to optimize the time available for in-water surveying. It consisted of collecting georeferenced photoquadrats (in-water images) along transect lines (Fig. 1). Due to the limitation of GPS signal penetration underwater, the GPS device was secured in a floating airtight bag at the surface. This device was programmed to record new geographic positions every second. Specifically, one diver captured benthic images every 3-5 meters, guided by a compass to maintain course, while a second diver at the surface maneuvered the GPS-containing bag, moving in synchronization with the diver below (Fig. 2). Each dive averaged 30 minutes, covering approximately 500 meters of transect where current conditions allowed; however, in instances of stronger currents, transects were shortened accordingly. Initial and final GPS coordinates of each transect were logged on a diving slate to facilitate subsequent GPS data integration. Approximately 250 photos were collected per transect, totaling 4187 photos across the 30 transects. Ground-truthing data collection was confined to depths shallower than 20 meters. Subsequently, the photos were linked to GPS data to assign specific geographic positions to each photoquadrat using the software GPS Photo Manager (Roelfsema *et al.*,

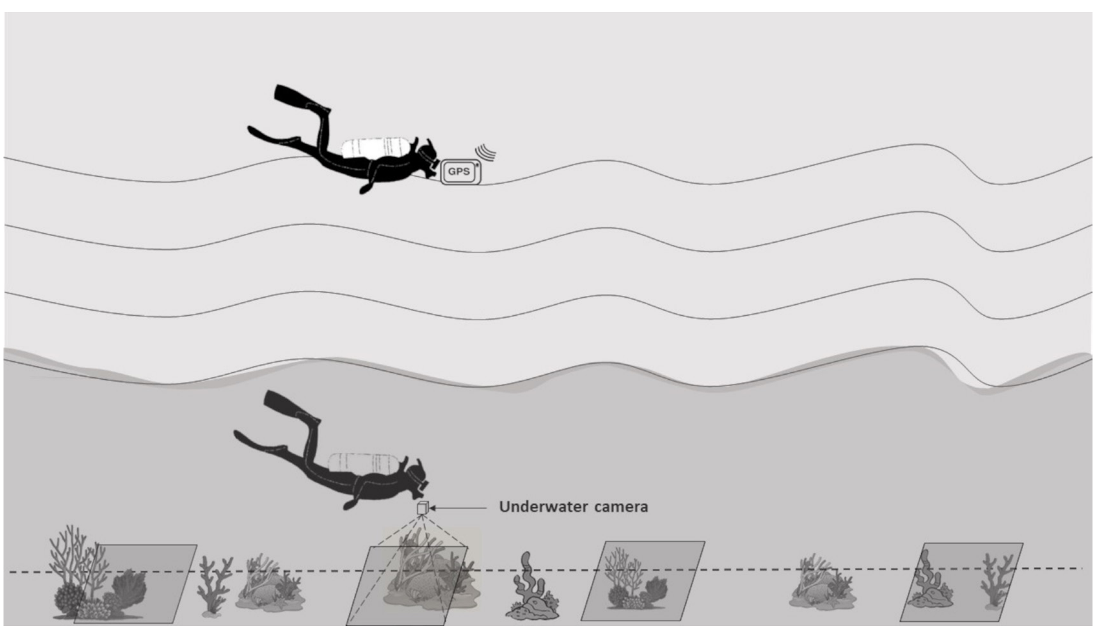

Figure 2. Method of photoquadrat data collection.

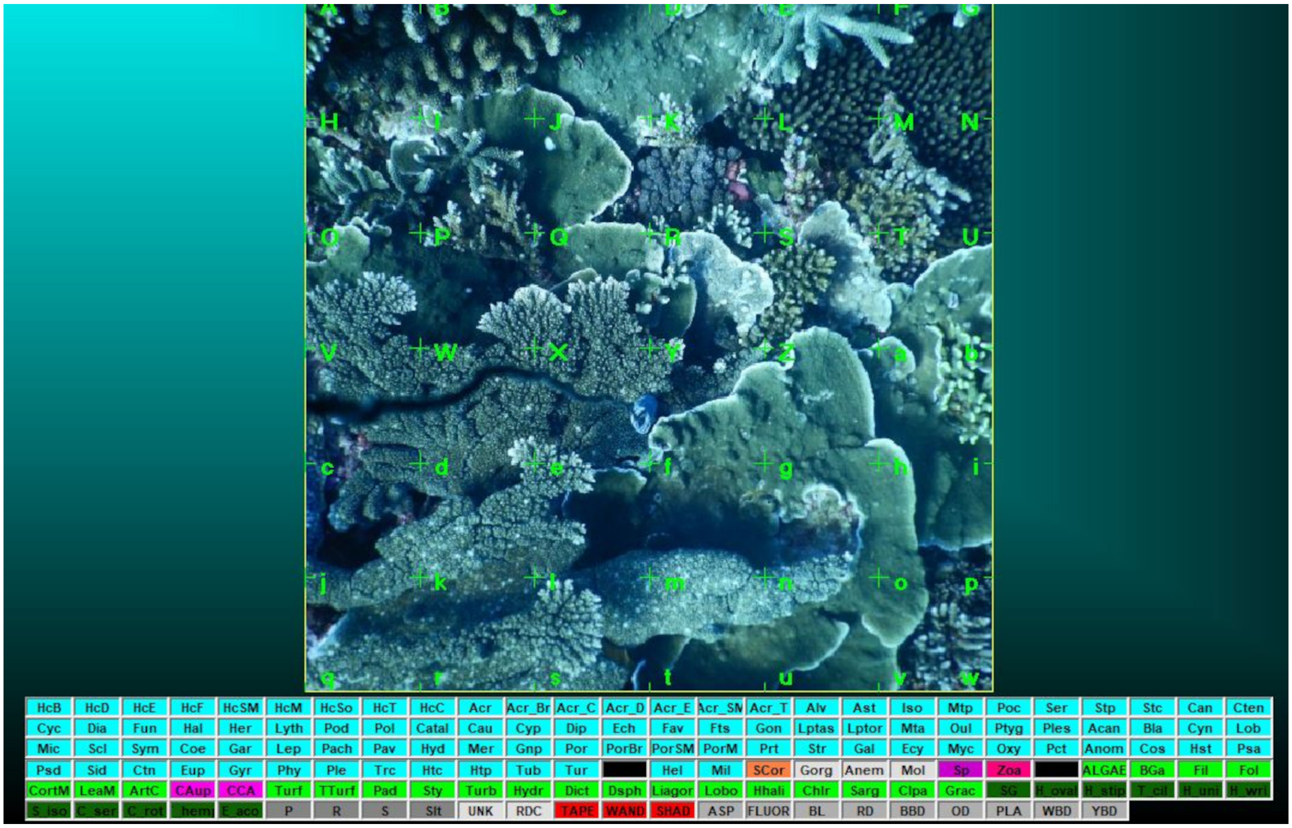

Figure 3. 49 stratified points (7 points per lines) on CPCE.

2019). This integration facilitated the visualization of photos and associated benthic attributes within a Geographic Information System (GIS) interface. All the GIS operations in this study were performed using the QGIS software, version 3.34.

Assessment of benthic cover from geotagged photos

Benthic cover within the transect lines was obtained from each geotagged photo by analyzing photoquadrats using the Coral Point Count with Excel extension (CPCe) software (Kohler and Gill, 2006). A total of 49 stratified points were distributed on each photo and the substrates corresponding to these points were identified using a predefined code name "CPCe benthic codes 41 Madagascar" provided with the software (Fig. 3). Based on their knowledge and the 41 CPCe benthic codes, the users attributed specific classes to each of the 49 stratified points. This number of stratified points per photoquadrat was sufficient to identify down to the level of benthic habitat classes. The benthic cover of the photoquadrat was then automatically estimated by the software as a function of the number of points occupied by each category of substrate. Once the analysis was finished, the software exported the data in an Excel file.

Conversion of the CPCE data into calibration and validation sample points

The data points produced by the CPCe software spans a 1x1 meter area, whereas the pixel size of a Sentinel-2 image measures 10 meters by 10 meters. To ensure comparability between CPCe data points and Sentinel pixels, the 41 benthic CPCe codes were refined into five classes: Coral (i.e., live corals), Algae, Rubbles, Sand, and Seagrass. These data points were then overlaid onto the Sentinel-2 image layer. A new sample was then manually created and assigned a pixel category based on the predominant benthic cover depicted in the CPCe data pie chart (Fig. 4). From the 4187 photos, 1243 control points were derived, where 75 % were used for calibrating the machine learning algorithms to classify the satellite image, while the remaining 25 % served as validation points to assess their accuracy.

Satellite image processing

Pre-processing

Sentinel-2A data were accessed from the Copernicus server (https://scihub.copernicus.eu/dhus/#/home). The Sentinel-2A image collected on 21-08-2021 was used in this study (Table 1). The Sentinel-2 satellite was launched in 2015 and offers 10 m spatial resolution for the visible and the Near Infra-Red (NIR)

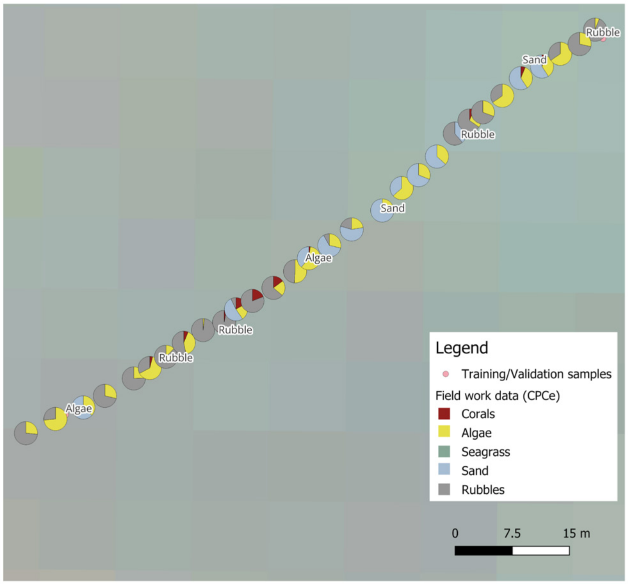

Figure 4. Calibration and validation sample points.

bands (ESA's Sentinel-2 team, 2015). The most appropriate images for coral reef habitat mapping are those that contain the least cloud cover and sun glint and are acquired during the spring low tide period. This last requirement is crucial in the analyses as the characterization of the benthos is complicated especially when they are submerged in the water column. The raw downloaded Sentinel-2 image was first corrected for the effect of the atmosphere. For this correction, the sen2cor atmospheric Correction Processor algorithm was used from the SNAP software or Sentinel Application Platform (Louis *et al.*, 2016). This approach consisted of transforming the Level-1C image (surface reflectance measured at the top of the atmosphere) into Level-2A (bottom-of-atmosphere reflectance). Next, as the benthos was submerged underwater, the water column effect needed to be corrected. The water column correction algorithm of Lyzenga (1981) was used for this purpose. This was processed using the sen2coral module of the SNAP software. As a result, three bands of the depth invariant index (DII) were generated, each composed of different combinations of spectral bands (Table 1): blue and green (DII_B2B3), blue and red (DII_B2B4), and green and red (DII_B3B4). Additionally, the Normalized Difference Vegetation Index (NDVI) was computed (Zoffoli *et al.*, 2020). These newly generated layers were added to the atmospheric corrected image for the geomorphic and benthic image classification.

Image classification

Coral reef habitat classification serves as an important tool for surveying and understanding these marine ecosystems. The process involves leveraging raw input data—such as videos or images—captured from coral sites (Nguyen *et al.*, 2021). Through this data, distinctive features of the seabed, referred to as classes, are extracted and categorized, including corals, sands, rubbles, seagrass, etc. To undertake this mapping,

Table 1. Sentinel-2 image specification used in this study.

Name	**Sentinel-2A**			
Correction level	Level 1A			
Date of acquisition	21-08-2021			
Radiometric resolution	12 bit/pixel			
Swath width	290 km at nadir			
	Band number	**Spectral Bands**	**Central wavelength (nm)**	**Spatial resolution (m)**
	1	Coastal Aerosol	442.7	60
	2	Blue	492.7	10
	3	Green	559.8	10
	4	Red	664.6	10
	5	Vegetation Red Edge	704.1	20
Multispectral Bands	6	Vegetation Red-Edge	740.5	20
	7	Vegetation Red-Edge	782.8	20
	8	Near Infra-Red (NIR)	832.8	10
	8A	Narrow NIR	864.7	20
	9	Water vapor	945.1	60
	10	Short Wave Infra-Red (SWIR)-Cirrus	1373.5	60
	11	Short Wave Infra-Red (SWIR)	1613.7	20
	12	Short Wave Infra-Red (SWIR)	2202.4	20

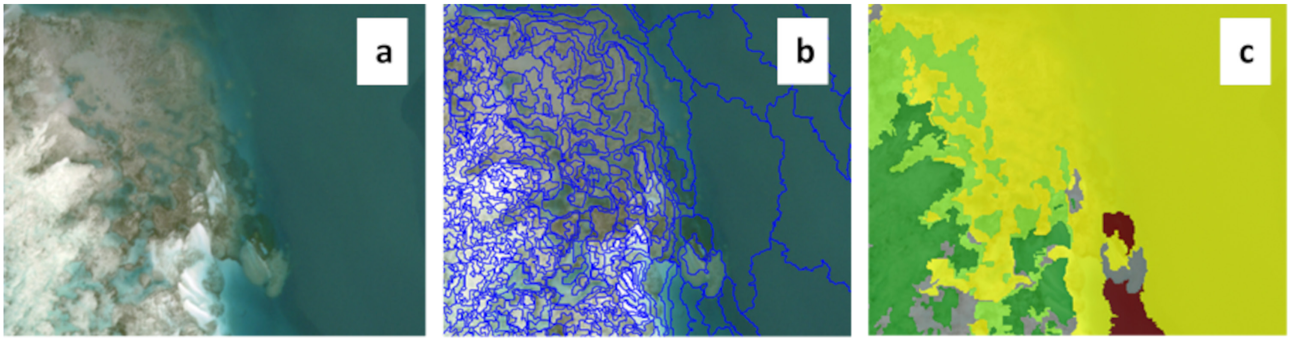

Figure 5. Object based image analysis (OBIA)process. a) Acquired image, b) Image segmentation, c) Image classification.

there are two approaches used: manual extraction of features, which offers high accuracy but is laborious and time-intensive, or the use of machine-learning algorithms for rapid processing and replicability, albeit with a potential risk of misclassification. While manual methods ensure precision, the advent of machine learning presents an efficient alternative, albeit with a need for careful validation and refinement to minimize errors in classification. For this, the OBIA method (Object-Based Image Analysis) was applied using eCognition Developer Software version 10.3 (Trimble Germany GmbH, 2022). This approach consists of grouping similar pixels (form, colour, texture, etc.) into segments (Fig.5), and then attributing classes to those segments (Hedley *et al.*, 2016).

Geomorphic zonation

The concept of geomorphic zonation in coral reefs involves identifying features that are relatively stable over time. In this study, the extraction of several key geomorphic features was focused on, including Internal Reefs, Lagoon, Reef Flat, Enclosed Basins, Reef Crest, Reef Front, and Reef Slope, following the definitions provided by Battistini *et al.* (1975). A lagoon is a naturally occurring depression with varying depths and sizes, typically found either behind a barrier reef or completely enclosed by reef structures. Enclosed basins, on the other hand, are smaller, shallower depressions or pools nestled within the reef structure on the reef flat. The reef slope constitutes the submerged front portion of a reef, sloping seaward with differing inclinations. It comprises coral formations and sedimentary deposits primarily of biogenic origin. The reef front delineates the outer edge of the reef flat at low tide, particularly during spring tides. Meanwhile, the reef flat is a horizontally oriented platform atop a reef structure, often reaching or surpassing sea level. It may exhibit material accumulations and surface incisions. The reef crest is primarily composed of coarse elements and is situated on the anterior part of the reef flat, manifesting in various shapes such as domes, ramparts, or scattered accumulations. Internal reefs are positioned within a lagoon, frequently separated from the open ocean by a barrier reef. They exhibit diverse sizes and shapes and are typically surrounded by shallow lagoon waters. These internal reefs consist of lagoonal coral patches, some of which extend to the surface and larger lagoon reefs, which are substantial coral formations within the lagoon, either partially exposed or submerged. These larger formations often display distinctive zoning patterns akin to those observed on reef flats.

After the image segmentation, these features were extracted by the visual photo-interpretation method using the built-in manual classification tools within the Ecognition Developer software.

Benthic image classification

For this purpose, the processed satellite image comprised: (i) the 10-meter resolution spectral bands from Sentinel-2 images (Table 1) which were atmospherically corrected, (ii) the calculated depth invariant bottom indexes (DII_B2B3, DII_B2B4, DII_B3B4), and (iii) the NDVI layer. Several satellite imageries, classification techniques, typologies and machine learning algorithms have been globally adopted to gather data on benthic coverage of coral reefs (Burns *et al.*, 2022). However, determining the most effective approach presents challenges due to inconsistencies in various factors, including the spectral and spatial resolutions of satellite images, methodologies for *in situ* reference data collection, the diversity and quantity of benthic classes mapped, and protocols for accuracy assessment. In this study, five prominent machine learning classifiers commonly employed in mapping coral reef benthic habitats were assessed with the goal of identifying the best performer tailored to environmental conditions, fieldwork data characteristics, and the particular benthic classes under study. These findings

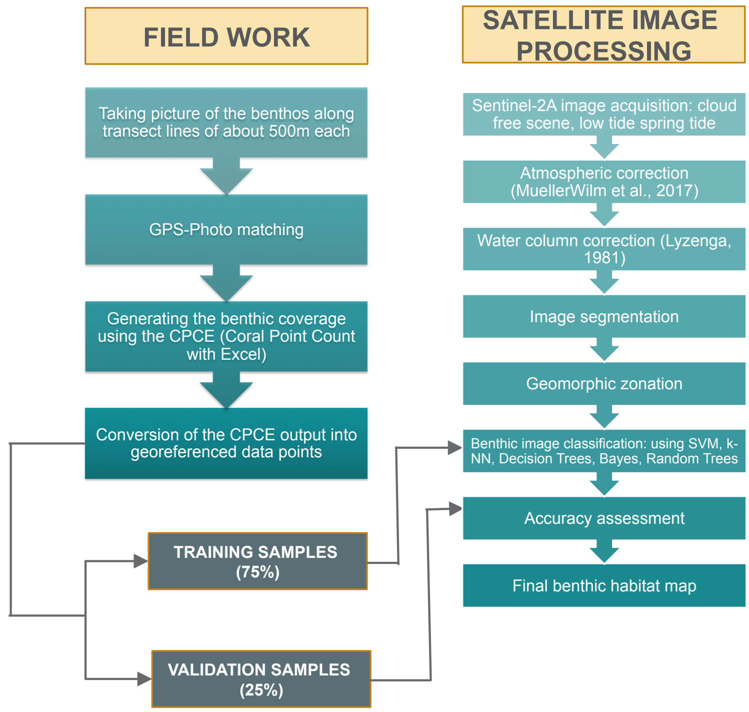

Figure 6. Image processing workflow.

will inform future investigations, guiding parameter choices, and streamlining classification processes, minimizing trial-and-error efforts. The five machine learning classifiers were:

Support Vector Machine (SVM). This assigns class labels to segmented objects by determining an optimal hyperplane, guided by feature vectors extracted from objects' attributes such as spectral values, texture, and spatial relationships. This hyperplane maximizes the margin between classes while minimizing misclassifications, with a focus on support vectors to define the boundary effectively (Mountrakis *et al.*, 2011).

Decision Tree (DT). The DT algorithm recursively partitions the dataset into subsets based on feature conditions, constructing a tree where each internal node represents a feature test and each leaf node denotes a class label. It employs a divide-and-conquer approach to classify instances, following a path from the root to a leaf node determined by feature conditions (Dietterich, 2000).

Random Trees (RT). This is a combination of multiple tree-based classifiers to produce a single classification, an ensemble of decision trees, where each single tree contributes a vote for the assignment of the most popular class to the input data (Xie and Niculescu, 2021).

k-Nearest Neighbour (k-NN). The k-NN algorithm classifies segmented objects based on the class most represented by their k nearest neighbours. K is a user-defined parameter that is the number of nearest neighbouring objects that are included in the majority voting process (Burns *et al.*, 2022) Bayes.

The Bayesian algorithm assigns classes to segmented objects by calculating the probability of each class given the object's features. It uses Bayes' theorem to compute the conditional probability of each class, incorporating prior knowledge and assuming feature independence to make informed decisions (Lewis, 1998). The flow chart of the image processing is provided in Fig. 6.

Benthic class description

The coral class (Fig. 7a) refers to a category with a hard underlying framework that is typically composed of coral-derived limestone, although non-carbonate materials can also be present. This class includes living corals. The rubble class (Fig. 7b) pertains to any area featuring loose, cylindrical to irregularly shaped fragments of bedrock or clasts of corals, bivalves, and coralline algae. This category encompasses limestone reef matrix and underlying areas of coral sand cemented together. The macroalgae class (Fig. 7c) is composed of large, multicellular marine plants that typically thrive in shallow waters surrounding coral reefs. Macroalgae are often observed on top of dead

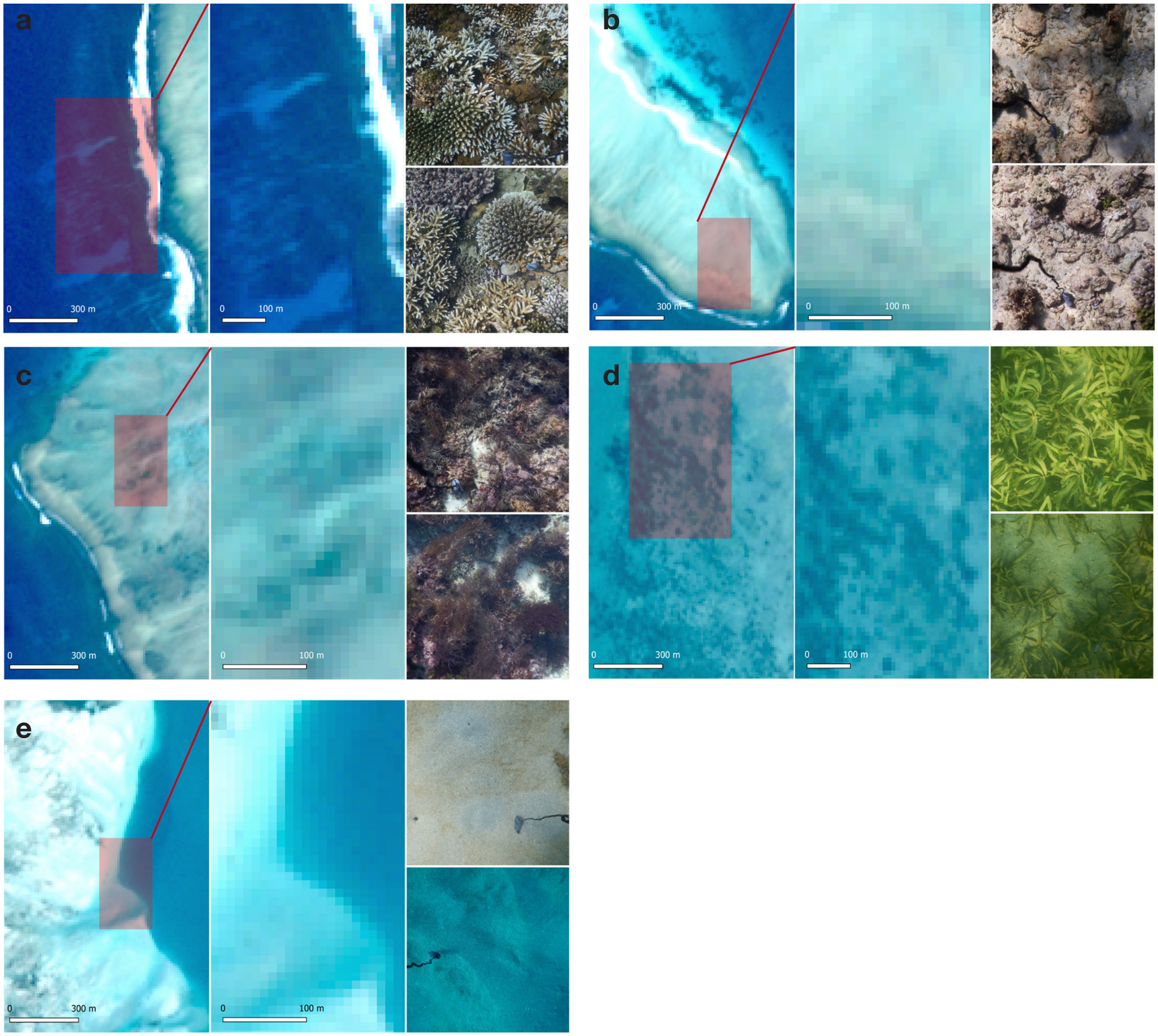

Figure 7. Benthic Classes Across Multiple Scales: 300 m (Sentinel-2 image), 100 m (Sentinel-2 image), and 1 m scale (photoquadrat). (a) Coral, (b) Rubbles, (c) Algae, (d) Seagrass, (e) Sand.

corals, in areas with clear water and abundant sunlight. The seagrass class (Fig. 7d) pertains to a soft-bottomed environment that is mainly characterized by the prevalence of a single or a combination of different species of seagrass. This classification also encompasses sparser or spatially confined seagrass as long as it forms the dominant benthic class. The sand class (Fig. 7e) pertains to soft-bottomed reef regions where fine unconsolidated granular material prevails. This granular material is finer than coral rubble but coarser than mud and thickly overlays any underlying bedrock. Sparse algae, scattered rocks, or small, isolated coral heads may also occur in the sand class. This class also encapsulates areas that are covered by a layer of fine-grained sediment that is mostly composed of organic matter and inorganic particles.

Accuracy assessment of the benthic habitat mapping

The validity, or usefulness, of any interpretation or classification map may be determined with an accuracy assessment that compares the created map with the field work data (Yamano, 2013). Accuracy assessment is commonly derived using an error matrix (also called confusion matrix), which tabulates the level of agreement between the thematic class at a location in the image-based map and the same location in the reference data (Yamano, 2013). The accuracy of each mapping category is described by the individual class accuracies, or according to the user's accuracy (UA) and producer's accuracies (PA) and Overall accuracy (OA), which are all derived from the error matrix. This is generated by using the built-in accuracy assessment tool in the eCognition Developer software. These metrics adhere to the definitions provided by Congalton and Green (2008). OA focuses on assessing the general performance of a classification algorithm across all classes in a dataset. It provides a holistic measure of the classifier's accuracy by considering all classes simultaneously. PA represents the proportion of correctly classified pixels or features for a specific class in relation to the total number of pixels or features that belong to that class on the ground. It focuses on how accurately the algorithm identifies and maps the pixels or features that truly belong to a specific class on the ground. UA represents the proportion of correctly classified pixels or features for a specific class in relation to the total number of pixels or features classified as that class by the classifier. It focuses on how accurately the algorithm assigns pixels or features to a particular class, regardless of whether they truly belong to that class or not. Five classifier algorithms (SVM, DT, RT, Bayes, k-NN) were executed on the image and their accuracy evaluated. Using the classifier algorithm that demonstrated the highest overall accuracy, the benthic classification outcome was refined by merging similar classes and eliminating small misclassified objects, thus improving the clarity and coherence of the final result.

Accuracy assessment of the Allen Coral Atlas

To assess the effectiveness of the benthic cover results for the reefs surrounding Toliara, a comparison with vector data from the Allen Coral Reef Atlas (ACA) (https://allencoralatlas.org) was conducted, acquired in August 2021. This data was cropped to match the scale of the current study, enabling meaningful comparisons. The ACA has the great advantage that it covers reefs around the world, so it is easy to refer to this atlas for a first approximation of benthic coverage of coastal and reef habitats (ACA, 2020).

Andréfouët (2008) mentioned the necessity of alignment of the extent of the ground truth data with the spatial resolution of the sensors. For a meaningful multi-sensor comparison accompanied by objective accuracy assessment, it is imperative that ground truth observations and typology align with the spatial resolution of the sensors. Given this, the raw CPCe data from this study, derived from 1 m x 1 m photoquadrat (with new pictures captured every 3-5 meters), offers a suitable basis for evaluating the accuracy of the ACA data rather than the readapted training and validation data (Fig. 4) that was used to match with the pixel size of the Sentinel-2 image. The ACA data is sourced from Planet Dove satellite images, a commercial satellite that provides a spatial resolution of 3-5meters (Safyan, 2020).

The benthic cover of the reefs of the ACA was composed by 6 categories: "Coral/Algae", "Microalgal Mats", "Rock", "Rubble", "Sand", and "Seagrass". To facilitate this comparison, the field work dataset was reorganized to mirror the typology of the ACA for the accuracy assessment. Specifically, classes such as "Coral" were changed to "Coral/Algae" and "Algae" to "Coral/Algae" in order to align with the ACA data. Classes such as "Rock" and "Microalgal mats" were also removed, as they were absent in the typology of raw data from the current study. However, classes like "Seagrass", "Sand", and "Rubbles" remained unaltered to maintain consistency across datasets.

Results

Accuracy of the image classifications

The overall accuracies of the classifier algorithms used for classifying benthic habitats of coral reefs in Toliara are depicted in Figure 8. The k-NN algorithm showed the highest Overall Accuracy (OA) at 83 %, followed by the Bayes classifier, DT, RT, and SVM, with OA values of 79 %, 68 %, 67 %, and 42 % respectively. Table 2 provides an in-depth analysis of classifier performance in categorizing various benthic habitat samples. It highlights the challenges faced by the Support Vector Machine (SVM) algorithm in accurately classifying Algae and Seagrass samples, with low values of Producer's Accuracy (PA) and User's Accuracy (UA). Other classifiers, such as Naive Bayes, DT, k-NN, RT, showed more robust performance. Naive Bayes and DT achieved high PA values for Rubble

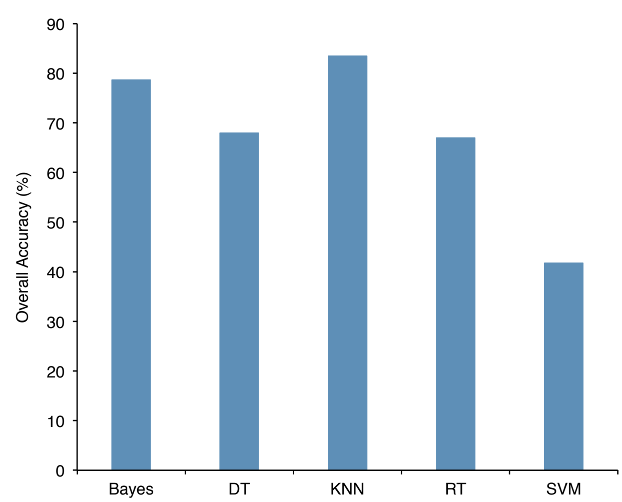

Figure 8. Overall accuracies of the classifiers algorithms used for benthic image classification.

Table 2. Confusion matrix of the five classifiers algorithms.

Classifier Algorithm	User Class \ Sample	Algae	Coral	Rubble	Sand	Seagrass
Bayes OA = 79%	Algae	9	0	0	0	0
	Coral	0	6	0	0	0
	Rubble	2	0	23	5	3
	Sand	1	0	1	20	2
	Seagrass	0	0	2	6	23
	PA	0,75	1	0,885	0,645	0,821
	UA	1	1	0,697	0,833	0,742
	User Class \ Sample	**Algae**	**Coral**	**Rubble**	**Sand**	**Seagrass**
DT OA = 68%	Algae	9	0	2	1	5
	Coral	0	6	0	0	0
	Rubble	0	0	22	7	9
	Sand	3	0	1	23	4
	Seagrass	0	0	1	0	10
	PA	0,75	1	0,846	0,742	0,357
	UA	0,529	1	0,579	0,742	0,909
	User Class \ Sample	**Algae**	**Coral**	**Rubble**	**Sand**	**Seagrass**
k-NN OA = 83%	Algae	11	0	2	2	2
	Coral	0	6	0	0	0
	Rubble	0	0	19	1	1
	Sand	1	0	4	28	3
	Seagrass	0	0	1	0	22
	PA	0,917	1	0,731	0,903	0,786
	UA	0,647	1	0,905	0,778	0,957
	User Class \ Sample	**Algae**	**Coral**	**Rubble**	**Sand**	**Seagrass**
RT OA = 67%	Algae	10	0	3	2	3
	Coral	0	6	0	0	0
	Rubble	0	0	16	6	6
	Sand	2	0	4	23	5
	Seagrass	0	0	3	0	14
	PA	0,833	1	0,615	0,742	0,5
	UA	0,556	1	0,571	0,676	0,824
	User Class \ Sample	**Algae**	**Coral**	**Rubble**	**Sand**	**Seagrass**
SVM OA = 42%	Algae	0	0	0	0	0
	Coral	3	6	0	2	0
	Rubble	4	0	12	1	0
	Sand	4	0	2	9	12
	Seagrass	1	0	12	19	16
	PA	0	1	0,462	0,290	0,571
	UA	undefined	0,545	0,706	0,333	0,333

and Seagrass, respectively, while k-NN demonstrated effectiveness in identifying Coral and Algae habitats. RF performed well in classifying Algae and Sand but struggled with Seagrass. In contrast, SVM struggled with Algae and Seagrass classification, with notably low PA values across multiple classes and variable UA values. Compared to other classifiers evaluated in this study, SVM exhibited lower OA, UA and PA for most habitat classes, indicating its potentially unsuitability for benthic habitat mapping applications under the given data calibration types and benthic environment. Figure 9 depicts various outputs of benthic habitat classification, highlighting the variability in results despite employing identical calibration data and image pre-processing techniques. This shows the importance of selecting the appropriate classifier algorithm, as it profoundly impacts the outcomes of benthic coverage assessment and the precise evaluation of each benthic habitat's surface area.

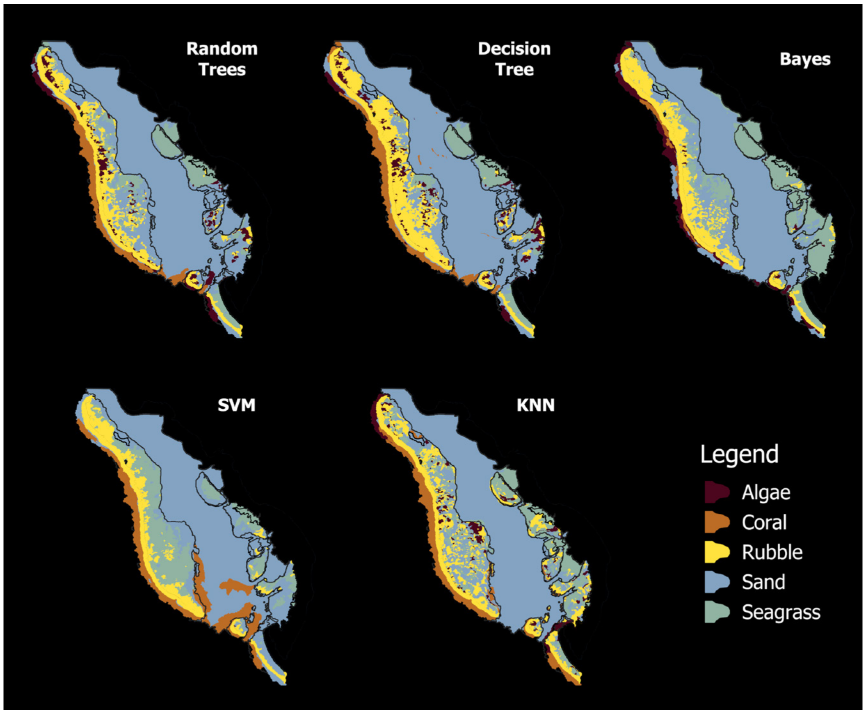

Figure 9. Outputs of different machine learning algorithms to map the benthic coverage of Toliara's coral reefs.

In reference to ACA, Table 3 presents the outcome of the accuracy assessment for its benthic coverage data. Compared with the 4187 CPCe data, the ACA achieves an overall accuracy of 50 %, failing to meet the 60 % minimum standard outlined in Yamano (2013). Additionally, a very low overall accuracy (OA) and user accuracy (UA) was noticed across all classes, particularly for rubbles and seagrass habitats.

Benthic coverage extent of the reefs of Toliara

The benthic coverage of the Toliara reefs was evaluated by using the k-NN classifier, which demonstrated the highest accuracy compared to the four other algorithms. Each habitat type is categorized based on its location within the reef system (Fig. 10). The internal reefs, along with their associated habitats, primarily span the area between Ankilibe and Sarodrano, as illustrated in Fig. 12c. The surface area measurements are presented for five benthic classes: Rubble (21 km^2), Coral (10 km^2), Algae (6 km^2), Seagrass (22 km^2), and Sand (73 km^2). Internal reefs exhibit substantial surface area coverage, particularly for Seagrass (8 km^2) and Rubble (4 km^2). Coral and Algae also contribute significantly to the internal reef ecosystem, with surface areas of 0.4 km^2 and 1.46 km^2 respectively. The Lagoon habitat features smaller surface areas compared to internal reefs. Notably, Sand is the predominant substrate in the lagoon, covering a substantial area of 57 km^2. Unlike Rubble (0.4 km^2), Coral (0.3 km^2) and Algae (0.5 km^2) that represent less significant surface coverage, Seagrass emerges as the next dominant feature in this habitat type, spanning an area of 5 km^2 indicating its importance as a habitat within the lagoon environment. Reef flat habitats showcase considerable surface area coverage, particularly for Seagrass (8 km^2) and Rubble (11 km^2) and Sand (12 km^2).

Table 3. Confusion matrix of the benthic coverage of the Allen Coral Atlas data.

OA= 50%	Coral/Algae	Rubble	Sand	Seagrass
Coral/Algae	596	35	155	54
Rubble	181	617	408	281
Sand	2	53	180	15
Seagrass	206	140	103	241
User accuracy (%)	0.6	0.7	0.2	0.4
Producer accuracy (%)	0.7	0.4	0.7	0.3

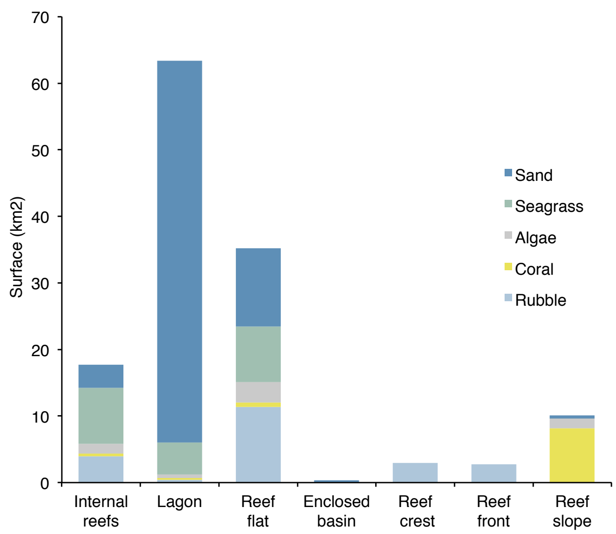

Figure 10. Surface of coral benthic cover per reef geomorphology.

Enclosed basin habitats exhibit relatively smaller surface area coverage compared to other habitat types. Sand dominates this habitat type, covering 0.2 km², while Coral (0.04 km²), Algae (0.004 km²) and Seagrass (0.006 km²) habitats exhibit minor surface area coverage. Both Reef Crest and Reef Front habitats exhibit minimal surface area coverage for all habitat classes, indicating their relatively limited extent within the reef system. These habitats are predominantly composed of Rubble which both represents about 3 km².

Reef slope habitats demonstrate unique characteristics with significant surface area coverage for Coral (8 km²). Algae (1 km²) and Sand (0.5 km²) classes also exhibit notable coverage, while Seagrass (0.017 km²) and Rubble classes are present in smaller amounts.

Discussion

Evaluating global coral reef mapping initiatives

Recent initiatives aimed at enhancing global coral reef mapping, such as the Allen Coral Atlas (Allen Coral Atlas, 2020) and the Global Distribution of Coral Reefs (UNEP-WCMC, WorldFish Centre, WRI, TNC, 2021), provide publicly accessible datasets facilitating easy access to information on coral reef geomorphology and benthic coverage. However, caution is warranted when using such data for national coral reef management strategies, restoration programmes, or economic valuations of these ecosystems. A comparison of the coral reef data generated by the UNEP-WCMC, the ACA geomorphology and the present study is provided in Figure 12. A total surface of 162 km² for the total extent of coral reef systems surrounding Toliara was calculated in the present study, while ACA provides a total of 126.8 km² and the UNEP-WCMC coral reef data totaled 61.1 km². It is also worth noting that this later dataset misses the fringing reef of Sarodrano (Fig. 12.a) which is present in both Figure 12.b and Figure 12.c. Figure 12.b also shows that the enclosed basin, locally

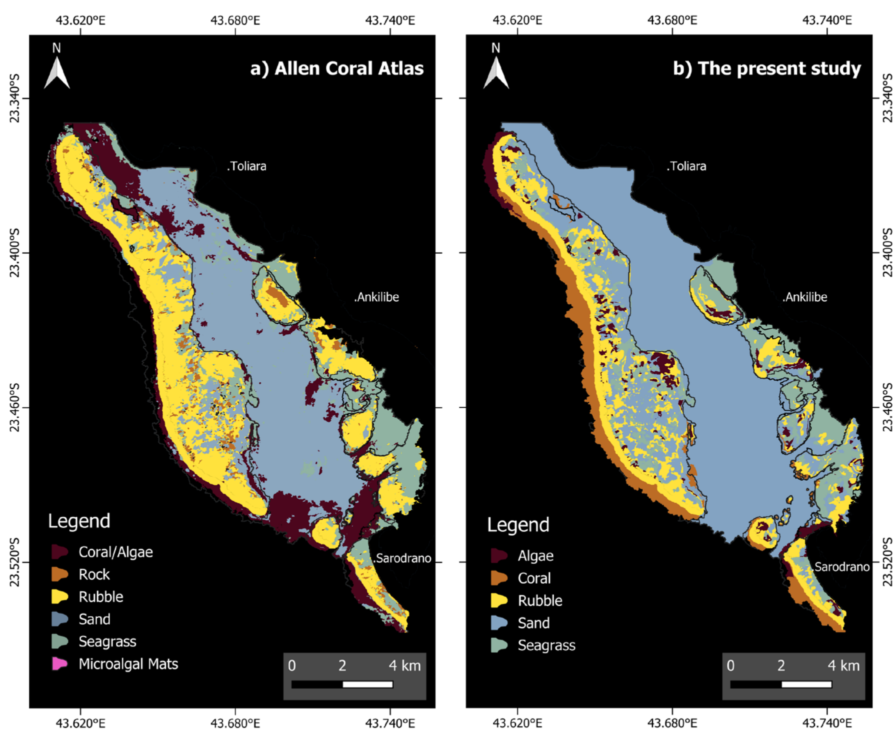

Figure 11. Comparison of benthic coverage provided by the ACA data and the present study.

known as "*Grande Vasque*" is erroneously interpreted as reef slope. This figure also showcases several hard reef structures within the lagoon which are misrepresented as reef slopes where it is clearly just a deep lagoon. Furthermore, concerning the benthic coverage of the coral reef system in Toliara, the current analysis reveals surface areas occupied by seagrass, sand, rubble, coral, and algae, amounting to 21 km², 73 km², 21 km², 10 km², and 6 km² respectively. Had marine scientists or policymakers used the ACA dataset, their calculations would have shown 22 km² for seagrass, and 48 km² for sand, 21 km² for Coral/Algae, and 28 km² for Rubble. The findings from the present study align with those reported by Botosoamananto *et al.* (2021), who conducted localized surveys within this coral reef system. The outer reef slope is predominantly characterized by robust hard corals, as highlighted, with the highest concentration of macroalgae observed in its northern section. Additionally, the authors noted substantial hard coral coverage within the reef patches of Sarodrano and the southern segment of the inner slope of the "*Grand Récif de Toliara*" (GRT). They also observed a significant presence of rubble on the reef flats as illustrated in Figure 11, a phenomenon documented by Bruggemann *et al.* (2012) and Andréfouët *et al.* (2013). These studies describe this particular section of the barrier reef, marked by an accumulation of dead coral and rubble on the reef flat.

Insight and considerations to enhance coral reef mapping

Assessment of benthic cover of coral reefs requires special attention, as variations in methodology can lead to inconsistencies in coral reef mapping and classification. The total surface area of coral reefs in Madagascar has been reported differently across studies, with UNEP-WCMC, WorldFish Centre, WRI, TNC (2021) reporting a total surface area of about 3,100 km² and the Allen Coral Atlas (Allen Coral Atlas, 2020) showing a total of 5,076 km² for the coral reefs of Madagascar. These variations do not indicate changes in coral reef extent but rather the different methodologies used to assess them. Coral reefs are complex and diverse ecosystems, with different species, morphology, and spatial arrangement depending on local environmental conditions. Inaccurate coral reef mapping and classification can have serious consequences for conservation efforts, leading to misinformed policy and management decisions. Overestimating coral cover can lead to inappropriate land-use decisions such as coastal development or tourism which can result in coral reef degradation and loss. Conversely, underestimating coral cover can result in inadequate protection or management measures, putting these important ecosystems at risk. Therefore, when using remote sensing techniques, a consistent and standardized methodology for assessing marine habitats is essential for accurate and effective conservation

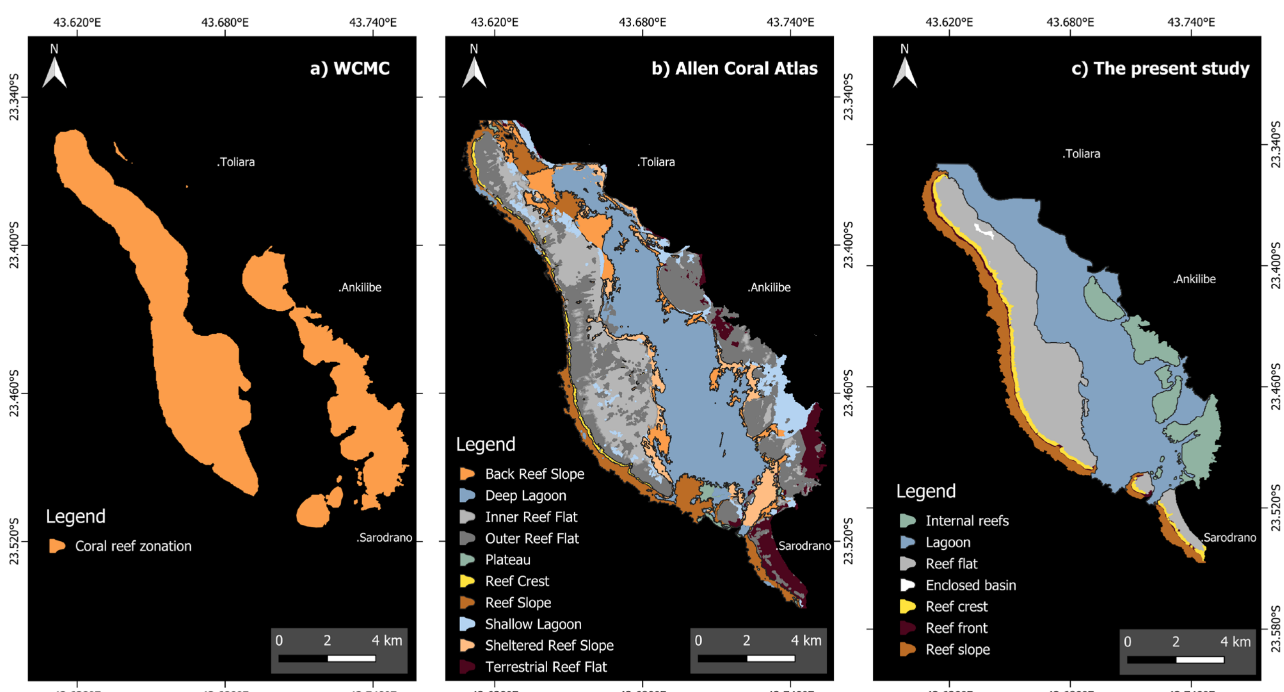

Figure 12. Comparison between the available data of coral reef geomorphology in Toliara.

and management. Validation and calibration of these techniques using *in-situ* data are necessary to ensure their accuracy and reliability. Overall, remote sensing can be a powerful tool for marine habitat assessment but it must be used with caution and care to avoid the negative consequences of methodology diversity. This study of the Toliara coral reef system underscores the significance of integrating multiple fieldwork datasets into the model training process, resulting in a notably enhanced mapping accuracy. This emphasizes the crucial role of *in situ* data collection in refining remote sensing techniques for more precise benthic habitat mapping. Across the globe, various classification algorithms are used for coral reef classification, with their effectiveness contingent upon regional characteristics, available field data, and the spectral and temporal resolution of satellite imagery. While SVM, k-NN, and RT algorithms are prevalent in object-based coral reef benthic mapping publications (Burns *et al.*, 2022), studies such as those by Mountrakis *et al.*, (2011) indicate that SVM often yields higher accuracy values compared to other techniques. However, the current research reveals that in this specific context, the k-NN algorithm outperforms SVM in terms of accuracy. This highlights the importance of selecting the most suitable algorithm tailored to local conditions when mapping coral habitats. Moreover, the performance of classification algorithms is subject to variations in image quality and resolution, training data set size, class types, and algorithm-specific parameter tuning. Hence, it is imperative to identify the most effective approach for this region based on these factors.

Conclusion

This study highlights the successful application of the OBIA method in conjunction with *in-situ* data to map coral reefs at local scales, using freely available high-resolution satellite images from Sentinel-2. This approach provides a replicable methodology for coral reef mapping projects using the same types of image and field data and lays the foundation to assess long-term changes in coral reef habitats, spatial observations of coral reef resilience, evaluation of seagrass distribution, and assessment of habitat health for herbivorous fishes. The use of georeferenced photographs not only establishes a formal linkage between the image and field data but also presents a valuable opportunity for informing stakeholders, managers, and other interested parties on the capabilities of satellite imaging for mapping and measuring reef features. Over 4,000 georeferenced photographs were used as reference data to produce a highly accurate map of the 18 km-long barrier reef of Toliara and the nearby reef systems. The efficacy of this mapping approach relies on both the quantity and quality of fieldwork data used to train the classifier algorithms for identifying features within satellite images. The greater the availability of comprehensive field data, the higher the accuracy of the resultant map. This explains why the ACA does not perform optimally in areas with limited field data. While global data may offer an initial reference for evaluating the extent or likelihood of coral reef occurrence, solely depending on such information for decision-making concerning coral reef management or restoration programmes entails considerable risks. Therefore, nations should prioritize local data collection and national-level satellite image processing to ensure precise assessments. This study illustrates that, even with freely available satellite imagery such as Sentinel-2 and basic logistical resources for field work data collection, sufficient accuracy can be attained to produce maps of coral reef geomorphology and benthic habitats.

Acknowledgements

We are sincerely grateful for the invaluable support extended by multiple institutions, whose contributions were instrumental in the success of this study. We extend our heartfelt appreciation to the Académie de Recherche et d'Enseignement Supérieur – Coopération Belge (ARES-CCD), the Western Indian Ocean Marine Sciences Association (WIOMSA), the Foundation for Future Generation (AETHER Fund), and the Station Marine de Belaza for their unwavering assistance in facilitating this research.

References

Allen Coral Atlas (2020) Imagery, maps and monitoring of the world's tropical coral reefs [doi: 10.5281/zenodo.3833242]

Andréfouët S (2008) Coral reef habitat mapping using remote sensing: A user vs producer perspective, implications for research, management and capacity building. Journal of Spatial Science 53: 113-129 [doi: 10.1080/14498596.2008.9635140]

Andréfouët S, Guillaume MMM, Delval A, Rasoamanendrika FMA, Blanchot J, Bruggemann, JH (2013) Fifty years of changes in reef flat habitats of the Grand Récif of Toliara (SW Madagascar) and the impact of gleaning. Coral Reefs 32: 757-768 [doi: 10.1007/s00338-013-1026-0]

Battistini R, Bourrouilh F, Chevalier JP, Coudray J, Denizot M, Faure G, Ficher JC, Guilcher, A, Harmelin-Vivien M, Jaubert J, Laborel J, Montaggioni L, Masse JP, Mauge LA, Peyrot-Clausade M, Pichon M, Plante R, Plaziat JC, Plessis YB, Richard G, Salvat B, Thomassin BA, Vasseur P, Weydert P (1975) Éléments de terminologie récifale indopacifique. Téthys 7: 1-111

Botosoamananto RL, Todinanahary G, Razakandrainy A, Randrianarivo M, Penin L, Adjeroud M, (2021) Spatial patterns of coral community structure in the Toliara region of southwest Madagascar and implications for conservation and management. Diversity 13: 486 [doi:10.3390/d13100486]

Bruggemann, JH, Rodier M, Guillaume MMM, Andréfouët S, Arfi R, Cinner JE, Pichon M, Ramahatratra F, Rasoamanendrika F, Zinke J, McClanahan TR (2012) Wicked social–ecological problems forcing unprecedented change on the latitudinal margins of coral reefs: the case of southwest Madagascar. Ecological Society 17: 4-21 [doi:10.5751/ES-05300-170447].

Burns C, Bollard B, Narayanan A (2022) Machine-learning for mapping and monitoring shallow coral reef habitats. Remote Sensing 14: 2666 [doi:10.3390/rs14112666]

Congalton RG, Green K (2008) Assessing the accuracy of remotely sensed data: Principles and practices. Mapping Science. CRC Press, Boca Rotan FL., Second Edition. 183 pp

Dietterich TG (2000) An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. Machine Learning 40: 139-157

Eakin CM, Nim CJ, Brainard R, Aubrecht C, Elvidge C, Gledhill DK, Muller-Karger F, Mumby PJ, Skirving W, Strong AE, Wang M, Weeks S, Wentz F, Ziskin D (2010) Monitoring coral reefs from space. Oceanography 23:119-133

ESA's Sentinel-2 team (2015) The story of Sentinel-2. European Space Agency 161: 4-9

Foo SA, Asner GP (2019) Scaling up coral reef restoration using remote sensing technology. Frontiers in Marine Science 6: 1-8 [doi:10.3389/fmars.2019.00079]

Hedley JD, Roelfsema CM, Phinn SR, Mumby PJ (2012) Environmental and sensor limitations in optical remote sensing of coral reefs: Implications for monitoring and sensor design. Remote Sensing 4: 271-302 [doi: doi:10.3390/rs4010271]

Hedley JD, Roelfsema CM, Chollet I, Harborne AR, Heron SF, Weeks S, Skirving W, Strong AE, Eakin CM, Christensen TRL, Ticzon V, Bejarano S, Mumby PJ (2016) Remote sensing of coral reefs for monitoring and management: A review. Remote Sensing 8: 1-40

Hedley JD, Roelfsema C, Brando V, Giardino C, Kutser T, Phinn S, Mumby PJ, Barrilero O, Laporte J, Koetz B (2018) Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8. Remote Sensing of the Environment 216: 598-614 [doi:10.1016/j.rse.2018.07.014]

Kennedy EV, Roelfsema CM, Lyons MB, Kovacs EM, Borrego-Acevedo R, Roe M, Phinn SR, Larsen K, Murray NJ, Yuwono D, Wolff J, Tudman P (2021) Reef Cover, a coral reef classification for global habitat mapping from remote sensing. Scientific Data 8: 1-20 [doi:10.1038/s41597-021-00958-z]

Kohler KE, Gill SM (2006) Coral point count with Excel extensions (CPCe): A Visual Basic programme for the determination of coral and substrate coverage using random point count methodology. Computers & Geosciences 32: 1259-1269 [doi:10.1016/j.cageo.2005.11.009]

Lewis DD (1998) Naive (Bayes) at forty: The independence assumption in information retrieval, In: Nédellec C, Rouveirol C (eds) Machine Learning: ECML-98, Lecture Notes in Computer Science. Springer Berlin, Heidelberg pp 4–15 [doi:10.1007/BFb0026666]

Louis J, Debaecker V, Pflug B, Main-Knorn M, Bieniarz J, Mueller-Wilm U, Cadau E, Gascon, F (2016) Sentinel-2 Sen2Cor: L2A processor for users. In: Proceedings of the Living Planet Symposium 2016. ESA SP-740, Prague, Czech Republic. 8 pp

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

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66: 247-259 [doi:10.1016/j.isprsjprs.2010.11.001]

Nguyen T, Liquet B, Mengersen K, Sous D (2021) Mapping of coral reefs with multispectral satellites: A review of recent papers. Remote Sensing 13: 4470 [doi:10.3390/rs13214470]

Nurlidiasari M, Budiman S (2010) Mapping coral reef habitat with and without water column correction using Quickbird image. International Journal of Remote Sensing and Earth Sciences 2:45-56 [doi:10.30536/j.ijreses.2005.v2.a1357]

Phinn SR, Roelfsema CM, Mumby PJ (2012) Multi-scale, object-based image analysis for mapping geomorphic and ecological zones on coral reefs. International Journal of Remote Sensing 33: 3768-3797 [doi:10.1080/01431161.2011.633122]

Pichon M (1972) Les peuplements à base de sléractiniaires dans les récifs coralliens de la Baie de Tuléar

(Sud-Ouest de Madagsacar), Station Marine d'Endoume. ed. Marseille, France. pp 135-154

Razakandrainy A (2018) Mécanismes de maintien des populations de coraux dans la région de Toliara, sud-ouest de Madagascar (Mémoire DEA en Océanographie Appliquée). Université de Toliara, Institut Halieutique et des Sciences Marines (IH.SM), Toliara, Madagascar. 54 pp

Roelfsema CM, Markey K, Kennedy EV, Kovacs EM, Borrego-Acevedo R, Fox HE, Bambic B, Free B, Rice K, Phinn SR (2019) Protocol for georeferenced benthic psotoquadrat surveys. Remote Sensing and Research Centre, School of Earth and Environmental Sciences. University of Queensland, Brisbane, Australia. 17 pp

Safyan M (2020) Planet's Dove Satellite Constellation. In: Pelton, JN (ed) Handbook of small satellites. Springer International Publishing, Cham. pp 1-17 [doi:10.1007/978-3-030-20707-6_64-1]

Todinanahary GGB (2016) Evaluation du potentiel biologique, économique et social de la coralliculture dans le sud-ouest de Madagascar (Thèse de Doctorat). Université de Mons (Belgique), Faculté des Sciences, Biologie des Organismes Marins et Biomimétisme.

Trimble Germany GmbH (2022) eCognition Developer.

UNEP-WCMC, WorldFish Centre, WRI, TNC (2021) Global distribution of coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 4.1, updated by UNEP-WCMC. Includes contributions from IMaRS-USF and IRD (2005), IMaRS-USF (2005) and Spalding *et al*. (2001). UN Environment Programme-World Conservation Monitoring Centre, Cambridge, UK [doi: 10.34892/t2wk-5t34]

van Hooidonk R, Maynard J, Tamelander J, Gove J, Ahmadia G, Raymundo L, Williams G, Heron, SF, Planes S (2016) Local-scale projections of coral reef futures and implications of the Paris Agreement. Scientific Reports 6:1-8 [doi:10.1038/srep39666]

Wilkinson C (2008) Status of coral reefs of the world. Global Coral Reef Monitoring Network (GCRMN). Townsville, Australia. 294 pp

Wouthuyzen S, Abrar M, Corvianawatie C, Salatalohi A, Kusumo S, Yanuar Y, Darmawan, Samsuardi Y, Arrafat MY (2019) The potency of Sentinel-2 satellite for monitoring during and after coral bleaching events of 2016 in the some islands of Marine Recreation Park (TWP) of Pieh, West Sumatra. Earth and Environmental Science 284: 012028 [doi: 10.1088/1755-1315/284/1/012028]

Xie G, Niculescu S (2021) Mapping and monitoring of land cover/land use (LCLU) changes in the Crozon Peninsula (Brittany, France) from 2007 to 2018 by machine learning algorithms (Support Vector Machine, Random Forest, and Convolutional Neural Network) and by Post-classification Comparison (PCC). Remote Sensing 13: 3899 [doi:10.3390/rs13193899]

Xu J, Zhao D (2014) Review of coral reef ecosystem remote sensing. Acta Ecologica Sinica 34: 19-25 [doi:10.1016/j.chnaes.2013.11.003]

Yamano H (2013) Chapter 3: Multispectral applications. In: Coral reef remote sensing: A guide for mapping, monitoring and management. Springer, Dordrecht, Netherlands. pp 51-78

Yunus AP, Jie D, Song X, Avtar R (2019) High resolution Sentinel-2 images for improved bathymetric mapping of coastal and lake environments (preprint). Earth Sciences [doi:10.20944/preprints201902.0270.v1]

Zoffoli ML, Gernez P, Rosa P, Le Bris A, Brando VE, Barillé A-L, Harin N, Peters S, Poser K, Spaias L, Peralta G, Barillé L (2020) Sentinel-2 remote sensing of *Zostera noltei*-dominated intertidal seagrass meadows. Remote Sensing of Environment 251: 112020 [doi:10.1016/j.rse.2020.112020]