Oakajee Habitat Mapping 2025

Technical Report

ENVIRONMENT
An O2Marine company

CLIENT: Mid West Ports Authority

STATUS: Rev0

REPORT NUMBER: R250285

ISSUE DATE: 6 January 2026

MARINE

Important Note

This report has been prepared for the sole use of Mid West Ports Authority (MWPA) (herein, 'the client'), for a specific site (herein 'the site', the specific purpose specified in Section 1 of this report (herein 'the purpose'). This report is strictly limited for use by the client, to the purpose and site and may not be used for any other purposes.

Third parties, excluding regulatory agencies assessing an application in relation to the purpose, may not rely on this report. O2 Marine waive all liability to any third-party loss, damage, liability or claim arising out of or incidental to a third-party publishing, using or relying on the facts, content, opinions or subject matter contained in this report.

O2 Marine waive all responsibility for loss or damage where the accuracy and effectiveness of information provided by the client or other third parties was inaccurate or not up to date and was relied upon, wholly or in part in reporting.

Maps are created in GDA94 MGA Zone 50 (EPSG:28350) coordinate reference system and are not to be used for navigational purposes. Positional accuracy should be considered as approximate.

Acknowledgement Of Country

O2 Marine acknowledges Aboriginal and Torres Strait Islander people as this land's first storytellers and holders of scientific knowledge through their ongoing and continued connection to land, sea and community. We pay our respect to Elders past and present for their custodianship of the land and sea over millennia, which inspires us daily in our collective responsibility to sustain the land and sea country which we live by, work in and dream about.

WA Marine Pty Ltd t/as O2 Marine
ACN 168 014 819
Originating Office – Western Australia
20 Mews Road FREMANTLE WA 6160
T 1300 219 801 | info@o2marine.com.au

Version Register

Version	Status	Author	Reviewer	Comment	Authorised for Release (signed and dated)
Rev A	Internal Review	M. Stacey	S. Preston		25/09/2025
Rev B	Issue to client	M. Stacey	MWPA		
Rev 0	Issued as final	M. Stacey			

Transmission Register

Controlled copies of this document are issued to the persons/companies listed below. Any copy of this report held by persons not listed in this register is deemed uncontrolled. Updated versions of this report if issued will be released to all parties listed below via the email address listed.

Name	Email Address
William Jones	william.jones@midwestports.com.au

Acronyms and Abbreviations

Term	Full term
AHD	Australian Height Datum
BCH	Benthic Communities and Habitat
CATAMI	Collaborative and Automated Tools for Analysis of Marine Imagery
DII	Depth Invariant Index
DoT	Department of Transport
GPS	Geographic Positioning System
ha	hectares
LAU	Local Assessment Unit
LiDAR	Light Detection and Ranging
m	Meters
MBES	Multibeam Echosounder
MSL	Mean Sea Level
MWPA	Mid West Ports Authority
O2M	O2 Marine
OBIA	Object Based Image Analysis
RF	Random Forest
SIA	Strategic Industrial Area
WA	Western Australia

Contents

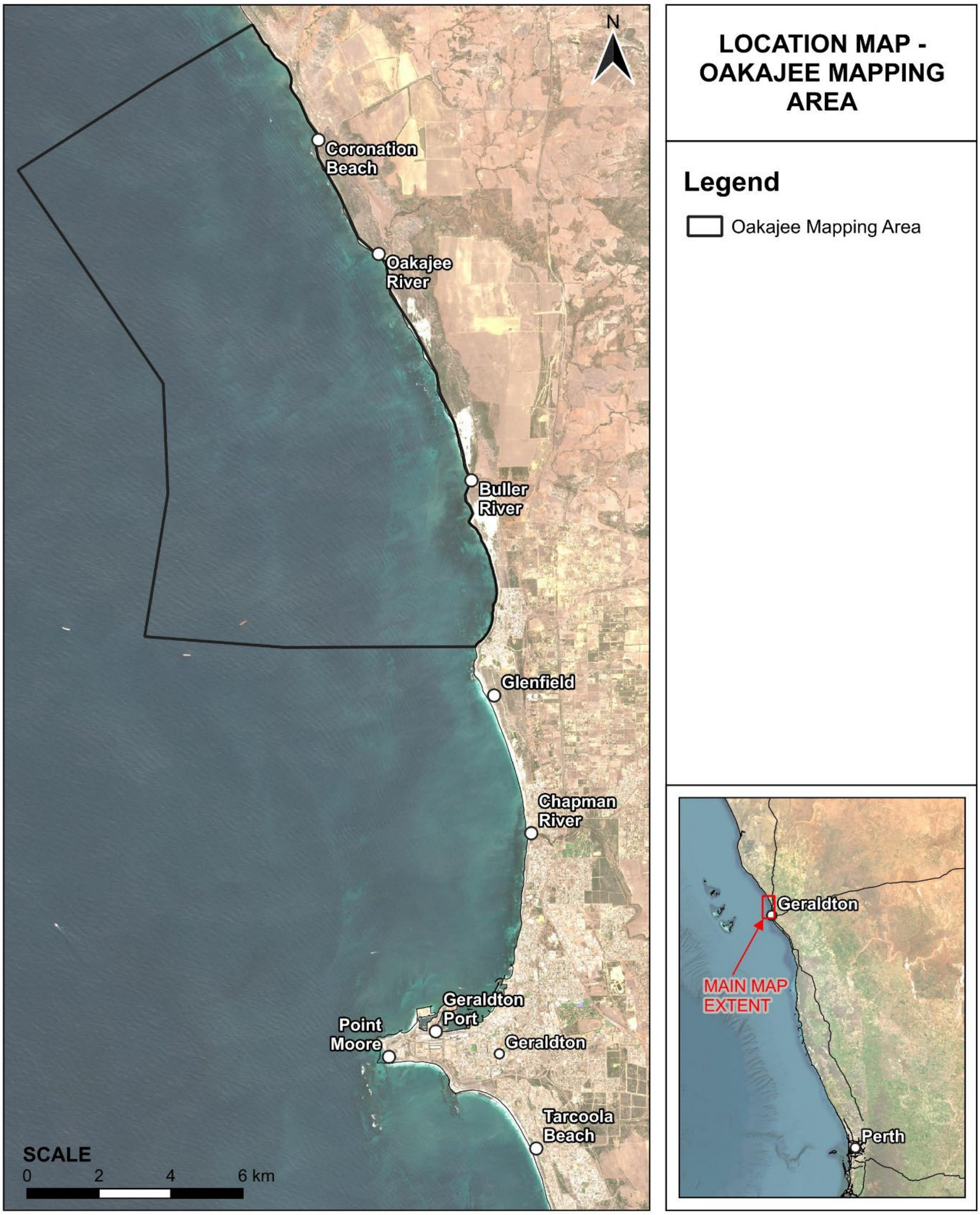
Tables

Figures

1. Introduction

The Oakajee Strategic Industrial Area (SIA), located approximately 20 km north of Geraldton in Western Australia, includes a proposed deepwater port and a desalination plant. The port is intended to serve mining operations in the Mid-West region, and the desalination plant will provide water for the port's construction and operation. The Oakajee SIA is envisioned as a clean energy hub, with the port facilitating the export of renewable hydrogen and ammonia.

In light of the above, MWPA engaged O2 Marine (O2M) to produce a subtidal benthic habitat map of an area of interest adjacent to the proposed SIA (Figure 1), to be informed by a ground-truthing dataset set to be acquired in June 2025. This area has been mapped previously in 2020 by AECOM (2020) (Figure 2).



Datum/Projection:	GDA94 / MGA zone 50
Date:	21/07/2025
Basemap:	Bing Virtual Earth
Map Creator:	M.S.
Map Number:	25ENV391_OKJ-01
Client:	MWPA

NOTE:

Positional accuracy is to be considered approximate. Not for navigation.

Figure 1: Location map showing area of interest adjacent to the proposed Oakajee Strategic Industrial Area (SIA) representing the targeted mapping area

2. Oakajee

Located north of Champion Bay, the Oakajee coastline of Western Australia is shaped by a high-energy marine environment driven by strong winds and swell-dominated currents. These dynamic oceanographic conditions play a central role in influencing sediment distribution, substrate composition, and the composition of benthic habitats throughout the region (Tecchiato et al. 2015).

A key geomorphological feature of the Oakajee area is the extensive limestone substrate, which forms the foundation for much of the nearshore marine environment. Variability in reef relief, together with the depth and mobility of overlying sand, strongly influences the distribution and diversity of benthic communities. Persistent exposure to south-westerly swells drives processes of sand transport, deposition, and erosion, creating shifting sediment regimes that affect which epibenthic assemblages can establish. As a result, habitats occurring under similar depth and substrate conditions often exhibit contrasting biological communities, reflecting differences in local hydrodynamic exposure.

A 2020 habitat mapping study conducted by AECOM (2020) classified the benthic environments of the Oakajee coastline into five distinct habitat types, based on factors such as substrate composition, water depth, and biological assemblages (Figure 2). Mosaic habitats of macroalgae, seagrass, and bare sand are common across the nearshore area, mainly associated with transitional zones where limestone pavement meets sandy substrate. As such, the constant movement of sediment throughout the bay means that these mosaic habitats can frequently vary in composition and cover (BMT, 2021). There are also areas where certain species are dominant, including seagrass species like *Amphibolis* spp. and *Posidonia* sp., or high-relief reef areas support dense macroalgal communities of *Sargassum* sp. or *Ecklonia* sp.

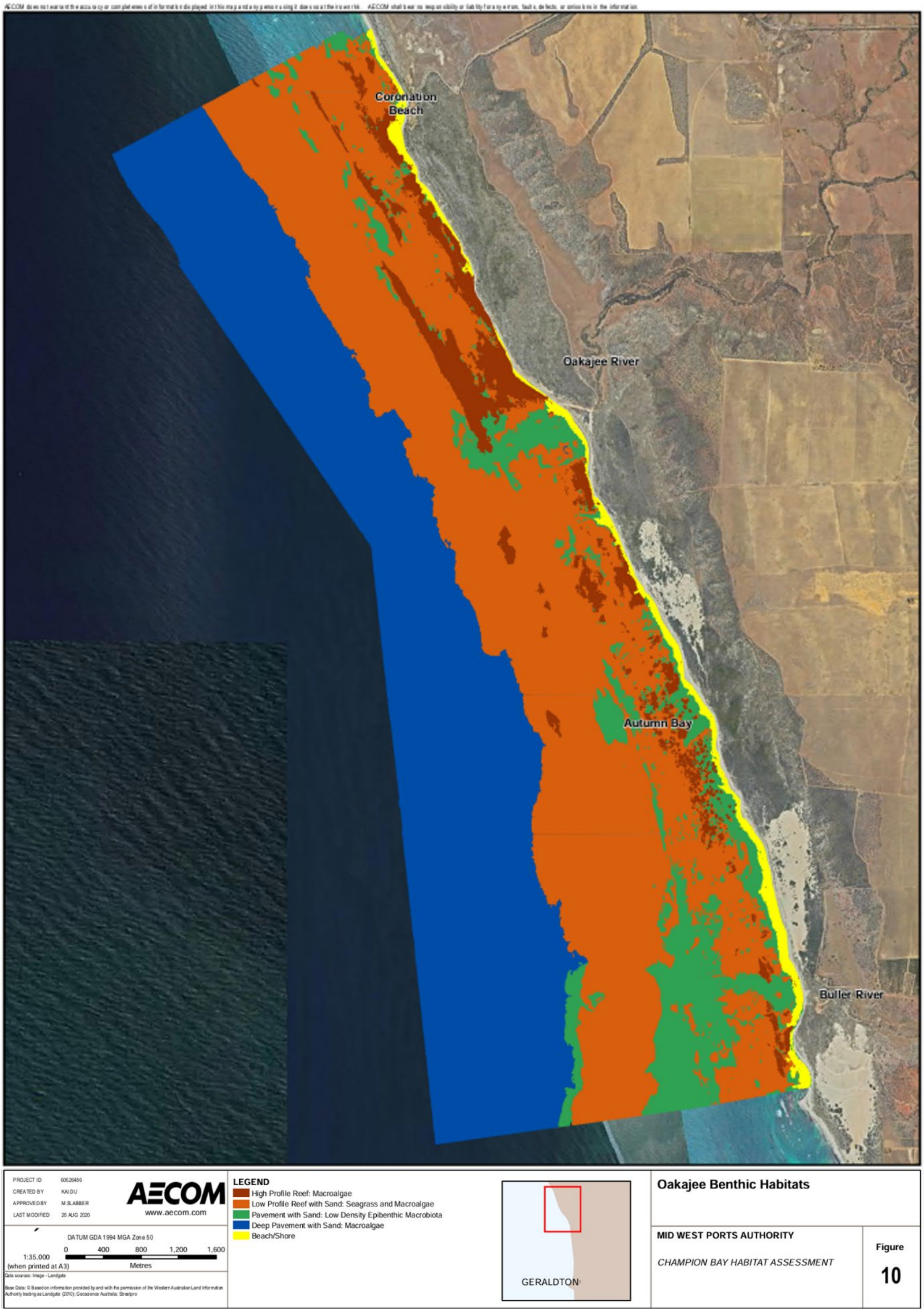


Figure 2: Oakajee benthic habitat map produced by AECOM (2020)

3. Approach and Methods

3.1. Overarching Approach / Survey Design

An overview of the approach and methods used to develop the 2025 Oakajee benthic habitat map, including the key stages from pre-field, the field survey campaign for collection of ground truth data, the post-field processes and reporting, is presented in Figure 3.

The pre-field methodologies associated with the acquisition and preparation of these datasets are described in Section 3.1 to 3.2, with the approach used to complete the ground-truthing survey described in Section 3.3, while post-field classification of ground truth data and the processes used to develop the final habitat map are described in Section 3.4 to 3.5.

To accommodate differences in spatial data availability and depth-related constraints, three distinct mapping zones (Figure 4) were defined:

- High-Resolution Zones
    - Shallow High-Resolution Zone (< 20 m depth) – Mapped using a machine learning supervised classification integrating high-resolution LiDAR-derived bathymetry, satellite imagery, and derived indices.
    - Deep High-Resolution Zone (> 20 m depth) – Mapped using machine learning classification based solely on high-resolution bathymetry.
- **Low-Resolution Zones**
    - Mapped using manual delineation with a coarse 250 m depth model.

This multi-method approach was necessary to account for the varying availability of spatial data and the limitations imposed by depth, ensuring complete coverage and maximising classification accuracy across the Oakajee mapping area.

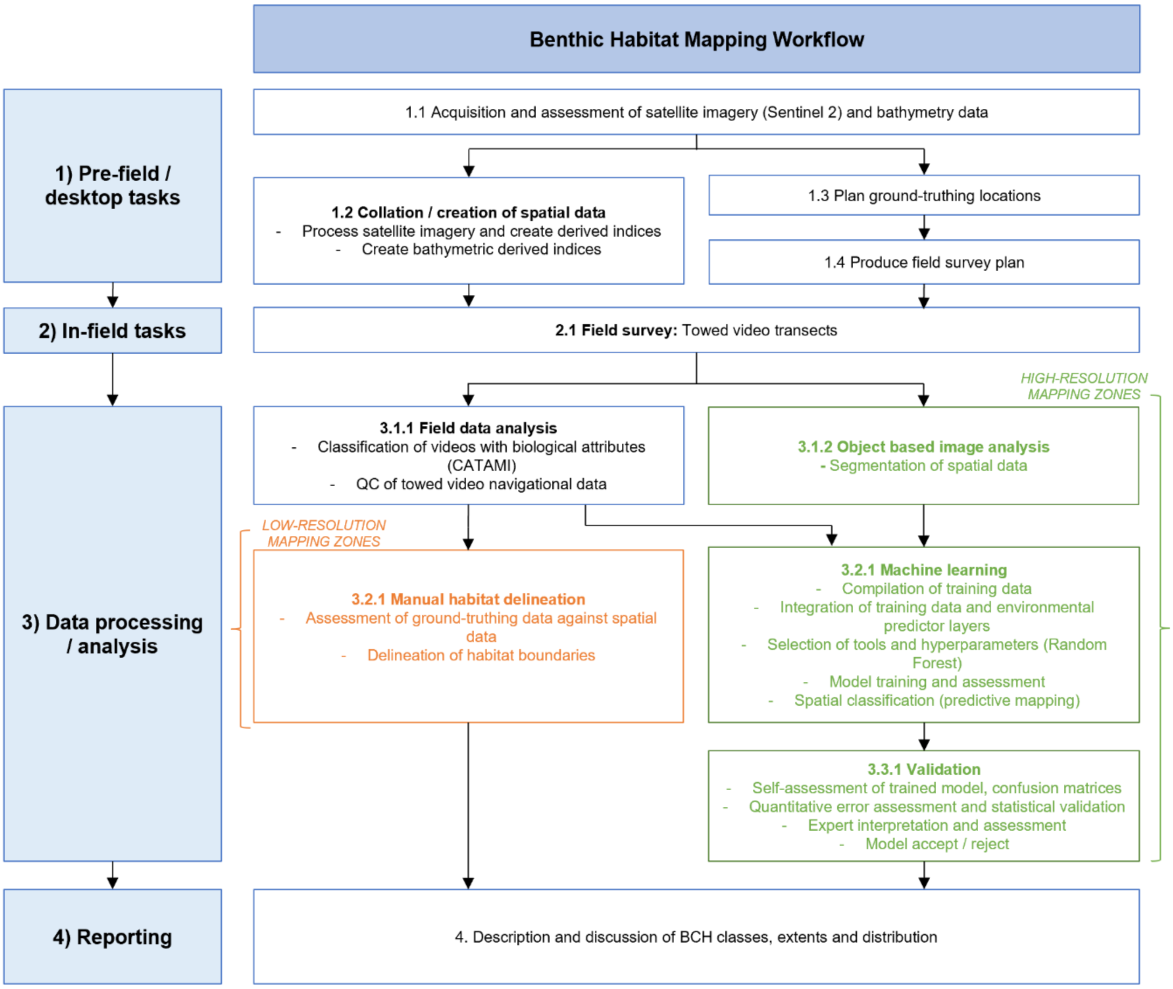


Figure 3: Workflow diagram of the habitat mapping processes

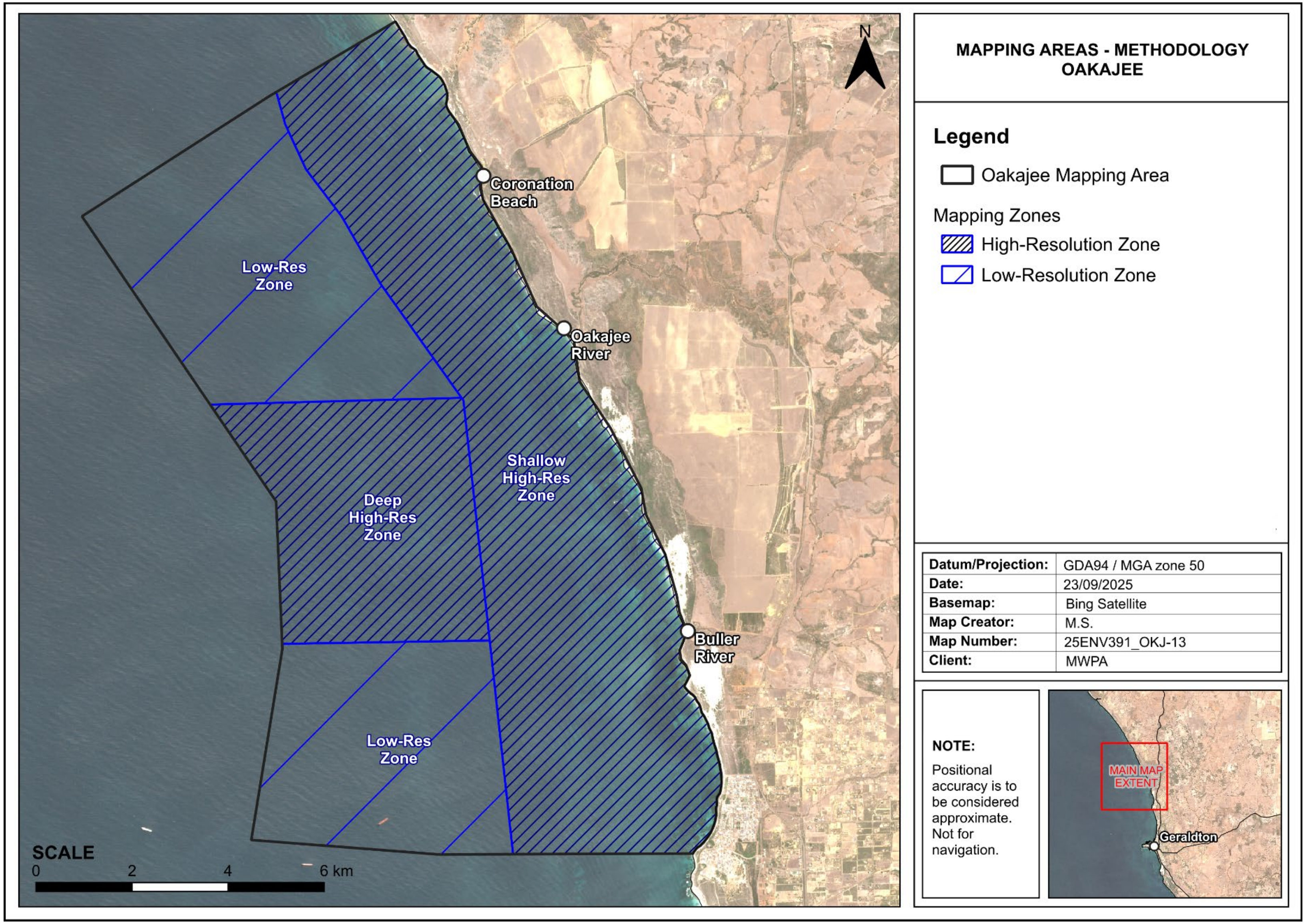


Figure 4: Oakajee mapping zones including; Shallow High-Resolution Zone, Deep High-Resolution Zone, Deep Low-Resolution Zones

3.2. Collation and Assessment of Spatial Data

3.2.1. High-Resolution Zones

Environmental predictor layers are datasets that act as proxies for habitat distribution. By sampling the values of these layers at the known locations of habitats (ground-truthing data), profiles of physical characteristics of each habitat type can be assembled and as such used to predict the distribution of these habitats across the area of interest.

Environmental predictor layers are derived from two main sources:

- Satellite imagery
- Bathymetric data

Satellite-derived Layers

The following section relates to the **Shallow High-Resolution Zone** only.

Images from the Sentinel-2 satellite system (10 x 10 m grid cell resolution) have been reliably used for benthic habitat mapping and have been shown to produce high accuracy in mapping seagrass habitats (Wicaksono et al. 2021).

An excellent quality Sentinel-2 image was identified, dating from 10/5/2025, selected from a large number of Sentinel 2A scenes of the Geraldton region. Image selection primarily focussed on identifying periods of low ocean turbidity and minimal sea state as close as possible to the period of 2025 ground truth data acquisition.

The selected image was deglinted, and depth invariant indices were calculated (Table 1), creating water column corrected imagery of the substrate useful to a depth of at least 16 m. The effects of sun glint were removed using the methods described by Hedley et al. (2005).

Table 1: Satellite bands and derived indices used as environmental predictor layers

Environmental Predictor Layer	Derivation	Comment	Reference
B02	Sentinel 2 band (490 nm)	Blue band	N/A
B03	Sentinel 2 band (560 nm)	Green band	N/A
B04	Sentinel 2 band (665 nm)	Red band	N/A
DII23	Sentinel 2	$\frac{\text{Blue}}{\text{Green}}$	Lyzenga (1978)
DII24	Depth Invariant Ratio	$\frac{\text{Blue}}{\text{Red}}$	Lyzenga (1978)
DII34	Sentinel 2	$\frac{\text{Green}}{\text{Red}}$	Lyzenga (1978)
Turbidity	Depth Invariant Ratio	$\frac{\text{Green} \times \text{Red}}{\text{Blue}}$	Pisanti et al. (2022)

Depth Invariant Index (DII)

The effect of water depth on benthic reflectance values was compensated for using a simple linear regression, following the methodology of Lyzenga (1978) and Green et al. (2000), using:

$X_i = -\ ln(R_i - R_i^{deep})$, where:

- R_i is the pixel reflectance in band i, and

- R_i^{deep}, is the deep-water reflectance in that band.

A ratio of Bands 2 and 3 was used to maximise water penetration, computed from sample pixels over the same bottom type at different depths, and a reference deep-water sample.

Turbidity

Satellite remote sensing instruments can obtain an optical measurement of water turbidity as it increases the backscattering of light (Pisanti et al. 2022). Multiple studies have found correlation between the in-situ measurements and the individual bands known to be most sensitive to water turbidity, namely Sentinel bands blue (B2), green (B3), red (B4). The index ratio showing the best correlation was:

$$\frac{(B3 \times B4)}{B2}$$

While a site-specific regression against field samples was not possible, this band ratio was applied to the 2025 Champion Bay data processing procedures, as it provides a suitable approximation of relative turbidity.

Bathymetric-derived layers

Datasets

LiDAR bathymetry – Shallow High-Resolution Zone

The primary dataset used was LiDAR bathymetry data collected of the Geraldton and Oakajee marine areas (Figure 4) by the Western Australian Department of Transport (DoT) in 2016 (Archive 17620702 and 17620703). The product has a grid cell resolution of 2 x 2 m, making it ideal for mapping large regions such as the study area. In addition to depth, a LiDAR intensity grid was also available for analysis. The dataset was cropped to cover the full extents of the study area.

Preliminary data analysis was undertaken using a topographic gridding and sun-shading tool, highlighting terrain features. The dataset was assessed for quality within the proposed mapping area. Significant artefacts in the bathymetry data can have an impact on benthic classification results, potentially causing misclassification and unreliable results.

MBES bathymetry – Deep High-Resolution Zone

A multibeam echosounder (MBES) survey output from PHS (2023) was available for part of the deeper area of the Oakajee mapping area (Figure 4). The product has a grid cell resolution of 1 x 1 m. Preliminary data analysis was undertaken using a topographic gridding and sun-shading tool, highlighting terrain features. The dataset was assessed for quality within the proposed mapping area, and the data was filtered to remove inter-swath stripe noise and merged with the LiDAR bathymetry (at a cell size of 2 m).

Derived indices

The integrated MBES and LiDAR dataset was used to create a large number of potential predictor layers, which were first compared to each other using a Pearsons correlation to eliminate redundancy. Seven layers were carried through for subsequent classification (Table 2). Derivatives are obtained by calculating, for each pixel in a primary data layer, a summary statistic from the values of all surrounding pixels within a defined neighbourhood (Olaya & Conrad 2009; Fisher et al. 2017; Wilson & Gallant 2000).

3.2.2. Low-Resolution Zones

In the absence of any available high-resolution bathymetry, a depth model from Geoscience Australia (2024) was used to provide depth and terrain information for the remaining deep-water parts of the mapping area (Figure 5). The product features a grid cell resolution of 250 m. At such low-resolution the smoothing and averaging of seabed features mean smaller structures like reef pinnacles, channels, or steep slopes may be lost or misrepresented. Depth and slope variability are reduced, and interpolation artefacts may create false patterns. This dataset was utilised in the absence of alternatives and was not integrated with the other datasets. Benthic habitat mapping of these areas followed a separate approach.

Table 2: Bathymetric-derived indices used as environmental predictor layers

Environmental Predictor Layer	Comment	Reference
Depth	Depth in mean sea level (MSL)	
Slope	Terrain slope (Horns method)	Florinsky (2016)
Aspect	Orientation of slope	Florinsky (2016)
Multiscale Roughness	Local topographic relief	Riley et al. (1999)
Topographic Position Index	Elevation of a cell relative to its neighbouring cells (e.g. peak or pit)	Lindsay et al. (2014, 2015); Newman et al. (2018)
Profile Curvature	Rate of change in slope along a flow line / downslope	Florinsky (2017)
Plan Curvature	Curvature of a contour line at a given point on the topographic surface	Florinsky (2016)

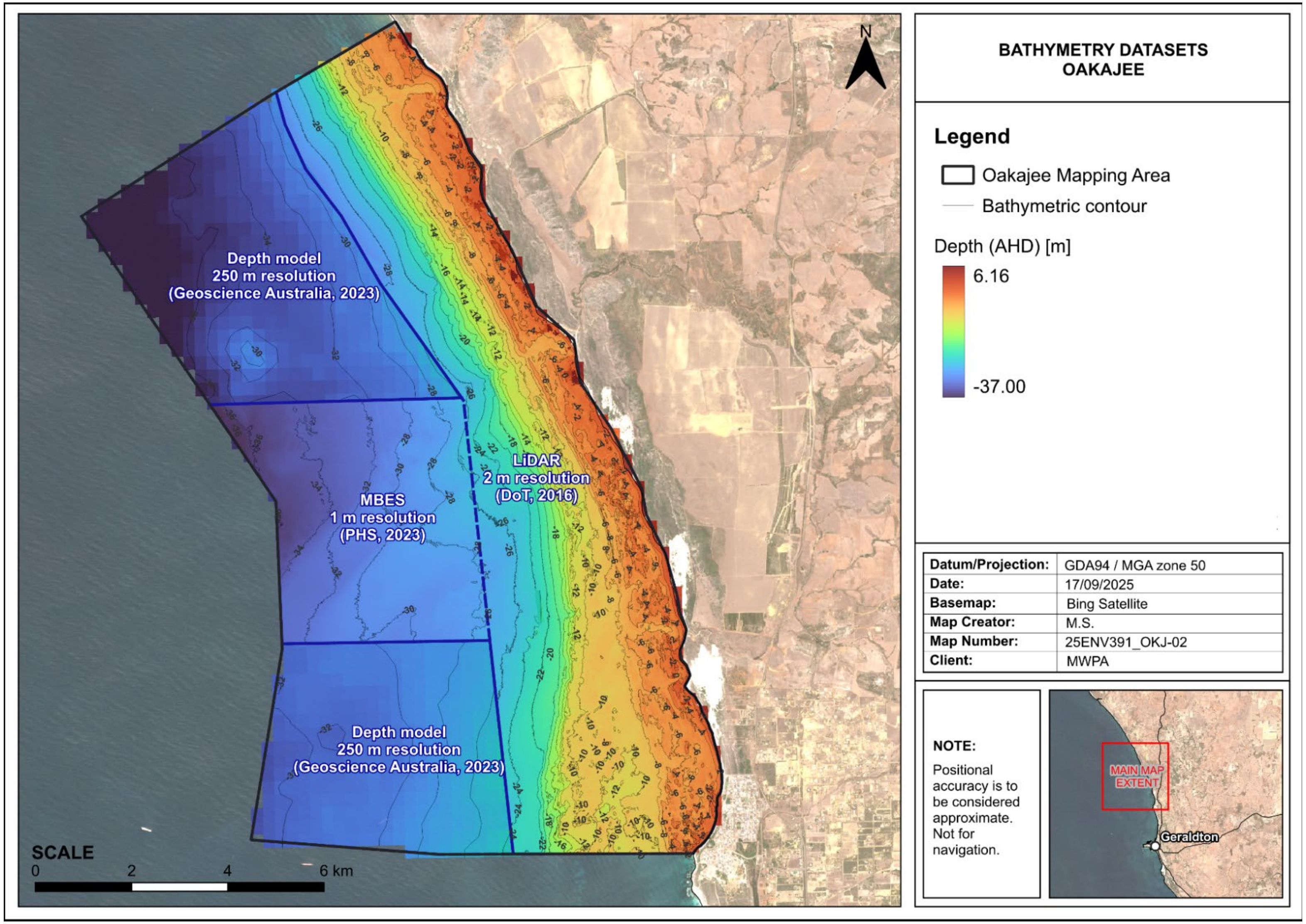


Figure 5: Bathymetry datasets used in analysis including a LiDAR dataset from DoT (2016), multibeam echosounder dataset from PHS (2023), and depth model from Geoscience Australia (2023)

3.3. Ground-Truth Survey

3.3.1. Survey Design

To obtain the primary ground-truth dataset for this study, the location of pre-determined transects were identified by a GIS analyst prior to the field survey (Figure 6), where the transects comprised both targeted and randomly generated locations (~50% of each):

- targeted transects: were positioned to verify the nature of benthic features of visible in satellite imagery and bathymetric data
- randomly generated transects: were haphazardly positioned to ensure that diverse habitat types are well represented, reducing potential bias from under- or over-sampling certain classes.

Each ground-truth transect was planned to be approximately 50 m in length.

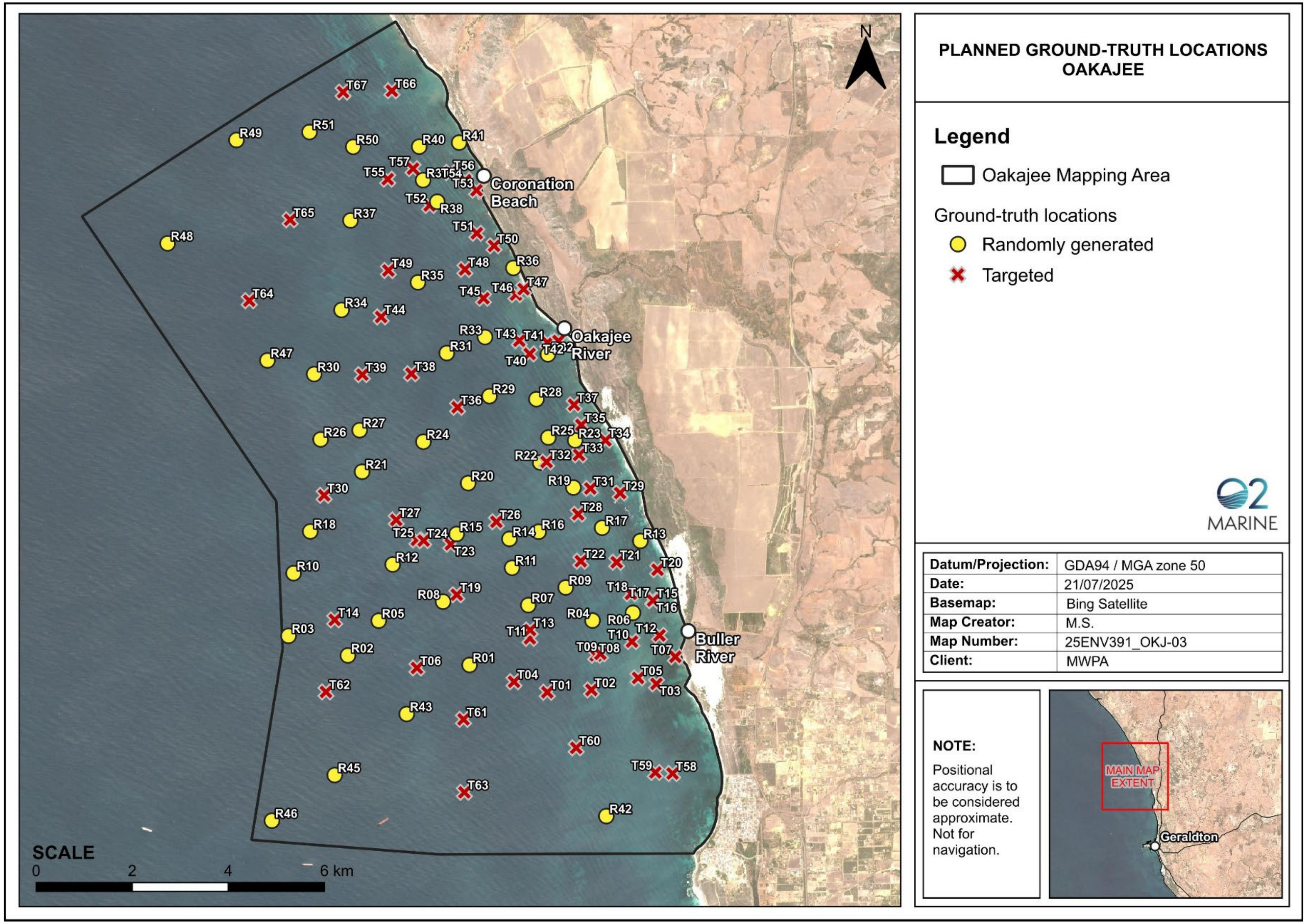


Figure 6: Planned ground-truth locations across the Oakajee mapping area

3.3.2. Equipment

Ground-truth survey acquisition was carried out across two days (12/06 - 13/06/2025). The survey took place onboard a locally operated 6.6 m charter vessel, *'Sarafore'*. The towed video camera system used for the ground truth survey was Technautics custom modular tow camera system (Figure 7). The system was operated using a topside video app, where a live video feed can be viewed, and settings can be adjusted. GPS positioning was recorded on a GlobalSat BU 353S4 receiver, and backup tracks were recorded on a handheld Garmin GPS. A DJI Action Camera was attached to provide a backup recording for each transect.

Figure 7: Technautics towed video system used for the ground-truth survey

During the survey, the camera was flown at a depth of approximately 0.5 m above the seabed, with the operator recording between 3 to 4 minutes (60 - 80 m) of benthic video footage at each transect. Vessel speed ranged between 1-2 knots to allow for the acquisition of good-quality imagery.

Key information was recorded on field sheets at each targeted location including date and time, GPS coordinates, water depth, dominant BCH type, and comments to aid post-processing.

3.3.3. Survey Effort

Daily survey effort is outlined in Table 3. Overall survey effort is illustrated in Figure 8. Not all planned sites were surveyed, as the survey design intentionally included a surplus of locations to ensure ample options were available during fieldwork.

Table 3: Daily survey effort for towed video

Survey Date	Towed Video Transects	Transect Distance (km)
12/06/2025	32	3.1
13/06/2025	39	3.8

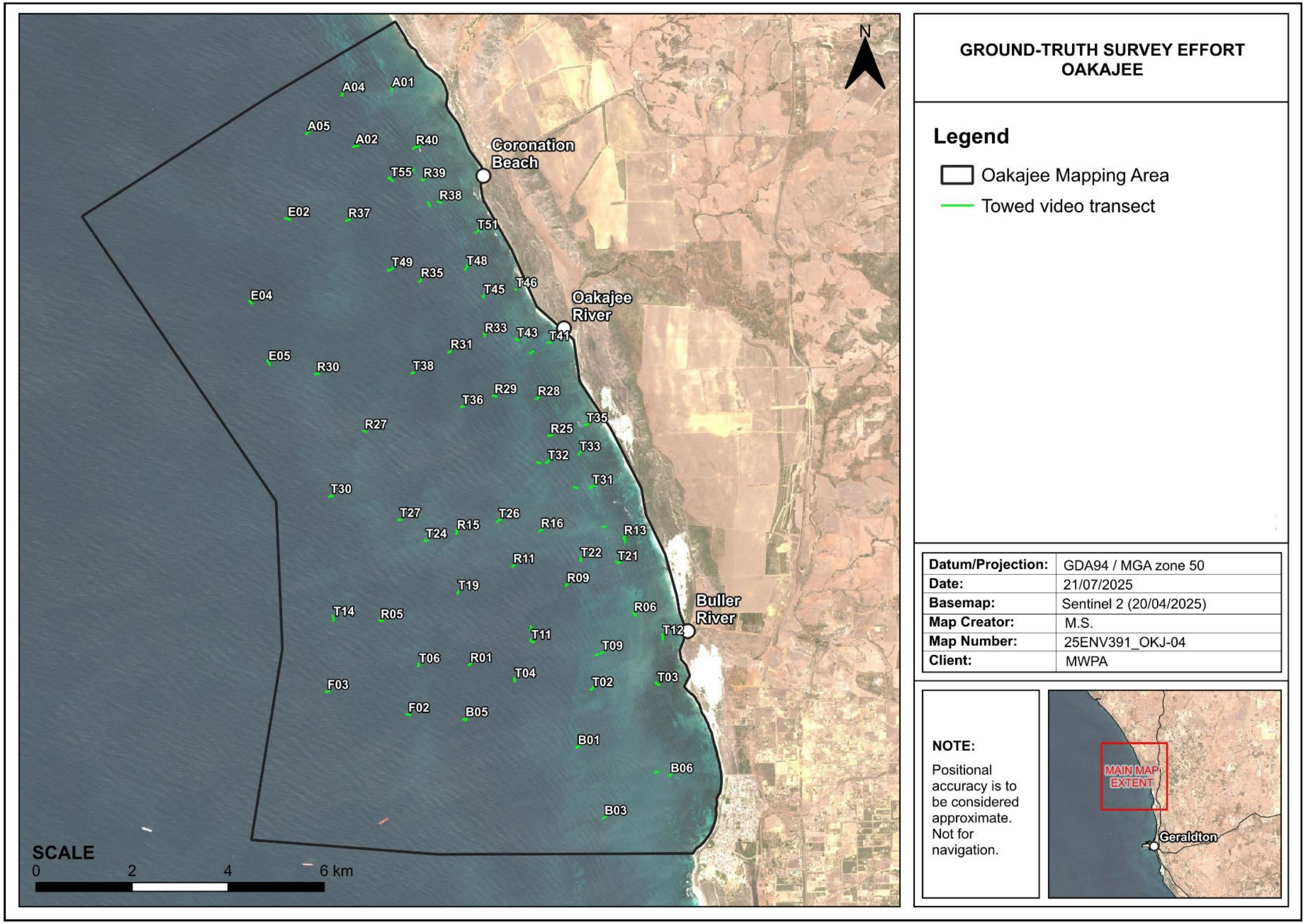


Figure 8: Ground-truthing survey effort

3.4. Classification of Towed Video Footage

Ground-truth videos were assessed for quality, and any videos with poor visibility (e.g. high turbidity obscuring the identification of biota) were discarded. Recordings were visually analysed by a suitably qualified marine scientist using TransectMeasure software and classified into habitat classes (Table 4) following the Collaborative and Automated Tools for Analysis of Marine Imagery (CATAMI) standard classification scheme for scoring marine biota and physical characteristics (Althaus et al. 2015) from underwater imagery.

A quality assurance and control check of the classifications was conducted by an experienced marine scientist specialising in benthic taxonomy and habitat classifications, including verification of percent cover estimates and species identification.

The TransectMeasure data output was synced with the GPS track log, as well as the corresponding auxiliary information (time, depth), to attribute the appropriate BCH characteristics at each point location. GPS offsets and cable lengths, which were recorded during the field survey, were applied to the transects to obtain the most accurate positioning of data. Navigation from each transect was checked for quality control in QGIS.

Table 4: Towed video classification

Substrate	Major Biota Class	Biota (minor subcategory)	Percent cover
Sand / mud Pebbles / gravel Cobbles Boulders Rock with flat relief Rock with low relief (<1 m) Rock with moderate relief (1 – 3 m)	Bare	N / A	Bare (< 1%) Sparse (1 - 10%) Low (10 - 20%) Moderate (20 – 50%) High (50 - 75%) Dense (>75%)
	Macroalgae	*Ecklonia* spp.	
		Sargassum spp.	
		Brown macroalgae	
		Green macroalgae	
		Red macroalgae	
		Filamentous macroalgae	
		Mixed macroalgae	
	Seagrass	*Amphibolis* spp.	
		Halophila spp.	
		Posidonia australis	
		Posidonia coriacea	
		Posidonia sinuosa	
		Syringodium spp.	
	Filter Feeders	Ascidians	
		Mixed	
		Octocorals	
		Sponges	

Substrate	Major Biota Class	Biota (minor subcategory)	Percent cover
		Soft Corals	
	Hard Coral	Free Living	
		Attached	
	Mixed Assemblage	Seagrass and macroalgae	
		Filter feeders and macroalgae	
		Filter feeders and seagrass	

3.5. Mapping Procedures

3.5.1. High-Resolution Zones

Combined-Area Classification Approach

To improve model performance, the classification was conducted over a combined spatial extent that included both Oakajee and the adjacent area of Champion Bay to the south. This approach provided access to a larger and more ecologically diverse training dataset, allowing the model to learn a broader range of habitat–spectral relationships, incorporate a higher representation of rare habitat types and reduce the risk of overfitting to local conditions around Oakajee.

Although the model was trained and run across the combined extent, this report focuses exclusively on mapping outputs for the Oakajee mapping area. Input datasets and results for the Champion Bay area are documented separately in O2 Marine (2025).

OBIA

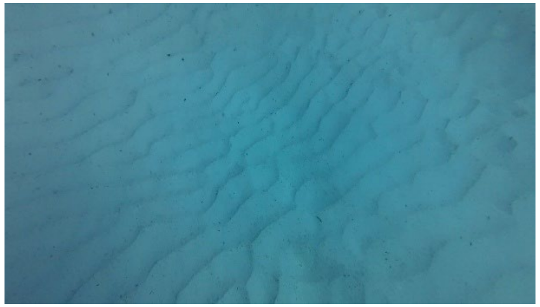
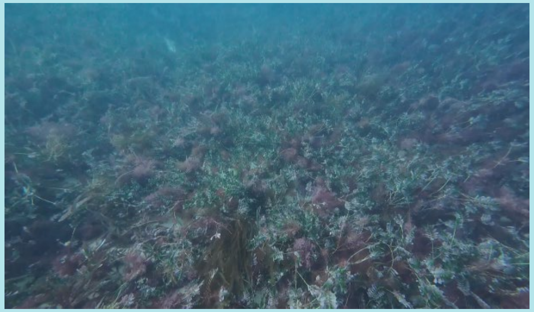
In order to integrate different scale ground truthing point observations, 2 m Lidar bathymetry data and 10 m Sentinel 2 data products, an object-based image analysis (OBIA) technique was employed. OBIA is a method of image analysis that groups pixels into meaningful objects (polygons) based on spectral, shape and neighbourhood properties (Hossain & Chen 2019). This allows integration of data of different scales, reduction of speckle noise, and faster processing times. For the Shallow High-Resolution Zone, segmentation was undertaken on a high-resolution image of the study area using a meanshift algorithm, resulting in the creation of polygons, ranging in surface area between approximately 20 – 80 m^2, adhering to the 'shapes' of visible seabed features. For the Deep High-Resolution Zone, segmentation was undertaken on the MBES dataset. The polygons are attributed with ground-truthing and environmental predictor layer statistics and subsequently subjected to classification techniques.

Training Data for Machine Learning

Training data refers to the input dataset used to train a machine learning model. For this study, the primary dataset comprised classified towed video observations collected during the 2025 Oakajee ground-truthing campaign. To improve class balance and model robustness, supplementary data were incorporated from the Champion Bay area, immediately south of Oakajee (O2 Marine, 2025). These additional data provided greater spectral and environmental variability, helping the model distinguish between classes with overlapping signatures. The Champion Bay dataset was well suited for integration with the primary dataset, as it was collected and classified using the same methodology.

The ground-truthing data were linked to the OBIA polygons which intersect spatially. Dominant and subdominant biota classes, and substrate classes were assigned to each polygon. The classes were assessed for significance, with many hundreds of classes possible through substrate and biota combinations. Only the most commonly occurring classes were retained, and rare classes merged with similar more common classes. Following assessment, ten mapping categories (Table 5) were decided and assigned to the training data points (Figure 9).

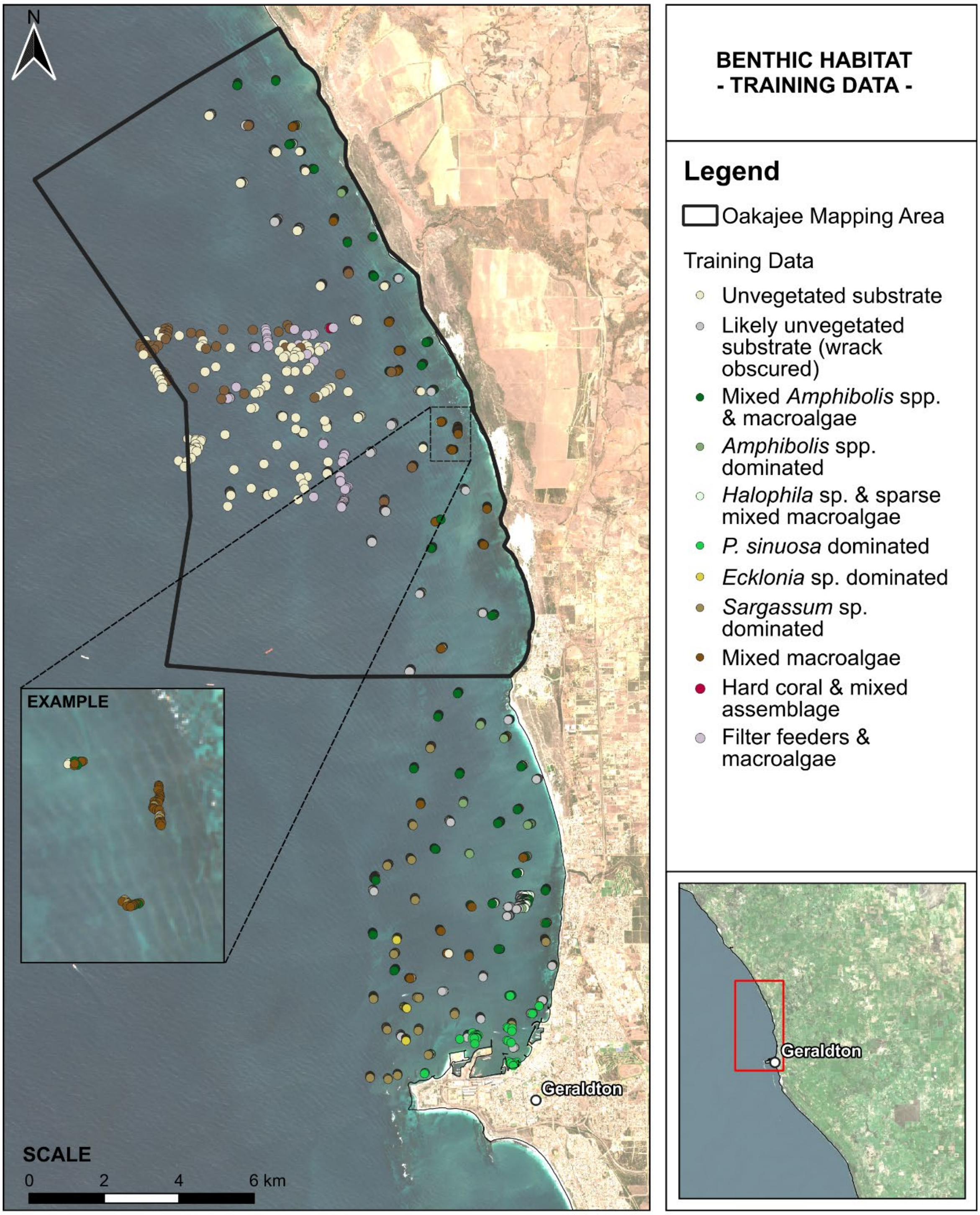
Table 5: Mapping classifications with examples from ground-truth video

BCH Classification	Description	Example Images From Ground-truth Video	
Unvegetated substrate	Bare sediment with flat relief or exhibiting 2D/3D ripple patterns, supporting sparse (<3%) or no visible biota.		
Mixed *Amphibolis* spp. & macroalgae	Heterogeneous cover of *Amphibolis* species and macroalgae, including but not limited to *Sargassum* sp., various red and brown macroalgae, and epiphytes. Typically found over pavement reef and stable sand veneer.		
Halophila sp. & sparse mixed macroalgae	Area of seagrass dominated by *Halophila* species. Typically found on unconsolidated sediment or patches of sediment within areas of pavement reef.		

BCH Classification	Description	Example Images From Ground-truth Video	
Posidonia sinuosa dominated	Area of seagrass dominated by *Posidonia sinuosa*. Typically growing on areas of unconsolidated sediment. Limited macroalgal cover, may feature some epiphytes and leaf detritus.		
Ecklonia sp. dominated	*Ecklonia* sp. dominated macroalgal assemblage on low- to moderate relief reef. Typically high – dense levels of cover with limited visibility of the substrate beneath. May co-occur encrusting red algae, or turfing species, but *Ecklonia* clearly dominates the vertical structure and biomass.		
Sargassum sp. dominated	Area of macroalgae dominated by *Sargassum* species on limestone pavement or low- to moderate relief reef. Some co-occurring biota may include filamentous or encrusting algae.		

BCH Classification	Description	Example Images From Ground-truth Video	
Mixed macroalgae	Macroalgal assemblage with no single dominant species or genus. May include *Sargassum* sp., *Ecklonia* sp., various red, brown and green filamentous macroalgae. Occurs over limestone pavement, rubble, or low-relief reef.		
Likely unvegetated substrate (wrack obscured)	Benthic areas that are likely to be bare or sparsely vegetated, but true substrate or biotic cover cannot be clearly identified due to the presence of drift algae, seagrass wrack. It is a provisional classification, used to account for spectral ambiguity in imagery caused by accumulated organic material on the seafloor.		
Filter feeders & macroalgae	A mix of filter feeders, such as sponges and soft corals, present at varying levels of coverage, with mixed macroalgae occurring as a subdominant biota.		

BCH Classification	Description	Example Images From Ground-truth Video
Hard coral & mixed assemblage	Habitat characterized by the presence of hard corals, commonly accompanied by mixed macroalgae and filter feeders.	 



Datum/Projection:	GDA94 / MGA zone 50
Date:	17/09/2025
Basemap:	Bing Virtual Earth
Map Creator:	M.S.
Map Number:	25ENV391_OKJ
Client:	MWPA

NOTE:

Positional accuracy is to be considered approximate. Not for navigation.

Figure 9: Benthic habitat mapping training data

Supervised Classification

The supervised classification method used utilises Random Forest; an ensemble learning method for supervised classification that operates by constructing a large number (500) of decision trees during training. Random Forest classification uses a combination of 'tree' predictors, where each tree depends on the values of a random vector sampled independently for all trees in the 'forest'. Multiple trees are generated at each node, with classes being assigned through a majority vote (Breiman 2001). The random forest classification technique has been successfully applied in numerous benthic habitat mapping studies involving the use of bathymetry and its derivatives, and other related work (Brown et al. 2011; Hasan et al. 2012).

Using the training data (Figure 9), the known locations of identified habitats are used to query the environmental predictor layers (Figure 10). Once a signature set has been developed for each confirmed habitat location, the machine learning algorithm then interrogates the entire dataset and attempts to identify other 'suitable' background signature combinations which might also indicate the existence of the habitat.

Supervised data classification was undertaken in a Python-based software implementation based on WhiteBoxTools (Lindsay 2014). Classes are outlined in Table 5. The classification was then applied to the entire dataset, allowing the algorithm to assess the band spectral values for each pixel cell. The classification was undertaken on every Sentinel 2 image (time slice) in order to capture variation in habitat distribution over time. Classified images were further integrated for analysis using a fusion of classes (majority vote) procedure to produce a single robust classification map. This procedure integrates all classification maps to obtain a majority vote to determine the final class assigned to each cell, providing the most rigorous assessment of habitat distribution. A 'Mixed' class is assigned when no majority can be found (indicating high variability in that cell).

Following classification, a quality control check of the output raster was carried out by an ecologist. Minor areas which were considered affected by noise or artefacts in predictor layers were manually corrected to the appropriate classifications.

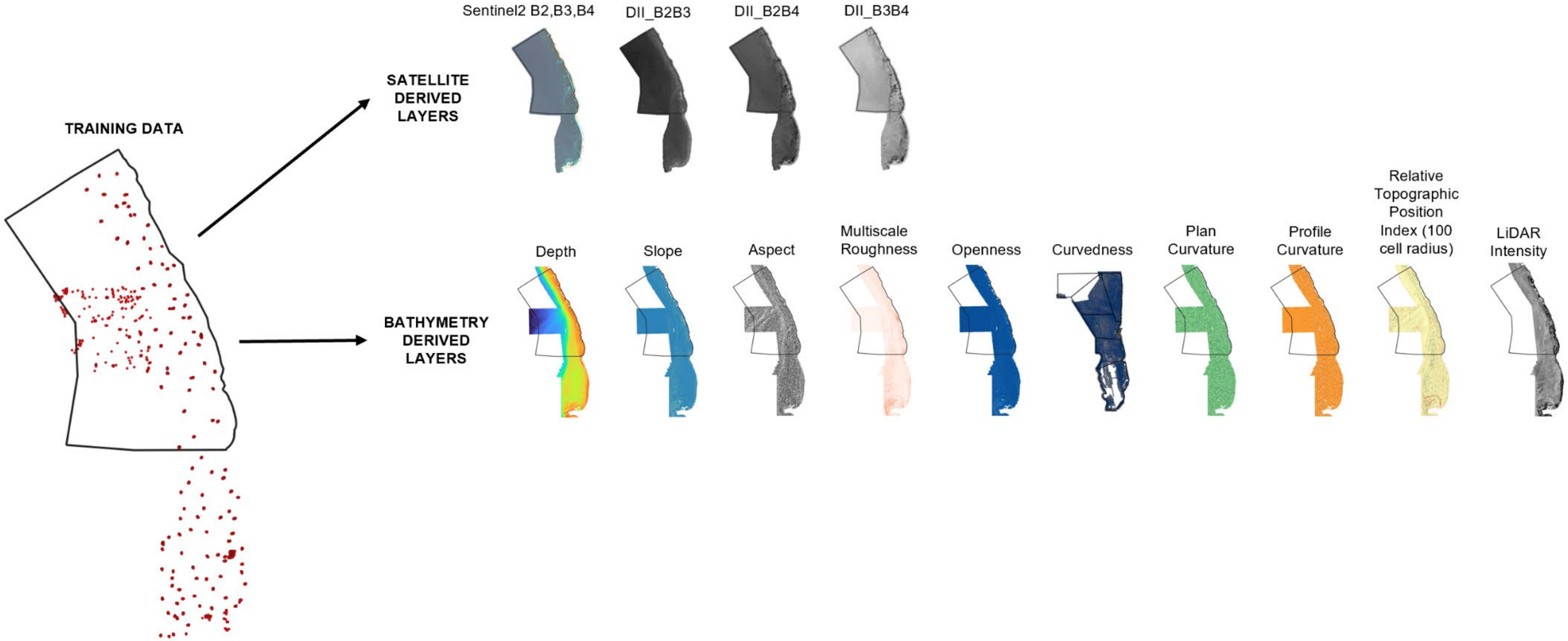


Figure 10: Schematic diagram illustrating the process of sampling the values of the environmental predictor layers at the locations of the training data

3.5.2. Low-Resolution Zones

The manual digitisation process was conducted in QGIS, where habitat boundaries were manually drawn by interpolating classified ground-truthing and referencing supporting spatial datasets, namely the 250 m depth model and derived indices, including slope and aspect. Digitisation was performed at a scale of 1:50,000.

Traditional accuracy assessments, such as error matrices and Kappa statistics, could not be applied for these areas, as these metrics are specific to supervised classification methods.

4. Results

4.1. Ground-truthing

4.1.1. Substrates

Using the 2025 towed video data, a total of 8,682 points across the Oakajee coastal region were classified with biological attributes and habitat classifications (Table 6, Figure 11). Note that due to the scale of the maps relative to the length of the transects, Figure 11, Figure 13, Figure 14 and Figure 15 do not reveal habitat and substrate changes across transects. The purpose of these figures is to provide an overview of the dataset, and classifications observed at each transect on these maps should not be considered completely representative of the broader BCH type at each location.

An overview of the substrate information classified along ground-truthing transects is presented in Figure 11, with the proportion of each substrate class presented in Table 6. The majority of points (41.47%) were assigned to Sand / Mud - Ripples, indicating widespread unconsolidated, dynamic sediments. Sand / Mud - Flat, including soft substrates, was the second most common class, comprising 18.21% of points. Rocky substrates were also well represented, with Rock - Moderate (1–3 m relief) and Rock - Low (<1 m relief) accounting for 11.30% and 10.90% of points respectively. Less common classes included Cobbles (5.51%), Flat Rock (4.37%), and Pebble / Gravel – Rubble (3.25%), while High Relief Rock (>3 m) and Boulders (>255 mm) were relatively rare, comprising just over 3% of classifications combined. Overall, the data indicate a benthic environment dominated by soft sediments, with patchy areas of rock and coarse substrate common across the shallow regions.

Table 6: Observed substrate types identified in ground-truthing data

Substrate	Number of points classified	Proportion of classified points
Sand / Mud - Ripples	3,600	41.47
Sand / Mud - Flat (inc Soft Substrate)	1,581	18.21
Rock - Moderate (1-3m)	981	11.30
Rock - Low (<1m)	946	10.90
Cobbles (>64mm)	478	5.51
Rock - Flat	379	4.37
Pebble / Gravel - Rubble (<64mm)	282	3.25
Rock - High (>3m)	186	2.14
Sand / Mud - Coarse / Shell	170	1.96
Boulders (>255mm)	79	0.91

Substrate	Number of points classified	Proportion of classified points
Total	8,682	100

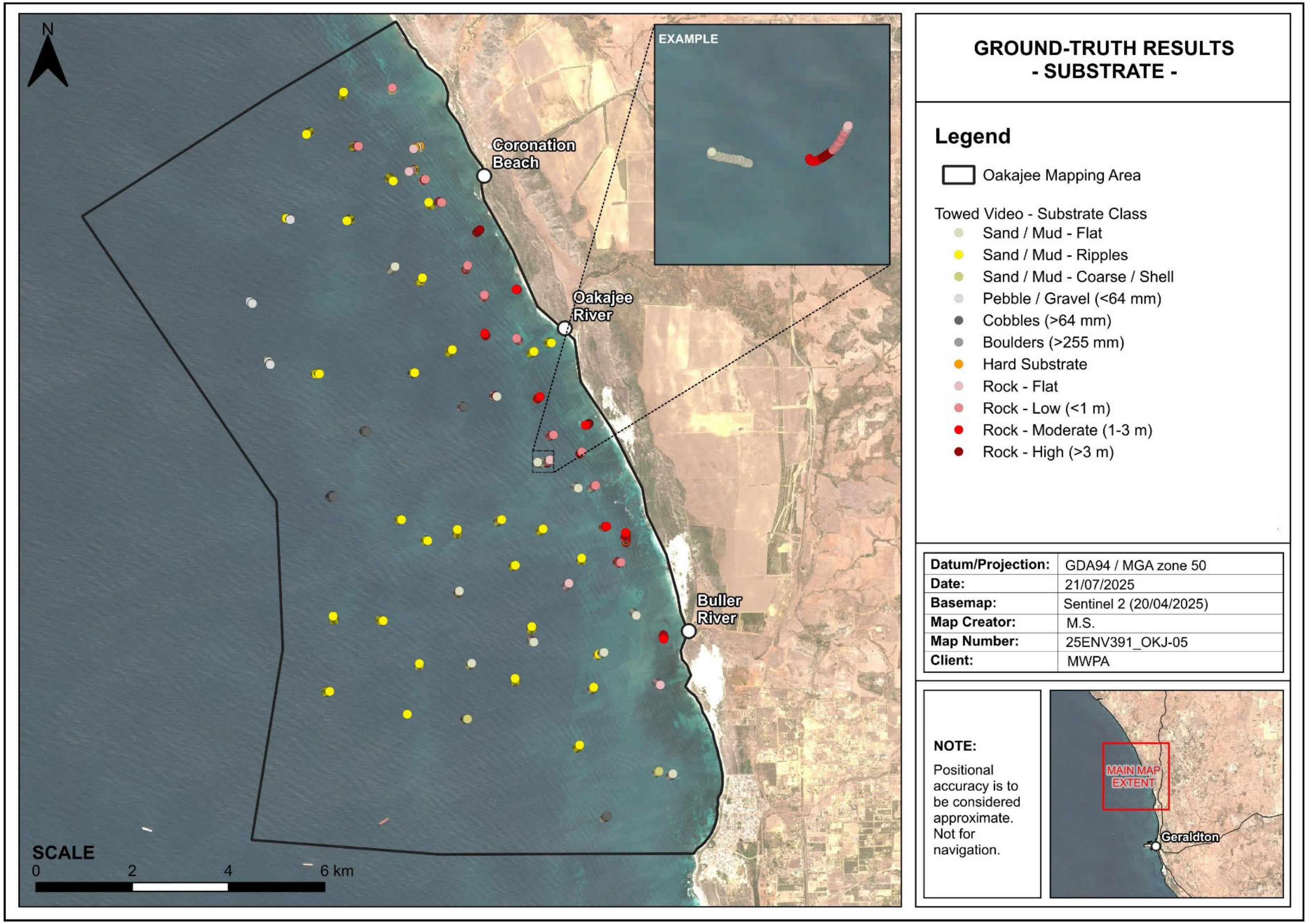


Figure 11: Ground-truthing results overview - substrate class

4.1.2. Biota

Figure 13 provides an overview of the major biota classes recorded along the ground-truthing transects, with a proportional representation of each class detailed in Table 7. Further detail can be seen in Figure 12, Figure 14, Figure 15.

Bare sediment was the most common category, comprising 29.9% of all points, indicating extensive areas with no visible biological cover. Overall, approximately 70% of the classified points demonstrated biological cover, with macroalgae being the most widespread, accounting for 22.8% of points with biota, predominantly in high (11.7%) and dense (10.0%) cover categories. Macroalgae was most frequently observed across inshore reef areas. 'Mixed Macroalgae 'comprised 93% of all points classified as macroalgae, with Sargassum spp. representing 4.1% and the remaining classes with less than 3% combined (Figure 12).

Mixed seagrass and macroalgae assemblages were also well represented, making up 14.1% of biota points, largely at a moderate level of cover. Seagrass alone contributed 6.9%, mostly in high (3.5%) and dense (2.1%) categories. Seagrass appeared only in nearshore areas (approximately <2 km from coastline), interspersed between areas of reef. Among seagrass taxa, *Amphibolis* spp. was the most prevalent (99.2% of all seagrass observations), while other seagrass types such as *Halophila* spp. and mixed *Amphibolis* & *Posidonia* were rare (<1%).

Mixed assemblages and filter feeders were present in lower proportions, comprising 5.2% and 4.6% of biota points respectively, primarily observed in the deeper waters of the western side of the mapping area. Mixed filter feeders were the most common (91.3% of observations of this class) followed by sponges (8.7%).

Although drift algae and rhodoliths were recorded across all cover classes, they were not included in the biota proportion calculations; however, they appear to be ecologically significant, particularly in low (12.3%) and moderate (3.9%) densities. Wrack was particularly prevalent, accounting for 14.9% of all dominant biota points, suggesting a strong presence of drift material.

Table 7: Number of ground-truth points classified per major biota and cover class

Biota Class	Cover	Number of points classified	Class total	Proportion of points classified (%)	Proportion of classified points with biota (%)
Macroalgae (MA)	MA – Sparse	95	2,127	1.1	2.2
	MA – Low	94		1.1	2.2
	MA - Moderate	126		1.5	2.9
	MA – High	977		11.7	22.8
	MA – Dense	835		10.0	19.5
Seagrass (SG)	SG - Sparse	14	533	0.2	0.3
	SG - Low	4		0.0	0.1
	SG - Moderate	40		0.5	0.9
	SG - High	295		3.5	6.9
	SG - Dense	180		2.1	4.2
Mixed Seagrass & Macroalgae (MSM)	MSM - Sparse		723	0.0	0.0
	MSM - Low	38		0.5	0.9
	MSM - Moderate	603		7.2	14.1

ENVIRONMENT
An O2Marine company

Biota Class	Cover	Number of points classified	Class total	Proportion of points classified (%)	Proportion of classified points with biota (%)
	MSM - High	82		1.0	1.9
	MSM - Dense			0.0	0.0
Filter Feeders (FF)	FF – Sparse	42	345	0.5	1.0
	FF – Low	198		2.4	4.6
	FF - Moderate	76		0.9	1.8
	FF – High	29		0.3	0.7
	FF – Dense	42		0.0	0.0
Mixed Assemblage (MX)	MX - Sparse	222	550	2.6	5.2
	MX - Low	127		1.5	3.0
	MX - Moderate	190		2.3	4.4
	MX - High	11		0.1	0.3
	MX - Dense	222		0.0	0.0
Bare Sediment (BS)	Bare	2,505	2505	29.9	
Drift Algae / Rhodoliths (OT)	OT - Sparse	1,034	1,603	0.0	
	OT - Low	331		12.3	
	OT - Moderate	125		3.9	
	OT - High	79		1.5	
	OT - Dense	34		0.9	
Total		8386	8,386	100	100

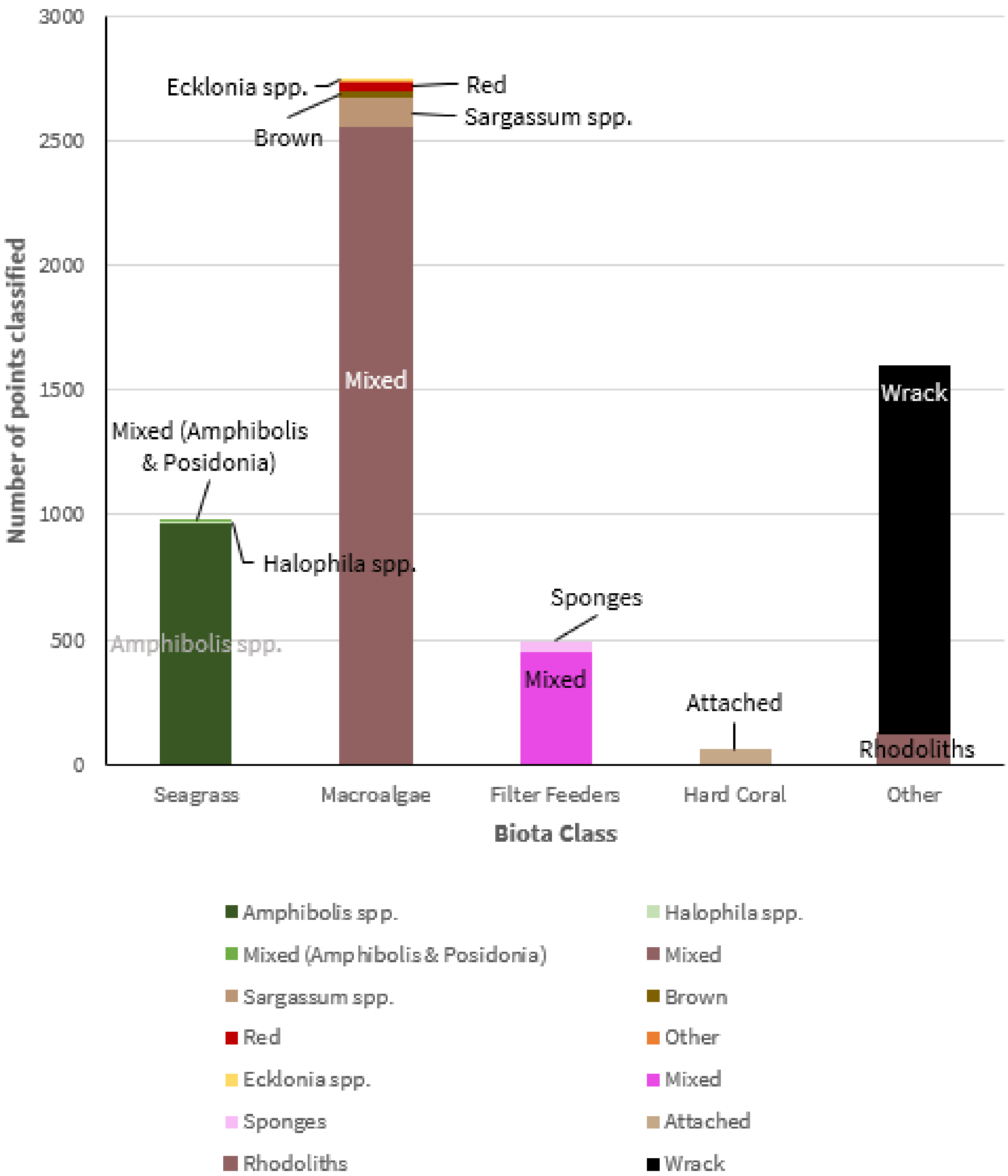


Figure 12: Number of ground-truth points classified as each dominant biota class

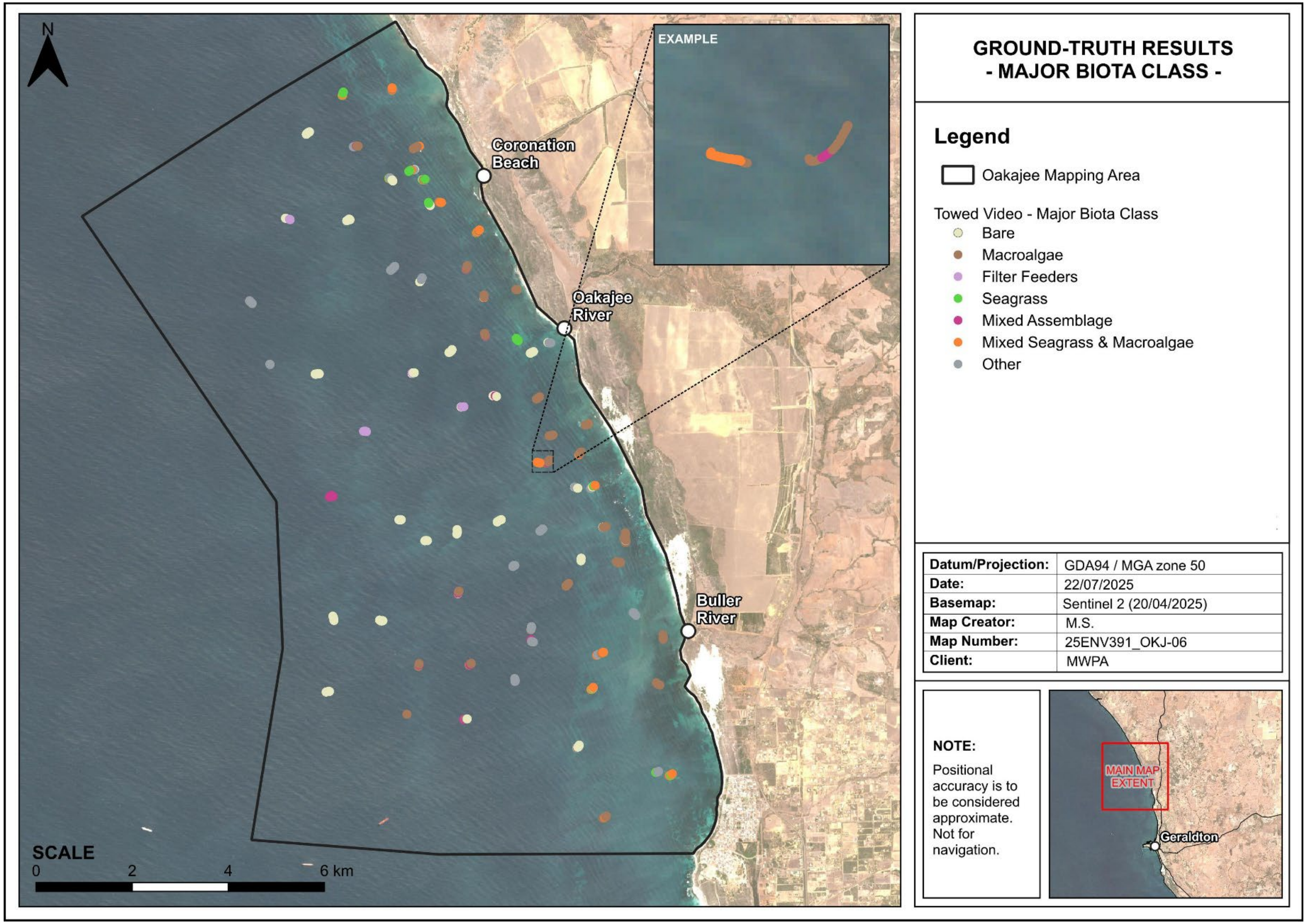


Figure 13: Ground-truthing results overview - major biota class

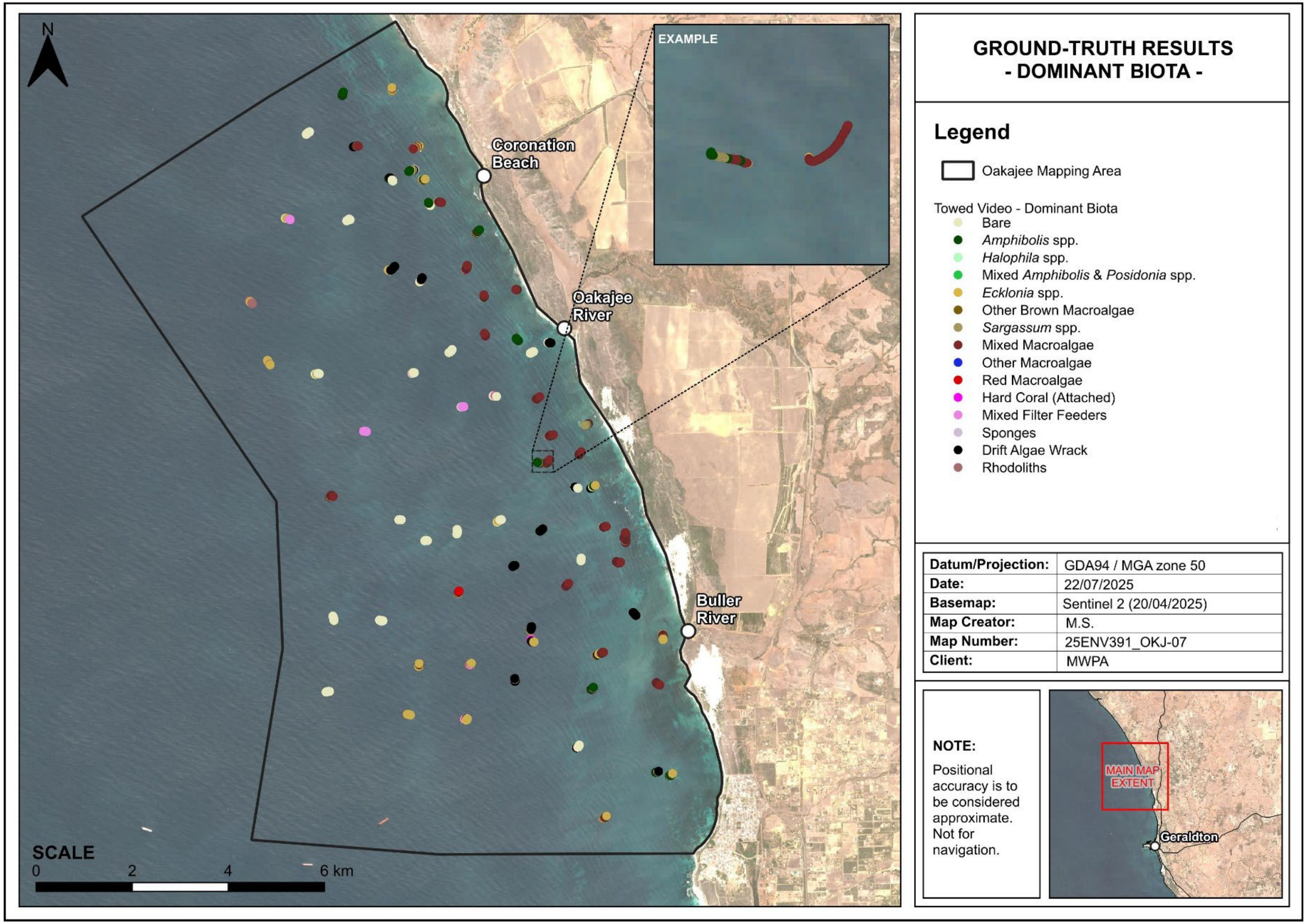


Figure 14: Ground-truthing results overview – dominant biota

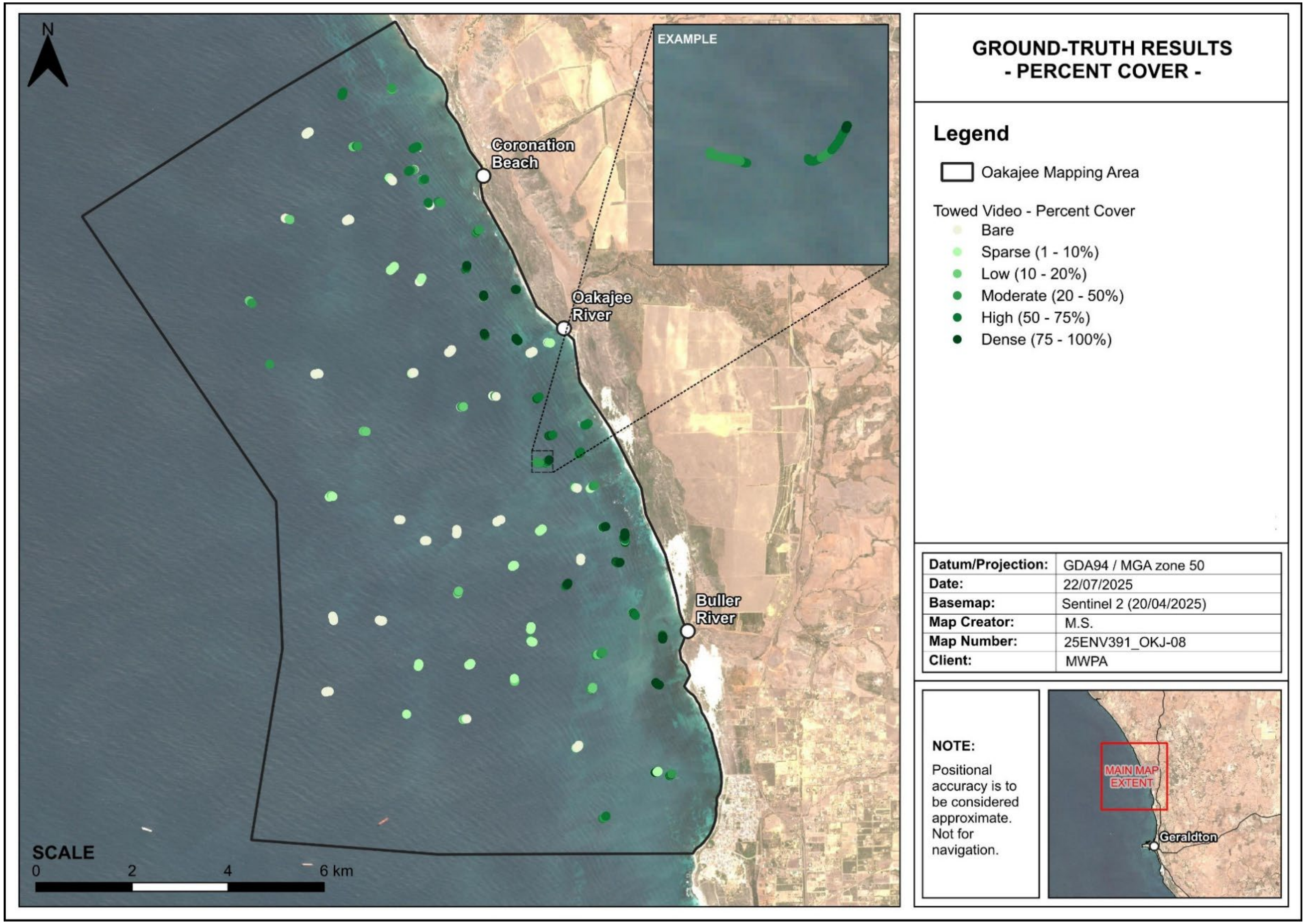


Figure 15: Ground-truthing results overview - percent cover

4.2. Mapping

4.2.1. Substrates

The substrate map of the Oakajee region presents a detailed classification of benthic habitats across a total mapped area of approximately 12,949.37 hectares (Table 8, Figure 16).

The substrate map of the Oakajee region shows that sedimentary substrates dominate the mapped area, with potential mobile sediment being the most widespread class. It covers a total of 7,668.70 hectares, comprising 3,680.07 hectares in high-resolution zones and 3,988.63 hectares in low-resolution zones, which together represent approximately 59.2% of the total mapped area. These sediments are prevalent throughout the offshore region and extend into nearshore areas, particularly in the central and southern portions of the mapping area. Stable sand veneer is the next most significant sediment type, covering 198.28 hectares (1.5%), primarily along the nearshore fringe, often overlying harder substrates. These sedimentary areas are mostly flat and unconsolidated, contributing to the generally low-relief characteristics of the seafloor in much of the offshore zone.

Reef substrates are concentrated mainly in the southern nearshore region, particularly between the Oakajee River and Buller River, where seafloor complexity increases. The most extensive reef class is pavement reef, which occupies 3,532.99 hectares, accounting for 27.3% of the total mapped area. It forms broad, continuous bands nearshore. Low relief reef covers 487.07 hectares (3.8%), and moderate relief reef occupies 177.12 hectares (1.4%), both appearing in patchy distributions and typically interspersed with pavement reef and sand veneer. In the low-resolution zones, potential pavement reef accounts for 836.97 hectares (6.5%), located in the offshore northwestern and southwestern areas. Additionally, a mixed substrate class covers 48.24 hectares (0.37%), representing areas where the model did not predict a single majority class. This class is primarily distributed across shallower zones, where substrate composition is more heterogeneous

Table 8: Substrate classes by mapped area at Oakajee

Substrate	Area (ha)	Percentage of mapped area (%)
High-Resolution Zones		
Pavement reef	3532.99	27.28
Stable sand veneer	198.28	1.53
Potential mobile sediment	3680.07	28.42
Low relief reef	487.07	3.76
Moderate relief reef	177.12	1.37
Mixed substrate	48.24	0.37
Low-Resolution Zones		
Potential pavement reef	836.97	6.46
Potential mobile sediment	3988.63	30.80
Total	12,949.37	100

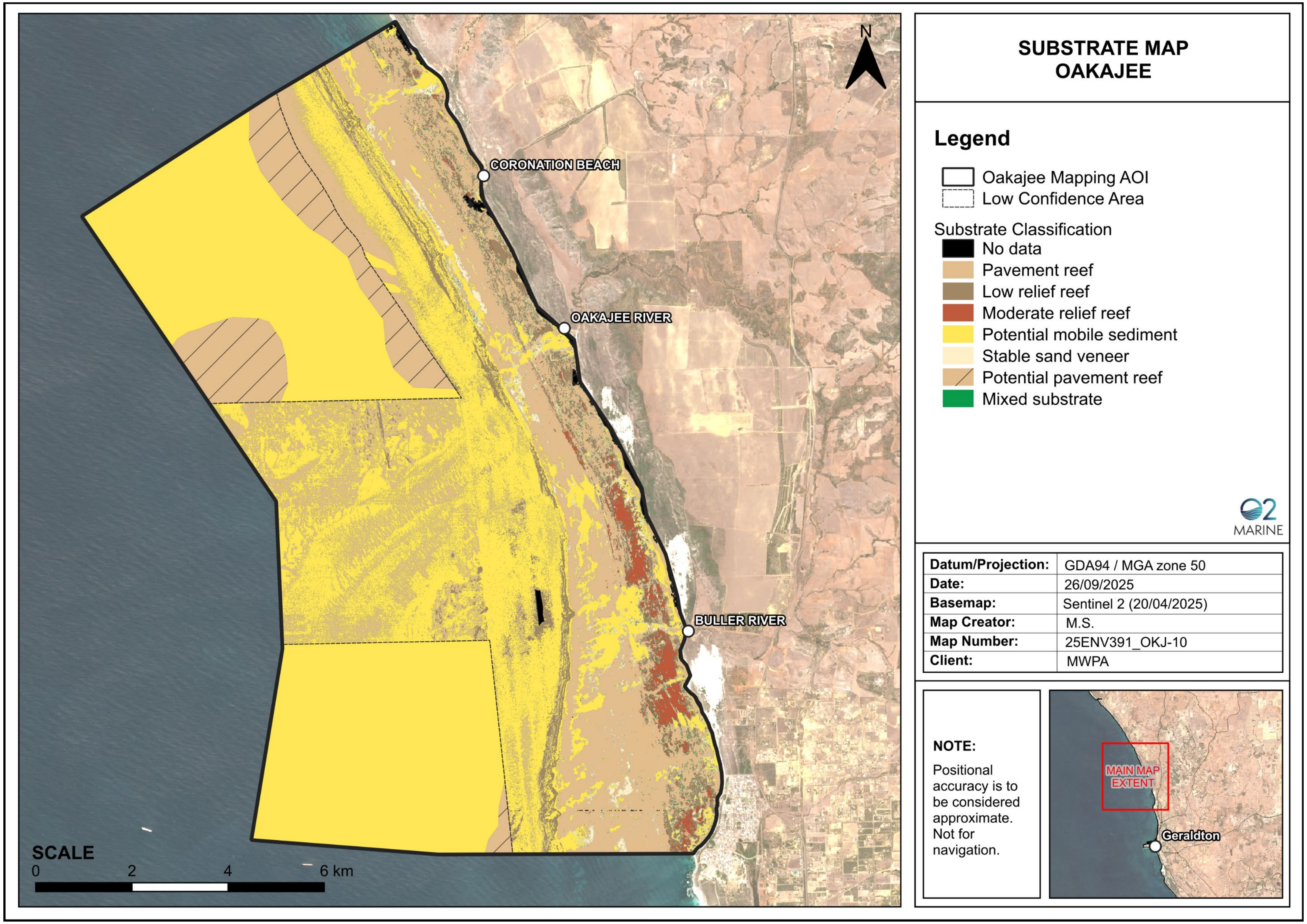


Figure 16: Substrate map of Oakajee

4.2.2. Benthic Habitats

The random forest classification procedure undertaken on the composite stack and training dataset provided a coherent 'dominant class' habitat map, which comprised fourteen final classes mapped across a total area of 12,955.2 ha (Table 9, Figure 17).

Unvegetated substrate was the most widespread habitat type, represented by three classes totaling 8,534.1 ha (65.9% of the mapped area). Half of this area lies within the low-confidence mapping zone. These substrates dominate depths greater than 20 m, forming broad expanses, while at shallower depths they occur only as scattered patches between reef systems. Some additional areas (646.9 ha; 5.0%) are likely unvegetated but remain uncertain due to wrack obscuring satellite and field observations.

Mixed *Amphibolis* spp. and macroalgae formed the most extensive biotic class, covering 1,725.0 ha (13.3%) primarily above 20 m depth. Substrate analysis (Figure 17) links this habitat to pavement reef and stable sand veneer. Three other seagrass-dominated habitats were mapped: *Halophila* sp. (44.3 ha, 0.3%), *Posidonia sinuosa* (24.2 ha, 0.19%), and *Amphibolis* spp. dominated seagrass (10.1 ha, 0.1%). These habitats occur as small, isolated patches in nearshore areas, interspersed within the more extensive *Amphibolis*–macroalgae assemblages. The *Amphibolis* spp. dominated seagrass habitat was the least common, found only in shallow areas south of Buller River.

Three macroalgae-dominated habitats were identified. Mixed macroalgae was the largest, occupying 1,341.2 ha (10.4%) of shallow reef (<20 m). *Sargassum* sp. dominated habitat (255.2 ha, 2.0%) was restricted to two north–south reef bands in the south near Buller River. *Ecklonia* sp. dominated habitat was the least extensive, limited to small, scattered patches on shallow reefs.

Filter feeders and macroalgae (614.5 ha, 4.7%) dominated deeper areas (>20 m), especially in the north, alternating with unvegetated substrate. In the deepest zones, they competed with mixed macroalgae across pavement reef. Low-confidence areas suggest further extensive patches of filter feeder–macroalgae communities in the northwest and south, associated with slope features in the depth model.

Finally, the Mixed biota class - areas without a single dominant habitat - covered 25.9 ha (0.2%), appearing as small (10 - 20 m) patches scattered throughout the map.

Table 9: Benthic habitat classes by mapped area at Oakajee

Benthic Habitat Class	Area (ha)	Percentage of mapped area (%)
High-Resolution Zones		
Unvegetated substrate	3,902.64	30.12
Mixed *Amphibolis* spp. & macroalgae	1,725.04	13.32
Mixed macroalgae	1,341.2	10.35
Likely unvegetated substrate (wrack obscured)	642.86	4.96
Sargassum sp. dominated	254.22	1.96
Filter feeders & macroalgae	152.12	1.17
Halophila sp. & sparse mixed macroalgae	44.3	0.34
Mixed biota	25.94	0.2
Posidonia sinuosa dominated	24.15	0.19
Amphibolis spp. dominated	10.07	0.08

Benthic Habitat Class	Area (ha)	Percentage of mapped area (%)
High-Resolution Zones		
Ecklonia sp. dominated	5.52	0.04
Hard coral & mixed assemblage	1.59	0.01
Low-Resolution Zones		
Potential unvegetated substrate	3,988.63	30.79
Potential filter feeders & macroalgae	836.89	6.46
Total	12,955.18	100

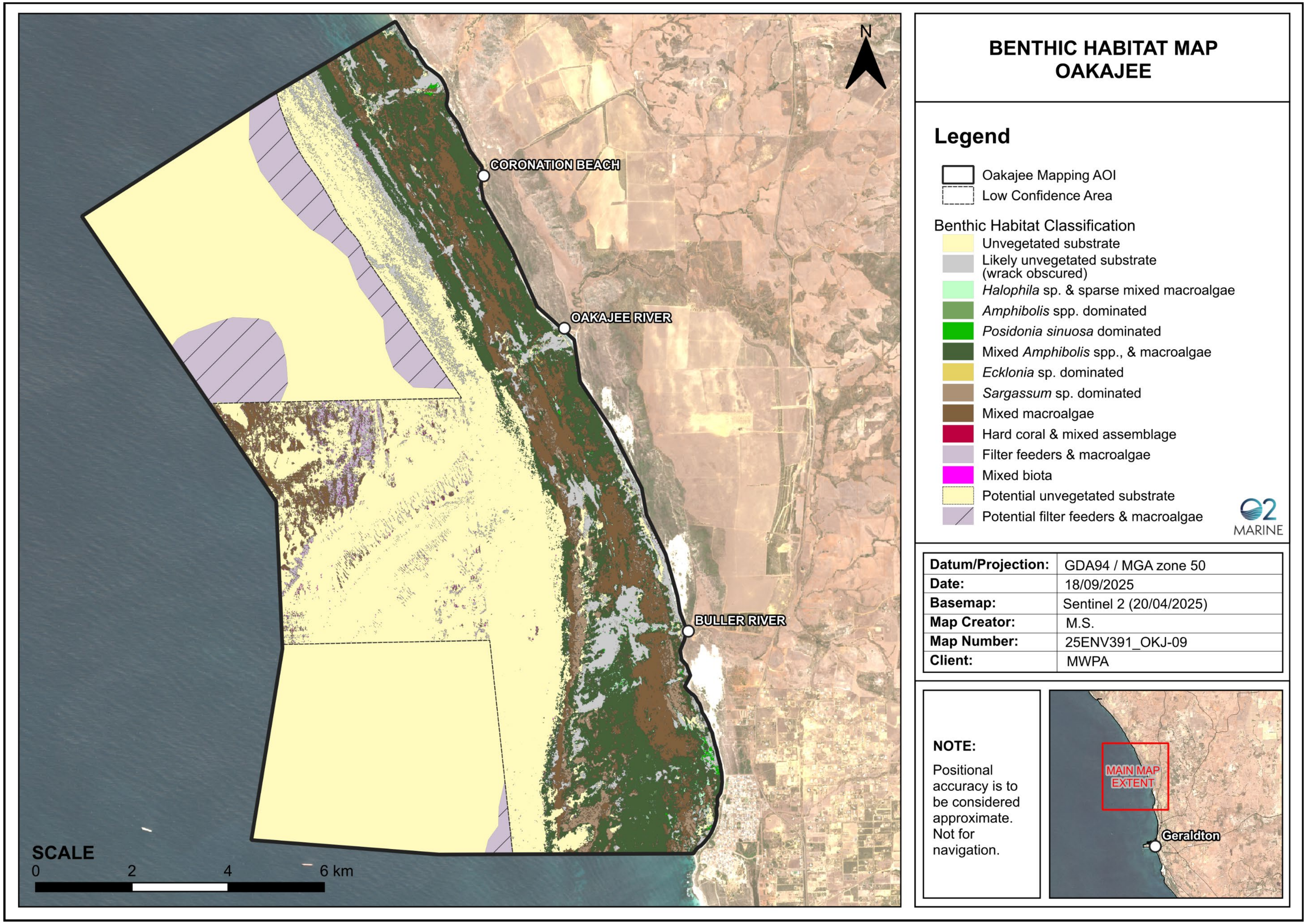


Figure 17: Benthic habitat map of Oakajee

4.3. Mapping Validation

On completion of the model training, validation metrics were generated to assess the performance of the classifier across different substrate and habitat types. The confusion matrices (Table 10, Table 11, Table 14) and class statistics (Table 12, Table 13, Table 15) present a detailed account of the model's predictive accuracy, with the rows indicating reference (true) labels and the columns depicting the labels predicted by the model. This premise is also visually depicted in Figure 18 and Figure 19.

The calculation of accuracy is derived from the confusion matrix, which compares actual vs. predicted classifications. Indicators of accuracy include key performance metrics used in classification tasks to evaluate the accuracy of a model:

- True Positives: correctly predicted as positive for each class.
- True Negatives: correctly predicted as negative i.e. not belonging to a particular class.
- False Positives: incorrectly predicted as positive.
- False Negatives: incorrectly predicted as negative.
- Precision: measures the proportion of true positive predictions out of all positive predictions made (True positives/(True Positives + False Positives)), indicating how many of the predicted positive instances are actually correct.
- Recall (also known as sensitivity): measures the proportion of true positive predictions out of all actual positive instances (True positives/(True Positives + False Negatives)), reflecting the model's ability to identify all relevant cases.
- F-score: the harmonic mean of precision and recall, providing a single metric that balances both aspects. It is particularly useful in scenarios where both false positives and false negatives are critical.

Validation statistics and confusion matrices were calculated separately for the shallow and deep water areas, owing to the different methodologies applied to classify each.

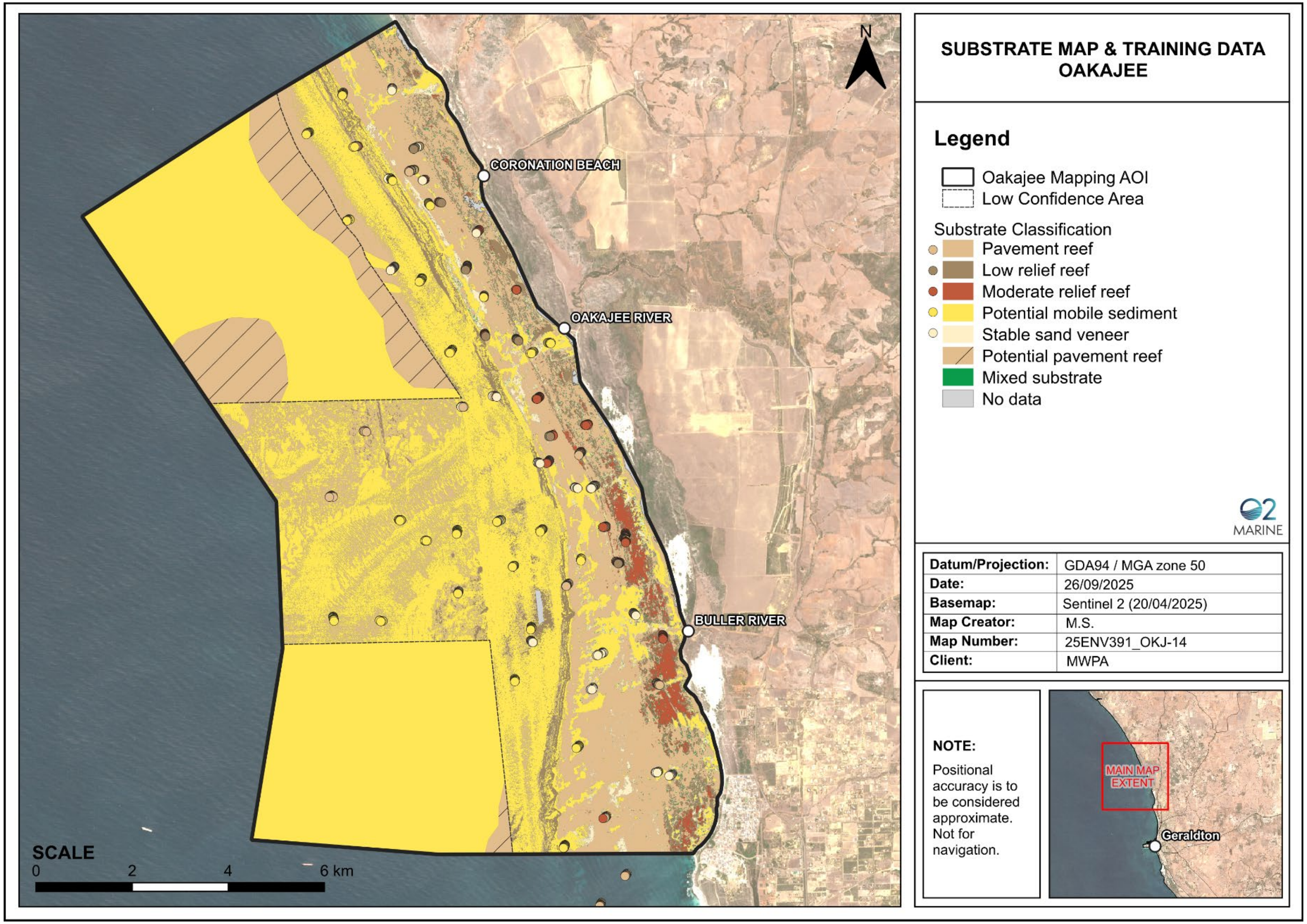


Figure 18: Model training data overlaid on the Oakajee substrate map

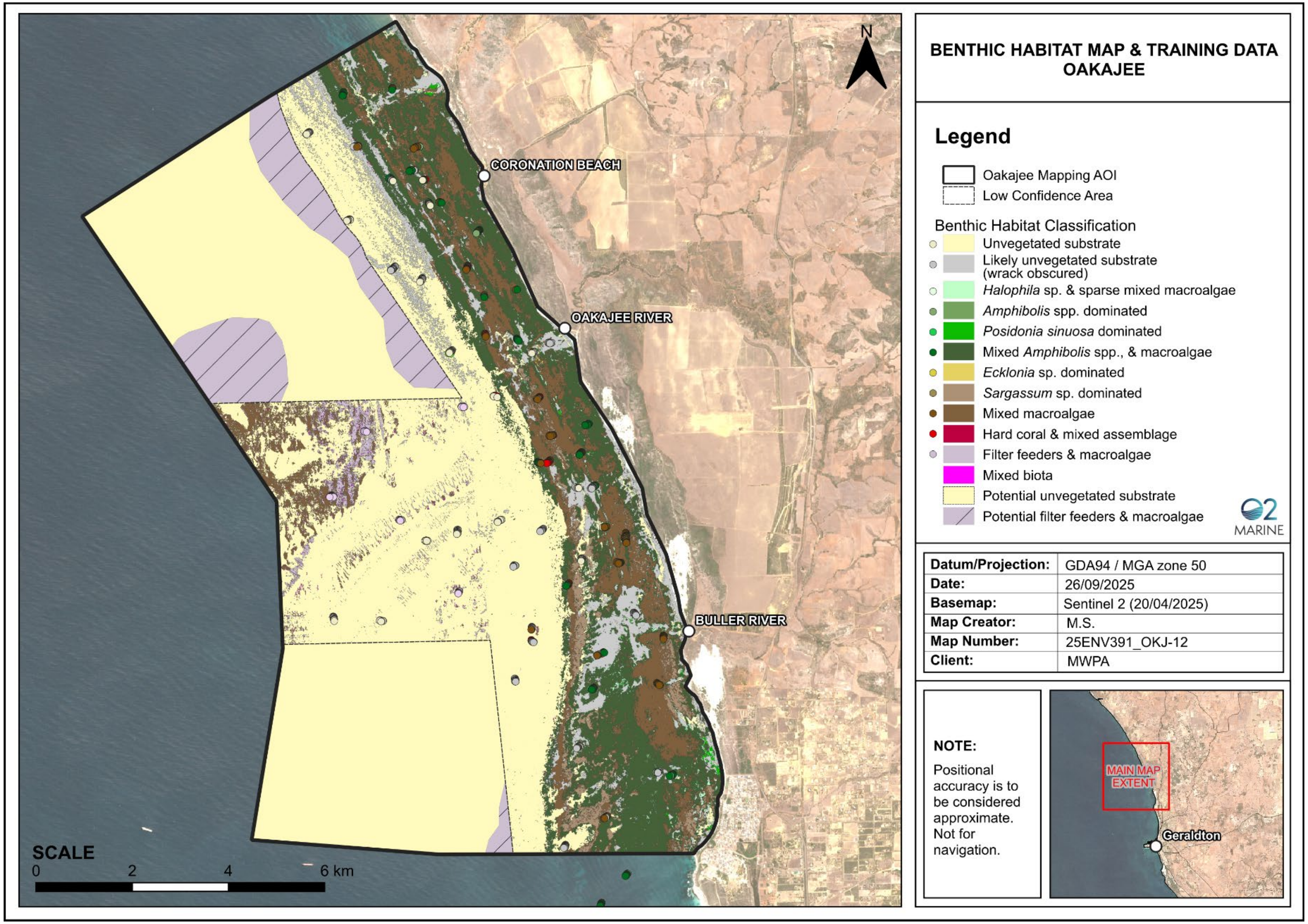


Figure 19: Model training data overlaid on the Oakajee benthic habitat map

4.3.1. Shallow High-Resolution Zone

Table 10: Confusion matrix for substrate classifications (misclassification assessment)

	MODEL PREDICTION					
TRAINING DATA	**CLASS**	Pavement reef	Low relief reef	Moderate relief reef	Potential mobile sediment	Stable sand veneer
	Pavement reef	19.0	0.5	0.0	0.0	0.5
	Low relief reef	9.0	9.0	1.5	0.5	0.0
	Moderate relief reef	5.0	0.5	14.0	0.5	0.0
	Potential mobile sediment	1.5	0.0	0.0	17.5	1.0
	Stable sand veneer	5.3	0.0	0.0	1.3	13.5

Table 11: Confusion matrix for benthic habitat classifications (misclassification assessment)

	MODEL PREDICTION										
TRAINING DATA	CLASS	Unvegetated substrate	Likely unvegetated substrate (wrack obscured)	*Halophila* sp. & sparse mixed macroalgae	*Ecklonia* sp. dominated	*Amphibolis* spp. dominated	Mixed *Amphibolis* spp. & macroalgae	Mixed macroalgae	*Sargassum* sp. dominated	*Hard coral & mixed assemblage*	*Posidonia sinuosa dominated*
	Unvegetated substrate	15.3	3.0	0.0	0.0	0.3	0.7	0.3	0.0	0.0	0.4
	Likely unvegetated substrate (wrack obscured)	0.1	19.3	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0
	Halophila sp. & sparse mixed macroalgae	0.0	1.7	12.9	0.0	0.7	2.7	1.6	0.0	0.0	0.4
	Ecklonia sp. dominated	0.6	0.1	0.0	11.0	0.0	0.6	0.4	7.3	0.0	0.0
	Amphibolis spp. dominated	0.3	3.7	2.9	0.0	5.1	6.4	0.9	0.7	0.0	0.0
	Mixed *Amphibolis*	0.4	0.7	0.6	0.1	0.1	16.6	1.3	0.0	0.0	0.1

	spp. & macroalgae										
	Mixed macroalgae	0.0	1.4	0.1	0.0	0.0	2.0	15.1	1.3	0.0	0.0
	Sargassum sp. dominated	0.0	0.1	0.0	0.7	0.0	0.6	1.1	17.0	0.4	0.0
	Hard coral & mixed assemblage	1.1	3.7	0.0	0.0	0.6	3.1	6.3	2.7	2.4	0.0
	Posidonia sinuosa dominated	0.3	1.1	0.1	0.0	0.0	3.0	0.1	0.0	0.0	15.3

Table 12: Class statistics for substrate types in the Shallow High-Resolution zone

Class	Precision	Recall	F-Score	Comment
Pavement reef	0.48	0.95	0.64	Moderate precision and very high recall, indicating that the model is able to capture most instances of this substrate type.
Low relief reef	0.90	0.45	0.59	High precision but relatively low recall, indicating that the model confuses some true instances of this habitat.
Moderate relief reef	0.92	0.70	0.79	High precision and moderate recall. Suggests that the model is generally accurate when it does predict this class, however there is a large degree of overlapping spectral signatures which likely reflects a combination of imperfect predictor layers as well as actual habitat overlap.
Potential mobile sediment	0.89	0.88	0.88	Strong overall performance with very high precision and recall, indicating that there are almost no false positives and the model effectively identifies most cases.
Stable sand veneer	0.90	0.68	0.77	High precision but moderate recall, reflecting some difficulty in capturing all true instances.

Table 13: Class statistics for benthic habitats in the Shallow High-Resolution Zone

Class	Precision	Recall	F-Score	Comment
Unvegetated substrate	0.82	0.83	0.83	Strong overall performance with very high precision and recall, indicating that there are almost no false positives and the model effectively identifies most cases.
Mixed *Amphibolis* spp. & macroalgae	0.53	0.84	0.65	Moderate precision with high recall. This indicates the model detects most actual habitats but at the cost of some false predictions.
Amphibolis spp. dominated	0.81	0.34	0.48	High precision with low recall. This means the model produces few false positives but missing many actual instances of this habitat.

Class	Precision	Recall	F-Score	Comment
Halophila sp. & sparse mixed macroalgae	0.80	0.64	0.71	High precision meaning most predicted instances are correct. Moderate recall, indicating that the model confuses some true instances of this habitat.
Posidonia sinuosa dominated	0.91	0.83	0.87	Strong overall performance with very high precision and recall, indicating that there are almost no false positives and the model effectively identifies most cases.
Ecklonia sp. dominated	0.90	0.63	0.74	High precision meaning most predicted instances are correct. However, recall is moderate, indicating that the model confuses some true instances of this habitat.
Sargassum sp. dominated	0.63	0.92	0.75	Strong overall performance with moderate precision and a high recall, indicating that the model is able to capture most instances of this habitat.
Mixed macroalgae	0.61	0.84	0.71	Strong overall performance with moderate precision and a high recall, indicating that the model is able to capture most instances of this habitat.
Likely unvegetated substrate (wrack obscured)	0.63	1.0	0.77	Moderate precision but perfect recall, meaning the model detects all actual instances of this habitat but generates some false positives.
Mixed biota	N/A	N/A	N/A	There are no statistics for this class as it is not predicted. This class serves to represents uncertain model predictions (no dominant class).

4.3.2. Deep High-Resolution Zone

Table 14: Confusion matrix for classification of the Deep High-Resolution Zone

		MODEL PREDICTION			
TRAINING DATA	CLASS	Unvegetated substrate	Likely unvegetated substrate (wrack obscured)	Mixed macroalgae	Filter feeders & macroalgae
	Unvegetated substrate	18.71	1.14	0.14	0.00
	Likely unvegetated substrate (wrack obscured)	0.00	20.00	0.00	0.00
	Mixed macroalgae	4.00	2.43	11.86	1.71
	Filter feeders & macroalgae	1.71	0.00	0.29	18.00

Table 15: Class statistics for benthic habitats in the Deep High-resolution Zone

Class	Precision	Recall	F-Score	Comment
Unvegetated substrate	0.61	0.94	0.74	Moderate precision with high recall. This indicates the model detects most actual habitats but at the cost of some false predictions.
Likely unvegetated substrate (wrack obscured)	0.59	1.00	0.74	Moderate precision with high recall. This indicates the model detects most actual habitats but at the cost of some false predictions.
Mixed macroalgae	0.85	0.59	0.70	High precision meaning most predicted instances are correct. However, recall is moderate, indicating that the model confuses some true instances of this habitat.
Filter feeders & macroalgae	0.86	0.90	0.87	Strong overall performance with very high precision and recall, indicating that there are almost no false positives and the model effectively identifies most cases.

5. Discussion

The distribution of fourteen benthic habitat classes were predicted across 12,955.2 ha at Oakajee using random forest supervised classification. Mapping results and associated validation statistics provide opportunity to assess the value of this mapping in contributing to an improved understanding of benthic habitat distribution within the area.

Habitat distribution

The habitat map was dominated by unvegetated substrate, which accounted for nearly two-thirds of the total mapped extent. This dominance reflects the prevalence of unconsolidated sediments across depths >20 m, a pattern consistent with sedimentary shelf environments in temperate regions (Jordan et al. 2005; Ierodiaconou et al. 2011). At shallower depths, unvegetated areas were restricted to small patches between reefs, a distribution pattern likely driven by substrate stability and wave exposure (Harris et al. 2013).

Biotic habitats were dominated by mixed *Amphibolis* spp. and macroalgae, covering more than 13% of the mapping extent. The association of this class with pavement reef and stable sand veneer substrates aligns with known preferences of *Amphibolis* spp. for consolidated or semi-consolidated substrates in shallow waters (Kendrick et al. 2000). This assemblage is particularly suited to higher-energy environments such as Oakajee, where wave action and current exposure favour species with robust anchoring structures and where reef and pavement substrates provide stable attachment points for persistent growth. Other seagrass-dominant habitats (*Halophila* sp., *Posidonia sinuosa*, and *Amphibolis* spp.) were mapped only in small patches. In particular, *Posidonia sinuosa* was observed only in isolated nearshore patches, consistent with its known preference for sheltered, sandy environments (Cambridge & Hocking 1997). The relative rarity of *Posidonia* and *Halophila* at Oakajee, compared with their greater prevalence in more protected environments such as Champion Bay (O2 Marine 2025), highlights the role of higher wave and current exposure in limiting the establishment and persistence of these less disturbance-tolerant seagrass types.

Among macroalgae-dominated habitats, mixed macroalgae and *Sargassum* sp. were the most extensive, occupying shallow reef zones. The confinement of *Sargassum* to two distinct north–south reef bands mirrors patterns reported in other southern Australian coastal systems, where *Sargassum* sp. forms dense monospecific stands under favourable hydrodynamic regimes (Phillips et al. 1997). In contrast, *Ecklonia* sp. was restricted to scattered shallow reef patches, consistent with its patchy distribution at the northern limits of its range (Wernberg et al. 2003).

Deeper areas were characterized by filter feeder and macroalgae assemblages, which occurred in association with slope features and pavement reef. These communities have been similarly reported in other temperate reef systems, where sponges and ascidians co-occur with macroalgae at the lower depth limits of light penetration (Butler et al. 2001).

Methodological considerations

Mapping the Oakajee benthic habitats required the use of three methodologies to account for varying spatial data availability and depth-related constraints. All areas were supported by a comprehensive ground-truth dataset collected via towed video, providing consistent reference information for habitat validation. In shallow waters, supervised machine learning was applied with multi-source data inputs; in deeper waters with reduced data availability, the same framework was adapted to function with fewer predictors; and in the deepest areas for which no high-resolution datasets exist, manual delineation was used where automated approaches were not feasible. The coarse resolution of the depth model in the deeper zones limited the ability to resolve habitat boundaries, making outputs highly uncertain. To reflect this caveat, classifications were explicitly labelled as "potential," underscoring their provisional nature and the need for caution in interpretation.

Overall, this rule-based approach, though less automated, enabled expert interpretation to address gaps where data were sparse. Collectively, the methods formed a hierarchical decision-making framework in which machine learning was prioritised where multi-source datasets were available, and manual expert judgement was employed where predictive modelling was constrained. This methodological design ensured consistent coverage while acknowledging spatial variation in data quality and mapping confidence.

Model performance

Validation results demonstrate that the random forest classifier performed strongly across several habitat classes. Unvegetated substrate and *Posidonia sinuosa* dominated habitat were classified with high precision and recall (>0.80), reflecting their distinct spectral signatures and relatively homogeneous spatial patterns. Similarly, *Ecklonia* sp. and *Halophila* sp. habitats were mapped with relatively strong performance (F-scores 0.71–0.74). Strong results were also achieved for *Sargassum* sp. dominated habitats and likely unvegetated substrate obscured by wrack (F-scores 0.75–0.77), demonstrating that the classifier effectively distinguished these classes despite potential structural complexity or transient cover.

Some classes showed lower classification accuracy, largely due to spectral and structural overlap inherent in complex benthic habitats. Mixed *Amphibolis* spp. and macroalgae exhibited moderate precision (0.53) but relatively high recall (0.84), indicating that the classifier successfully identified most true occurrences even if some areas were misassigned. This pattern is consistent with other optical habitat mapping studies where canopy-forming macroalgae can obscure underlying seagrass (Roelfsema et al. 2013). Similarly, *Amphibolis* spp. dominated seagrass habitat showed lower recall (0.34), reflecting the challenge of detecting small, patchy areas (<1 ha) that can fall below the spatial resolution of the input data. Despite these limitations, the classifier still provided valuable information about the general distribution of these habitats and highlighted areas where further targeted observation could improve mapping confidence.

The overall class-specific performance aligns with benthic habitat mapping studies elsewhere, where mixed assemblages (particularly seagrass–macroalgae mosaics) typically represent the greatest source of error (Mumby et al. 2004; Lyons et al. 2011). Machine learning approaches such as random forest have proven effective for habitat discrimination in optically complex waters, often outperforming traditional classifiers (Cutler et al. 2007; Pu & Bell, 2017). Nevertheless, the challenges encountered here - particularly confusion between mixed habitats and underestimation of small, patchy classes - mirror findings from previous seagrass and macroalgae mapping efforts across Australia (Ierodiaconou et al. 2011; Roelfsema et al. 2014).

The model achieved an overall Kappa of 0.64, indicating substantial agreement between predicted and observed habitats and supporting the robustness of the random forest approach for mapping complex benthic communities across a heterogeneous, high-energy coastal environment.

6. Conclusion

The random forest classification mapped fourteen benthic habitat classes across 12,955.2 ha at Oakajee. Unvegetated substrate dominated depths >20 m, while mixed *Amphibolis* spp. and macroalgae were the most extensive biotic habitat, well-suited to higher-energy conditions. Rarer seagrass types (*Halophila* sp., *Posidonia sinuosa*, and *Amphibolis*-dominated assemblages) occurred only in small patches, reflecting their lower tolerance to hydrodynamic stress and contrasting with more sheltered areas such as Champion Bay.

Validation showed high accuracy for distinct habitats, moderate for mixed or patchy assemblages, with an overall Kappa of 0.64, indicating substantial agreement between predicted and observed classes. Accuracy was highest in shallow areas with multi-source data, moderate in deeper high-resolution zones, and lowest in the deepest, low-resolution zones mapped using a 250 m depth model. These areas were classified as “potential” to reflect low confidence, as fine-scale or patchy habitats may be underrepresented and spatial boundaries less precise. Despite these limitations, the map provides complete coverage and a robust understanding of benthic habitat patterns across Oakajee.

7. References

AECOM. (2020). 2020 Benthic Habitat Mapping Report – Champion Bay & Surrounds. Prepared for Mid-West Ports. 01-Sep-2020.

Althaus, F., Hill, N., Ferrari, R., Edwards, L., Przeslawski, R., Schönberg, C.H.L., Stuart-Smith, R., Barrett, N., Edgar, G., Colquhoun, J., Tran, M., Jordan, A., Rees, T. & Gowlett-Holmes, K. (2015). A standardised vocabulary for identifying benthic biota and substrata from underwater imagery: The CATAMI classification scheme. PLOS ONE, 10(10).

BMT (2021). Assessment of Potential Dredge Material Placement Areas within Champion Bay, WA. Report prepared for Mid West Ports Authority. Report No. R-10708-1.

Breiman, L. (2001). Random Forests. Machine Learning, 45, 5-32.

Brown, C.J., Smith, S.J., Lawton, P., & Anderson, J.T. (2011). Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuarine, Coastal and Shelf Science, 92, 502-520.

Butler, A. J., Jernakoff, P., & Connolly, R. M. (2001). Distribution and abundance of sessile filter feeders and macroalgae on temperate subtidal reefs. Marine Ecology Progress Series, 215, 27–41.

Cambridge, M. L., & Hocking, P. J. (1997). Annual primary production of Posidonia sinuosa meadows in temperate Australian waters. Aquatic Botany, 56(3–4), 223–237.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783–2792.

Fisher, R., Hobgen, S., Mandaya, I., Riwu K., & Zulkarnain N. (2017). Satellite image analysis and terrain modelling: A practical manual for natural resource management, disaster risk and development planning using free geospatial data and software. SAGA GIS 4+ Version 2.

Florinsky, I. (2016). Digital terrain analysis in soil science and geology. Academic Press.

Florinsky, I. (2017). An illustrated introduction to general geomorphometry. Progress in Physical Geography, 41(6), 723-752.

Geoscience Australia. (2024). AusBathyTopo (Australia) 250m 2024 - A national-scale depth model (20240011C). Retrieved from: https://ecat.ga.gov.au/geonetwork/srv/api/records/880dbb6c-411d-487e-bfe0-40ef167afbeb

Green, E.P., Mumby, P.J., Edwards, E.D., Clark, C.D. (2000). Remote sensing handbook for tropical coastal management. UNESCO Publishing, Paris.

Harris, D. L., Kendrick, G. A., & Walker, D. I. (2013). Substrate stability and the distribution of shallow temperate seagrass beds. Marine Ecology Progress Series, 476, 115–128.

Hasan, R.C., Lerodiaconou, D., & Monk, J. (2012). Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. Remote Sensing, 4(11), 3427-3443.

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

Hossain, M.D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. ISPRS Journal of Photogrammetry and Remote Sensing, 150, 115-134.

Ierodiaconou, D., Laurenson, L., Harvey, E., Heyward, A., & Kendrick, G. (2011). Benthic habitat mapping of temperate reefs using remote sensing: validation and challenges. Estuarine, Coastal and Shelf Science, 92(2), 234–243.

Jordan, A., Ierodiaconou, D., Laurenson, L., & Heyward, A. (2005). Sedimentary environments of temperate continental shelves: implications for benthic habitat distributions. Continental Shelf Research, 25(16), 2020–2038.

Kendrick, G. A., Waycott, M., & Hovey, R. (2000). The ecology and distribution of Amphibolis spp. in Western Australian coastal waters. Aquatic Botany, 66(2), 123–140.

Lindsay, J.B. (2014). The Whitebox Geospatial Analysis Tools project and open-access GIS. Proceedings of the GIS Research UK 22nd Annual Conference. The University of Glasgow.

Lindsay, J.B., Cockburn, J.M., & Russell, H.A. (2015). An integral image approach to performing multi-scale topographic position analysis. Geomorphology, 245, 51-61.

Lyons, M., Roelfsema, C., Phinn, S., & Dennison, W. (2011). Remote sensing of marine benthic habitats: a review of classification approaches. International Journal of Remote Sensing, 32(21), 6613–6650.

Lyzenga, D. R. (1978). Passive Remote Sensing Techniques for Mapping Water Depth and Bottom Features. Applied Optics, 17, 379–383.

Mumby, P. J., Skirving, W., Strong, A. E., Hardy, J. T., LeDrew, E. F., Hochberg, E. J., ... & Brown, E. (2004). Remote sensing of coral reefs and their physical environment. Marine Pollution Bulletin, 48(3–4), 219–228.

Newman, R., Lindsay, J.B., Cockburn, J.M.H. (2018). Evaluating metrics of local topographic position for multiscale geomorphometric analysis. Geomorphology, 312, 40-50.

O2 Marine. (2025). Champion Bay Habitat Mapping 2025 – Technical Report. Report prepared for Mid West Ports Authority. O2 Marine Pty Ltd, Perth, WA.

Olaya, V, & Conrad, O. (2009). Geomorphometry in SAGA. Developments in Soil Science, 33, 293-308.

Phillips, J. A., Kendrick, G. A., & Walker, D. I. (1997). Patterns of distribution of Sargassum spp. along temperate Western Australian reefs. Marine Ecology Progress Series, 146, 157–166.

PHS. (2023). Multibeam echosounder bathymetry dataset, 14/08/2023. Dataset provided by MWPA.

Pisanti, A., Magri, S., Ferrando, I., & Federici, B. (2022). Sea water turbidity analysis from Sentinel 2 images: atmospheric correction and bands correlation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 371-378.

Pu, R., & Bell, S. (2017). Machine learning approaches for mapping aquatic habitats in optically complex waters. Remote Sensing, 9(5), 475.

Riley, S.J., DeGloria, S.D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. Intermountain Journal of Sciences, 5 (1-4), 23-27.

Roelfsema, C., Lyons, M., & Phinn, S. (2013). Mapping seagrass, macroalgae, and substrate in coastal waters using high-resolution satellite imagery. Estuarine, Coastal and Shelf Science, 117, 45–55.

Roelfsema, C., Lyons, M., Phinn, S., & Mumby, P. (2014). Coral reef habitat mapping using remote sensing: an operational approach. *Remote Sensing*, 6(1), 1–19.

RoyalHaskoningDHV. (2023). Geraldton Coastal Process Study. Stage 2 Hydrodynamic Modelling, Stage 3 Sediment Transport Understanding. Report prepared for Mid West Ports Authority (MWPA).

Tecchiato, S., Collins, L., Stevens, A. (2015). The influence of geomorphology and sedimentary processes on benthic habitat distribution and littoral sediment dynamics: Geraldton, Western Australia. Marine Geology. 359, 148-162.

Wernberg, T., Thomsen, M., Tuya, F., Kendrick, G. A., & Staehr, P. A. (2003). Distribution and biomass of Ecklonia radiata at its northern limit in Western Australia. Marine and Freshwater Research, 54(2), 215–223.

Wicaksono, P., Wulandari, S. A., Lazuardi, W., & Munir, M. (2021). Sentinel-2 images deliver possibilities for accurate and consistent multi-temporal benthic habitat maps in optically shallow water. Remote Sensing Applications: Society and Environment, 23, 100572.

Wilson, J.P. & Gallant, J.C. (2000). Primary Topographic Attributes. In Terrain Analysis: Principles and Applications (pp. 51-85). John Wiley & Sons.