Norsk institutt for vannforskning

7995-2024

Method development for mapping kelp using drones and satellite images: Results from the KELPMAP-Vega project

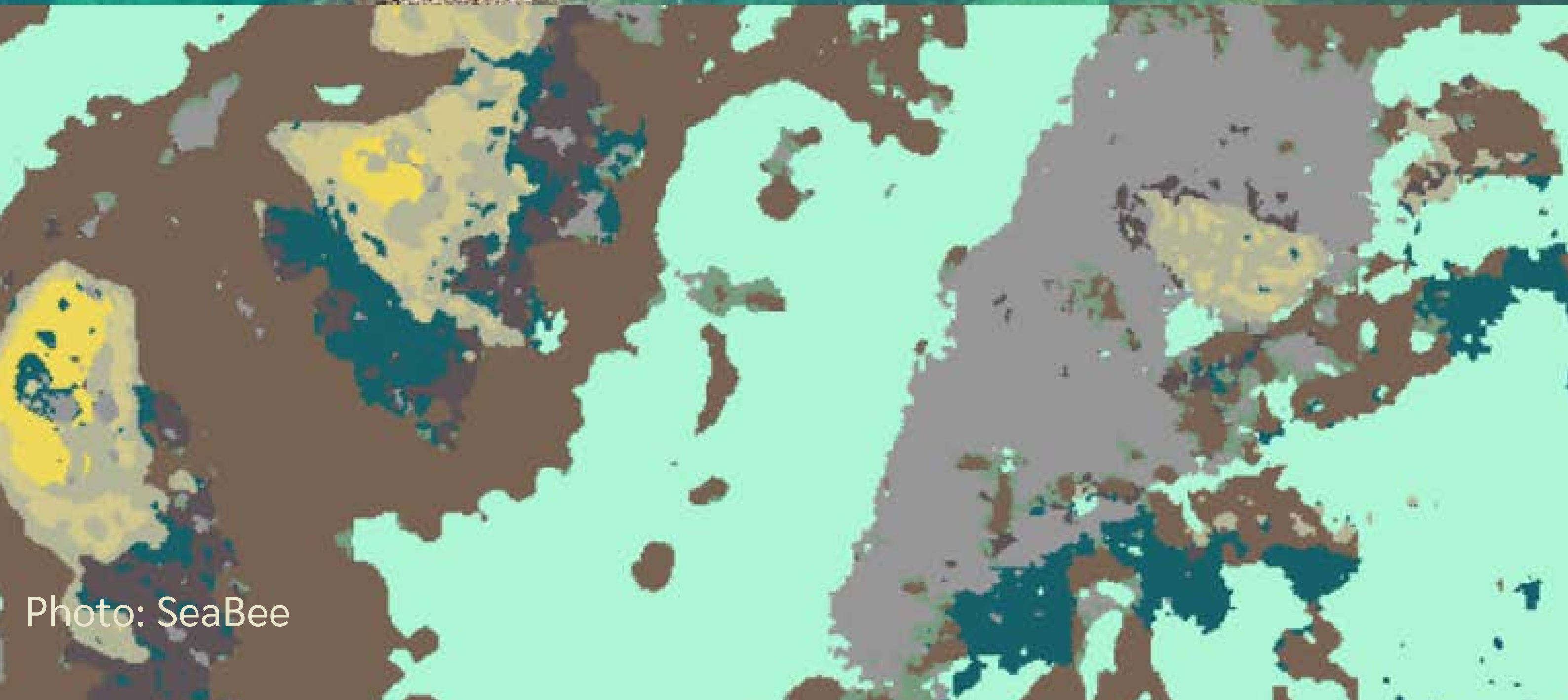


Photo: SeaBee

Report

Norwegian Institute for Water Research

Serial no: 7995-2024

ISBN 978-82-577- 7732-6
NIVA report
ISSN 1894-7948

This report has been quality assured according to NIVA's quality system and has been approved by:

Hege Gundersen
Project Manager/Lead Author

Paul R. Berg
Research Manager / Quality Assurer

www.niva.no

Suggested citation:
Gundersen H, Salberg A-B, Kvile KØ, Poulsen RN, Buls T, Liu I, Ghareeb M, Christie H, Kile MR, Bekkby T, Arvidsson KS, Hancke K. 2024. Method development for mapping kelp using drones and satellite images: Results from the KELPMAP-Vega project. NIVA report 7995-2024. 58 pp.

Title
Method development for mapping kelp using drones and satellite images: Results from the KELPMAP-Vega project

Metodeutvikling for kartlegging av tareskog ved hjelp av droner og satellittbilder: Resultater fra KELPMAP-Vega-prosjektet

Pages
53 + Appendices

Date
06.08.2024

Authors
Hege Gundersen, Kasper Hancke, Arnt-Børre Salberg, Robert N. Poulsen, Toms Buls, Izzie Liu, Medyan Ghareeb, Hartvig Christie, Maia R. Kile, Trine Bekkby, Karoline S. Arvidsson, Kristina Ø. Kvile.

Topic group
Marine biology

Distribution
Open

Client
Norwegian Environment Agency

Client's contact person
Ragnvald Larsen

Published by NIVA
Prosjektnummer 220083

Client's publication:
M-2757|2024

Abstract
The KELPMAP study demonstrated that high-resolution multispectral data from drones and satellites, combined with AI-based image analysis, can efficiently map kelp forests and other coastal habitats. The field campaign, conducted in 2022 within Vega and Herøy municipalities, produced orthomosaics with 9 cm GSD for multispectral and 5 cm for RGB images. Based on drone data and AI, between 11% and 65% of the study area was identified as brown algae. Satellites overpredicted kelp forests but aligned with drone data after removing uncertain predictions. In Helgeland's clear waters, benthic species and habitats were identified down to 10 meters. Using NIVA's statistical model, drones were estimated to map almost 60% of Norwegian kelp forests and 80-90% of total kelp biomass, despite only reaching 10 meters. Upscaling habitat maps using satellite images is possible but limited by satellite resolution. Drone-based training data enhances satellite-derived map accuracy. High-resolution drone maps are ideal for local marine spatial planning, while satellite maps are suitable for national-level applications like carbon accounting. More ground truth data are needed for improved species-level mapping and validation of upscaled products. The study also assessed mapping benthic habitats according to NiN 3.0, identifying kelp forests, seaweed beds, eelgrass, and various seabed types.

Keywords: Uncrewed/Unmanned Aerial Vehicles (UAVs), Convolutional Neural Network (CNN), Blue Carbon Habitats, , Copernicus, Blue forests, Submerged Aquatic Vegetation (SAV), Macroalgae, *Laminaria hyperborea*, *Saccharina latissima*, *Laminaria digitata*, Seaweed, Rockweed, Artificial intelligence (AI), Machine Learning (ML) analyses, Sentinel-2.
Emneord: UAV, Kunstig intelligens (KI), Naturtyper, , Copernicus, blå skog, Natur i Norge (NiN), makroalger, blått karbon, tare, tang, stortare, sukkertare, fingertare, droner, Maskinlæring, Habitatkartlegging, Satellittbilder, Sentinel-2.

Table of contents

Preface

This report represents the final delivery in the project "Method development for mapping kelp using drones and satellite images" for the Norwegian Environment Agency's (NEA) (contract number 22047043). The internal project name at NIVA is KELPMAP-Vega (project number 220083). The project has been carried out as a pilot under the "Copernicus fund" and has aimed to build competence in the use of data from the Copernicus program and other types of remote sensing in the NEA monitoring and mapping work. The client has been NEA by Ragnvald Larsen (technical responsible), as well as Ida Maria Evensen and Ingrid Holtan Søbstad (main contacts). Other resource persons at the NEA have been Agnès Moquet-Stenback and Ida Egge Johnsen.

The assignment has been divided into four activities, where the fieldwork and post-processing of drone images (activity 1) was carried out during 2022, while image analysis (activity 2), validation (activity 3) and upscaling (activity 4) has been completed in 2023. Activity 1 has previously been reported in a field report (Gundersen et al. 2022) that includes the collection of high-resolution imagery using drones and the generation of coherent georeferenced orthophotos, as well as the collection of ground truth data and their compilation. The present report also covers much of this, so the entire study should be understood from this report, but with some references to Gundersen et al. (2022).

The project is carried out as an assignment performed by, and in close collaboration with the national drone infrastructure SeaBee (Norwegian infrastructure for drone-based research, mapping and monitoring in the coastal zone, https://seabee.no/), led by NIVA/Kasper Hancke and funded by the Research Council of Norway. Drone flying and post-processing of drone images have been carried out by Medyan Ghareeb (NIVA/SeaBee), as well as Robert N. Poulsen and Toms Buls from the subcontractor and SeaBee partner SpectroFly (www.spectrofly.dk). Hartvig Christie and Maia R. Kile were responsible for collection of ground truth data in the field, while administration, planning, and field design were done by project leader Hege Gundersen, Kasper Hancke, Karoline S. Arvidson, and Kristina Ø. Kvile (all from NIVAs SeaBee team, see Appendix A). Subcontractor (and SeaBee partner) Norwegian Computing Center by Arnt-Børre Salberg and Izzie Liu, has performed the machine learning analyses, as well as the upscaling via satellite images and the model validation. All project participants have contributed to the writing of the report.

Data synthesis work and writing of the report has been co-financed by SeaBee (Research Council of Norway grant #296478), OBAMA-NEXT (Observing and Mapping Marine Ecosystems – Next Generation Tools, Horizon EU grant #101081642), and BLUEFORESTING (Climate Resilient Marine Forests for a Sustainable Future, EEA Blue Growth Program grants #PT-INNOVATION-0077).

Drone images, ground truth data, analysis products, and relevant metadata have been shared internally via NEA's *Rasterportal* and externally via SeaBee's data visualization platform, at Sigma2. NIVA thanks the Norwegian Space Center for funding through the Copernicus funds, and Norwegian Environment Agency for the assignment and for good contact with Ida Maria Evensen, Agnès Moquet-Stenback, and Ragnvald Larsen along the way.

Hege Gundersen

Project lead of KELPMAP-Vega

Oslo, 21. June 2024

Summary

The KELPMAP study demonstrated that high-resolution multispectral data from drones and satellites, combined with AI-based image analysis, can be used as a cost-effective method for mapping kelp forests and other coastal habitats.

The field campaign, including drone missions and ground truth data collection, was conducted in August 2022 on the Helgeland coast within Vega (3.5 km^2) and Herøy (2.0 km^2) municipalities. By using modern sensors and flying at 120 meters, orthomosaics with up to 9 cm ground sampling distance for multispectral and 5 cm for RGB images were created.

Based on the drone data and machine learning analyses, 2.41 km^2 (65%) of the total study area was identified as brown algae (including various species of kelp and rockweed). A species-level model yielded lower certainty but still enabled mapping of many kelp and other macroalgae at the species level. Due to the lower certainty of this map, where uncertain pixels were masked out, the more conservative area estimate for kelp was 0.39 km^2 (11%). For satellites, the initial unmasked estimates significantly overpredicted the extent of kelp forests, but after removing uncertain predictions, the results aligned more closely with the drone-based data, demonstrating the potential for using satellite data to support large-scale coastal habitat mapping.

In the clear water of the Helgeland coast, benthic species and habitats were identified down to a depth of approximately 10 meters, demonstrating the potential for drone and satellite mapping at least to this depth, given the same water transparency.

Using NIVAs statistical spatial distribution model of kelp forests, which extends down to more than 20 meters, we estimated that drones are able to map almost 60% of the Norwegian kelp forest areas and 80-90% of the total kelp biomass, despite only reaching depths down to 10 meters.

Upscaling habitat maps from drones to larger areas using satellite images is possible but limited by the spatial and spectral resolution of the satellite data. Here we tested open-source data (10 x 10 meters, Sentinel-2) and commercial data (2 x 2 meters, Pleiades VHS) for mapping of kelp and macroalgae, though species-level mapping is rarely feasible due to the heterogeneous nature of macroalgae.

Using drone-based training data for satellite upscaling enhances the accuracy of resulting habitat maps compared to those derived solely from satellite images, though this conclusion is based on subjective comparisons and not fully quantitatively evaluated.

High-resolution drone maps are ideal for local and regional marine spatial planning and fulfilling Nature Agreement (CBD) obligations, while semi high-resolution satellite maps are suitable for national-level applications like broad-scale carbon accounting.

Although the present models are a great advancement over previous studies that categorized submerged vegetation less diversely, still more training data is needed for more accurate species-level mapping and for improving the AI algorithms to make them relevant also to other regions. High precision ground truth data are needed to validate classification maps from drones and satellites. To validate upscaled products, a comprehensive ground truth dataset covering large regions and many habitat classes is necessary. Existing data from other projects could potentially be used for validation of satellite-based maps, since they require lower spatial precision than drone-based products.

The study includes an assessment of mapping benthic habitats according to the NiN 3.0 classification system using drone-based data and discuss options and limitations for identifying basic types (*grunntyper*) and the variable system (*variabelsystemet*) from drone data. This included identification of kelp forests, seaweed beds, eelgrass and different types of soft and hard substrate seabed.

Sammendrag

I KELPMAP har vi vist at høyoppløselige multispektrale data fra droner og satellitter, kombinert med AI-basert bildeanalyse, er en kostnadseffektiv metode for kartlegging av tareskog og andre kysthabitater.

Feltarbeidet, inkludert droneflyging og innsamling av bakkesannhetsdata, ble gjennomført i august 2022 på Helgelandskysten i Vega (3,5 km²) og Herøy (2,0 km²) kommuner. Ved bruk av moderne sensorer og flyvning på 120 meters høyde, ble ortomosaikker med en oppløsning på opptil 9 cm for multispektrale bilder og 5 cm for RGB-bilder opprettet.

Basert på dronedata og maskinlæringsanalyser ble 2,41 km² (65%) av det totale studieområdet kartlagt som brunalger (inkludert ulike arter av tang og tare). En modell på artsnivå ga lavere prediksjonsevne, men gjorde det likevel mulig å kartlegge mange tare- og andre makroalger på artsnivå. På grunn av den lavere prediksjonsevnen i dette kartet, hvor usikre piksler ble maskert ut, gav dette et mer konservativt arealestimat for tare på 0,39 km² (11%). Oppskalerte prediksjonskart, ved bruk av satellitter, overpredikerte arealet av tareskog betraktelig, men samsvarte godt med dronebaserte data etter å ha fjernet usikre prediksjoner, noe som viste potensialet i å bruke satellittdata til å støtte storskala kartlegging av kysthabitater.

I det klare vannet langs Helgelandskysten ble bentiske arter og habitater identifisert ned til et dyp på ca. 10 meter, noe som lover godt for muligheten til å kartlegge enda dypere, gitt samme vanngjennomsiktighet. Ved å bruke NIVAs statistiske romlige utbredelsesmodell for tareskog, som strekker seg ned til mer enn 20 meter, estimerte vi at droner er i stand til å kartlegge bortimot 60% av de norske tareskogsområdene og 80-90% av den totale tarebiomassen, til tross for at de bare når ned til 10 meter.

Oppskalering av habitatkart fra droner til større områder ved hjelp av satellittbilder er mulig, men begrenses av den romlige og spektrale oppløsningen til satellittdataene. Her testet vi åpent tilgjengelige data (10 x 10 meter, Sentinel-2) og kommersielle data (2 x 2 meter, Pleiades VHS) for kartlegging av tare og andre makroalger. Vi konkluderer med at kartlegging på artsnivå sjelden er gjennomførbart på grunn satellittbildenes relativt lave oppløsning, og makroalgenes heterogene natur.

Bruk av dronebaserte treningsdata for satellittoppskalering øker nøyaktigheten til de resulterende habitatkartene sammenlignet med de som kun er avledet fra satellittbilder. Denne slutningen er imidlertid basert på subjektive sammenligninger og ikke kvantitativt evaluert.

Høyoppløselige dronekart er ideelle for lokal og regional marin arealplanlegging og for oppfyllelse av Naturavtalen (CBD)-forpliktelser, mens satellittkart med middels god oppløsning er egnet for nasjonale anvendelser som storskala karbonregnskap.

Selv om de nåværende modellene viser et stort fremskritt sammenlignet med tidligere studier som har kartlagt marin vegetasjon, trengs det fortsatt mer treningsdata for mer nøyaktig kartlegging på artsnivå og for å forbedre AI-algoritmene for å gjøre dem relevante også for andre regioner. Bakkesannhetsdata med høy romlig presisjon er nødvendig for å validere klassifikasjonskart fra droner og satellitter. For å validere oppskalerte produkter er det nødvendig med et omfattende bakkesannhetsdatasett som dekker store regioner og mange habitatklasser. Eksisterende data fra andre prosjekter kan muligens brukes for validering av satellittbaserte kart, men ikke for de høyoppløselige dronebaserte kartene, siden de krever høyere romlig presisjon.

Studien inkluderer en vurdering av kartlegging av bentiske habitater i henhold til NiN 3.0-klassifikasjonssystemet ved bruk av dronebaserte data, og ser på muligheter og begrensninger for å identifisere arter og habitater etter NiN's grunntyper og variabelsystemet. Identifikasjon av tareskoger, tangbelter, ålegras og ulike typer bløtbunn og hardbunn bør være mulig å kartlegge med droner i henhold til NiN.

1 Introduction

1.1 The assignment

The assignment is one of the Norwegian Environment Agency's (NEA) pilot projects for remote sensing, mapping, and environmental monitoring. NEA aims to explore the possibilities of developing methods for more cost-efficient and systematic mapping of marine habitats by using imaging data from flying drones and satellites and compare these to more traditional and manual monitoring techniques.

The project has been conducted by researchers involved in the SeaBee drone infrastructure and aims to develop cost-effective and reproducible methods for mapping kelp forests and other marine vegetation in coastal shallow areas (from the surface down to 5-10 meters depth). This is achieved using high-resolution multispectral sensor data captured from flying drones, combined with automated image analysis based on Machine Learning (ML) / Artificial Intelligence (AI) for improved benthic habitat classification. Additionally, this work is an initial step in evaluating the potential of using satellite remote sensing data for broad scale benthic habitat mapping by upscaling drone data products to satellite-based habitat maps.

1.2 Important coastal habitats

Coastal vegetated habitats such as kelp forests, rockweed beds, and seagrass meadows, often called “blue forests”, covers more than 10,000 km^2 along the Norwegian coast and colonize large areas globally (Hancke et al. 2022, Krause-Jensen et al. 2022). These habitats provide several ecosystem services that supports both ecosystem resilience and marine biodiversity, as well as coastal societies through commercial and recreational values, including nursery and feeding grounds for fish and sea birds (Gundersen et al. 2016). Understanding the distribution and status of these key coastal habitats is both ecologically and economically valuable and plays a key role for a sustainable coastal management, contributing to climate regulation, carbon storage and sequestration, regulation of oxygen availability and ocean acidification. On a global scale, coastal vegetated habitats create the basis for several of the United Nations Sustainable Development Goals (SDGs), e.g., zero hunger (SDG #2), climate action (SDG #13), and life below water (SDG #14).

Species and habitat types of particular interest in this study were different species of kelp, such as tangle kelp (*stortare*), sugar kelp (*sukkertare*), and oar weed (*fingertare*), and rockweed species such as toothed wrack (*sagtang*), knotted wrack (*grisetang*), and bladder wrack (*blæretang*). See Appendix B for a complete list of species, with scientific names, and habitats observed and classified in the study.

Throughout the report we have tried to use consistent terminology when describing various species and groups of species. Generally, the term “macroalgae” encompasses all large algae, including red, green, and brown varieties. Synonymous with macroalgae and equally general is the term “seaweed”. When referring to “rockweed”, we mean the Norwegian term “tang”, which includes all littoral brown macroalgae. “Kelp”, known as “tare” in Norwegian, is part of the “macroalgae” group and typically lives in somewhat deeper waters than rockweed.

1.3 The technology

Traditional techniques for coastal habitat mapping typically require extensive manual work and resource-demanding fieldwork campaigns. Using remote sensing technology for benthic mapping and monitoring is an effective approach which will significantly increase areal coverage and spatial resolution, as well as improve data reproducibility relative to conventional methods.

Flying drones, or Unmanned/Uncrewed Aerial Vehicles (UAVs), are increasingly being used for different applications within coastal zone management, such as mapping and monitoring of species and habitats (e.g., Duffy et al. 2018, Cavanaugh et al. 2021). At the habitat level, and for species that are identifiable from a distance, the use of drones in coastal areas allows for time- and cost-efficient mapping covering larger areas than traditional mapping and monitoring techniques such as underwater cameras and aquascopes (Joyce et al. 2023). Also, drone images give much more detailed information than what is available from airplanes and satellites, enabling to separate different types of coastal vegetation, based on their individual spectral reflectance and structural patterns. For example, flying at 120 m above ground, drones can collect multi-spectral images (MSI) with a ground sampling distance (GSD) of better than 10 x 10 cm, or RGB better than 2 x 2 cm, which is much higher than what is available from satellites (typically on meters to tens of meter scale). Still drones can cover areas up to one km^2 per hour, and capacity is increasing markedly every year due to rapid technological developments.

The use of drone imaging in combination with sophisticated Machine Learning (ML) / Artificial Intelligence (AI) analyses, such as Convolutional Neural Network (CNN), provides a novel and unique tool for labor-saving and reproduceable image analysis (Liu et al. 2022). ML methods will ultimately allow us to map and classify coastal habitats without the need to collect *in situ* field data. However, collection of ground-truth data is still required when training new algorithms and for validation of ML results, and as such an important component of this development project. In this study, we have further included developments of ML for analyzing satellite remote sensing data as a novel approach. Especially novel is the attempt to train ML algorithms on drone data habitat classification maps and apply these for classification of satellite remote sensing images with the aim to upscale drone-based habitat maps to regional scales.

1.4 Strategic initiatives, programs, and policies

Currently, it is not standard to use remote sensing techniques for mapping and monitoring benthic habitats in the coastal zone in Norway, neither are they included in environmental monitoring protocols within the European Union, such as for the Water Framework Directive (WFD), Marine Strategy Framework Directive (MSFD), or other. However, the technology developed in KELPMAP, and the national infrastructure for drone research, SeaBee, as such, has perspectives relevant for several major national initiatives of high priority, some of which are listed below, along with international management goals.

1.4.1. The Marine Base Maps

The *Marine Base Maps in the Coastal Zone* (*Marine grunnkart i kystsonen*, e.g., Schimel et al. 2023) aims to map the seabed along the Norwegian coast, focusing on the shallow parts, where people interact with the marine environment. The outcome and lessons learned from the KELPMAP project could contribute to the realization of a large-scale, and more cost-efficient mapping of Norwegian coastal habitats. We envision that the techniques presented here could be further developed and adapted for habitat and underwater vegetation mapping in collaboration with the Norwegian Mapping Authority (Kartverket), Geological Survey of Norway (NGU), and the Institute of Marine Research (IMR), to ensure synergies and technological developments needed for an efficient and sustainable management of Norwegian coastal ecosystems, their biodiversity and ecosystem services.

1.4.2. Nature Types in Norway (NiN)

The *Nature Types in Norway* (*Natur i Norge*, NiN, Halvorsen et al. 2024) is the established system for classifying and describing all Norwegian nature, on land, in freshwater and in the sea. The aim of the NiN system is to provide a tool for standardized mapping and a common conceptual framework for describing nature. As part of this system, it is possible to describe everything from large landscapes and landforms to small-scale habitats. The nature system is the natural level for comprehensive mapping and can be divided into Major Type Groups (*Hovedtypegrupper*), Major Types (*Hovedtyper*) and Basic Types (*Grunntyper*), with a variable system to describe nature beyond what is covered by the type-classification (Halvorsen et al. 2024). The Norwegian Parliament has established that publicly funded mapping of habitat types in Norway must be carried out in accordance with NiN (Meld. St. 14, 2015-2016). Therefore, for the application of drones, ML, and satellite images to be effective for Norwegian mapping and monitoring programs, the development of these techniques should, as far as possible, be in accordance with the NiN system.

1.4.3. The Norwegian Red List for Habitats

According to the Norwegian Red List for Habitats (Gundersen et al. 2018), kelp forests as nature types (habitats) are threatened at different levels across Norway. In Northern Norway, sugar kelp forests are classified as highly threatened (EN), while tangle kelp forests are near threatened (NT), and oar weed forests are vulnerable (VU). The same high threat status (EN) applies to sugar kelp forests in southern Norway. Therefore, these kelp ecosystems are identified as management-relevant nature types (Bekkby et al. 2021), suggested prioritized for mapping in accordance with the criteria in the white paper "Nature for life" (Meld. St. 14, 2015-2016). This document describes the Norwegian Government's policy for safeguarding biodiversity. Through the KELPMAP project, SeaBee has developed a useful tool to map and monitor these threatened kelp habitat types.

1.4.4. The Nature Agreement (CBD)

Norway also has obligations through the Kunming-Montreal Global Biodiversity Framework (CBD or *Naturavtalen or Nature agreement*), where Norway has committed to contribute to conserve 30% of the land and sea areas for biodiversity by 2030. Important then, is to know where the most valuable coastal species and habitats are, its distribution, coverage, and composition. In this context, it is of high importance to develop accurate mapping tools and protocols to ensure sustainable management using reproduceable and cost-efficient methods. The Norwegian coast is 100,000 km long, and historically, the stretch of the coastal zone from the top of the intertidal to about 10 meters of water depth has been understudied. This is mainly due to methodological and logistical challenges for data collection. Remote sensing using drone and high-resolution satellite data offers significant opportunities to enhance research and monitoring in the coastal zone, and closing many current knowledge gaps regarding habitats, species, biodiversity, and ecosystem functions. In turn, this technology can help support sustainable and ecosystem-based management strategies and assist in guiding conservation and restoration efforts in the coastal zone (Meld. St. 21, 2023–2024).

2 Materials and methods

2.1 Study area

This study was carried out in the Vega and Herøy municipalities on the Helgeland coast, central Nordland County (Figure 1). Two study areas were predefined by the client (NEA) with specified location data. These were Steinskjæran (southern area) located in Vega municipality covering 3.5 km^2, and Sandværodden (northern area) in Herøy municipality covering 2.0 km^2. Neither of the two areas are located within protected areas or covered by specific restrictions related to travel, drone flying, or else. The nearest nature reserve is the Lånan/Skjærvær bird sanctuary located 2 km north-west of the southern study area (Figure 1). All fieldwork activities were conducted in August, which was outside the nesting season and maintained at a safe distance from the nature reserve and other precious wildlife.

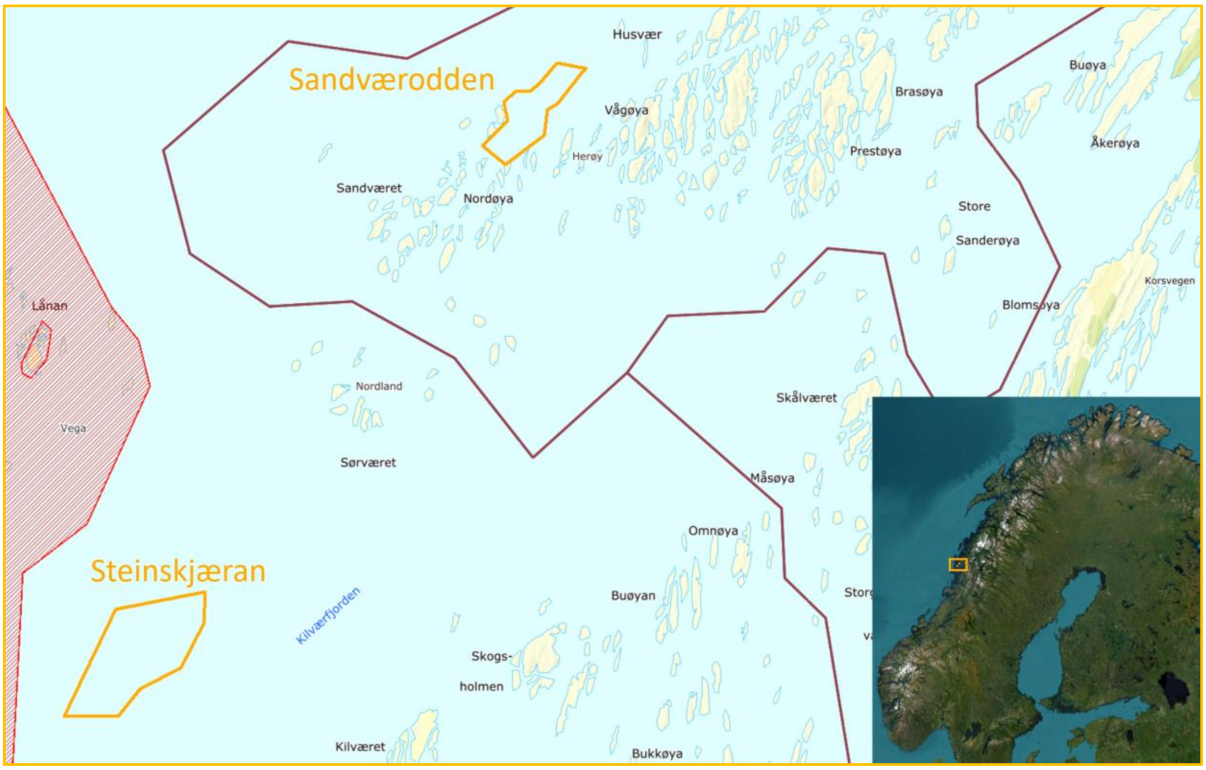


Figure 1. The two study areas (orange polygons) that were mapped with high resolution drone images, i.e., Steinskjæran in the south and Sandværodden in the north, in the municipalities of Vega and Herøy, respectively, at the Helgeland coast, western Norway. The shaded red area to the left is the Lånan/Skjærvær bird sanctuary.

The Helgeland archipelago is a complex network of islands, islets, and skerries, with a rich marine vegetation colonizing the relatively shallow waters in between. Rockweed (*tang*) is the dominating habitat type in the littoral zone (tidal zone) stretching into the sublittoral zone, where kelp forests take over as the key habitat to depths at which light supports photosynthesis for growth, which is around 20, or even 30 meters, in this region. Typical benthic vegetation in the region includes kelp forest forming species (Figure 2), i.e. *Laminaria hyperborea* (tangle kelp / stortare), and *Saccharina latissima* (sugar kelp / sukkertare), a number of rockweed species in the fucoid family including *Fucus serratus* (toothed wrack / sagtang), *Fucus vesiculosus* (bladder wrack / blæretang), *Ascophyllum nodosum* (knotted wrack / grisetang), miscellaneous filamentous algae (also referred to as turf algae, or "lurv" in Norwegian), and *Zostera marina* (eelgrass / ålegras). Abiotic habitat types are typically soft sediments and different sized rocks (boulders, cobble, and gravel). See Appendix B for a more complete list of potential species and habitat types to be found in the Helgeland archipelago.

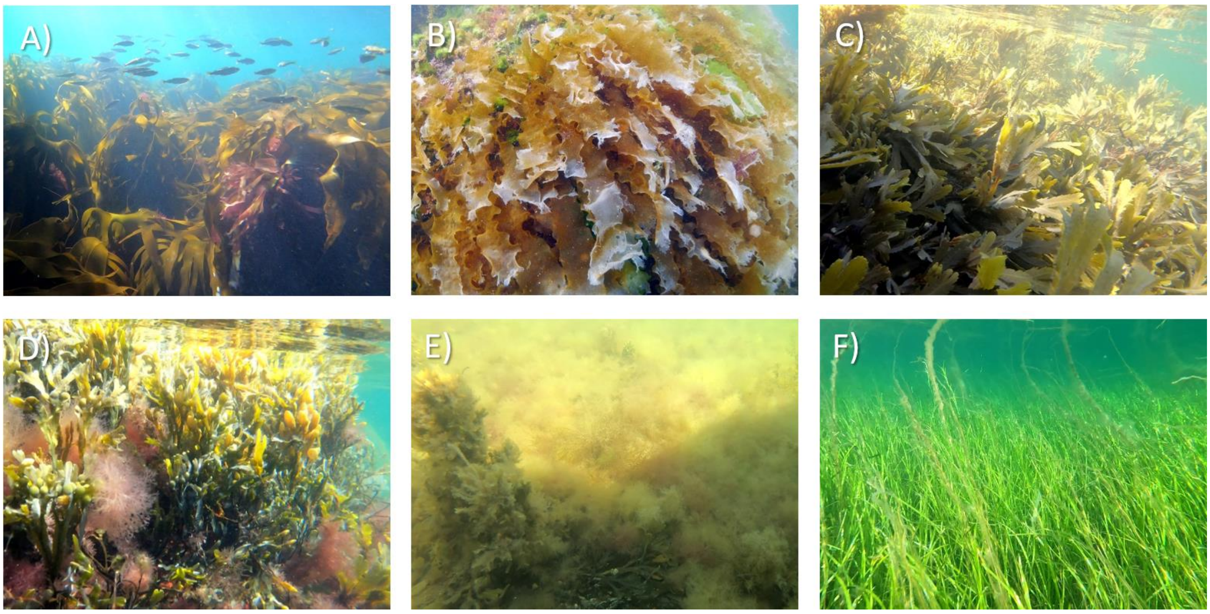


*Figure 2. Some examples of the typical species and habitat types found in the coastal zone of the Helgeland archipelago. A) tangle kelp (*Laminaria hyperborea*), B) sugar kelp (*Saccharina latissima*), C) toothed wrack (*Fucus serratus*), D) bladder wrack (*Fucus vesiculosus*), E) misc. turf algae, and F) eelgrass (*Zostera marina*). Photos (CC): Kasper Hancke (A, E), Janne K. Gitmark (B), Hege Gundersen (C, D, F).*

2.2 Field work

Fieldwork was carried out from 15th to 19th of August 2022. Two teams were working in parallel during this period – a drone flying team (three experienced drone pilots) and a ground data sampling team (two skilled marine biologists with experience from working in the region). August was chosen for fieldwork to avoid the birds' breeding season and to coincide with a period of extra low tide amplitudes. The latter enabled good conditions and visualizations of the marine vegetation both for drone imaging and ground data collection. All drone flights were carried out around low tides and missions were scheduled to fit the tidal rhythm. In addition, the following criteria were evaluated for drone mission planning; 1) good light conditions over the sea surface (sun angle and intensity); 2) good visibility in the water (low concentration of phytoplankton, few particles from e.g., runoff from land, and low concentration of dissolved organic matter, also called yellow matter, in the water); 3) small to moderate wave movements; and 4) moderate wind conditions (<8 m/s). Collecting ground data is less sensitive to weather and light conditions but still sampling was coordinated with drone flying for best practice and data collection (Kvile et al. submitted).

A comprehensive description of the field work carried out including data collection of drone images and ground truth data are given in Gundersen et al. (2022). The report also gives a thorough summary of criteria and issues to consider when collecting drone data, such as logistics, weather, tidal water, etc., to consider when flying drones for coastal mapping. See also Kvile et al. (submitted) for a detailed protocol for drone and ground-truth data collection and image annotation for coastal habitat classification and mapping.

2.2.1. Drone data collection and mission planning

All drone flights were carried out in accordance with current national rules and regulations and in compliance with the Norwegian Aviation Authority guidelines. Drone imaging data were collected using a fixed-wing drone (eBee X mapping drone, AgEagle Aerial Systems, Switzerland [formerly SenseFly], Figure 3). Fixed-wing drones are beneficial for covering larger areas in a time-efficient manner, as they fly with high speed and have long battery capacity relative to rotor drones. The drone was equipped with a multispectral imaging (MSI) camera (Micasense, Rededge-MX) that collects data in 5 specific and independent wavebands. The five color bands are blue, green, red, dark red and infrared. The spectral sensitivity and waveband width is given in Table 1. A multispectral camera was chosen for the study as these provide quantitative data specific of each color band (radiometrically corrected reflectance data) in contrast to RGB sensors, that have only three colors bands (red, green, blue) with overlapping color bands and in addition are hard to obtain quantitatively (not radiometrically calibrated). This makes MSI sensors superior to RGB sensors when it comes to habitat mapping and ML classification protocols. On the other hand, RGB images are superior for visual species identification in the true-color images as used for annotation of MSI datasets, that again are used for training of machine learning algorithms (Section 2.3.2), as they better reproduce the natural color of vegetation and landscapes. Due to budget limitations, the KELPMAP project only had the opportunity to collect one data type and MSI was chosen. However, due to collaborations with other drone projects at NIVA (including the SeaBee project) it was possible to fund a partial coverage with RGB images of the study area (Figure 7). See Kvile et al. (submitted) for a more detailed discussion on sensors.

Figure 3. The fixed-wing drone used in the project, an eBee X mapping drone from AgEagle Aerial Systems. The drone was equipped with a multispectral imaging sensor (Micasense, Rededge-MX) that collected data in 5 specific color bands (see Table 1 for details). Photos (CC): Kasper Hancke.

All drone missions were preplanned and detailed flight plans were designed prior to the flights based on the area to cover, image specifications (altitude, desired pixel resolution, and image overlap), weather condition, and with respect to animal life and human activities. For all missions, the flight altitude was 120 m with 75/75 image overlap, i.e., all images overlapped 75% with the next image both frontal and sideways. This resulted in a pixel resolution on the ground of 8.5 - 9.0 cm / pixel for MSI image data. A summary of the metadata related to the drone mission planning and completion are listed in Table 1. Additional information on drone flights is given in Appendix C.

The completed flight plans were uploaded to the drone that later was launched from a suitable place that facilitated visual-line-of-sight (VLOS) missions in all cases. Since the total area was larger than what could be obtained for VLOS range for one battery set, the study area was covered by carefully planning and executing multiple flight missions. These missions were designed to complement each other with sufficient overlap. The actual flight paths flown by the drones are shown in Appendix D.

The assigned study included coverage of tidal zones and the water between islands as deep as it was possible to identify benthic vegetation from drones and satellites. In this area it was down to a depth of ~10 meters. Also, islands, islets, reefs, and beach zones in the area were mapped.

Table 1. Summary of the metadata related to the drone mapping missions at the two study areas, Steinskjæran and Sandværodden.

	Southern area (Steinskjæran)	**Northern area (Sandværodden)**
Flown area covered	350 ha (3.5 km^2)	200 ha (2.0 km^2)
Date of drone data collection	18. August 2022	19. August 2022
Time of day	09:45-13:54, in total 5 missions incl. battery replacements	10:32-13:31 in total 2 missions incl. battery replacements
Time of low tide	11:00	11:50
Weather condition	First overcast with 6-8 m/s wind, then sunny and 3-4 m/s wind	Overcast with 6-8 m/s wind
Visibility depth	~10 m	~10 m
Drone and sensor	AgEagle eBee X equipped with Micasense Rededge-MX MSI sensors	AgEagle eBee X equipped with Micasense Rededge-MX MSI sensors
Flight altitude	120 m	120 m
Image overlap, frontal/sideways	75/75	75/75
GSD for MSI data	8.5-9 cm	8.5-9 cm
Total number of images collected	~6000 (x 5 wavebands). Equivalent to ~70 GB raw data (prior to processing)	~3000 (x 5 wavebands). Equivalent to ~35 GB raw data (prior to processing)
Number of ground truth (GT) observations	403 positions at deep water with handheld GPS (precision +/- 2 m), 332 positions at shallow water with Leica GS18T (precision +/- 2 cm)	146 positions at deep water with handheld GPS (precision +/- 2 m), 168 positions at shallow water with Leica GS18T (precision +/- 2 cm)

2.2.2. Drone image positioning

Accurate positioning of the drone images was secured using three different approaches, depending on which performed the best at the time and location. The drone/imaging sensors were supported by a Real Time Kinematic (RTK) positioning system that was used as default and was priority when it gave good results. When RTK data quality was insufficient Post Processing Kinetic (PPK) technology was applied. In addition, several Ground Control Points (GCPs) were placed on the ground for a post processing quality check of image positioning (Figure 4).

Drones are commonly equipped with "regular" GPS receivers that provide positioning with a precision of less than 2 meters. However, when equipped with RTK or PPK technology, the positioning precision of images typically can be enhanced to better than 5 cm. This is essential when several drone image datasets are merged or if compared to other drone or ground truth dataset, as is the case during image annotation (Section 2.3.2). RTK corrections entail communication between a receiver on the drone and a ground station, which adjusts for atmospheric interference. Factors like cloud cover, ionospheric disturbances, and satellite positions can affect the coverage and precision of RTK. Incidents of high ionospheric activity can be monitored through online services. While drones with good satellite coverage can make real-time corrections during flight, those with poor coverage must undergo post-processing (PPK) for corrections. Post-processing is also necessary when using GCPs for positioning, or when used for quality check of the RTK or PPK georeferencing of drone images. This process is particularly important in areas with limited coverage, such as remote coastal regions.

Ground Control Points (GCPs) were used to improve and evaluate the accuracy of georeferencing drone orthomosaics. These points were placed on the ground in the mapped area, using crosses made of wood or degradable paint (Figure 4). The position of these points was then determined with a high precision Global Navigation Satellite System (GNSS) instrument (Leica GS18T), with an accuracy level better than 3 cm. A total of 15 control points were placed in the southern area and 5 in the northern area.

Figure 4. Ground control points (GCPs) were used to ensure correct geopositioning of the drone images. For the marking, either an environmentally friendly, soluble chalk paint (left) that disappears in less than 30 days or a portable and reusable wooden cross painted in white (right) was used. Photos (CC): Hege Gundersen (left) and Kasper Hancke (right).

2.2.3. Ground truth data collection

Ground truth (GT) data were collected as a central part of the vegetation mapping and was collected by trained marine biologists that sampled information about the present habitats and species through on-site ground observations. Collection of GT data had two purposes, 1) to ensure georeferenced accurate and expert-identified species and habitat observations which could be used to guide the annotation of the drone images (i.e., identification of species and habitat types in the images); and 2) to validate the quality of the habitat maps created from the ML-trained habitat classification algorithms. Data for the two purposes were collected as part of the same fieldwork, but the GT data were separated after the fieldwork to avoid that no GT data were used for both purposes, and thus ensure an unbiased validation. GT data observations were positioned either with a regular GPS sensor (549 observations, Garmin GPSmap 66, precision +/- 2 m) or with a high-resolution GNSS receiver (500 observations, Leica GS18T, precision +/- 2 cm) (Figure 5).

Observations on land and in shallow waters, in the littoral and sublittoral zones to a water depth of approx. 0.5 meters positions were registered with a high-precision GNSS receiver, while at deeper waters a boat was used and observations were registered with either underwater aquascopes (down to about 2 meters) or a video-transmitting drop camera (under 2 meters depth, Figure 5). In total, 1049 observations were collected over two consecutive days at Steinskjæran (735 observations) and at Sandværodden (314 observations) respectively (see details in Table 1).

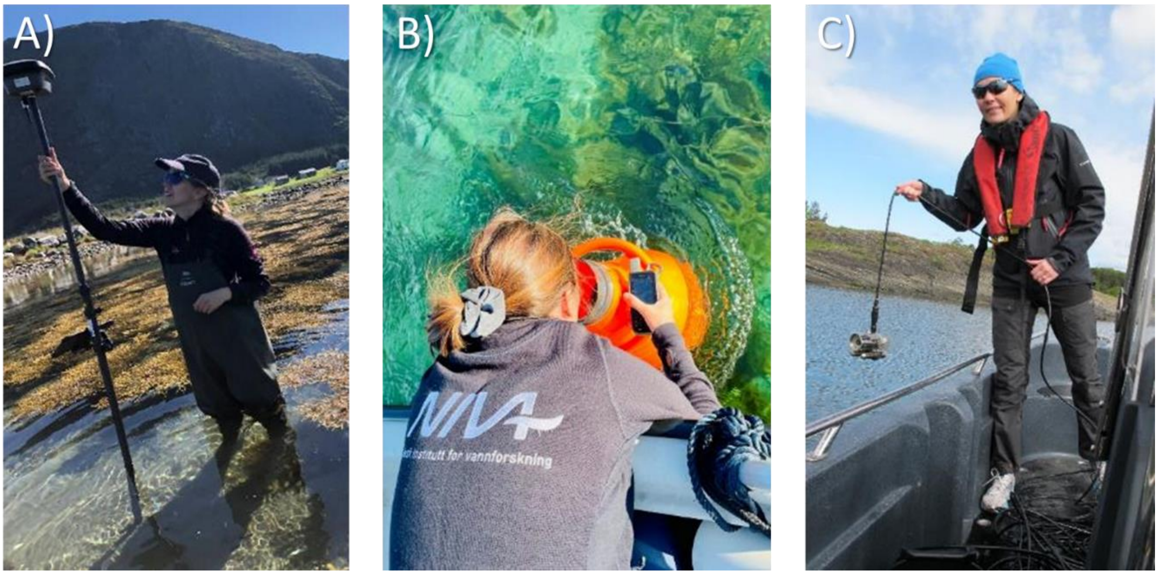


Figure 5. Equipment used for ground truth data collection. A) high-precision GNSS positioning receiver (Leica GS18T GNSS/GPS), B) traditional aquascopes, and C) a video-transmitting underwater drop camera with depth sensor and recording capability (Atlantis 5080) which is commonly used in traditional underwater mapping applications. Photos: Kasper Hancke (A, B) and Janne K. Gitmark (C).

Ground truth observations were registered according to a pre-designed classification system developed by the SeaBee drone infrastructure project (Kvile et al. submitted) and customized for KELPMAP. The classification system includes both biotic and abiotic species, habitat- and nature types ordered into three hierarchical classes. The hierarchical structure contains 11 level 1 classes, 17 level 2 classes and 39 level 3 classes (Table 2). Each GT observation was registered to the highest possible precision level (lowest class) and observations organized accordingly.

2.2.4. Satellite data

The satellite images used were VHR images obtained from the Pleiades service, provided by the client (NEA) and open available Sentinel-2 satellite images. The VHR images have a resolution of 2 m, whereas Sentinel-2 has a resolution of 10 m (Figure 6). The selected VHR image was an orthomosaic from 1. July 2018 and were considered to be the best one among an incomplete patchwork of available images of various quality. The criteria used to select the images was the amounts of clouds, waves, and other distortions in the water, as well as the spatial completeness. The chosen image represented the orthomosaic of best quality, based on these criteria, and the fact that it covered both study areas. Despite being more than four years old, we assumed that the benthic vegetation has not changed significantly, at this spatial resolution, over this period. The VHR satellite data set consisted of four bands (red, green, blue and near infrared), at a 2-meter ground sampling distance.

The Pleiades orthomosaic is processed at remote sensing Level 3 (ortho). This product is a georeferenced orthomosaic in Earth geometry, corrected from acquisition and terrain off-nadir effects, and produced using fully automatic processing and proprietary algorithms. The orthomosaic has been corrected so that it may be superimposed on a map (for viewing angle and ground effects). On top of radiometric and geometric adjustments, a geometric process using a relief model (known as orthorectification) eliminates the perspective effect on the ground, restoring the geometry of a ‘perfect’ vertical shot (see ESA webpage and Perko et al. 2019 for additional information on the image data, orthorectification, and processed steps).

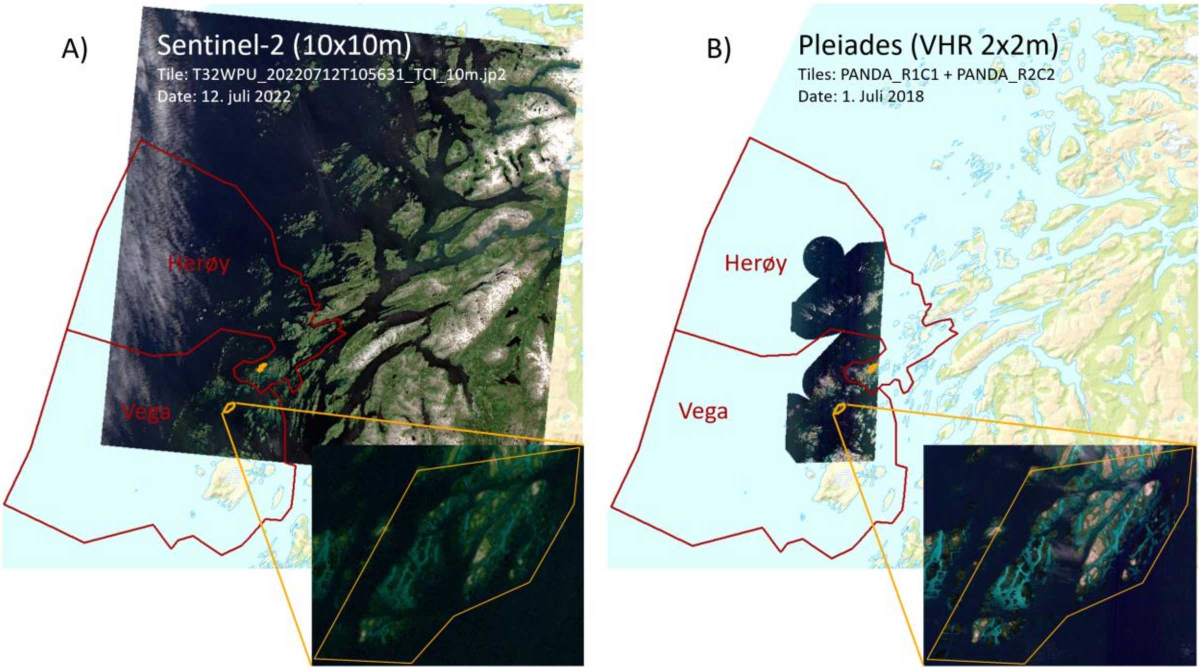


Figure 6. The satellite images, covering the southern and northern study areas (small orange polygons) along with large parts of the Vega and Herøy municipalities (red outlines). In the background shows "Kartdata3 WMS" from the Norwegian Mapping Authority (Kartverket). A) shows the Sentinel-2 tile T32WPU from 12. July 2022, at a 10 m resolution, downloaded from Copernicus Open Access Hub. B) shows Pleiades satellite orthomosaic images (tiles R1C1 and R2C2) at a 2 m resolution. The Pleiades satellite data were provided by the client (NEA) and the selected orthomosaic is from 1. July 2018, downloaded from the PANDA database. The presented satellite images are regarded as the best ones available, based on the cover of clouds, waves, and other distortions in the image, as well as the spatial completeness (covering both study areas). Credits: Copernicus Sentinel-2 data (2022) and © CCME (2018), provided under COPERNICUS by the European Union and ESA, all rights reserved.

2.3 Data analysis

2.3.1. Post-processing of drone images

The drone missions resulted in many individual images that were post-processed through several steps necessary to complete prior to any habitat classification analysis. First, image quality was assured through orthorectification and photogrammetry (a process where images are converted to a mapping format by removing geometric distortions originating from the drone, sensor, terrain, and other factors). Secondly, individual images were stitched together to a composite image orthomosaic, which then was georeferenced. All drone image processing was performed using the commercial drone image analysis software Pix4D (www.pix4d.com), as described in Li et al. (2023).

In total, approximately 45,000 drone images (>100 GB of data) were collected from the two areas (Table 1). From the southern area, the data material consisted of approximately 6,000 images for each of the five color bands, i.e. a total of 30,000 images. From the northern area, around 3,000 images were collected resulting in around 15,000 multispectral images. For more detailed information on the image post-processing methodology, we refer to the field report by Gundersen et al. (2022).

2.3.2. Annotation of drone images for ML analysis

Annotation is the process of labeling drone images to distinguish and classify various elements of interest in the image, such as objects or segments. ArcGIS Pro version 3.0.3 with the *Image Classification* tool in the *Image Analyst* extension was used for this purpose. The drone images were divided into 17 different subareas, where subarea 1 to 10 covered the southern study site in Vega, and subarea 11 to 17 covered the northern study site in Herøy (Figure 7 and Figure 8). The division into subareas was done to facilitate dividing the dataset into training and validation data, in the analyses. Polygons where the vegetation or ground cover could be identified were annotated by biology experts at NIVA. Each annotated polygon was labelled with a thematic class selected from a class hierarchy of three different levels (Table 2). The annotator aimed to annotate using a class from the highest level (finest detail) as possible. In the southern area, subarea 1 to 5 were annotated, whereas in the northern area, subarea 11 to 14 were annotated. See more details about the annotation process in Kvile et al. (submitted).

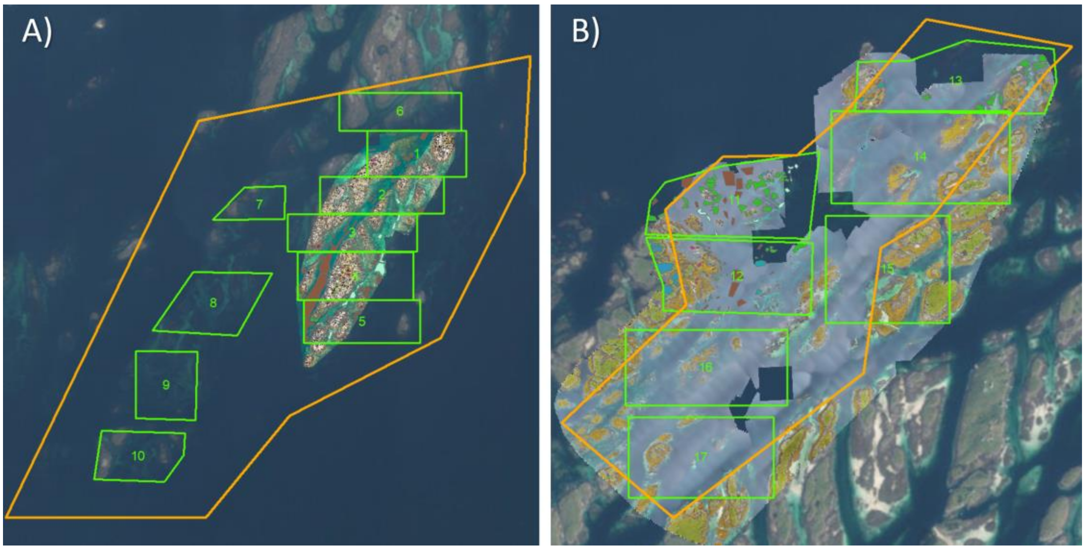


Figure 7. Overview of the 17 subareas used for annotations, validation and testing of the ML analysis in the southern (A) and northern (B) study area. RGB drone images in the background, overlayed aerial images by the Norwegian Mapping Authority (Geocache, Kartverket).

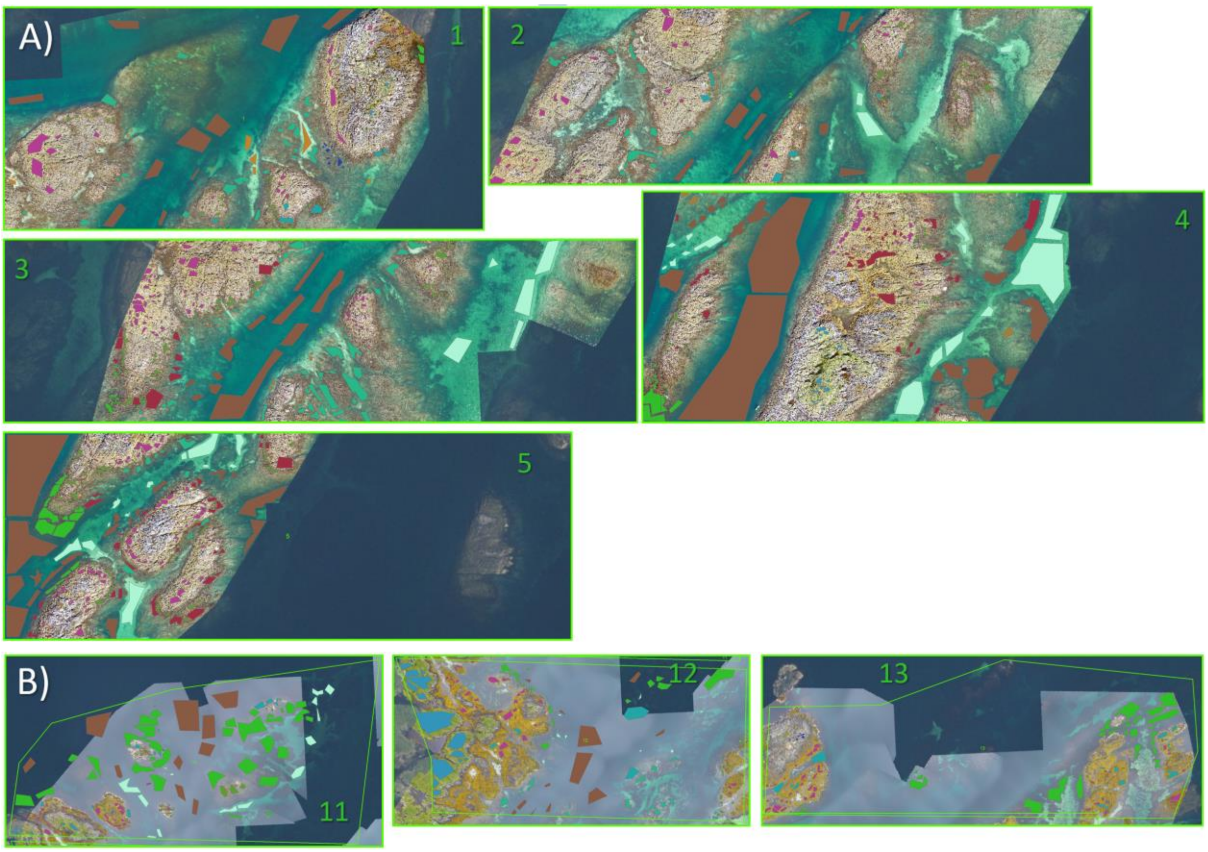


Figure 8. Close-up images of some of the subareas used for annotations from the southern (A) and northern (B) study area. RGB drone images in the background, overlayed aerial images by the Norwegian Mapping Authority (Geocache, Kartverket).

Table 2. The hierarchical class structure contains 11 classes at level 1, 17 classes at level 2 and 39 classes at level 3. This is version 0.1 of the SeaBee habitat classification, which is in continuous development (see latest version at SeaBee GitHub). Not all the classes in this list were found in the present study.

LEVEL 1	LEVEL 2	LEVEL 3	Scientific name	Norwegian name
ALGAE	BROWN	ALAES	*Alaria esculenta*	Butare
ALGAE	BROWN	CHOFI	*Chorda filum*	Martaum
ALGAE	BROWN	DESVI	*Desmarestia viridis*	Mykt kjerringhår
ALGAE	BROWN	LAMHY	*Laminaria hyperborea*	Stortare
ALGAE	BROWN	LAMDI	*Laminaria digitata*	Fingertare
ALGAE	BROWN	PELCA	*Pelvetia canaliculata*	Sauetang
ALGAE	BROWN	SACLA	*Saccharina latissima*	Sukkertare
ALGAE	BROWN	FUCSE	*Fucus serratus*	Sagtang
ALGAE	BROWN	ASCNO	*Ascophyllum nodosum*	Grisetang
ALGAE	BROWN	FUCSP	*Fucus spiralis*	Spiraltang
ALGAE	BROWN	FUCVE	*Fucus vesiculosus*	Blæretang
ALGAE	BROWN	HALSI	*Halidrys siliquosa*	Skulpetang
ALGAE	BROWN	MACROMIX	*Mixed macroalgae*	Blandet makroalger
ALGAE	GREEN	CLARU	*Cladophora rupestris*	Grønndusk
ALGAE	GREEN	GREENROCK	*Green algae on rock*	Grønnalge på stein
ALGAE	GREEN	PRASP	*Prasiola spp.*	Mosegrønske
ALGAE	GREEN	ULVIN	*Ulva intestinalis*	Tarmgrønske
ALGAE	RED	PORUM	*Porphyra umbilicalis*	Fjærehinne
ALGAE	RED	CHOCR	*Chondrus crispus*	Krusflik
ALGAE	RED	FURLU	*Furcellaria lumbricalis*	Svartkluft
ALGAE	RED	HILRU	*Hildenbrandia rubra*	Fjæreblod
ALGAE	RED	PALPA	*Palmaria palmata*	Søl
ALGAE	RED	VERLA	*Vertebrata lanosa*	Trøffeltang
ALGAE	TURF	TURF	Turf unspec.	Lurv uspes.
MAERL	MAERL	MAERL	Maerl unspec.	Rugl uspes.
URCHIN	URCHIN	ECHES	*Echinus esculentus*	Rød kråkebolle
BEACHCAST	BEACHCAST	BEACHCAST	Beachcast	Tangvoll
ANGIO	ANGIO	ZOSMA	*Zostera marina*	Ålegras
WOOD	WOOD	WOOD	Driftwood	Drivved
TERR.VEG.	TERR.VEG.	TERR.VEG.	Terrestrial vegetation	Terrestrisk vegetasjon
ROCK	BOULDER	BOULDER	Boulder	Blokk
ROCK	BOULDER	BOULDER_RUR	Boulder with barnacles	Blokk med rur
ROCK	BOULDER	BOULDER_LAV	Boulder with lichen	Blokk med lav
ROCK	COBBLE	COBBLE	Cobble stones	Stein
ROCK	GRAVEL	GRAVEL	Gravel stones	Grus
SEDIMENT	SAND	SAND	Sandy bottom	Sandbunn
SEDIMENT	MUD	MUD	Muddy bottom	Mudderbunn
ANTHRO	ANTHRO	ANTHRO	Anthropogenic	Antropogent
DEEP	DEEP	DEEP	Deep sea	Dypt hav

2.3.3. ML analysis and habitat classification

In our study, two different convolutional neural network (CNN) solutions were integrated into the SeaBee processing pipeline. These were the U-Net (Ronneberger et al. 2015) and the Dense Dilated Convolutions Merging (DDCM) Network (Liu et al. 2020). Both networks were trained using a loss function consisting of a DICE + Cross Entropy (CE) loss functions and an ADAM optimizer. The learning rate was 0.00012 for both networks and the ADAM related hyper-parameters were selected to $\beta_1 = 0.9$, $\beta_2 = 0.99$ and $\epsilon = 1e^{-8}$ and a weight decay of 0.00002. See Fact Box 2.1 for a simplified explanation of the CNN analysis.

2.3.4. Out-of-distribution (OOD) detection and confidence estimation

A commonly used approach for estimating the per-pixel confidence is to use the normalized score values after the softmax at the output of the network. Each class is assigned a score value between 0 and 1, which sums up to one over all classes. As an estimate for confidence, we may consider using e.g., the largest score value of all classes, the difference between the largest and second largest score values, or entropy. However, for deep neural networks the score values are not calibrated, i.e., they do not reflect the exact probability of each class. Hence, classifier calibration may be necessary. Moreover, the score values provided by the softmax function may not handle out-of-distribution (OOD) samples properly, as an OOD sample may be given a score value of one if the observation is in the tail of all class conditional distributions.

In this study, we therefore applied a confidence estimation algorithm that estimated the per-pixel p-value of the assigned class. We considered the values, $\boldsymbol{y}$, in the layer prior to the softmax in the network. During training, we used an estimate of the corresponding class-wise mean vector, $\boldsymbol{\mu}_c$, and covariance matrix, $\boldsymbol{\Sigma}_c$, using validation data calculated as:

$$\boldsymbol{\mu}_c = \frac{1}{N_v} \sum_{n=1}^{N_v} \boldsymbol{y}_n$$

$$\boldsymbol{\Sigma}_c = \frac{1}{N_v} \sum_{n=1}^{N_v} (\boldsymbol{y}_n - \boldsymbol{\mu}_c)(\boldsymbol{y}_n - \boldsymbol{\mu}_c)^T$$

where N_v is the number of samples and T denotes the matrix transpose. In the inference (production) phase, we applied the estimated per-class mean vectors and covariance matrices in the equations above to estimate the p-value of the response for the selected class by assuming a Gaussian distribution, i.e.,

$$p_c = 1 - \Phi^{-1}_{\chi^2_{C-1}}\left((\boldsymbol{y} - \boldsymbol{\mu}_c)^T \, \boldsymbol{\Sigma}_c^{-1} (\boldsymbol{y} - \boldsymbol{\mu}_c)\right),$$

where c is the selected class and $\Phi^{-1}(\cdot)$ is the cumulative distribution function of the χ^2_{C-1}-distribution. If the response $\boldsymbol{y}$ is an out-of-distribution value, the p_c will be very small, indicating that the response corresponds to an OOD sample (i.e., not reliable). Hence, the p-value of the selected class, p_c, serves as a confidence measure, indicating how confident we are about the decision. Moreover, we may also use the p-value to perform OOD detection, in a similar way as p-values are used to perform hypothesis tests in traditional statistical modelling. Thus, the generated p-values can be interpreted as the probability that a sample (pixel) is from the same distribution as the training data. If the p-value is lower than a pre-defined level of significance, we reject the null hypothesis and classify the corresponding pixel as OOD.

2.3.5. Creating masks to exclude uncertain predictions

A habitat classification map should always be accompanied by a map of uncertainty, which displays the reliability of the predictions for each pixel. The final mapping product should be refined based on p-values, filtering out highly uncertain pixels. Generally, pixels with p-values below 0.05 should be excluded, although this threshold might be adjusted depending on the specific management purpose in question (Bryn et al. 2021). In our study, we used a threshold of $p < 0.05$ to create masks for producing classification maps that exclude predictions of low certainty.

2.3.6. Upscaling to satellite data

Two sources of satellite data were used for the spatial upscaling of drone algorithms – both accessed through the Copernicus repository (www.copernicus.eu). These were Sentinel-2, with a spatial resolution of 10 m, and commercial Very-High-Resolution (Pleiades VHR) data from the PANDA database, with a resolution of 2 m. Sentinel-2 images are freely accessible, while access to the commercial dataset was granted to NIVA and NR through NEA.

In contrast to the MSI drone images with approximately 9 cm resolution, the satellite images have a significantly lower spatial resolution. Therefore, the annotated polygons from the drone data translate into substantially fewer pixels in the satellite images, presenting a challenge in the upscaling process from high to low resolution. Due to this limitation, the annotated data from the drones are not used directly as training data for the satellite images. Instead, the aggregated classification maps produced by the CNN analysis of drone images served as training data for the satellite images, in the same way that annotated data were used for the drone image.

To ensure the reliability of the satellite training data, only predictions from the drone maps with high p-values (non-OOD pixels) were used. This training data was then applied to train an ML model for the satellite data, aiming to predict habitat classes from the satellite images using the same procedure as the drone-data ML training. Given the potential risk of including erroneous classified pixels into the satellite training data, we applied a label smoothing using a value of 0.1. Label smoothing is a form of output distribution regularization that prevents overfitting of a neural network by softening the ground-truth labels in the training data and penalizing overconfident outputs.

The upscaling process was limited to the classification maps from habitat class level 1, which has the lowest level of detail, and thus largest uniform habitat polygons. For instance, at this level all species of kelp and rockweed are commonly classified as ALGAE (Table 2).

FACT BOX 2.1 A simplified explanation of habitat classification using Artificial Intelligence

Convolutional Neural Networks (CNN) are a type of deep learning model used for image analysis tasks. **U-Net** and **Dense Dilated Convolutions Merging** (DDCM) are two types of networks, especially suitable for picking out patterns in images.

During the training of these networks, a so-called **loss function** measures how well the model's predictions match the actual outcomes. By minimizing this loss through various strategies, the model learns the specific task. An **optimizer** is an algorithm that tweaks the model parameters, like weights, to reduce loss. One such optimizer is **ADAM**, which is especially suited to handling large and complex datasets.

One aspect of model reliability is handling **Out-of-Distribution** (OOD) data, which is new or different data types not seen during training. Through **confidence estimation**, we can assess how certain the model is about its **predictions**, often using statistical methods to calculate the **likelihood** of predictions being correct.

When using CNN to analyze images, **p-values** help us understand how confident we can be in the predictions from the ML analysis. P-values ranges between 0 and 1, where low p-values (usually below 0.05) suggest that the prediction might not be reliable because it appears unusual or out of the ordinary compared to what the ML (or AI) usually sees. High p-values mean the prediction fits well with the CNN's previous knowledge, indicating that the higher the p-value, the more confident we are in the prediction's **accuracy**.

The local high-resolution drone products can be applied to a larger area by upscaling. Two types of satellite data from the Copernicus repository were used for this purpose, these are Sentinel-2 and the commercial PANDA, which has a spatial resolution on 10 m and 2 m, respectively. The classification maps made in CNN analysis was used as training data for the satellite images. Only pixels with p-values > 0.05 was used as training data.

3 Results

3.1 Environmental conditions and logistics

Numerous precautions and considerations were taken to conduct a safe and successful field campaign. The drone missions were scheduled over two successive days, 18-19 August 2022, when weather conditions were suitable for drone flights (Table 1, Appendix C). The flight missions were aimed at being conducted at times when the tide was at its lowest, thus they occurred within a timeframe that extended a few hours before and after the lowest tide, resulting in tidal waters being max 50 cm above absolute minimum. Several other factors had to be aligned for a successful drone mission, such as those regarding precipitation, wind speed, cloud cover, visibility, sunrise, and sunset times. See Gundersen et al. (2022) for a more detailed overview of the environmental parameters and criteria related to the drone fieldwork for the Vega campaign, and Kvile et al. (submitted) for more general specifications.

3.2 Drone data collection

Both multispectral imaging (MSI) and RGB images were collected for both areas, however, only the MSI images covered both areas completely. This is because MSI images were part of the assignment, whereas RGB images were collected opportunistically funded by NIVA/SeaBee. Nevertheless, all collected images were used in the annotation (RGB) and analyses (MSI). In total, approximately 9,000 individual MSI drone images were collected (6,000 and 3,000 in the southern and northern areas, respectively), each with five bands, totaling approximately 105 GB (70 + 35) of raw material. The corresponding numbers for RGB were 2,731 individual images and ~32 GB (Table 1, Figure 9).

The final stitched images, i.e., the georeferenced orthomosaics, covered close to 100% (north) and approximately 55% (south) of the surveyed polygons (Figure 9C and D). The reason for the rather low stitching percentage in the southern area is because this polygon contained large areas of deep water without any physical features, which largely hampered the capability of stitching single images into orthomosaics. The northern area, however, included more islands, skerries, and visible structures on the seafloor, making the stitching of individual images into orthomosaics more successful.

3.3 Ground truth data collection

The ground truth (GT) data were collected for two purposes – to train the ML analysis and to validate the final model. The maps in Figure 10 show all the GT data collected and indicate the transects that were followed when collecting ground data. Following such transects is done both to cover depth gradients in the area and because it is an efficient way of making registrations (Gundersen et al. 2022, Kvile et al. submitted).

The resulting dataset contained 425 (41%) records of kelp, of which the priority species tangle kelp (*stortare*) accounted for 26%, while sugar kelp (*sukkertare*), winged kelp (*butare*), and oar weed (*fingertare*) accounted for 7%, 4% and 4% respectively (Figure 11). Seaweed (*tang*) species accounted for 28% (288 records), consisting of toothed wrack (*sagtang*) (9%), knotted wrack (*grisetang*) (7%), bladder wrack (*blæretang*) (4%), channeled wrack (*sauetang*) (3%), spiral wrack (*spiraltang*) (2%) and sea oak (skulpetang) (2%). Note that the number of different species is not representative of what is present in the area, as this methodology is tailored towards providing the best possible dataset for training of ML algorithms for habitat classification and is not necessarily restricted by the same constraints as statistical analyses. See also Appendix B for the full list of species and habitat types observed in the area.

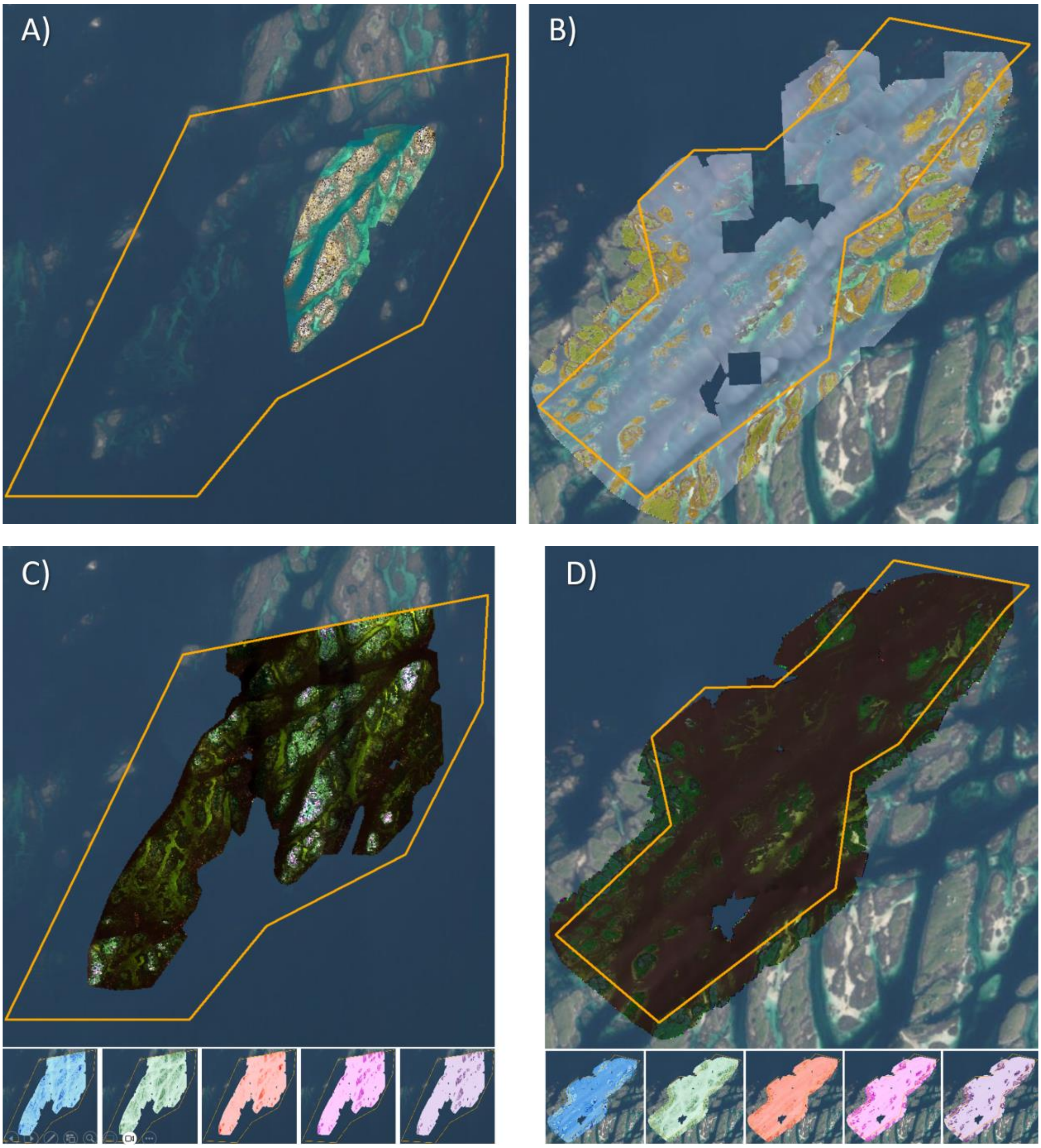


Figure 9. Georeferenced orthomosaics of drone images collected for the southern (A and C) and northern (B and D) areas. RGB images (upper) represent true colors, however, with some slight differences between A and B due to different camera/sensor types and settings on NIVA's (A) and SpectroFly's (B) drones. MSI images (lower) are composite images that combine data from the five spectral bands to create a "false" RGB, also called RGB composite image. Individually colored images from each of the five spectrums are displayed in smaller images below. Aerial images by the Norwegian Mapping Authority (Geocache, Kartverket) in the background.

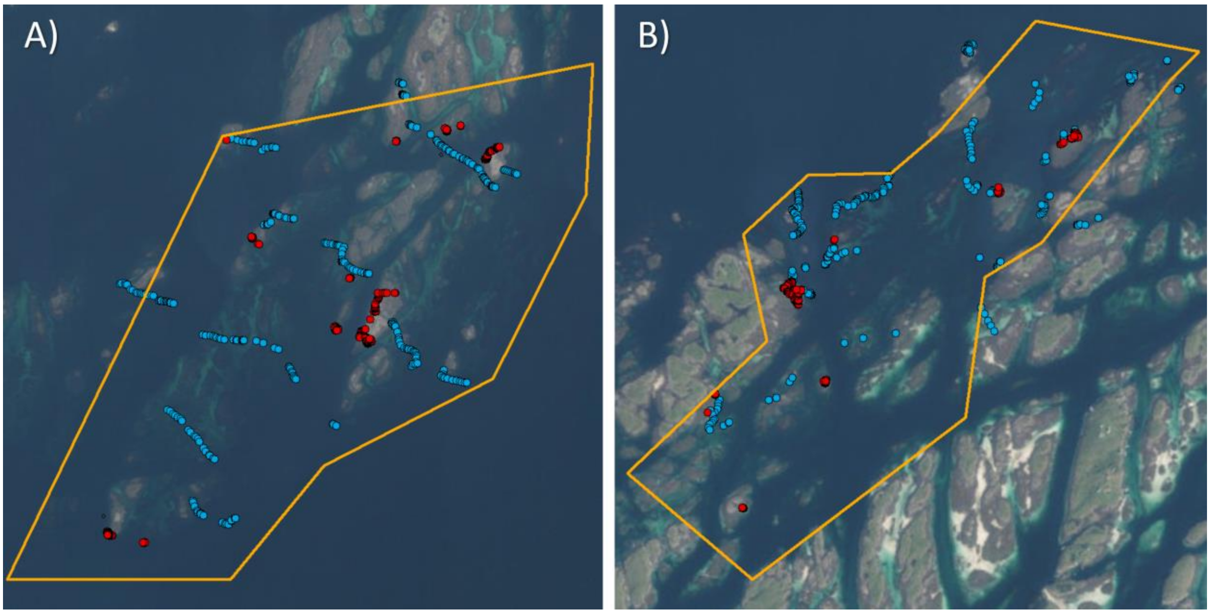


Figure 10. Map of data collected in the southern (A) and northern (B) areas. Red dots show recordings made in shallow water (< 1 m) using wader pants and high-precision Leica positioning systems (Figure 5A). Blue dots represent recordings from deeper areas (from 1 m, down to approximately 12 m), made from a boat with either aquascopes or an underwater camera (Figure 5B, C). Aerial images by the Norwegian Mapping Authority (Geocache, Kartverket) in the background.

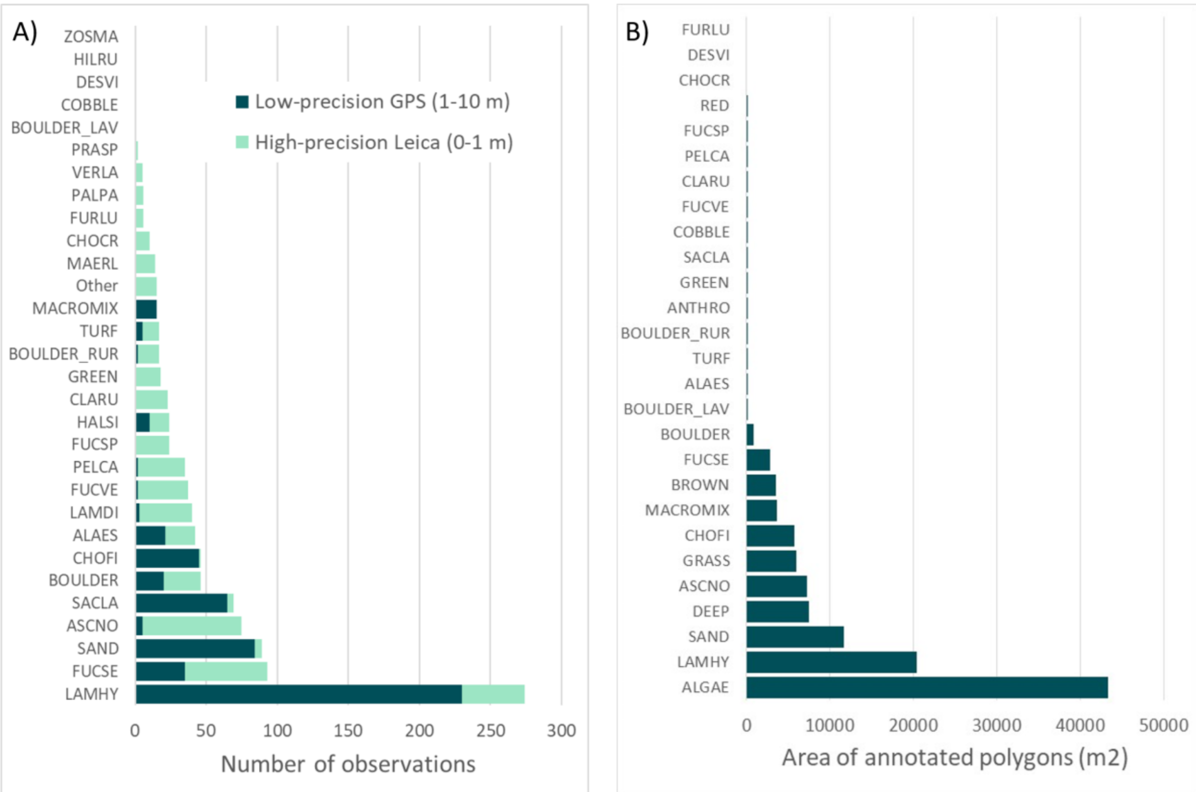


Figure 11. Histograms showing A) the number of field observations per habitat class, sampled at deep (1-12 m; dark green) and shallow (< 1 m; light green) water, using low-precision GPS (dropdown underwater camera) and high-resolution Leica GNSS/GPS, respectively, and B) the total area (m^2) of annotated polygons used to train the ML algorithms. The reason for the very small areas for several of the habitat classes is due to the species' limited size and distribution. Habitat class codes are a mix of classes from different levels. See Table 2 and Appendix B for explanation of habitat class codes. Note that the relative number of observations for each class is not necessarily representative for their areal coverage of the respective habitats in the field.

3.4 Habitat classification and maps of uncertainty

The habitat classification was performed at all three levels for both areas. The total predicted area was 3.68 km^2, corresponding to 1.98 km^2 for the southern area and 1.70 km^2 for the northern area. In total, 8 level-specific habitat classes were identified/predicted at level 1, whereas 12 and 16 level-specific classes were identified at level 2 and 3, respectively (Figure 12). ALGAE dominated at level 1 with 2.39 km^2 in total, followed by DEEP (0.56 km^2), SEDIMENT (0.50 km^2), ROCK (0.18 km^2) and TERR.VEG. (0.05 km^2) (Table 3). The top-most abundant habitat classes at level 2 were BROWN (2.44 km^2), DEEP (0.51 km^2), BOULDER (0.38 km^2), SAND (0.29 km^2) and TERR.VEG. (0.06 km^2). Whereas for level 3 it was SAND_SEA (0.98), DEEP (0.60 km^2), CHOFI (0.58 km^2), LAMHY (0.42 km^2), MACROMIX (0.38 km^2), BOULDER (0.28 km^2), ASCNO (0.14 km^2), FUCSE (0.09 km^2), TERR.VEG. (0.07 km^2), ALAES (0.06 km^2), TURF (0.04 km^2), and BOULDER_LAV (0.02 km^2).

From the analyses, also maps of uncertainty was derived, that shows a p-value ranging from 0 to 1 at the pixel level, for each of the six classification maps (Figure 13). Here we defined a cut-off at $p > 0.05$ to be reliable predictions for a specific habitat class at the pixel level (see Section 0 and Box 2.1 for details).

The uncertainty maps were used to mask-out uncertain predictions, to create classification maps with high prediction certainty. These results are shown in Figure 14. In Table 3 the results are summed up as predicted areas (km^2) for each habitat class for each of the three levels, before and after correction by masking out the most uncertain pixels. At level 1, only 1% of the area classified as DEEP was removed in the masking, whereas the areal coverage for all other habitat classes at this level were reduced by more than 50%, including ALGAE that was reduced by 55%. At level 2, however, as much as 99% of the area classified as BROWN was kept after the masking. DEEP and SAND followed by 72% and 50%, respectively. For GREEN (1%), TURF (0%), and COBBLE (0%), practically the entire area predicted as these habitat classes were removed in the masking. At level 3, most of the ANTHRO (82%) and DEEP (71%) classifications were kept, whereas the rest of the habitat classes at this level, where less than 50% was kept (Table 3). This is also reflected in a slightly more red-colored appearance in the uncertainty map at level 3, especially visible for the southern area, in Figure 13.

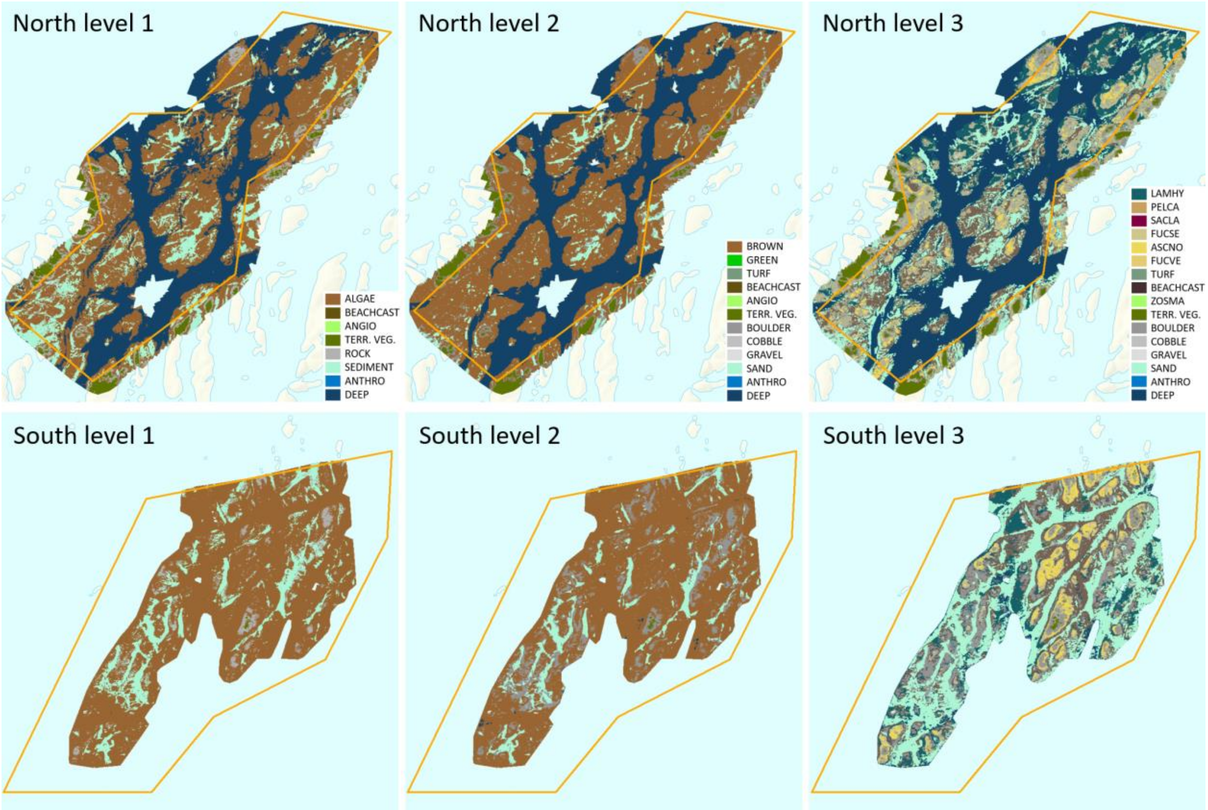


Figure 12. Habitat classification maps of the two study areas, at habitat class levels 1, 2 and 3.

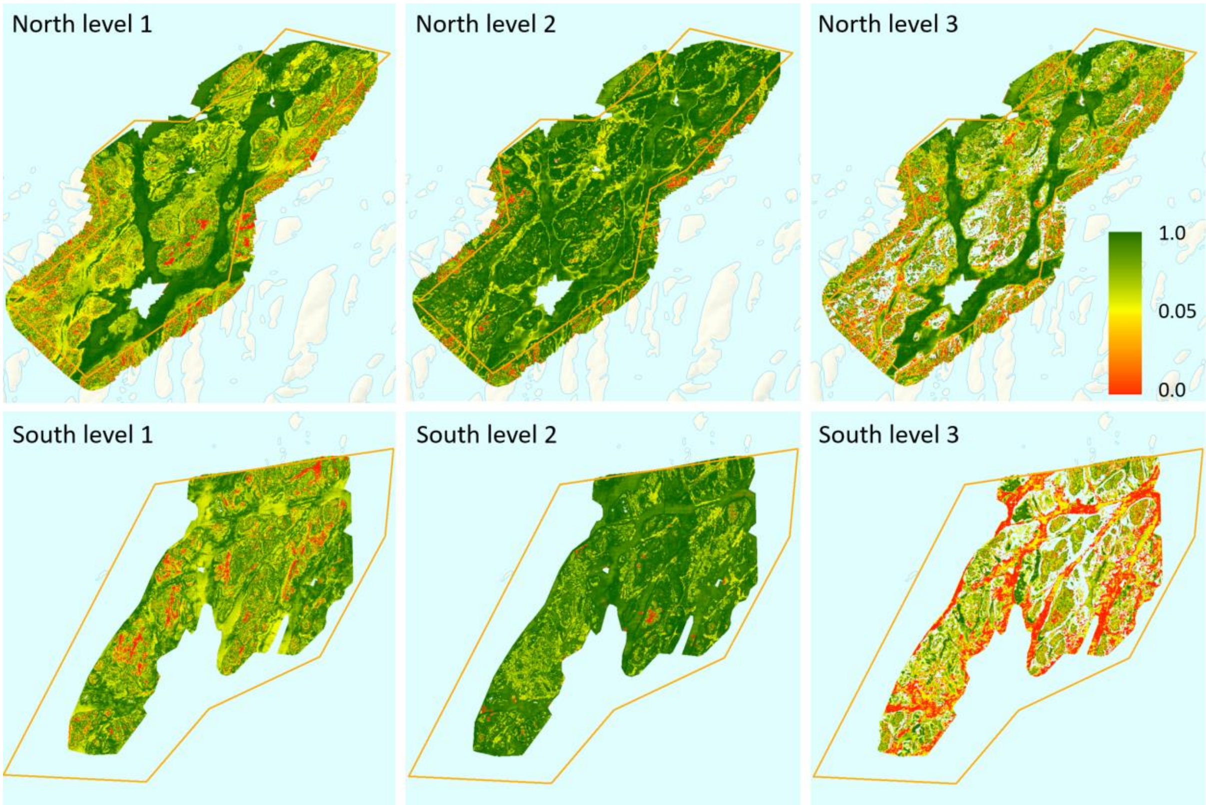


Figure 13. Maps of uncertainty of the predictions shown in Figure 12, where p-values above 0.05 (green) show more reliable predictions, and values below 0.05 (red) show less reliable predictions. The color range is centered around 0.05 (yellow) to indicate the threshold for masking out uncertain pixels to produce the corrected classification maps (Figure 14).

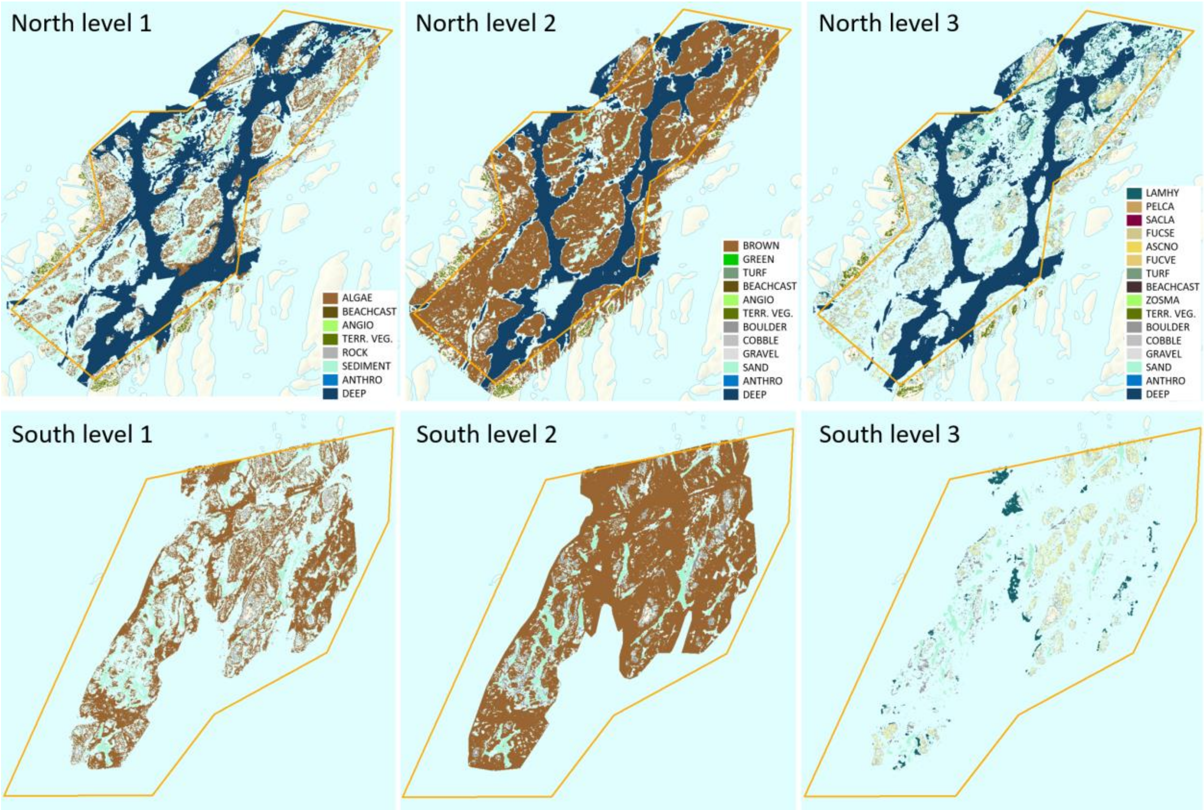


Figure 14. The corrected classification maps for level 1, 2 and 3, after masking out uncertain predictions ($p < 0.05$, Figure 13) from the original maps (Figure 12).

Table 3. Summary of the predicted areal coverage for each habitat class from the ML analyses at the three applied levels. The results summarize the northern and southern sites. The level-specific (predicted) area is given for each habitat class, both before and after masking out (correcting) the most uncertain pixels, using a threshold at $p < 0.05$. The tables are sorted and colored according to the area of each habitat class that is kept after the masking process, where green depicts high predictability, and red depicts low.

LEVEL 1	Area before correction (km²)	% of total area	Area after correction (km²)	% of corrected area	Area removed (km²)	% of original model kept
DEEP	0.56	15 %	0.56	30 %	0.00	99 %
ALGAE	2.39	65 %	1.08	58 %	1.30	45 %
SEDIMENT	0.50	14 %	0.19	10 %	0.31	37 %
TERR.VEG.	0.05	1 %	0.02	1 %	0.04	32 %
ROCK	0.18	5 %	0.04	2 %	0.14	21 %
Total	3.68	100 %	1.88	100 %	1.80	

LEVEL 2	Area before correction (km²)	% of total area	Area after correction (km²)	% of corrected area	Area removed (km²)	% of original model kept
BROWN	2.44	66 %	2.41	79 %	0.03	99 %
DEEP	0.51	14 %	0.37	12 %	0.14	72 %
SAND	0.29	8 %	0.14	5 %	0.14	50 %
TERR.VEG.	0.06	2 %	0.02	1 %	0.04	30 %
BOULDER	0.38	10 %	0.10	3 %	0.29	25 %
GREEN	0.00	0 %	0.00	0 %	0.00	1 %
TURF	0.00	0 %	0.00	0 %	0.00	0 %
COBBLE	0.00	0 %	0.00	0 %	0.00	0 %
Total	3.68	100 %	3.04	100 %	0.64	

LEVEL 3	Area before correction (km²)	% of total area	Area after correction (km²)	% of corrected area	Area removed (km²)	% of original model kept
ANTHRO	0.00	0 %	0.00	0 %	0.00	82 %
DEEP	0.60	16 %	0.43	43 %	0.17	71 %
BOULDER_RUR	0.00	0 %	0.00	0 %	0.00	48 %
MACROMIX	0.38	10 %	0.17	16 %	0.22	44 %
BOULDER_LAV	0.02	1 %	0.01	1 %	0.01	42 %
ASCNO	0.14	4 %	0.05	5 %	0.09	38 %
FUCSE	0.09	2 %	0.03	3 %	0.05	38 %
LAMHY	0.42	11 %	0.14	14 %	0.28	33 %
TERR.VEG.	0.07	2 %	0.02	2 %	0.05	24 %
BOULDER	0.28	8 %	0.04	4 %	0.24	16 %
SAND_SEA	0.98	27 %	0.11	11 %	0.87	11 %
CHOFI	0.58	16 %	0.00	0 %	0.58	0 %
ALAES	0.06	2 %	0.00	0 %	0.06	0 %
TURF	0.04	1 %	0.00	0 %	0.04	0 %
COBBLE	0.00	0 %	0.00	0 %	0.00	0 %
Total	3.68	100 %	1.01	100 %	2.67	

For a better visual comparison of the details in the predicted maps from level 1, 2 and 3, we show in Figure 15 some close-up images of a 9 ha (0.09 km^2) of the southern study area. The details in the level 3 model are mostly due to the separation into different species of brown macroalgae. Level 2 and level 1 predicted maps are, however, quite similar except for a larger proportion of BOULDER at level 2, which is classified as SAND in level 3.

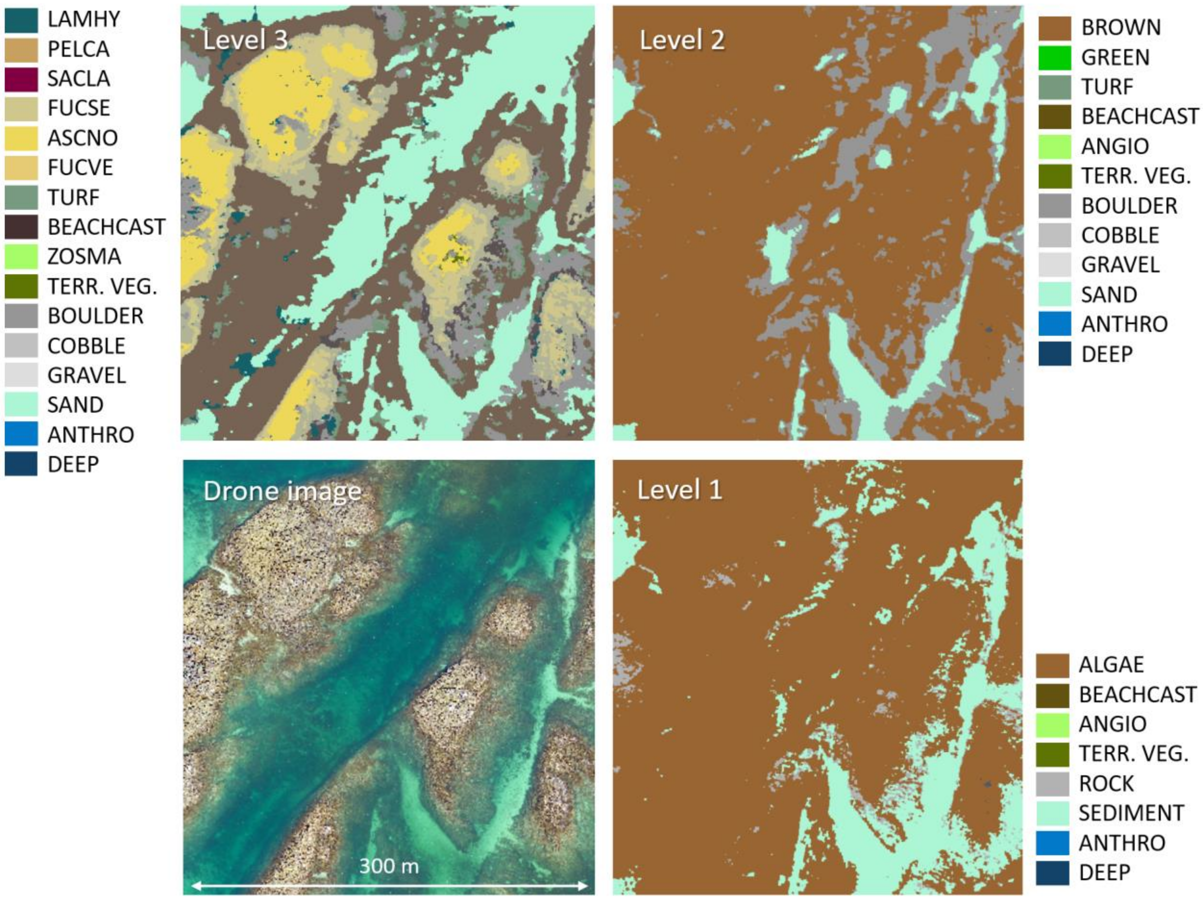


Figure 15. The level 1, 2 and 3 predictive maps of the southern study area shown in Figure 12, zoomed in to compare between levels at a finer scale, and with the details in the RGB drone image (lower left).

3.5 Validation of drone-based habitat classification

To validate the ML models, a set of validation data was held out from the annotation dataset. These were compared to the predicted habitat classes and resulted in the confusion matrixes shown in Figure 16. In general, the diagonal patterns shown in the matrixes from all three analysis levels indicate a high degree of "true positives" (TPs), meaning an overall good match between the predicted (X-axis) and the observed (y-axis) data.

Especially the predictions at level 1 (Figure 16A), resulted in almost full match for TERR.VEG. (TP = 0.98) and ROCK (0.97), and very good for DEEP (0.85), ALGAE (0.81) and SEDIMENT (0.75). However, there is a striking misprediction of ANTHRO that are classified as ROCK (0.78). This is very likely due to very few annotations in the ANTHRO class, and the fact that this is a very heterogeneous habitat class covering all sort of anthropogenic items in the images (houses, infrastructure, etc.).

The predictions at level 2 (Figure 16B) also show full match for TERR.VEG. (1.00) and SAND (0.99), and satisfactory also for DEEP (0.94), BROWN (0.87) and BOULDER (0.84). However, at this level there are some serious mispredictions where GREEN is predicted as BOULDER (0.86), and ANTHRO as TERR.VEG. (0.63). Also, there is a total misprediction of TURF being classified as BROWN (1.00), however this is not surprising since TURF is a collective term for filamentous macroalgae, and a large fraction of these are in fact brown algae.

Overall, the predictions at level 3 (Figure 16C) show relatively good predictions for many of the habitat classes at this detailed level, including several of the kelp and rockweed species, which were of primary focus in this study: CHOFI (1.00), ASCNO (0.92), LAMHY (0.70) and FUCSE (0.65). However, the kelp species ALAES was misinterpreted as another kelp, LAMHY (1.00), which probably have similar spectral signals. And SACLA was largely taken for DEEP (0.72). The latter which might be due to few training and validation data for SACLA, despite a decent number of ground truth observations (Table 3). Also, the fact that SACLA was found quite deep made it difficult to make confident annotations of this species in the drone images. And again, TURF was misinterpreted as CHOFI (0.94), which is a brown macroalgae with similar thread-patterned appearance as CHOFI. Also, ANTHRO is misinterpreted as either TERR.VEG. (0.62), BOULDER_RUR (0.22) or DEEP (0.16), and clearly needs much more annotations to be improved.

A)

OBSERVED / PREDICTED

LEVEL 1	ALGAE	ANTHRO	DEEP	TERR.VEG.	ROCK	SEDIM.
ALGAE	0,81	0,00	0,19	0,00	0,00	0,00
ANTHRO	0,08	0,00	0,09	0,06	0,78	0,00
DEEP	0,08	0,00	0,85	0,00	0,07	0,00
TERR.VEG.	0,02	0,00	0,00	0,98	0,00	0,00
ROCK	0,03	0,00	0,00	0,00	0,97	0,00
SEDIMENT	0,22	0,00	0,00	0,00	0,03	0,75

B)

OBSERVED / PREDICTED

LEVEL 2	ANTHRO	BOULDER	BROWN	DEEP	TERR.VEG.	GREEN	RED	SAND	TURF
ANTHRO	0,00	0,20	0,00	0,16	0,63	0,00	0,00	0,00	0,00
BOULDER	0,00	0,84	0,13	0,00	0,00	0,00	0,00	0,03	0,00
BROWN	0,00	0,00	0,87	0,13	0,00	0,00	0,00	0,00	0,00
DEEP	0,00	0,00	0,00	0,94	0,02	0,00	0,01	0,03	0,00
TERR.VEG.	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00
GREEN	0,00	0,86	0,10	0,00	0,00	0,04	0,00	0,00	0,00
RED	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SAND	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,99	0,00
TURF	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00

C)

OBSERVED / PREDICTED

LEVEL 3	ALAES	ANTHRO	ASCNO	BOULDER	BOULDER_LAV	BOULDER_RUR	CHOFI	COBBLE	DEEP	FUCSE	TERR.VEG.	LAMHY	MACROMIX	SACLA	SAND	TURF
ALAES	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00
ANTHRO	0,00	0,00	0,00	0,00	0,00	0,22	0,00	0,00	0,16	0,00	0,62	0,00	0,00	0,00	0,00	0,00
ASCNO	0,00	0,00	0,92	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,00
BOULDER	0,02	0,00	0,00	0,97	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00
BOULDER_LAV	0,00	0,00	0,00	0,49	0,51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
BOULDER_RUR	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
CHOFI	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
COBBLE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
DEEP	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,92	0,00	0,00	0,00	0,00	0,00	0,00	0,00
FUCSE	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,65	0,00	0,02	0,31	0,00	0,00	0,00
TERR.VEG.	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00
LAMHY	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,00	0,70	0,21	0,00	0,00	0,00
MACROMIX	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,30	0,45	0,00	0,00	0,00
SACLA	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,72	0,00	0,00	0,23	0,02	0,00	0,02	0,00
SAND	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
TURF	0,00	0,00	0,00	0,00	0,00	0,00	0,94	0,00	0,00	0,04	0,00	0,02	0,00	0,00	0,00	0,00

Figure 16. Confusion matrices for validation of the ML analyses at level 1 (A), level 2 (B) and level 3 (C). The diagonal dark green pattern indicates a high number of true positives (TPs), i.e., the relative amount of correctly predicted cases, but also some mispredictions diverging from this pattern. For explanation of the habitat classes see Table 2.

3.6 Upscaling drone outputs with satellite images

Two types of satellite images were used for the upscaling of the drone-based habitat classifications at larger scales. These were Very High Resolution (VHR) images at 2 m resolution from the Pleiades service, and Sentinel-2 at 10 m resolution from the Copernicus Open Access Hub both covering a larger area with the two study areas included (Figure 6). For training of ML classification algorithms for this upscaling, the corrected classification maps from the drone-based predictions were used as training data (Figure 14). This was done only at level 1, and in separate processes for each of the two satellite image sources. Unfortunately, there was not enough data available to create a confusion matrix for the upscaled habitat classifications, so the validation of these maps is solely based on visual inspection.

3.6.1. Upscaling based on VHR (2x2 m)

The output from the upscaling analyses using VHR images are shown in Figure 17 and Figure 18, respectively, the first showing the entire upscaled area and the second a subsection of the areas corresponding to the southern and northern study areas. Very large areas were predicted as kelp in the upscaled prediction map (Figure 17B), also in areas assumed to be too deep for kelp to grow. Whereas these areas were masked out in the masking process, where uncertain predictions (p-values <0.05) were removed. The resulting prediction map (Figure 17D) looks more reliable and realistic as seen by a trained eye with knowledge of the local conditions.

3.6.1. Upscaling based on Sentinel-2 (10x10 m)

Figure 19 and Figure 20 show the same results as in Figure 17 and Figure 18, but based on Sentinel-2 data instead of VHR. The tendency to over-predict kelp forest is evident also for this upscaling. A likely explanation for this is the relatively few annotated data for the DEEP-water class in the drone-based classification maps (see Discussion for details). The reason for this again is the fact that the drone images are difficult to stitch in deep areas, so that deep areas of a study region are sparsely represented in the stitched drone orthomosaic (Figure 9C).

The predicted area covered of each habitat class, at each analysis level, and using different upscaling methods, is shown in Table 4 and discussed in Section 4.2.4. In Section 4.4 we also discuss the size of our predictions compared to other models of macroalgae, such as NIVAs statistical kelp models (Frigstad et al. 2021) and satellite-based models by NORCE (Haarpaintner and Davids 2020).

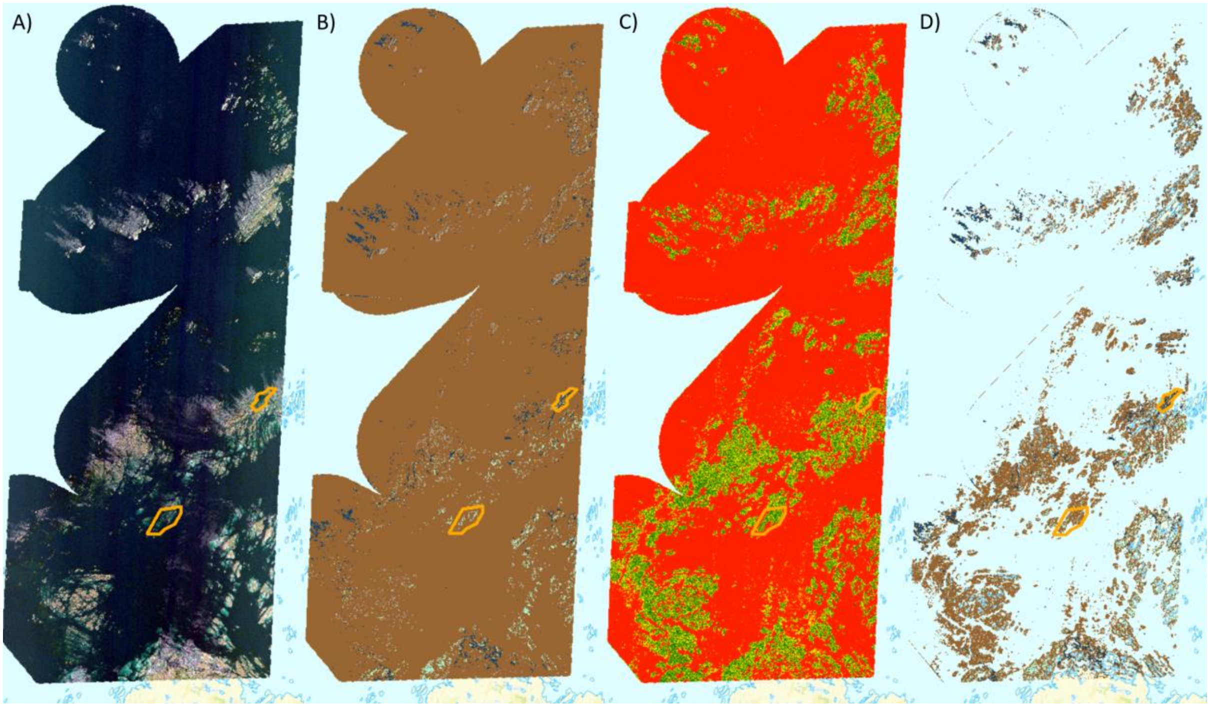


Figure 17. A) The VHR satellite image (2 m resolution) used for upscaling (© CCME (2018), provided under COPERNICUS by the European Union and ESA, all rights reserved), B) the resulting prediction map where kelp forest probably is largely overpredicted, C) the map of uncertainty, showing large areas with uncertain predictions (p < 0.05, in red) typically restricted to deeper areas, but also considerable areas with predictions above the 0.05 cutoff (yellow and green), and D) the resulting prediction map after masking out uncertain predictions (p < 0.05).

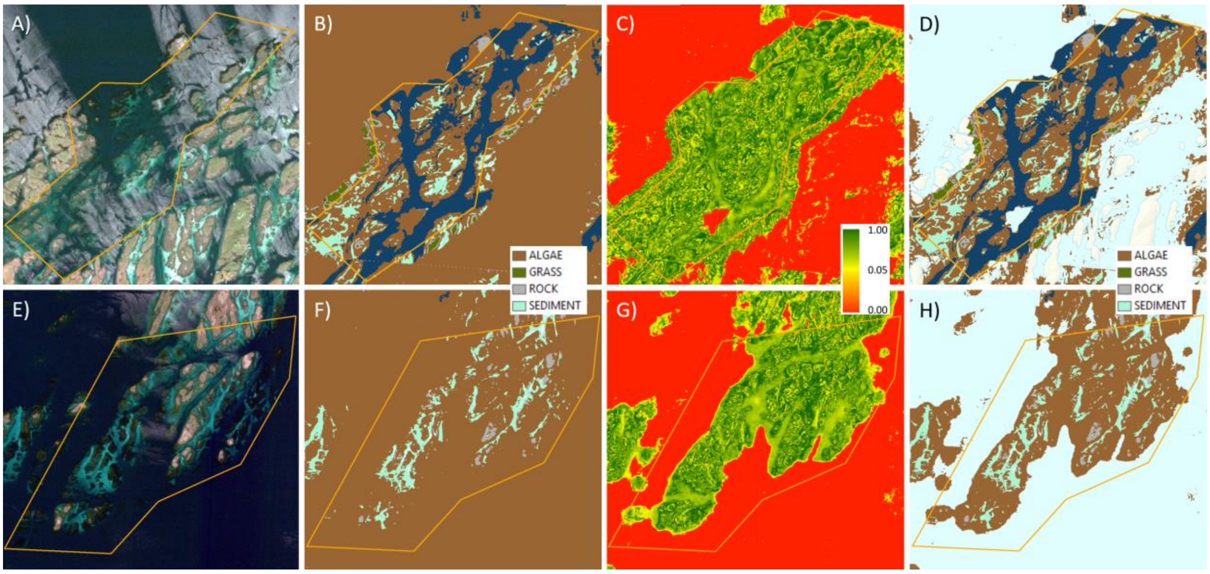


Figure 18. Images and maps in the upscaling process from VHR satellite images (A, E) (© CCME (2018), provided under COPERNICUS by the European Union and ESA, all rights reserved), through classification maps (B, F), uncertainty maps (C, G), showing large areas with uncertain predictions (p < 0.05, in red) typically restricted to deeper areas, but also considerable areas with predictions above the 0.05 cutoff (yellow and green), and the resulting corrected classification maps masked for uncertain predictions (D, H), shown for the Northern (upper) and Southern (lower) study areas, respectively.

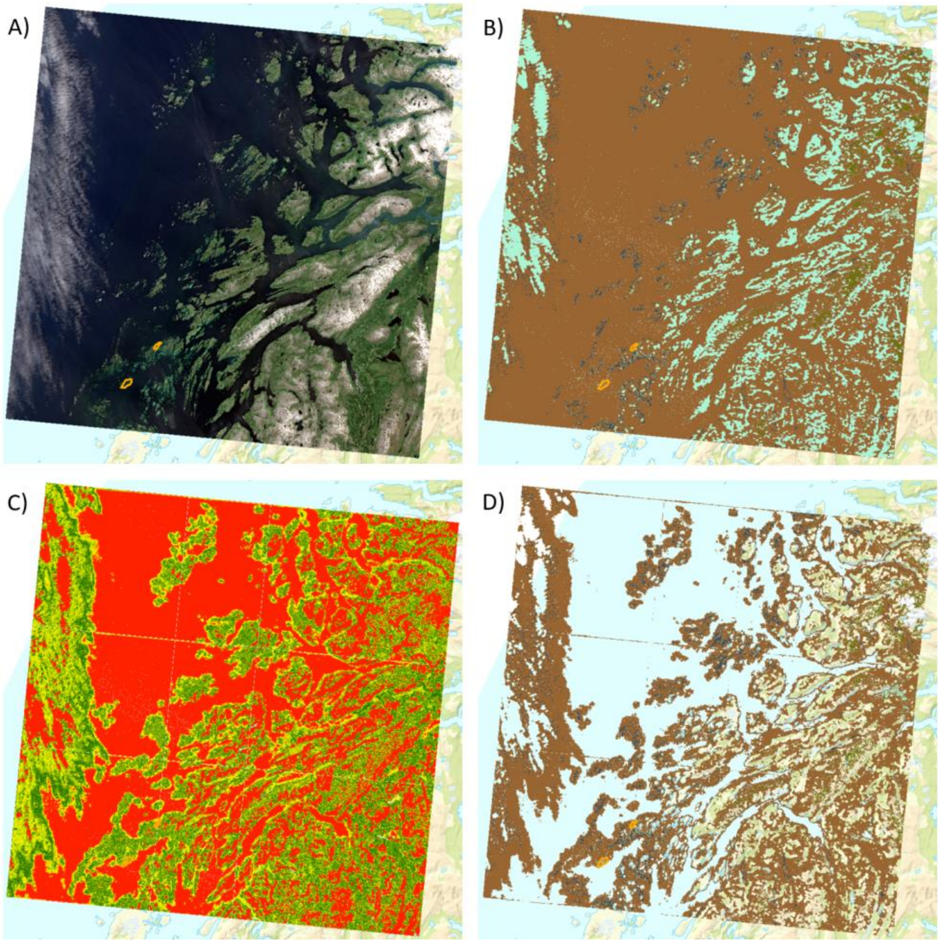


Figure 19. A) The Sentinel-2 satellite image (2022, resolution 10 m) used for upscaling, B) the resulting prediction map (also here, as with the VHR upscaling, kelp forest is largely overpredicted), C) the map of uncertainty, showing large areas with uncertain predictions (p < 0.05, in red) typically restricted to deeper areas, but also considerable areas with predictions above the 0.05 cutoff (yellow and green), and D) the resulting prediction map after masking out uncertain predictions.

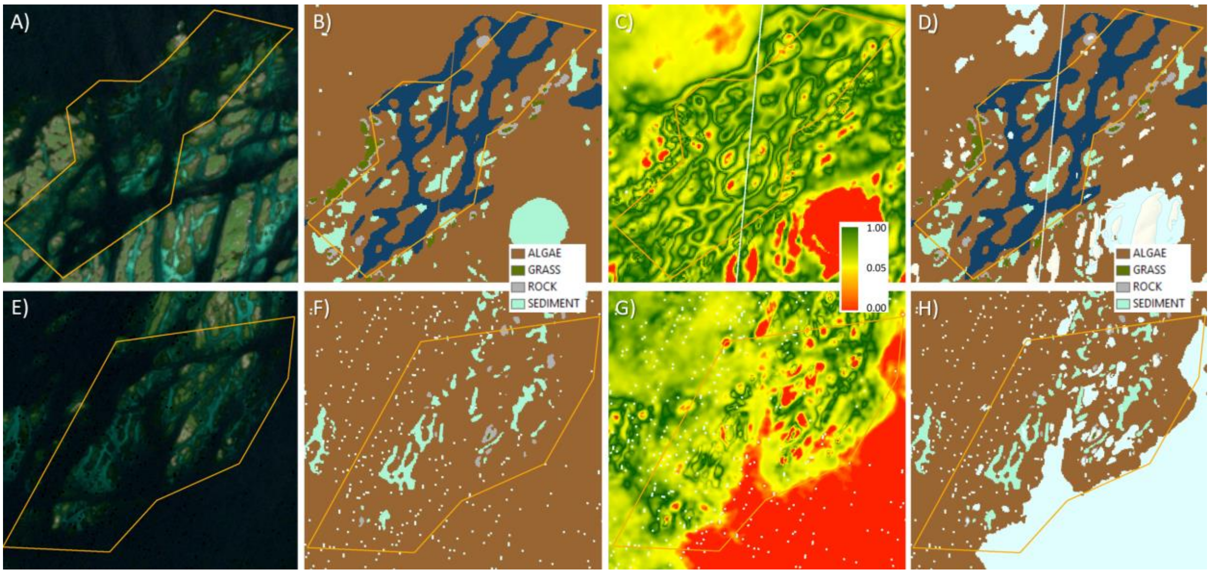


Figure 20. Images and maps in the upscaling process from Sentinel-2 (2022) satellite image (A, E), through classification maps (B, F), uncertainty maps (C, G), showing areas with uncertain predictions in red colors (p < 0.05), typically restricted to deeper areas, and predictions above the 0.05 cutoff in yellow and green, and the resulting corrected classification maps with the uncertain predictions masked out (D, H), shown for the Northern (upper) and Southern (lower) study areas, respectively.

4 Discussion

4.1 The KELPMAP study

In this study, we evaluated the results of the first Norwegian attempt to map kelp forest and associated submerged marine vegetation from flying drones equipped with high-resolution optical multispectral sensors (MSI), combined with machine learning (ML) image analysis for high-resolution seafloor habitat classification. Additionally, this work represents the initial steps toward assessing the use of satellite remote sensing for benthic habitat mapping by upscaling drone-generated data products to satellite-based habitat maps with the potential for national coverage.

We consider the current project successful as a pilot study and a comprehensive test of the potential for using flying drones for mapping of submerged kelp forests in complex coastal waters, i.e. with myriads of skerries and islets, steep gradients of water depth, fractionated and complex vegetation, and weather-exposed conditions, as typical for the Norwegian coast (Figure 21). The collected raw images were of good quality and provided a useful basis for production of georeferenced orthomosaics, accurate habitat annotations, and further quality assured habitat classification maps from ML-based analysis. These data, then again showed useful for training of new algorithms based on satellite images at two different spatial resolutions demonstrating the potential for using satellite remote sensing data for coastal habitat classification guided by high resolution drone data (Figure 22).

Figure 21. The Helgeland archipelago with thousands of skerries and islands and rich kelp forests underneath the water surface, from a few meters down to approximately 30 m depth.

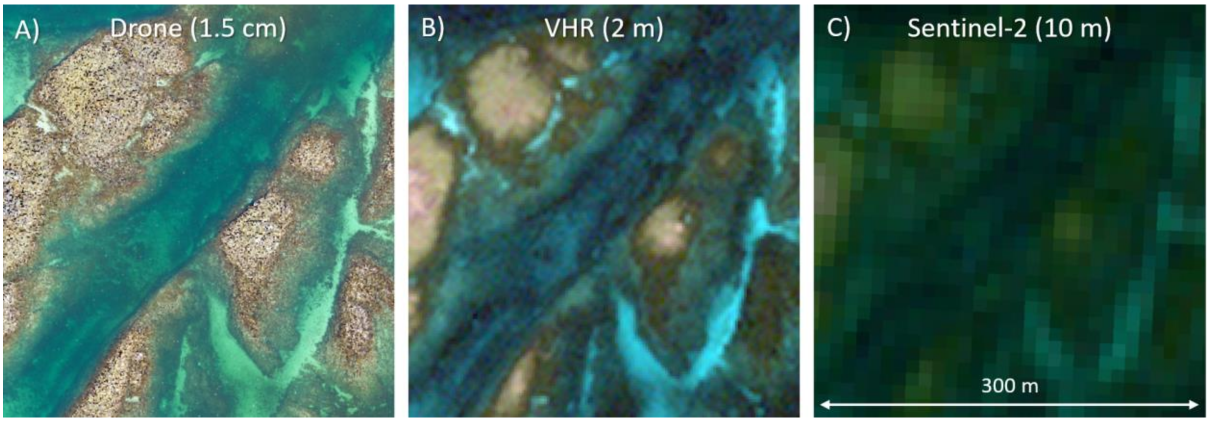


Figure 22. Comparing resolution and image quality from three different data sources A) drone multispectral image data (here illustrated with a RGB image, 1.5 cm pixel resolution), B) VHR satellite remote sensing data (2 m pixel resolution) (© CCME (2018), provided under COPERNICUS by the European Union and ESA, all rights reserved), and C) Sentinel-2 satellite data (2022, 10 m pixel resolution), from a close-up section of the southern study area at size 0.09 km² (9 ha).

4.2 Mapping kelp forests and benthic habitats from drones

Over the past few years, there has been an explosion in the technical development of flying drones and a following steep development of applications in research, mapping, and monitoring of aquatic habitats (Joyce et al. 2023). Flying drones combined with high-resolution optical sensors represents a novel technology and a strong complement to existing methodologies for shallow water monitoring. The approach has the potential to transition from labor-intensive, high-cost vessel or boat-based manual data collection and low-resolution remote sensing data sampling from satellite remote sensing platforms. Today, flying drones can provide high-resolution image data at centimeter-scales with spatial coverage at kilometer scale per sampling day, as we have demonstrated in this study.

Already, drones have been proven to be efficient tools for coastal habitat mapping and the study of various ecosystems and species. This includes mapping of seagrass meadows (Duffy et al. 2019), coastal wetlands (Doughty and Cavanaugh 2019), and assessing the abundance and community dynamics of seabirds (Edney et al. 2023) and marine mammals (Álvarez-González et al. 2023). Smaller commercial drones have also been used to investigate anthropogenic substances such as plastic pollution (Torsvik et al. 2020) and oil spills (Adade et al. 2021). While a few studies have focused on kelp forests, these primarily targeted kelp plants that reach the sea surface, allowing identification from the out-of-the-water canopy top (Cavanaugh et al. 2021).

It is worth noting that advancements in drone technology and its applications in ecosystem research and conservation management are currently progressing at an extremely rapid pace (e.g. Borja et al. 2024). Since the completion of the data collection in this study (in August 2022), the aerial capacity of drones, quality of sensors, and the performance of the applied ML algorithms have already improved. Nevertheless, we consider the results, conclusion and recommendations presented here still to represent the forefront of drone technology and its applications in coastal habitat mapping and monitoring. This is further reflected by the ongoing developments in the SeaBee Research Infrastructure, which are evolving in concert with this project.

4.2.1. Habitat classification

In this study, we classified benthic habitats and associated vegetation at three hierarchical levels, as explained in Section 2.3 and shown in Table 2. This classification design aimed to evaluate how detailed habitat classes could be mapped and categorized using ML and data from MSI sensors on drones. The class structure included 11 classes at level 1, 17 classes at level 2, and 39 classes at level 3. In the simplest model (level 1), all vegetation was combined into one single class, ALGAE, where brown, red and green algae were pooled together (except a few occurrences of eelgrass). The second level differentiated the algae class into three classes based on their photosynthetic pigments: BROWN, GREEN, and RED. The third level provided species-level classification, such as sugar kelp, tangle kelp, knotted wrack, and bladder wrack (see Table 2 for a detailed list of the classes at each level).

4.2.2. Classification model performance

All three hierarchical models performed well in classifying the marine vegetation and differentiating these from other habitats, as evaluated from the confusion matrices, which present the relative number of correctly predicted pixels when compared to ground truth observations (Figure 16). For broadly classifying and mapping kelp forests, the level 2 model performed the best. At level 2, brown algae (BROWN) were classified with 87% confidence and were clearly separated from SAND (99%), BOULDER (84%), TERR.VEG. (100%) and DEEP areas (94%). Maps of uncertainty supported these results, showing a low percentage of red-colored pixels at level 2 (Figure 13). This was also verified by the most accurate areal coverage predictions with 99% of the pixels classified as BROWN (Table 3) passing the masking threshold ($p > 0.05$) that indicated representation of brown algae species.

The class BROWN included all brown algae, not strictly constrained to kelp forests, and encompasses other seaweeds species such as knotted wrack (*Ascophyllum nodosum*, grisetang), Fucus species (e.g., bladder wrack (*Fucus vesiculosus*, blæretang), and the mixed algae group MACROMIX (containing a heterogeneous mix of different brown algae species).

4.2.3. Water visibility and impact of water depth

Water depth and optical properties significantly affect drone habitat mapping and classification performance, as for optical remote sensing in general (Rowan and Kalacska 2021). In this study, benthic vegetation and habitat classes were accurately identified down to a depth of approximately 10 meters in drone images, as well as in satellite images (see Section 4.3). This conclusion was valid for the prevailing conditions, and based on an evaluation of the image data, p-values, and confusion matrixes. However, the effectiveness of benthic habitat identification depends on water depth and water optical properties at each location and time.

In general, the optical properties, which dictate light attenuation and thus visibility and ultimately classification depth, are governed by four main factors: 1) chlorophyll concentration (mostly microalgae in the water), 2) colored dissolved organic matter concentrations (from terrestrial runoff and produced by micro and macroalgae), 3) inorganic particles concentrations (from terrestrial runoff and sediment resuspension), and 4) water itself. It is beyond the scope of this report to delve into each factor in detail, but it is worth mentioning that water optical properties are crucial for the performance of drone and satellite remote sensing. In general terms, remote sensing-based habitat mapping conditions are better in exposed areas far from land, as concentrations of chlorophyll, organic, and inorganic matter are often higher in fjords and sheltered regions than in more exposed coastal areas. Thus, the Helgeland coast is considered an ideal location for testing drone and satellite remote sensing for benthic habitat mapping.

Regarding kelp forest mapping, the visibility limitation means that areas mapped using drones and satellites only represent the upper part of the kelp forest, here at water depths from 0 to 10 meters. The extent of kelp forests growing below 10 meters varies by region, exposure, bathymetry, and bottom substrate, and can be calculated or modeled if such data is available (e.g., Bekkby et al. 2019). A bulk calculation based on the NIVA kelp distribution model (and the underlaying 630 field observations of kelp forests, Gundersen et al. 2021) revealed that 80 to 90 percent of the kelp forest biomass grows at depths of 10 meters or shallower and thus can be observed from drones and satellites. These numbers include two standard errors on the biomass estimate but might be overestimated as the model is limited to the upper 20-meter depth. In contrast, it is conservative as the percentage of hard bottom substrate is assumed to be evenly distributed with depth, though it is likely overrepresented in shallower areas (<10 m). In the Helgeland area, kelp forests have been observed at depths of 25 to 30 meters, but with biomass less than 5% of that at 3 to 5 meters depth, per meter square. Conservatively, it can be argued that the techniques presented here cover approximately 30% of the depth range for kelp forests, however a much larger part of the biomass can assumedly be detected using drone imagery.

In a management context, it is important to note that the lower growth limit of kelp forests (a parameter in the Norwegian management plans for the marine areas) cannot be quantified using current drone or satellite image sensors for most of Norway. However, future developments in, for instance, laser-based technology might overcome these limitations and could be implemented in the future.

This study does not include a dedicated evaluation of the effects of depth on classification performance, as the dataset and funding available were insufficient for a detailed quantitative analysis. High-resolution bathymetry data was not collected, and NIVA's 25x 25 m bathymetry model (e.g., Bekkby et al. 2019), was considered too coarse relative to the drone image resolution. However, a simple test correlating prediction uncertainty (p-values) at each ground truth point with water depth (collected at all ground truth points) showed no significant trend ($p < 0.05$) within a depth range of 0 to 10 meters. This suggests that, within this range, depth did not significantly affect predictability at any level (level 1 $R^2 < 0.01$; level 2 $R^2 = 0.02$; level 3 $R^2 = 0.02$; Figure 23). This correlation is based on a small subsample (n = 1 039) and may not represent the entire dataset, which includes millions of pixels. In theory, classification success is expected to decrease with increasing water depth due to reduced light penetration, increased turbidity, and difficulty in distinguishing habitat features at greater depths.

Overall, while challenges remain, particularly in deeper waters, this study demonstrates the potential for using drones and machine learning to enhance the mapping and monitoring of marine vegetation in complex coastal environments, down to at least 10 meters of water depth.

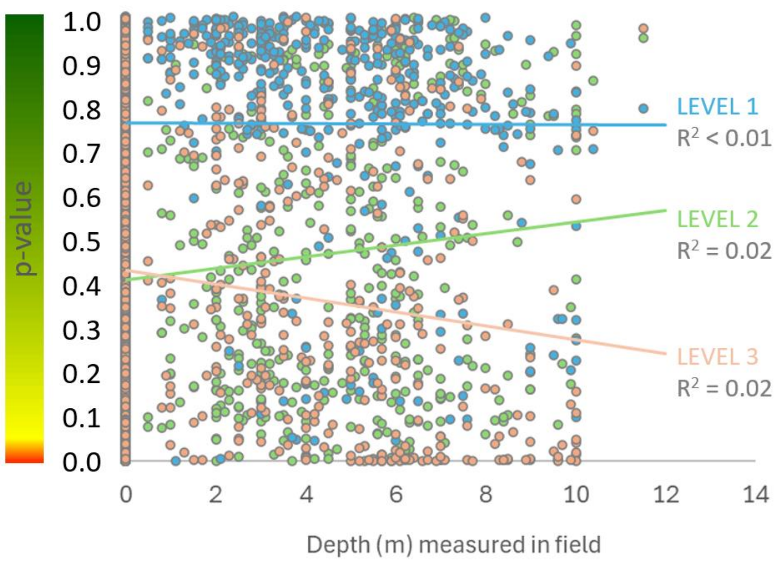


Figure 23. Correlating depth (meters) against the prediction uncertainty (p-values) at each ground truth observation point, where exact depth measures were available (n = 1 039). Colors depict the three analysis levels, and R^2 values showing the lack of correlation, indicating depth is not an important limiting factor within the depth range (10m) of this study.

4.2.4. Estimating kelp areal coverage and distribution

The total areal coverage of brown macroalgae (including both kelp and rockweed, referred to as "kelp" for simplicity in this section) in the study region varied depending on the analysis level and whether uncertain predictions were included or masked out. The total area of kelp was estimated to be between 0.39 and 2.44 km^2. The lower estimate comes from the masked level 3 model (summing all brown algae species), while the higher estimate is from the unmasked level 2 model (BROWN). This corresponds to a conservative kelp coverage of 0.39 km^2 (11% of the total study area), potentially reaching up to 2.44 km^2 (66%) if the uncertain predictions are accurate (Table 3). These estimates represent the part of the kelp forests growing shallower than approximately 10 meters only, which is the depth limit for the drone image data (see Section 4.2.3).

In more detail, kelp forest classification at level 1 was represented by the broad ALGAE category, which was predicted with 81% certainty based on the confusion matrix validation (Figure 16). This resulted in a conservative area coverage of 1.08 km^2 out of 3.68 km^2 after masking, but may also be as high as 2.39 km^2, represented by the uncorrected model.

At level 2, brown algae (BROWN) were predicted with 87% certainty (Figure 16), covering an estimated 2.41 km^2 after masking, compared to 2.44 km^2 before masking, leaving out only 1% of the classified pixels. The masked brown algae coverage closely matched the unmasked ALGAE coverage at level 1, supporting confidence in the highest estimate.

At Level 3, the dominant kelp species *Laminaria hyperborea* (LAMHY) was predicted with 70% certainty (Figure 16), covering an estimate 0.14 km^2. This represents 14% of the masked habitat coverage and only 4% of the entire study area (Table 3), which is less than the estimates for kelp forests at level 1 and 2. However, LAMHY is a single kelp species, unlike the larger groups of brown macroalgae that combine signals from similar algae species. Given that *L hyperborea*, *S. latissima*, Alaria spp., and other macroalgae have very similar colors, it is remarkable that the ML algorithm can distinguish LAMHY from the rest. In contrast, sugar kelp (*S. latissima*) was barely classified in the predictions, with a 0% prediction certainty in the confusion matrix. The physical and visual differences between the two species, along with annotation and ML processing issues discussed below, may partly explain this variation.

In conclusion, mapping brown algae as a group distinct from other habitats was successful, with an 87% prediction accuracy and 99% of the pixels classified as brown algae passing the confidence threshold. However, as expected, the specific classification of submerged kelp species (*Laminaria hyperborea* and *Saccharina latissima*) proved to be challenging. Validation from the confusion matrix showed a prediction accuracy of 70% and 0% for the two species, respectively. Furthermore, only 33% of the predicted pixels for *L. hyperborea* surpassed the cutoff threshold, and none for *S. latissima*.

In an ecological context, it is important to note that non-kelp brown algae generally provide many of the same ecosystem services as “true” kelp forests. Therefore, the ability to map this group as "brown algae” offers significant scientific and management advantages with wide-ranging applications.

4.2.5. Machine learning as a tool for habitat classification

The outcome of machine learning analyses is often complex, and there can be numerous reasons behind the model predictions. We consider the following points to be the main drivers for the lack of accurate predictions at different habitat classification levels:

1. **Insufficient annotation data:** The performance of machine learning models relies heavily on the availability of substantial input data. For habitat classes with limited annotation and training data, model performance is expected to be low. In this study, this issue is evident for the ANTHRO class (level1), TURF (level 2), and SACLA (level 3, Figure 11 and 16). The deep-water class (DEEP) presents a unique challenge, as it is inherently difficult to sample and annotate adequately within the current workflow. Orthomosaics, which form the foundation for both training and validation of the ML model, depend on the successful stitching of raw drone images. This process requires

clear and visually distinct features for accurate bottom recognition. Consequently, orthomosaics often fail to stitch areas of optically deep water, resulting in these areas being underrepresented or totally missing in the annotation and training data. In the present study, this issue is particularly pronounced in the southernmost study area.

2. **Incorrect annotation**: With increasing water depth, it becomes gradually more difficult for the human eye to distinguish between specific habitat types and optically deep waters. Consequently, it is likely that some data annotated as DEEP may in fact be kelp or macroalgae. During the manual annotation process, it was challenging to differentiate these from deep water, leading to misclassification of DEEP. This misclassification can result in underestimating the presence of brown algae or kelp species across all hierarchical levels, thereby lowering the p-values and affecting the confusion matrix statistics for the relevant pixels. This issue could also explain the small area of annotated polygons for sugar kelp (SACLA) despite numerous in situ observations (Figure 11).
3. **Uneven distribution of annotation data:** High prediction success requires an equal and well-represented distribution of annotation data for all considered classes. Achieving this necessitates substantial time and effort in the field to collect ground truth data and ensure the presence of species included in the classification system. In marine ecosystems, some species are of course naturally less frequent than others, which can easily introduce an imbalance in the available data to guide the image annotation process (Table 2). This imbalance likely contributes to the relatively high number of uncertain predictions at level 3.

Habitat predictions and classification models would likely improve with more training data from larger areas and additional ground truth data collected in the field to guide the annotation. Species and habitat types that are similar or less common will need more extensive sampling and annotation. With sufficient data, level 3 classification could reach the accuracy of level 2 shown here. Although this requires considerable effort from marine biologists, making it costly, developing generic models to recognize kelp forests and other marine habitats could eventually lead to robust algorithms for broader and more general coastal habitat mapping in Norway and internationally.

Note that the habitat classes developed by SeaBee, and used in this study, are designed to be representative of the entire Norwegian coast. Therefore, some species are poorly represented in this relatively restricted area, with implications particularly evident at level 3.

4.2.6. Precision in geopositioning

With increased spatial resolution and the need to combine multiple datasets of different sources for improved mapping and management, high accuracy on geopositioning of data products is essential. In this study, we used RTK and PPK techniques to georeference all image data collected by drones (see Section 2.2.2 for method details) and similar high-resolution GNSS labeling for ground truth data. This approach allowed us to stack RGB and MSI images during post-processing using GIS software, along with the ground truth data. This integration enabled accurate annotation guided by collected ground truth data and RGB images for subsequent use with MSI data and training of ML algorithms. However, errors or inaccurate georeferencing at any of the involved levels would introduce errors that could impact the quality of the annotation layers, the training data for ML and/or the output and validation data calculations.

4.2.7. Systematic and Repetitive Data Acquisition

The use of drones for benthic mapping facilitates systematic and repetitive data acquisition, creating possibilities for research and monitoring of shallow water habitats and tracking changes over time (Borja et al. 2024). This methodology provides an efficient tool to assess ecosystem changes, study natural dynamics, assess climate impacts, and monitor vegetation and ecosystem health in response to anthropogenic influences. In the present study, we combined drone data with similarly accurate ground truth data to supervise polygon-based annotation. This workflow resulted in multiple data layers that aligned precisely, facilitating the production of output data such as benthic habitat maps of biotic and abiotic environments, including submerged vegetation like kelp forests, smaller seaweeds, sand, and rock. This approach enabled the prediction of benthic habitats at three independent hierarchical levels, as shown in Table 2.

4.3 Mapping kelp forests and benthic habitats from satellite data

This study includes preliminary tests of using satellite remote sensing data to upscale drone products for mapping and classifying benthic habitats, with a particular focus on evaluating the potential for mapping kelp forests along the Norwegian coast.

There has long been a desire to apply satellite remote sensing data to mapping of shallow water habitats and submerged vegetation (e.g. Rowan & Kalacska 2021, Stæhr et al. 2024). However, this has been largely hampered by several core factors, including: 1) inadequate spatial resolution of satellite remote sensing data to resolve the complexity of coastal vegetation, 2) insufficient spectral resolution to distinguish between vegetation classes (algae, seagrass, etc.), and 3) optical distortion from neighboring pixels with terrestrial origin, which have much stronger reflectance signals. These challenges have been explored and mitigated through numerous studies and technological developments, such as increased spatial and spectral resolution, along with improvements in satellite calibration procedures and data post processing (Rowan & Kalacska 2021).

Recently, a coastal habitat model was published and tested for estimating Swedish shallow water vegetation using satellite remote sensing (Berglund et al. 2023). It successfully differentiated coastal vegetation from non-vegetative habitats; however, the model did not distinguish between different groups and classes of underwater vegetation, such as macroalgae and seagrass. Similar work has been published for Danish (Stæhr et al. 2024) and Norwegian (Rinde et al. 2021) shallow water benthic habitats.

In clear and less complex waters with large patches of uniform vegetation classes, satellite remote sensing data have been successfully applied to map distribution and areal cover of seagrass meadows (Gullström et al. 2006, Hossain et al. 2015), shallow water macroalgae (Duffy et al. 2019), and floating macroalgae (Qi et al. 2020). However, still there is no single approach that is capable of accurately measuring and distinguishing between all submerged vegetation and its relevant parameters (presence/absence, coverage, species distribution, and biomass).

Also, satellite remote sensing has been used to map exposed and surfaced kelp canopies by deriving traditional vegetation index information, such as the Normalized Difference Vegetation Index (NDVI) (Mora-Soto et al. 2020). However, these approaches are only reliable for non-submerged vegetation, as they typically involve reflectance data in the near-infrared (NIR) band. This NIR band is attenuated by just a few centimeters of water, thus ineffective for submerged vegetation (Johnsen et al. 2013, Rowan & Kalacska 2021).

4.3.1. Upscaling drone-based classification maps using satellite images

In this study, we tested an alternative approach that uses high-resolution (centimeter-scale) habitat classification maps from drone data to train ML algorithms for predicting costal habitats from satellite data. This approach offers several advantages over traditional methods: 1) It eliminates the limitations of traditional satellite remote sensing indexes, such as the use of NIR data, which is ineffective for submerged vegetation; 2) It allows the use of extensive training data, feeding millions or even billions of quality-assured and corrected (masked) pixel-based habitat classifications, enhancing algorithm robustness; and 3) It utilizes the full spectral resolution of remote sensing data, beyond the few bands typically used in satellite indexes (Dierssen et al. 2015, Skjelvareid et al. 2023). However, in this study, we only applied RGB satellite data, leaving the last point untested.

This approach allows for the upscaling of high-resolution data from limited areas to enhance large-scale mapping using lower spatial resolution data. We consider the current results as a preliminary proof of concept.

The unmasked upscaled maps of kelp forests from satellite data initially showed a strong dominance of algae (ALGAE) for both VHR (2 x 2 m resolution, Figure 17) and Sentinel-2 (10 x 10 m resolution, Figure 19) data. These results likely overpredicted the actual extent of kelp forests, which was supported by the low statistical confidence for most areas. Note that habitat classification was only performed at level 1 for satellite data, so the ALGAE class includes all types of algae, also 'true' kelp forest species.

After removing the uncertain predictions ($p < 0.05$), the resulting vegetation maps improved and aligned well with the expected distribution and coverage for the region. When compared to the drone-based habitat maps and studied in detailed, the visual similarity of these classification maps was highly evident (Figure 18 and Figure 20). Especially, the VHR-based prediction map (Figure 18D,H) showed remarkable similarities to the drone-based level 1 prediction (Figure 12 and Figure 14). Unfortunately, there was insufficient data to create confusion matrixes for the satellite-based habitat classifications, so the validation of these maps is based on visual inspection and comparison of the derived areal cover of the submerged vegetation, as described in Section 4.4.

4.4 Comparing drone and satellite-based habitat classifications

To evaluate the analysis output and compare it with other mapping products and models for the same area, we calculated the total estimated area for each habitat class in both study areas (Table 4, Figure 24). The first three columns of Table 4 show the habitat-specific predicted areas from the mapping products before masking, followed by the corrected area estimates after masking out uncertain predictions. Finally, these results are compared with the satellite-based model provided by NORCE (Haarpaintner and Davids 2020) and the statistical kelp model developed by NIVA (Frigstad et al. 2021).

There was a remarkably good agreement between the unmasked estimated areas of the drone-based model and the two satellite-based models for both the northern and southern areas across all habitat classes (Table 4 and Figure 24). This was particularly evident for the submerged vegetation class "ALGAE", where the estimated area ranged from 0.80 to 0.94 km^2 for the northern site and from 1.59 to 1.66 km^2 for the southern site. (Note that the sum of these areas is shown in Table 3). This indicated that the developed AI algorithm, trained on drone image data, was approximately equally effective in classifying habitats at centimeter-scale from drone data and at meter-scale from satellite data when all vegetation was pooled into one class (i.e. ALGAE) and predictions were not evaluated.

However, this interpretation changed considerably when examining the areal coverage of the data that passed the masking. Only about 30 to 50% of the drone data passed the quality control (masking), while only a minor fraction of the satellite-based data was masked out within the specific study area. This discrepancy reduces the confidence in the satellite-based areal coverage compared to the drone-based estimates.

When compared to the statistical distribution models from NORCE and NIVA, the remotely sensed area was largely within the range of the two models. The satellite products overshot the NORCE estimate by 2-3 times but were still below the NIVA model estimate for the northern site. A similar trend was observed for the southern site, with the satellite products overestimating the NORCE estimate by 6-7 times but matching the NIVA model estimate. It's important to note that the remote-sensing areas only cover the upper 10 meters of the macroalgae depth range, representing approximately 30% of the total forest area (see Section 4.2.3).

Considering this, the drone-based ALGAE estimate is more likely to be close to the 'true' areal coverage than the satellite-based estimate for both the northern site (23-71% of the NIVA and NORCE models, respectively) and the southern site (57-345% of the NIVA and NORCE models, respectively). The low area estimates by the NORCE model for the southern site seem to be largely due to a larger area classified as DEEP in this model.

In conclusion, while satellite remote sensing is applicable for mapping benthic habitats and classifying submerged vegetation, the satellite-based estimates appear to be less robust than the drone-based results. This is based on the semi-quantitative evaluation comparing areal kelp forest estimates and the lack of sufficient data for a true quantitative evaluation using confusion matrix statistics.

Considering the spatial resolution and precision of drone- and satellite-generated maps, each serves different managemental purposes. At the local or regional level, such as municipalities, high-resolution, diverse, and precise classification maps are highly useful for marine spatial planning (MSP) and for meeting obligations under the Nature Agreement (CBD). Whereas semi high-resolution satellite-based maps are suitable for national-level applications, such as broad-scale carbon accounting.

Table 4. Area- (km^2) and class-specific area estimates (km^2) based on the mapping products from the present study, labeled as Drone, VHR and S2, compared with their corrected versions, where pixels with high p-values have been masked out (labeled "corr"). The two rightmost columns show the same habitat classes predicted by a satellite-based model from NORCE (Haarpaintner and Davids 2020) and the areas predicted by NIVA's kelp model (Frigstad et al. 2021). See corresponding visualizations in Figure 24.

NORTH	Drone	VHR	S2	Drone_corr	VHR_corr	S2_corr	NORCE [1]	NIVA kelp model [2]
Resolution	9 cm	2 m	10 m	9 cm	2 m	10 m	10 m	25 m
ALGAE	0.80	0.85	0.94	0.25	0.80	0.91	0.34	1.06
SAND	0.20	0.20	0.14	0.07	0.19	0.14	0.00	
ROCK	0.09	0.07	0.05	0.02	0.07	0.05	0.04	
DEEP	0.56	0.55	0.53	0.56	0.54	0.52	1.21	
TERR.VEG.	0.05	0.03	0.04	0.02	0.03	0.04	0.04	
Sum	1.70	1.70	1.70	0.91	1.64	1.67	1.64	

SOUTH	Drone	VHR	S2	Drone_corr	VHR_corr	S2_corr	NORCE [1]	NIVA kelp model [2]
Resolution	9 cm	2 m	10 m	9 cm	2 m	10 m	10 m	25 m
ALGAE	1.59	1.62	1.66	0.83	1.60	1.47	0.24	1.45
SAND	0.29	0.29	0.25	0.12	0.29	0.25	0.00	
ROCK	0.09	0.07	0.05	0.02	0.07	0.03	0.03	
DEEP	0.00	0.00	0.00	0.00	0.00	0.00	1.71	
TERR.VEG.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Sum	1.97	1.98	1.96	0.96	1.95	1.74	1.98	

[1] For comparison, the habitat classes of the NORCE model were converted to the KELPMAP habitat classes, assuming LAND = TERR.VEG., MUD and SAND = SEDIMENT, SEAWEED = ALGAE, WATER = DEEP, and ROCK = ROCK. Some very small areas (0.046 km^2) classified as SHALLOW were removed from the NORCE map, as it didn't have any good match in the KELPMAP classes.

[2] NIVAs kelp model is originally two separate models, one for *Saccharina latissima* and one for *Laminaria hyperborea*. These two are combined in this estimate.

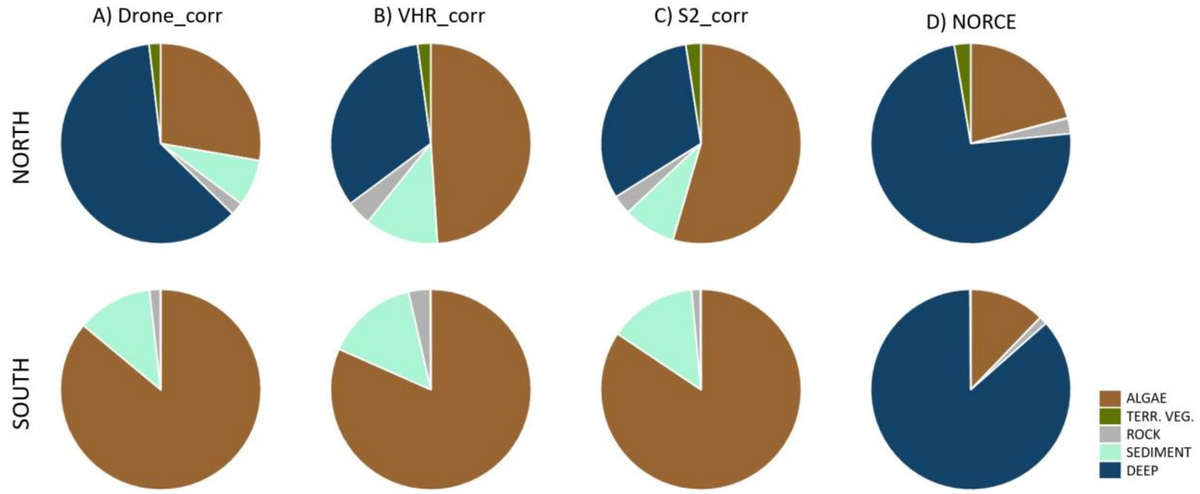


Figure 24. The relative differences in the amount of the different habitat classes of level 1, shown for the classification maps based on A) the drone algorithm (9 cm), B) the upscaled VHR satellite image (2 m), C) the upscaled Sentinel-2 image (10 m), and D) the satellite-based prediction model by NORCE (Haarpaintner and Davids 2020). The calculations are presented for the northern (upper) and southern (lower) study areas. Only corrected outputs (masking out uncertain predictions) are used in this comparison, except for the NORCE model, for which only an uncorrected map was available. NIVA's kelp model is not shown here, as it does not provide estimates for other habitat types.

4.5 Traditional versus emerging technologies for mapping coastal habitats

Accurate, reproducible, and cost-effective mapping of kelp forests and other shallow water submerged marine vegetation is essential for sustainable management and ecosystem preservation. The marine coastal zone is often overlooked and under-sampled in both Norwegian and international monitoring and management programs. Traditional marine monitoring methods include large research vessels, small boats and diving operations, and to a limited degree satellite remote sensing. Larger research vessels are restricted to depths greater than 10 meters and are expensive to operate. Smaller boats are suitable for shallow water mapping but offer limited spatial and temporal coverage and require significant time investment. Similarly, SCUBA diving is costly, time-intensive, and often lacks spatial precision (Murphy and Jenkins 2010, Féral and Norro 2023). While satellite remote sensing provides extensive spatial coverage, it typically has low spatial resolution, with RGB and spectral data acquisition at resolutions of tens to hundreds of meters (Lønborg et al. 2022).

Emerging technologies, such as flying drones and high-resolution satellite remote sensing, are revolutionizing the field by enhancing both spatial and temporal resolution. These technologies capture high-quality data, enable reproducible sampling designs, and offer cost-efficient monitoring solutions.

Flying drones provide on-demand access to regions of interest, delivering the requested data precisely when and where needed, unlike satellite imaging, which relies on predefined overpass times and clear weather conditions. These constrains often limit data collection in the Norwegian coastal region. Additionally, drones operate using preprogrammed flight trajectories, carefully planned to ensure high coverage and necessary image resolution, resulting in reproducible data. Overall, this methodology offers a cost-effective solution for collecting high-quality spatial and temporal data in coastal regions.

4.6 Using drones for mapping according to the NiN system

This text is written in Norwegian, as it contains numerous Norwegian terms, that would become imprecise and cumbersome if translated into English.

I SeaBee har vi identifisert det vi kaller *habitatklasser* (Table 2) i henhold til hva vi mener er mulig å identifisere ved bruk av droner. Dette kan være arter, som for eksempel sagtang, sukkertare og ålegras, eller samlebetegnelser for flere ulike arter, som rugl og tangvoll, eller klasser som beskriver en helsetilstand, f.eks. lurv (denne er under utarbeidelse i SeaBee, og vil muligens registreres som et attributt knyttet til andre klasser i fremtidige versjoner av klassesystemet). Det kan også være substrattyper, som identifiseres ut fra sin kornstørrelse (grovhet), nemlig sandbunn, grus, blokk, og nakent fjell. Se siste versjon av liste over SeaBee's habitatklasser på GitHub.

Systemet Natur i Norge (NiN 3.0) er overordnet beskrevet i kapittel 1.4.2 (se også Halvorsen et al. 2024). For mange av SeaBee's habitatklasser er det ikke en en-til-en sammenheng med NiNs naturtyper. Nedenfor gjør vi en vurdering av hvilke elementer som lar seg kartlegge i henhold til NiN-systemet. Siden oppdraget dreiet seg om tareskog, har vi først og fremst fokusert på marin vegetasjon, som makroalger (tang og tare) og ålegras. Imidlertid diskuterer vi også en del arter og naturtyper som ikke nødvendigvis er identifisert og kartlagt i KELPMAP, men som vi har jobbet med i andre SeaBee-aktiviteter.

Grunnlaget for habitatklassifisering ved bruk av droner er unike spektrale signaler og strukturelle mønstre hos arter og objekter på sjøbunnen. Derfor, så sant en naturtype gir et slikt unikt signal, og det ikke er for dypt, skal det i prinsippet være mulig å kartlegge naturtypen. Og siden NiNs grunntyper ofte er definert av dominerende og karakteristiske arter, og derfor ofte (men ikke alltid) fremstår som mer homogene enn hovedtypene, vil det være mer aktuelt å kartlegge på grunntypenivå enn etter en grovere inndeling (unntak er f.eks. littoralbassenger og trålt tareskog, som er hovedtyper i NiN). Det sagt, kan man selvfølgelig lage akkumulerte kart over hovedtyper, basert på kartlagte grunntyper.

4.6.1. NiN hovedtypegrupper og hovedtyper

I NiN-systemet er natursystemet det mest naturlige nivået å kartlegge på, også med drone. Her har vi et hierarki fra hovedtypegruppe til hovedtype og grunntype. For habitatkartlegging i kystsonen ved bruk av drone er det først og fremst hovedtypegruppen *Marine bunnsystemer* (NA-M) som er aktuell. Her finner vi 23 hovedtyper, men ikke alle disse kan ses fra lufta. Siden dronens sensor er avhengig av godt lys helt ned på havbunnen, er det kun den øvre delen av de eufotiske hovedtypene som i utgangspunktet kan kartlegges med flyvende droner. Disse er i hovedsak dekket av *Fast saltvanns-fjærebeltebunn* (MA01), *Eufotisk fast saltvannsbunn* (MA02), *Fjærebelte-sedimentbunn* (MA04), *Eufotisk saltvanns-sedimentbunn* (MA05), *Marin helofyttsump* (MB01), *Saltvanns-undervannseng* (MB02), *Fast brakkvanns-fjærebeltebunn* (MC01), *Fast brakkvannsbunn* (MC02), *Littoralbassengbunn* (MC03), *Brakkvanns-sedimentbunn* (MC04) og *Brakkvanns-undervannseng* (MF01). Men på langt nær alle disse er forsøkt kartlagt med droner, og det vil stort sett kun være de grunneste delene av de eufotiske bunnene som kan sees med drone. De naturtypene som defineres av sedimenter av ulike typer, som sand og stein med ulike kornstørrelser, ligger primært innen hovedtypene fjærebelte-sedimentbunn (MA04), Eufotisk saltvanns-sedimentbunn (MA05), og Brakkvanns-sedimentbunn (MC04) (Table 5). Disse typene er definert av blant annet finmaterialinnhold (variabelen FI) og dominerende kornstørrelse (variabelen DK). I NiN er DK delt inn i leire, silt, sand, grus og stein.

Rinde et al. (2021) gjennomførte en studie der satellitt-genererte habitatkart ble brukt til å forsøke å identifisere visse naturtyper i fjæresonen og den grunne delen av sjøsonen/sublittoralen, med en romlig oppløsning på 10 meter. Dette arbeidet ble utført i henhold til NiN versjon 2.0. Når kartlegging utføres på en målestokk som er grovere enn det som er egnet for å kartlegge NiNs grunntyper, har man utviklet målestokktilpassede kartleggingsenheter. Disse enhetene består av sammenslåtte grunntyper, der sammenslåingen innenfor hovedtypen gir større og grovere typer etter hvert som målestokken øker (Bryn et al. 2023). Den mest relevante målestokken bestemmes av oppløsningen på satellittbildene. Bekkby et al. 2023) viser sammenhengen mellom målestokk og oppløsning. For både VHR og Sentinel-2 er målestokken 1:5000 den mest relevante.

4.6.2. NiN grunntyper

To av hovedtypene det er knyttet mye kartleggingsaktiviteter til og der vi finner tang og tare er hardbunn i fjæresonen (MA-01) og sonen like nedenfor (MA-02). Her finnes mange grunntyper som i prinsippet burde kunne kartlegges, det vil si identifiseres på naturtypenivå og areal-avgrenses, ved bruk av drone. I KELPMAP-prosjektet, og SeaBee generelt, har vi foreløpig hatt suksess med å kartlegge følgende naturtyper på grunntypenivå innen disse to hovedtypene: *Grisetangbunn* (MA01-01), *Blæretangbunn* (MA01-04), *Spiraltangbunn* (MA01-05), *Sagtang-saltvannsbunn* (MA02-02), *Sukkertarebunn* (MA02-05) og *Fingertarebunn* (MA02-03), som vanskelig lar seg skille fra *Stortarebunn* (MA02-06). Merk at sukkertarebunn og stortarebunn ofte strekker seg dypere enn hva som er mulig å identifisere med drone (se 4.2.3). I tillegg er *Saltvanns-undervannseng* (MB02), altså ålegraseng, spesielt godt egnet for dronekartlegging, da den lett skiller seg ut på grunn av sitt spektrale signal. Se Appendix B for en oversikt over naturtyper som ifølge våre erfaringer egner seg for dronekartlegging.

«Tidevannseng og -sump» er ikke en naturtype i henhold til NiN-systemet, men en sammenslåing av ulike typer fra både marine og terrestriske miljøer, definert av Miljødirektoratet, og er ment å fange opp den mye brukte internasjonale betegnelsen «saltmarshes» (Borgersen et al. 2020). Tidevannseng og -sump kan muligens kartlegges ved bruk av droner, men det finnes foreløpig ingen studier på dette. Disse typene strekker seg fra øvre sublittoral til øvre hydrolitoral. NiN-typer som faller innenfor begrepet saltmarshes er oppgitt i Borgersen et al. (2020). Merk at disse er basert på NiN 2.0, ikke NiN 3.0. Borgersen et al. (2020) vurderte hele hovedtypen *Helofytt-saltvannssump* (M8) til å falle innenfor saltmarsh-begrepet, i tillegg til grunntypene *Saltanrikningsmark på silt og leire i geolitoral* (T11-3), *Nedre strandeng* (T12-1), *Midtre strandeng* (T12-2), *Øvre strandeng* (T12-3) og *Nedre Semi-naturlig Strandeng* (T33-C-1). I NiN 3.0 er *Marin helofyttsump* (MB01) delt inn etter marin salinitet og tørrleggingsvarighet (soneringen fra sublittoral og opp).

4.6.3. Artskoder i NiN

Mange av observasjonene av natur man kan gjøre i felt oppfyller ikke kriteriene til å være naturtyper. NiN har derfor et veldig fleksibelt variabelsystem. Variablene i dette systemet kan også kodes ved hjelp av NiN-systemet. Dekningsgrad av samfunn og dekningsgrad av arter kodes forskjellig. F. eks. så kan dekningsgrad av tangsamfunn registreres med tetthet SA-ET-Samfunn-P9a-# (samfunn kan f.eks. være «Fucales»), der # erstattes med tall for dekningsgrad i henhold til NiN. Hvis det er enkeltarter som registreres og ikke samfunn, er kode på formen RA-MB-Art-P9a-#. Dersom kun tilstedeværelse eller fravær av arten registreres, blir koden RA-MB-Species-B-#, der 1 = tilstedeværelse og 0 = fravær. Kommentarfeltet kan benyttes til å skrive inn kommentarer knyttet til dekningsgrad hvis det er nødvendig. Fremmede arter, som japansk drivtang, kan kodes på tilsvarende måte, for eksempel som relativ fremmedartsdekning, det vil si AD-FD-Species-P9a-#. Andre eksempler er hvis man vil beskrive arter som opptrer innenfor naturtyper, for eksempel hvis man vil beskrive mengden arter, fremmede eller andre, innenfor en tareskog. Eller hvis man vil identifisere blåskjell på andre steder enn i de grunntypene der blåskjell er definert. Arter som med sannsynlighet kan kartlegges ved bruk av droner i henhold til artskodene i NiN er gitt i Table 6. I praktisk feltarbeid vil det være naturlig å jobbe med tall, som antall, % dekning eller en semi-kvantitativ skala. Men det er viktig at man kjenner til inndelingen i NiN, slik at tallene for f.eks. % dekningsgrad som benyttes i felt, i etterkant kan konverteres til NiN-koder, slik at dataene fra kartlegging av forvaltningsrelevante naturtyper også kan gå inn i baser for NiN-kartlegging.

For å kunne klassifisere etter artskodene i NiN, kreves en viss kjennskap til kontekst. For eksempel, om arten er funnet/predikert i fjæresonen, i den lyse sublittoralen (eufotisk sone) eller den mørke sublittoralen, og om arten er på hard- eller bløtbunn. I praksis vil dette være informasjon som legges til etter at man har et analyseprodukt (prediksjonskart), f.eks. ved bruk av en overlay analyse i GIS.

4.6.4. Annet

Nytt i NiN versjon 3.0 er taretrålingsbunn (NA-MJ01) som en egen hovedtype, og noe som bør kunne identifiseres ved droner, i hvert fall de grunneste deler av trålegatene. Det pågående MASSIMAL-prosjektet (NFR #301317, 2020-2024) ser blant annet på hvordan droner kan brukes til å kartlegge vegetasjon i trålegater etter tarehøsting, og har sett at trålgatene er godt synlige fra droner.

Littoralbassenger (NA-MC03) er også en egen hovedtype etter NiN 3.0, som burde egne seg veldig godt for kartlegging med drone, siden de befinner seg ovenfor lavvannsbeltet, og med særtrekk som burde være lett identifiserbare med drone.

Habitatklassen *Lurv* er verken en naturtype eller kun én art, men en sammensatt gruppe av trådformede alger, der mange ulike arter kan forekomme (Rinde et al. 2024). Allikevel har den ofte et lett gjenkjennelig signal, på grunn av sitt «lurvete», teppedannende, uttrykk. Lurv har stort forvaltningsmessig fokus, og dermed viktig å finne metoder for å kartlegge og overvåke. I NiN kan lurv kodes som SA-ET-*Samfunn*-P9a-# (der samfunn kan være «lurv»), der # erstattes med tall for dekningsgrad i h.h.t. NiN. Hvis kun tilstedeværelse av lurv skal registreres kan dette kodes SA-ET-*Samfunn*-B-1.

Table 5. Naturtyper i NiN-3.0 (se Artsdatabanken) der vi gjennom KELPMAP og andre SeaBee-aktiviteter har erfaring med at det er mulig (eller sannsynlig) å kartlegge ved bruk av droner. Alle er under hovedtypegruppe Marine bunnsystemer.

Hovedtyper	Grunntype
Fast saltvanns-fjærebunn (MA01)	Grisetangbunn (MA01-01)
	Blæretangbunn (MA01-04)
	Spiraltangbunn (MA01-05) *
	Sauetang-blåskjell-rurbunn (MA01-06) *
Eufotisk fast saltvannsbunn (MA02)	Sagtang-saltvannsbunn (MA02-02)
	Sukkertarebunn (MA02-05)
	Stortarebunn (MA02-03) / Fingertarebunn (MA02-06)
	Butarebunn (MA02-07) *
Fjærebelte-sedimentbunn (MA04) / Eufotisk saltvanns-sedimentbunn (MA05)	Ruglbunn i tidevannssonen (MA04-11) / Sublittoral ruglbunn (MA05-13)
Saltvanns-undervannseng (MB02)	«Ålegras» (MB02-01 / MB02-02) **
Ulike klasser av dominerende kornstørrelse, kan ligge under ulike hovedtyper, inkludert både fjæresonen og sublittoralen, i salt- eller brakkvann***	«Nakent fjell» (DK-H, dvs. kjempeblokk og fast fjell)
	«Blokk» (DK-FG, dvs. stor og liten blokk)
	«Stein» (DK-E)
	«Grus» (DK-D)
	«Sandbunn» (DK-C)

* På grunn av manglende feltdata (ground truth data) har vi ikke erfaring med å kartlegge denne naturtypen ved bruk av droner, men man tror dette skal være mulig.

** Dette er ålegras i enten tidevannssonen (MB02-01) eller i sublittoralen (det vil si under tidevannssonen, MB02-02 – denne er vanligst).

*** Kan beskrives ved bruk av variabelen «Dominerende kornstørrelse», DK.

Table 6. Arter som ikke er representert av en naturtype, men som muligens kan kartlegges ved bruk av droner med bruk av NiN variabelsystem for tilstedeværelse av arten. NiN åpner også for å kode artene ut fra deres dekningsgrad, da ved hjelp av en annen type NiN-koding.

Species (latin)	Norsk navn	NiN artskode
Chorda filum	Martaum	RA-MB-CHfi-B-1
Halidrys siliquosa	Skulpetang	RA-MB-HAsi-B-1
Desmarestia aculeata	Kjerringhår	RA-MB-DEac-B-1
Cladophora rupestris	Grønndusk	RA-MB-CLru-B-1
Ulva spp. (e.g. *U. intestinalis*)	Tarmgrønske	RA-MB-Ulva-B-1
Porphyra spp. (e.g. *P. umbilicalis*)	Fjærehinne	RA-MB-Porphyra-B-1
Chondrus crispus	Krusflik	RA-MB-CHcr-B-1
Furcellaria lumbricalis	Svartkluft	RA-MB-FUlu-B-1
Palmaria palmata	Søl	RA-MB-PApa-B-1
Vertebrata lanosa	Grisetangdokke	RA-MB-VEla-B-1
Echinus esculentus	Rød kråkebolle	RA-MB-ECes-B-1
Strongylocentrotus droebachiensis	Grønn kråkebolle	RA-MB-GRes-B-1
Gracilechinus acutus	Langpiggsjøpiggsvin	RA-MB-ECes-B-1
Ophiocomina nigra	Svart slangestjerne	RA-MB-OPni-B-1
Asterias rubens	Vanlig korstroll	RA-MB-ASru-B-1
Crassostrea gigas	Stillehavsøsters	RA-MB-CRgi-B-1

4.7 Logistics, weather, and environmental conditions

Overall, the fieldwork and data collection were successful. The two mapped areas, 2.0 and 3.5 km^2 respectively, were larger than previously mapped by SeaBee/NIVA in a single campaign, adding some logistical and practical challenges.

Both areas were imaged from an altitude of 120 meters, the maximum permitted for open category operations, maximizing coverage but compromising pixel resolution, since there is a trade-off between the size of the area you want to cover and the resolution of the images. We believe an acceptable balance was reached between these considerations in this study.

Weather is always a crucial factor for fieldwork, especially for drone data collection. The Norwegian coast, and the Helgeland coast in particular, is exposed to rough weather, which was carefully considered during planning. Retrospective weather analysis for August 2022 showed 15 windows of opportunity with suitable flying conditions (elaborated in the KELPMAP field report by Gundersen et al. 2022). In this statistic, there were periods of up to 6 days where flying was not feasible, highlighting the importance of flexible planning when aiming for high-quality drone data products.

In practice, ideal fly time is a rare occasion. Effective mission planning balances favorable and challenging conditions towards the mission aim and image quality issues. Gundersen et al. (2022) discuss specific weather conditions during this fieldwork in detail, while Kvile et al. (submitted) provide a more general overview. In short, key factors impacting drone data quality include wind, waves, water transparency, cloud cover, precipitation, fog, and air temperature, all being partly unpredictable. Predictable conditions include sun angle, sun intensity, and tide.

This dataset was collected in a rather exposed location, but the water surface was partly sheltered by numerous small islands. During the field campaign, wind speeds were consistently below 8 m/s, wave heights were under 40 cm, and water transparency was high (Secchi depth > 10 m). Overall, winds, waves, and water transparency did not hamper image analysis or habitat classification down to approximately 10 meters. It is worth noting that we did not quantitatively assess water transparency or quality in this study and recommend including measurements of optical water quality in future studies.

The two areas selected by the Norwegian Environment Agency showed to be well fit for the study, hosting kelp forests and various macroalgae species. However, the southern area, with few islands and reefs and large deep-water areas (>10 meters), was less ideal, partly hampering the generation of orthomosaics and accurate georeferencing. Consequently, unnecessary fly-time was spent over areas that yielded little information, and only 55% of the southern area could be stitched for orthomosaics and further data analysis due to the dominance of optically deep water. We recommend close collaboration with drone pilots and marine biologists for future studies to optimize fieldwork.

5 Conclusion

5.1 Summary of the most important results and findings

1. Drones equipped with high-quality MSI sensors, combined with state-of-the-art machine learning techniques can **effectively map coastal marine vegetation and distinguish it from abiotic features** such as sand, rock, and terrestrial vegetation.
2. The method can also **differentiate brown algae from seagrass and probably distinguish brown algae from red and green algae** based on their spectral signature. However, the latter distinction may be limited if ground truth data and image annotation data are sparse.
3. To some extent, drone-based habitat classification can also **differentiate between several species of kelp and other brown algae**, though more training data is needed to map brown algae accurately at the species level. Nevertheless, this result represents a significant advancement over previous studies, which were limited to less diverge categorization of submerged aquatic vegetation, or only mapping intertidal or floating/surface occurrences.
4. In relatively clear waters, such as those along the Helgeland coast, drones equipped with MSI sensors can **map the seafloor at least down to 10 m depths**. The precision of mapping did not appear to reduce with depth within the range of 0-10 meters, suggesting that it may be possible to map benthic vegetation at even greater depths. However, the data available for this analysis was limited and more ground truth data and higher resolution bathymetry data is necessary to confirm this.
5. **More ground truth data is essential for both the training and validation of habitat maps**. In particular, the habitat classes DEEP, ANTHRO and TURF were wrongly classified in our study, and would benefit from an improved dataset. Ideally, ground truth data should be specifically collected to ensure that all relevant habitats and species are represented in the dataset and at a spatial precision that matches the predictive maps, i.e., at a centimeter scale. Unfortunately, this requirement excludes most existing data from earlier monitoring and mapping projects for this purpose.
6. By using modern sensors and flying at an altitude of 120 meters, orthomosaics with a **spatial resolution (GSD) of up to 9 cm for MSI and 5 cm for RGB images** were created, which was reflected in the resulting habitat classification maps. Resolution can be further improved by lowering the flight altitude. Sensor technology has advanced significantly since the data collection, with current SeaBee drones now equipped with sensors of 48 and 62 MP, compared to the 24 MP sensors used in 2022.
7. **Some brown algae are still difficult to identify and differentiate from each other**, most likely due to their similar spectral signals. Misclassification between kelp species (such as *L. hyperborea* and *A. esculenta*) is not critical for most purposes, as they provide similar functions and services in coastal ecosystems.
8. The identification of **turf algae is not particularly elaborated in this study**, however, with more targeted data, turf algae coverage should be possible to identify, including its seasonal and annual dynamics. A task that the SeaBee Infrastructure is currently working on.
9. **It is possible to upscale habitat maps from drones to much larger areas using satellite images**. However, the precision of the resulting large-scale maps is inherently limited by the spatial resolution of the satellite images. For open-source data, the best resolution is 10 x 10 meters (Sentinel-2). Commercial data can offer resolutions of 2 x 2 meters (Pleiades VHS) or even higher in some cases. This allows for **mapping of kelp and other macroalgae at a generalized level using satellite images**. However, because macroalgae often occur in a heterogeneous patchwork of different species and habitat types, **species-level mapping based on satellite images is rarely feasible**.
10. Our study suggests that **using drone-based training data for satellite upscaling enhances accuracy** of the resulting habitat map compared to those from other studies, derived solely from satellite

images. This conclusion is merely based on subjective comparisons of classification maps from different studies, and not quantitatively evaluated.

11. To validate our upscaled products (and those of other studies), a comprehensive ground truth dataset covering large regions and encompassing as many habitat classes as possible is necessary. **Existing data from other mapping and monitoring projects could potentially be utilized for validation of satellite-based classification maps**, as the spatial precision required is significantly lower than that needed to validate drone-based products.
12. Considering the spatial resolution and precision of **drone- and satellite-generated maps, each serves different managemental purposes**. High-resolution drone maps are ideal for local and regional marine spatial planning (MSP) and fulfilling Nature Agreement (CBD) obligations, whereas semi high-resolution satellite maps are suitable for national-level applications, such as broad-scale carbon accounting.
13. **Several NiN types can be mapped with drones, defined as separate basic types (grunntyper)**, particularly those that are defined by macroalgae (including seaweed beds and kelp forests), eelgrass and seabed of different types of soft and hard substrates.

5.2 Future recommendations

5.2.1. One broad algorithm for the entire coast of Norway

The datasets and algorithms created in the KELPMAP project for the Helgeland region should be integrated with data from other projects, as for instance other SeaBee and drone data collection projects, to create a **comprehensive and generic habitat classification model, suitable for all of Norway's coastal zone**. SeaBee holds drone data and associated ground truth data from various regions along the coast, including Finnmark, the Møre coast, and the Oslofjord. By combining these datasets, all regions will benefit from a larger dataset for training and validation, including species that are less frequently found in some locations. Once a robust algorithm is developed and its robustness is tested across regions, new high-resolution classification maps can be generated on drone data from new areas without the need for dedicated training data, which is often the costliest part of a mapping campaign.

5.2.1. Synergies between surveys and monitoring of different purposes

One should encourage better coordination and linkage between drone data collected for different purposes. For instance, within the SeaBee project a large number of drone images are captured for **seabird counting** on islands and skerries in several regions along the coast. These images can be explored for improving benthic habitat mapping as well. Additionally, **marine and terrestrial mapping** efforts should be similarly coordinated. With thorough planning, drone images can be collected for multi-purpose usage. For example, including more terrestrial zones in the drone images for marine mapping, to include **helophytes** and other saltmarshes ("tidevannseng og -sump"), or incorporating more marine areas around localities surveyed for bird colonies. This approach is already in progress for being integrated in this year's (2024) bird mapping census organized by the Norwegian Institute for Nature Research (NINA), where SeaBee's long-range drones (MYGIN-2 PRO) are deployed, covering extensive areas of coastal zone areas.

5.2.2. Upscaling to large-scale mapping of coastal habitats using Sentinel-2

The current study demonstrates the potential for scaling up drone-based habitat maps using satellite remote sensing data, such as Sentinel-2. Further development of drone-based habitat classification using AI towards robust prediction maps will enable the use of local and regional maps to guide an upscaling to national levels, achieved by training AI models on drone-based habitat predictions.

However, one must be careful when training one AI model with the output of another. Carefully collected ground truth data from fieldwork is essential for evaluating and assessing the quality of upscaled prediction maps, which was not possible in this study due to lack of data. To some extent, existing data from Norwegian national mapping programs can be used for validation.

Currently, only a few selected species and habitats have been prioritized for full-scale mapping in the national mapping programs. Among these are **tangle kelp** and a few other important nature types for commercial exploitation (Bekkby et al. 2013). These spatial distribution models are unsurpassed in terms of coverage and resolution (25 x 25 m on a national scale) and have been useful for national estimates of areal coverage and **carbon storage** of tangle kelp and sugar kelp (Frigstad et al. 2021, Kvile et al. 2022, Hancke et al. 2022). These models are applicable to depths down to 30 meters. However, compared to the upscaled multi-class classification maps presented in this study, these models have more limited spatial resolution and is applicable only for a few species.

The amount and distribution of **rockweed** ("tang") in Norway represent a significant knowledge gap (Frigstad et al. 2021, Hancke et al. 2022, Krause-Jensen et al. 2022). An upscaled model for the intertidal zone would be an important advancement in quantifying the distribution, biomass, and carbon content of this overlooked species group in Norway.

The national distribution of **saltmarshes** ("tidevannseng og -sump") in Norway is also greatly understudied (Borgersen et al. 2020), as it falls between terrestrial and marine monitoring. This habitat type would significantly benefit from a dedicated mapping effort. This heterogeneous group of species is not yet included in SeaBee's list of habitat classes, but this is currently a work in progress. In particular, **helophytes** – a species group characterized by tall, erect stems and uniform appearance – should, in theory, be relatively easy to recognize and classify from drone imagery.

Another application of an upscaled national habitat map could be to create a rough map **distinguishing soft sediments from hard substrates**. This has long been the one critical missing variable in the spatial distribution models of kelp and seagrass. Such a model would greatly improve national estimates of carbon storage, both in living stocks of kelp (hard substrates) and in the long-term storage in coastal sediments.

5.2.1. Advancing the use of drones in national and international monitoring of distribution and health status

Drones offer the advantage of pre-programming for systematic and repeated surveys, allowing for the exact same area to be monitored multiple times. This capability is particularly useful for observing **seasonal dynamics or temporal changes** in coverage and distribution, for instance after a seagrass restoration effort or following the expansion of kelp forests from sea urchin barrens or after sea ice withdrawal. SeaBee is currently also exploring and developing various descriptors for assessing the **health status** of macroalgae and seagrass beds.

Of particular relevance is the potential for incorporating drone-gathered data into monitoring methods like the "**Miljødirektoratets instruks**" (Bekkby et al. 2023, 2024) and programs such as the Water Framework Directive (**WFD** – implemented in Norway) and the Marine Strategy Framework Directive (**MSFD** – not implemented in Norway). Key parameters for these programs include coverage, densities, lower growth limits, and the turf situation for coastal species of macroalgae and seagrass, all of which drone methodology is currently in the development stage. See also Sections 1.4.3 and 1.4.4 for considerations on **Marine Base Maps** and the Nature Agreement (**CBD**).

6 References

Adade R, Aibinu AM, Ekumah B, Asaana J. 2021. Unmanned Aerial Vehicle (UAV) applications in coastal zone management – a review. Environmental Monitoring and Assessment, 193: 154.

Álvarez-González M, Suarez-Bregua P, Pierce GJ, Saavedra C. 2023. Unmanned Aerial Vehicles (UAVs) in Marine Mammal Research: A Review of Current Applications and Challenges. Drones 7: 667. Doi: 10.3390/drones7110667

Bekkby T, Andersen GS, Dolan M, Rinde E, Thormar J, Haugland BT, van Son TC, Bryn A. 2023. Metodehåndbok marint – Kartleggingsmetodikk og variabler (NiN 3.0), Artsdatabanken, Trondheim. 36 pp.

Bekkby T, Moy FE, Olsen H, Rinde E, Bodvin T, Bøe R, et al. 2013. The Norwegian Programme for Mapping of Marine Habitats - Providing Knowledge and Maps for ICZMP. In: Global Challenges in Integrated Coastal Zone Management. Eds. Moksness E, Dahl E, Støttrup J. (Oxford, UK: John Wiley & Sons, Ltd), 19–30. Doi: 10.1002/9781118496480.ch2

Bekkby T, Rinde E, Kvile KØ, Brkljacic MS, Thormar J, Mjelde M, Gitmark JK, Moy SR, Schneider S, Oug E. 2022. Forslag til variabler for økologisk kvalitet for lokaliteter av forvaltningsrelevant marin natur. NIVA report 7797-2022 (M-2430|2022), 116 pp + attachments. (In Norwegian)

Bekkby T, Rinde E, Oug E, Buhl-Mortensen P, Thormar J, Dolan M, Mjelde M, Gitmark JK, Moy SR, Schneider S, Gonzales-Mirelis G, Systad G, van Son TC. 2021. Forslag til forvaltningsrelevante marine naturenheter. NIVA report 7372-2021.

Bekkby T, Smit C, Gundersen H, Rinde E, Steen H, Tveiten L, Gitmark JK, Fredriksen S, Albretsen J, Christie H. 2019. The Abundance of Kelp Is Modified by the Combined Impact of Depth, Waves and Currents. Frontiers in Marine Science, 6: 475. Doi: 10.3389/fmars.2019.00475.

Berglund J, Thomasdotter A, von Friesen CP, Åkerholm M, Saarinen A, Nordling P, Kraft E, Danbolt M, Wall A, Infantes E, Envall M, Möller H, Rasmussen M, Moksnes P-O. 2023. Marin Fjärranalys – Övervakning av vegetationsutbredning med satelliter och drönare. Länsstyrelsen Västerbotten. (In Swedish with English summary). Doi: 10.13140/RG.2.2.29791.74403

Borgersen G, Rinde E, Moy S, Gundersen H. 2020. Are there saltmarshes in Norway? An assessment of the term against Norwegian habitat types (NiN). NIVA report 7558-2020, 50 pp. (In Norwegian with English summary)

Borja A, Berg T, Gundersen H, Hagen AG, Hancke K, Korpinen S, Leal MC, Luisetti T, Menchaca I, Murray C, Piet G, Pitois S, Rodríguez-Ezpelata N, Sample JE, Talbot E, Uyarra MC. 2024. Innovative and practical tools for monitoring and assessing biodiversity status and impacts of multiple human pressures in marine systems. Environmental Monitoring and Assessment (In press).

Bryn A, Bekkby T, Rinde E, Gundersen H, Halvorsen R. 2021. Reliability in distribution modeling – a synthesis and step-by-step guidelines for improved practice. Frontiers in Ecology and Evolution. Doi: 10.3389/fevo.2021.658713

Bryn A, Haugland BT, Dolan M, Bekkby T, Naas AE. 2023. Feltveileder marint (NiN 3.0) - Regler, typetabeller og praktiske råd, Artsdatabanken, 2023. https://artsdatabanken.no/Files/55164/Marin_feltveileder_NiN_3.0.pdf

Cavanaugh KC, Cavanaugh KC, Bell TW, Hockridge EG. 2021. An automated method for mapping giant kelp canopy dynamics from UAV. Front. Environ. Sci. 8:587354. Doi: 10.3389/fenvs.2020.587354

Dierssen HM, Chlus A, Russell B. 2015. Hyperspectral discrimination of floating mats of seagrass wrack and the macroalgae Sargassum in coastal waters of Greater Florida Bay using airborne remote sensing. Remote Sensing of Environment 167: 247-258.

Doughty CL, Cavanaugh KC. 2019. Mapping Coastal Wetland Biomass from High Resolution Unmanned Aerial Vehicle (UAV) Imagery. Remote Sens. 11: 540. Doi: doi.org/10.3390/rs11050540

Duffy JE, Benedetti-Cecchi L, Trinanes J, Muller-Karger FE, Ambo-Rappe R, Boström C, Buschmann AH, Byrnes J, Coles RG, Creed J, Cullen-Unsworth LC, Diaz-Pulido G, Duarte CM, Edgar GJ, Fortes M, Goni G, et al. 2019. Toward a Coordinated Global Observing System for Seagrasses and Marine Macroalgae. Frontiers in Marine Science 6. Doi: 10.3389/fmars.2019.00317

Duffy JP, Pratt L, Anderson K, Land PE, Shutler JD. 2018. Spatial assessment of intertidal seagrass meadows using optical imaging systems and a lightweight drone, Estuarine, Coastal and Shelf Science, 200: 169-180. Doi: 10.1016/j.ecss.2017.11.001

Edney AJ, Hart T, Jessopp MJ, Banks A, Clarke LE, Cugnière L, Elliot KH, Juarez Martinez I, Kilcoyne A, Murphy M, Nager RG, Ratcliffe N, Thompson DL, Ward RM, Wood MJ. 2023. Best practices for using drones in seabird monitoring and research. Marine Ornithology 51: 265-280.

Féral J-P, Norro A. 2023. Specific initial training standards are needed to dive for science in Europe, Occupational vs. Citizen Science Diving. Frontiers in Marine Science, 10: Doi: 10.3389/fmars.2023.1134494

Frigstad H, Gundersen H, Andersen GS, Borgersen G, Kvile KØ, Krause-Jensen D, Boström C, Bekkby T, d'Auriac MA, Ruus A, Thormar J, Asdal K, Hancke K. 2021. Blue carbon – climate adaptation, CO2 uptake and sequestration of carbon in Nordic Blue forests. Results from the Nordic Blue Carbon Project. TemaNord 2020:541. 139 pp.

Gullström M, Lundén B, Bodin M, Kangwe J, Öhman MC, Mtolera MSP, Björk M. 2006. Assessment of changes in the seagrass-dominated submerged vegetation of tropical Chwaka Bay (Zanzibar) using satellite remote sensing. Estuarine, Coastal and Shelf Science 67: 399-408. Doi: 10.1016/j.ecss.2005.11.020.

Gundersen H, Bekkby T, Norderhaug KM, Oug E, Rinde E, Fredriksen F. 2018. Eufotisk fast saltvannsbunn, Marint gruntvann. Norwegian red list for nature types 2018. Artsdatabanken, Trondheim. (In Norwegian). Downloaded 04.04.2024 from https://artsdatabanken.no/RLN2018/15

Gundersen H, Bryan T, Chen W, Moy FE, Sandman AN, Sundblad G, Schneider S, Andersen JH, Langaas S, Walday MG. 2016. Ecosystem Services in the Coastal Zone of the Nordic Countries. TemaNord 2016:552 Copenhagen: Nordic Council of Ministers. 130 pp. https://norden.diva-portal.org/smash/record.jsf?pid=diva2:1067839

Gundersen H, Hancke K, Poulsen RN, Buls T, Ghareeb M, Christie H, Kile MR, Arvidsson KS, Kvile KØ. 2022. KELPMAP-Vega: Metodeutvikling for kartlegging av tareskog ved bruk av droner og satellittbilder (feltrapport). NIVA report 7790-2022. 29 pp. (In Norwegian)

Gundersen H, Rinde E, Bekkby T, Hancke K, Gitmark JK, Christie H. 2021. Variation in population structure and standing stock of kelp along multiple environmental gradients and implications for ecosystem services. Frontiers in Marine Science. Doi: 10.3389/fmars.2021.578629

Haarpaintner J, Davids C. 2020. Satellite Based Intertidal-Zone Mapping from Sentinel-1&2. Final Report, NORCE Klima Report nr. 2-2020. Norwegian Environment Agency report M-1646. https://www.miljodirektoratet.no/globalassets/publikasjoner/m1646/m1646.pdf

Halvorsen R, Bratli H, Bryn A, Erikstad L, Skarpaas O, Wollan AK. 2024. NiN 3 Systemdokumentasjon 1 (NiN3SD1). NiN versjon 3.0 systemkjerne: teori, generelle prinsipper og systemarkitektur (Preliminary edition 16. November 2023). Artsdatabanken report. (In Norwegian). 197 pp.

Hancke K, Andersen GS, Gundersen H, Kvile KØ, Trannum HC, Borgersen G. 2022. Status and outlook on marine areas important for carbon storage in Norway. NIVA report 7788-2022. 69 pp. (In Norwegian with English summary)

Hossain MS, Bujang JS, Zakaria MH, Hashim M. 2014. The application of remote sensing to seagrass ecosystems: an overview and future research prospects. International Journal of Remote Sensing, 36: 61–114. Doi: 10.1080/01431161.2014.990649

Infantes E. 2018. Coastal mapping in Kosterhavet National Park using drone technology: Eelgrass distribution 2018. Rapport Länsstyrelsen i Västra Götalands län, 2018.

Johnsen G, Volent Z, Dierssen H, Pettersen R, Ardelan MV, Søreide F, Fearns P, Ludvigsen M, Moline M, 2013. Underwater hyperspectral imagery to create biogeochemical maps of seafloor properties. In: Subsea optics and imaging; Watson J, Zielinski O (Eds.). Woodhead Publishing: Oxford, UK, pp. 508–540e.

Joyce KE, Fickas KC, Kalamandeen M. 2023. The unique value proposition for using drones to map coastal ecosystems. Cambridge Prisms: Coastal Futures, 1, e6, 1–10. Doi: 10.1017/cft.2022.7.

Krause-Jensen D, Gundersen H, Björk M, Gullström M, Dahl M, Asplund ME, Boström C, Holmer M, Banta G, Graversen AEL, Pedersen MF, Bekkby T, Frigstad H, Skjellum SF, Thormar J, Gyldenkærne S, Howard J, Pidgeon E, Ragnarsdóttir SB, Mols-Mortensen A, Hancke K. 2022. Nordic Blue Carbon ecosystems: Status and outlook. Frontiers in Marine Science, Review article. Doi: 10.3389/fmars.2022.847544

Kvile KØ, Andersen GS, Baden SP, Bekkby T, Bruhn A, Geertz-Hansen O, Hancke K, Hansen JLS, Krause-Jensen D, Rinde E, Steen H, Wegeberg S, Gundersen H. 2022. Kelp forest distribution in the Nordic region. Frontiers in Marine Science. doi: 10.3389/fmars.2022.850359

Kvile KØ, Gundersen H, Sample JE, Buls T, Poulsen RN, Bekkby T, Ghareeb ME, Salberg A-B, Hancke K. A protocol for drone and ground-truth data collection and image annotation for coastal habitat classification and mapping. Submitted to MethodsX Journal.

Li Y, Xie L, Gundersen H, Ge Z, Hancke K. 2023. Quantifying seaweed and seagrass beach deposits using high-resolution UAV imagery. Journal of Environmental Management 331. Doi: 10.1016/j.jenvman.2022.117171

Liu Q, Kampffmeyer M, Jenssen R, Salberg A B. 2020. Dense dilated convolutions' merging network for land cover classification. IEEE Trans. Geosci. Remote Sensing, vol. 58: 6309-6320.

Liu Y, Liu Q, Sample JE, Hancke K, Salberg A-B. 2022. Coastal habitat mapping with UAV multi-sensor data: an experiment among DCNN-based approaches. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-3-2022, 439–445. Doi: 10.5194/isprs-annals-V-3-2022-439-2022.

Lønborg C, Thomasberger A, Stæhr PAU, Stockmarr A, Sengupta S, Rasmussen ML, Nielsen LT, Hansen LB, Timmermann K. 2022 Submerged aquatic vegetation - overview of monitoring techniques used for the identification and determination of spatial distribution in European coastal waters. Integrated Environmental Assessment and Management 18:892-908.

Meld. St. 14 (2015–2016). Nature for life — Norway's national biodiversity action plan. https://www.regjeringen.no/en/dokumenter/meld.-st.-14-20152016/id2468099/

Meld. St. 21 (2023-2024). Helhetlige forvaltningsplaner for de norske havområdene — Barentshavet og havområdene utenfor Lofoten, Norskehavet, og Nordsjøen og Skagerrak. https://www.regjeringen.no/no/dokumenter/meld.-st.-21-20232024/id3032474/

Mora-Soto A, Palacios M, Macaya EC, Gómez I, Huovinen P, Pérez-Matus A, Young M, Golding N, Toro M, Yaqub M, Macias-Fauria M. 2022 A high resolution global map of giant kelp (Macrocystic pyrifera) forests and intertidal green algae (Ulvophyceae) with Sentinel-2 imagery. Remote sensing 12, 694. Doi: 10.3390/rs12040694.

Murphy H, Jenkins G. 2010. Observational methods used in marine spatial monitoring of fishes and associated habitats: a review. Marine and Freshwater Research 61: 236-252. Doi:10.1071/MF09068

Perko R, Raggam H, Roth PM. 2019. Mapping with Pléiades – end-to-end workflow. Remote sensing 2019 11, 2052. Doi: 10.3390/rs11172052.

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

Rinde E, Bekkby T, Kvile KØ, Andersen GS, Brkljacic M, Anglès d'Auriac M, Christie H, Fagerli CW, Fredriksen S, Moy S, Staalstrøm A, Tveiten L. 2021. Kartlegging av et utvalg marine naturtyper i Oslofjorden. NIVA report 7605-2021. (In Norwegian with English summary)

Rinde E, Gitmark JK, Kile MR, Moy S, Fagerli CW, Bekkby T. 2024. What is "lurv"? Is "lurv" always an indicator of poor ecological state? NIVA report 7968-2024 / M-2756|2024. (In Norwegian with English summary)

Ronneberger O, Fischer P, Brox T. 2015. U-Net: Convolutional networks for biomedical image segmentation. Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent, pp. 234-241, 2015.lo

Rowan GSL, Kalacska M. 2021. A review of remote sensing of submerged aquatic vegetation for non-specialists. Remote Sens. 13, 623. https://doi.org/10.3390/ rs13040623

Schimel ACG, Thormar J, Oveland I, Thorsnes T, Elvenes S, Kurz T, Welde H. 2023. Summary report of datasets acquired for the purpose of method development within the Marine Base Maps for the Coastal Zone (Marine grunnkart i kystsonen) project at Fjøløy and Klosterøy, Stavanger, Norway. NGU report 2023.008. 35 pp.

Skjelvareid MH, Rinde E, Hancke E, Blix K, Hoarau GG. 2023. Mapping Marine Macroalgae along the Norwegian Coast Using Hyperspectral UAV Imaging and Convolutional Nets for Semantic Segmentation. IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 4002-4005. Doi: 10.1109/IGARSS52108.2023.10282809.

Stæhr PAU, Nielsen LT, Azhar M, Göke C, Huber S, Christensen JPA, Hansen LB, Stæhr SU, Krause-Jensen D. 2024. Assessment of the areal distribution of submerged aquatic vegetation using remote sensing in Danish coastal waters. Integrated Marine Monitoring - analytical phase. Aarhus University, DCE, Scientific Report no. 596, 61 pp.

Torsvik BM, Poulsen RN, van Bavel B, Gundersen H, Hancke K. 2020. Detection of macroplastic on beaches using drones and object-based image analysis. NIVA report 7553-2020, 47 pp.

7 Appendices

Appendix A – Project participants and responsibilities

Name	Institution	Activity
Hege Gundersen	NIVA	Project manager, administration, main responsibility for field planning and reporting
Kasper Hancke	NIVA	Co-manager, leader of the SeaBee infrastructure, field planning and reporting
Medyan Ghareeb	NIVA	Drone pilot, field planning, post-processing of drone images, reporting
Hartvig Christie	NIVA	Collection of ground-truth data
Maia R. Kile	NIVA	Collection of ground-truth data, field planning
Karoline S. Arvidsson	NIVA	Administration, financial responsibility, field planning and reporting
Kristina Ø. Kvile	NIVA	Reporting
Robert N. Poulsen	SpectroFly	Drone pilot, field planning, post-processing of drone images, reporting
Toms Buls	SpectroFly	Drone pilot, post-processing of drone images, reporting
Arnt-Børre Salberg	Norwegian Computing Center	Convolutional neural network analyses, creating habitat maps, upscaling, reporting
Izzie Liu	Norwegian Computing Center	Convolutional neural network analyses

Appendix B – All species and habitat types

List of SeaBee habitat classes, which of several, such as many of the macroalga species and sediment types, are already proven to be predicted using drones and ML analyses. Whereas several are also yet to be tested out, such as faunal species like sea urchins, brittle stars, and mussels. All photos (CC-BY-NC-SA 4.0): Janne K. Gitmark.

1. *Alaria esculenta* (butare)

2. *Chorda filum* (martaum)

3. *Laminaria hyperborea* (stortare) / *L. digitata* (fingertare)

4. *Pelvetia canaliculata* (sauetang)

5. *Saccharina latissima* (sukkertare)

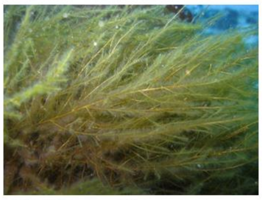
6. *Fucus serratus* (sagtang)

7. *Ascophyllum nodosum* (grisetang)

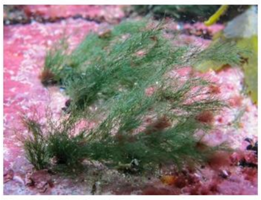
8. *Fucus spiralis* (spiraltang)

9. *Fucus vesiculosus* (blæretang)

10. *Halidrys siliquosa* (skulpetang)



11. *Desmarestia aculeata* (kjerringhår)



12. *Cladophora rupestris* (grønndusk)

13. Ulva spp. e.g. U. intestinalis (tarmgrønske)

14. Porphyra spp. e.g. P. umbilicalis (fjærehinne)

15. Chondrus crispus (krusflik)

16. Furcellaria lumbricalis (svartkluft)

17. Palmaria palmata (søl)

18. Vertebrata lanosa (trøffeltang)

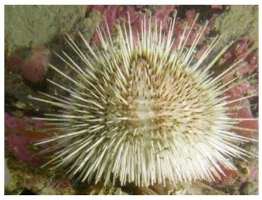
19. Unspecified turf (lurv)

20. *Zostera marina* (ålegras)

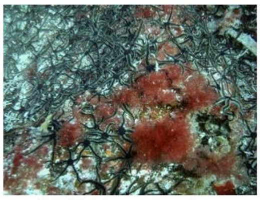
21.Maerl (løstliggende rugl)

22. *Echinus esculentus* (rød kråkebolle)

23. *Strongylocentrotus droebachiensis* (grønn kråkebolle)



24.*Gracilechinus acutus* (langpiggsjøpiggsvin)



25. Ophiocomina nigra
(svart slangestjerne)

26. Asterias rubens
(vanlig korstroll)

27. Crassostrea gigas
(stillehavsøsters)

28. Mytilus edulis
(blåskjell)

29. Unspecified, dried seaweed
(tangvoll)

30. Unspecified, dried seagrass
(sjøgressvoll)

31. Terrestrial vegetation

32. Various wood incl. driftwood
(drivved)

33. Bedrock (nakent fjell)

34. Boulder (blokk)

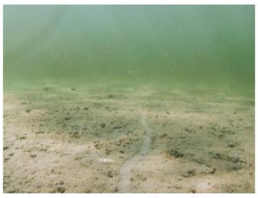
35. Cobble (stein)

36. Gravel (grus)

37. Sandy seafloor (sandbunn)

38. Sandy beach (sandstrand)

39. Sandy seafloor with microalgae
(sandbunn med mikroalger)



40. Muddy seafloor (mudderbunn)

41. Deep sea (deeper than Secchi
deep) (dypt hav)

42. Anthropogenic (antropogent),
e.g., house, litter, plastics

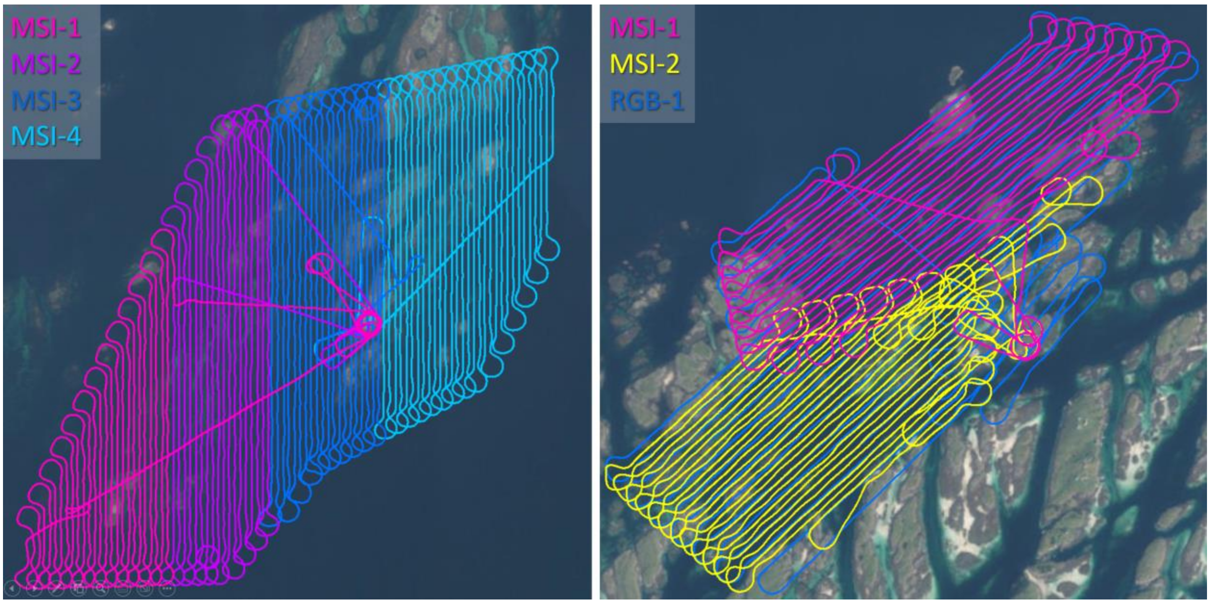
Appendix C – Drone mission specifications

Specifications for drone missions at Steinskjæran (southern area in Vega) and Sandværodden (northern area in Herøy), covering 3.0 and 1.5 km^2, respectively. Different missions represent flights separated by a battery shift.

Northern area (18. August 2022)						
Mission	Take-off	Landing	Height (m)	Sensor	Cloud coverage	Wind (m/s)
1	09:45	10:33	120	MSI	40%	6-7 Ø
2	10:38	11:30	120	MSI	10%	7-8 Ø
3	11:39	12:34	120	MSI	10%	5-6 Ø
4	12:38	13:43	120	MSI	10%	5-6 Ø
5	13:47	13:54	120	MSI	10%	5-6 Ø
Southern area (19. August 2022)						
Mission	Take-off	Landing	Height (m)	Sensor	Cloud coverage	Wind (m/s)
1	10:32	11:27	120	MSI	100%	6-8 SØ
2	11:31	12:25	120	MSI	100%	6-8 SØ
3	12:30	13:31	120	RGB	100%	6-8 SØ

Appendix D – Flight paths

Flight paths of the southern (left) and northern (right) areas. The colors represent different missions (i.e., between each battery shift) (Appendix C). Credit: Copernicus Sentinel-2 data (2022).



Appendix E – Data delivery

All data collected in KELPMAP, including drone orthomosaics and habitat classification maps, are openly shared under the CC-BY-4.0 license on SeaBee's GeoNode platform. Individual (unstitched) drone images and all ground truth data collected in the study are shared with the Norwegian Environment Agency (NEA) through their file sharing system. Also, all photos displayed in Appendix B have been made available through the same system and can be used by the NEA under the CC-BY-NC-SA 4.0 license, which permits use with attribution and prohibits commercial use.

The Norwegian Institute for Water Research

We are Norway's premier research institute in the fields of water and the environment. We are experts on ecosystems in both freshwater and marine environments, from mountains, lakes and rivers, to fjords, coasts and oceans. We develop science-based knowledge and solutions to challenges related to the interaction between water and climate, the environment, nature, people, resources and society.