Mapping at the Boundary: Simultaneous Above- and Underwater Surveying of Rocky Coastal Environments with an Uncrewed Surface Vehicle

Sergey Khokhlov [1], Fabio Menna [2,3], Erica Nocerino [4]

[1] PhD programme in Culture, Literature, Rights, Tourism and Territory, Department of Humanities and Social Sciences, University of Sassari, Sassari, Italy – s.khokhlov@studenti.uniss.it
[2] Department of Chemical, Physical, Mathematical and Natural Sciences, University of Sassari, Sassari, Italy – fmenna@uniss.it
[3] Department of Science and Technology, University of Napoli Parthenope, Napoli, Italy
[4] Department of Humanities and Social Sciences, University of Sassari, Sassari, Italy - enocerino@uniss.it

Keywords: Uncrewed Surface Vehicle, Coastal Monitoring, Habitat Mapping, Underwater Photogrammetry, Differential GNSS

Abstract

Mapping at the air–water interface in shallow coastal environments remains challenging due to the need to integrate heterogeneous datasets acquired under different geometric and operational conditions. This study presents a modular uncrewed surface vehicle (USV)-based system for simultaneous above- and underwater photogrammetric surveying supported by differential GNSS positioning. The system integrates a rigid multi-camera configuration, GNSS time synchronization, and a direct georeferencing workflow based on trajectory interpolation and lever-arm calibration. Experimental results from a rocky coastal site in Sardinia (Italy) show that underwater photogrammetry can achieve centimetric absolute accuracy (2–4 cm horizontally and ~8 cm vertically) without underwater ground control points. The USV enables controlled and repeatable acquisition in very shallow environments, while UAV photogrammetry complements the reconstruction of the emerged area. Limitations related to image quality and refraction effects are discussed. The system represents a flexible and scalable solution for integrated coastal mapping and monitoring.

1. Introduction

Rocky coastal environments are targeted for protection and monitoring by the European Union through several policies. They are listed as priority habitats within the Natura 2000 Network, a European ecological network of protected areas established in accordance with the "Habitat" Directive 92/43 / EEC. The Marine Strategy Framework Directive (MSFD) (2008/56/EC) enforces the monitoring of physical damages of seabed habitats and biodiversity and calls for the achievement of "Good Environmental Status" for coastal and marine ecosystems. The EU Biodiversity Strategy for 2030, part of the European Green Deal, promotes restoration and conservation of natural coastal processes and ecosystems, to strictly protect at least 30% of EU marine areas. The role of natural coastal barriers (including rocky shores) in reducing flood risks is recognised by the Floods Directive (2007/60/EC) and Climate Adaptation Strategy (2021). The European Environmental Agency (EEA) has acknowledged that rocky shorelines, although more resilient compared to sandy beaches, are subject to increasing stressors due to sea-level rise and human induced pressure, thus recommended continuous monitoring.

In the framework of the Italy's National Recovery and Resilience Plan (PNRR), the National Biodiversity Future Centre (NBFC) and e.INS – Ecosystem of Innovation for Next Generation Sardinia have set specific objectives on the monitoring, restoration and sustainable management of rocky coastal habitats and ecosystems. Within their scope, we are developing a methodology based on a customised uncrewed surface vehicle for high-resolution environmental mapping in shallow water and at the water–land interface, with the specific aim of supporting the monitoring objectives associated with rocky coastal habitats.

Surveying and mapping at the air–water boundary remains a challenging task because of the different physical and operational conditions of the emerged and submerged domains. Conventional bathymetric surveys are often impractical in very shallow waters because of vessel draft limitations, while terrestrial surveying techniques cannot capture submerged features. UAV photogrammetry can provide rapid coverage of emerged coastal sectors and, under favourable environmental conditions, also contribute to the reconstruction of shallow submerged areas, but its effectiveness is constrained by water transparency, sea-surface conditions, and refraction effects at the air–water interface. In contrast, close-range underwater photogrammetry can provide very high geometric detail but is generally limited in spatial productivity and often depends on ground control or dedicated positioning solutions.

For these reasons, shallow coastal mapping should not be interpreted as a problem solvable by a single platform or a single sensing technique. Recent literature increasingly supports the idea of a multi-platform workflow in which different survey tools are combined according to their operational strengths. In such a framework, UAVs provide rapid large-area reconnaissance, USVs enable systematic and safe surveying in very shallow coastal waters, and underwater photogrammetry supports detailed three-dimensional reconstruction of selected sectors.

Within this context, uncrewed surface vehicles – USV are particularly promising because they can operate in areas that are inaccessible or unsafe for conventional vessels while supporting multiple sensors in an integrated survey configuration.

The system presented in this work is based on this concept and exploits a catamaran-type USV equipped with a modular, low-cost, and largely cable-less payload designed for simultaneous above- and underwater data acquisition. The payload supports optical sensors as well as additional geophysical instruments, including side-scan sonar (SSS), magnetometers, and other optional survey sensors. The platform integrates differential

GNSS positioning, synchronized optical sensors, and a rigid multi-camera rig that allows direct georeferencing of underwater imagery through time synchronization, GNSS trajectory interpolation, and lever-arm calibration.
The main contribution of this study lies in the development and first validation of a modular USV-based system that combines low-cost sensors, differential GNSS positioning, synchronized image acquisition, and direct georeferencing into a single operational workflow for simultaneous mapping across the air–water interface.

2. Related works

Surveying across the air–water interface remains a challenging task due to the different physical properties of the two media and the geometric discontinuity introduced by refraction. Previous studies (Nocerino and Menna, 2020; Nocerino et al., 2019; Menna et al., 2018, 2013) demonstrated that integrated photogrammetric modelling across this boundary is feasible when appropriate calibration strategies and geometric constraints are applied.

Beyond purely optical approaches, several works explored multi-sensor integration to improve reconstruction of semi-submerged environments. For instance, Moisan et al. (2015) combined cameras and multibeam sonar for tunnel reconstruction, while Van der Lucht et al. (2018) proposed underwater laser scanning with refraction correction. These studies highlight the importance of combining complementary sensors to achieve complete geometric documentation across the water interface.

From an operational perspective, low-cost solutions have been developed for shallow-water surveys. The Geoswim project (Furlani and Antonioli, 2023) introduced a raft-based system for large-scale coastal mapping, while Casella et al. (2024) and Karaki et al. (2024) proposed portable or hybrid UAV–underwater approaches. These works emphasize the need for flexible and cost-effective systems capable of integrating data across the air–water interface.

Within this context, recent studies have focused on GNSS-supported underwater photogrammetry. Jaud et al. (2024) demonstrated that centimetric georeferencing can be achieved using a floating platform equipped with RTK GNSS and underwater cameras. Similarly, Lo et al. (2024) showed that accurate bathymetric reconstruction is possible without dense underwater ground control, provided that synchronization and the geometric relationship between sensors are well defined. Hatcher et al. (2020, 2023) further demonstrated that this approach can be scaled to large areas, confirming the potential of image-based underwater mapping.

These studies identify key requirements for shallow-water photogrammetry, including accurate synchronization, controlled acquisition geometry, and well-defined spatial relationships between sensors. In particular, systems based on a rigid lever-arm connection between underwater cameras and a surface GNSS platform are effective for enabling direct or quasi-direct georeferencing.

Within this framework, unmanned surface vehicles (USVs) play a crucial role by enabling safe and efficient surveys in shallow areas inaccessible to conventional vessels. Cao et al. (2023) showed that USVs can provide reliable bathymetric data and can be integrated with aerial data to generate seamless coastal models. More broadly, recent literature supports a multi-platform survey approach, where UAVs provide rapid large-scale coverage, USVs enable shallow-water access, and underwater photogrammetry allows high-resolution mapping of selected areas.
Taken together, these studies suggest that no single platform can address all survey requirements. Instead, an integrated workflow combining UAV, USV, and underwater photogrammetry represents an effective strategy for shallow coastal mapping. The system presented in this work is developed within this framework and focuses on the implementation of a modular USV-based platform for GNSS-supported, integrated above- and underwater photogrammetric surveys.

3. Material and methods

3.1 System Architecture

The developed system is based on a modular architecture designed to integrate positioning sensors, optical sensors, and navigation systems on a small uncrewed surface vehicle (USV) operating in shallow coastal environments. The main design requirements were modularity, low cost, ease of deployment, and the possibility to perform synchronized above- and underwater surveys with centimetric positioning accuracy.

The platform used in this work is a catamaran-type USV (BlueBoat, Blue Robotics, Fig.1), selected for its stability, shallow draft, and payload capacity. The platform is equipped with a Pixhawk autopilot running ArduPilot firmware, which enables autonomous navigation, waypoint tracking, and mission planning. The USV is normally controlled from a ground station located onshore via Wi-Fi communication, which provides a reliable connection up to approximately 100–150 m. To extend the operational range, an LTE (Long-Term Evolution) modem is integrated on board.

To improve reliability of mobile connectivity, a dual-SIM cellular router is used, allowing automatic switching between network providers in areas with unstable coverage. This is particularly important in coastal environments with rocky cliffs, where both radio links to RTK base stations and LTE connectivity may be unreliable or intermittently blocked. In areas without LTE coverage, the system can be extended with satellite communication (e.g., Starlink or long-range radio links), following the same network architecture.

The USV is accessed through a VPN connection (e.g., WireGuard) via a remote virtual server (Virtual Dedicated Server, VDS), to which both the onboard router and one or multiple operators can connect. This architecture allows distributed operation, where different operators can remotely monitor navigation, sensor data, and live video streams from any location.

A key aspect of the system is the development of a modular payload architecture, designed to be flexible and easily reconfigurable. Sensors primarily communicate wirelessly (e.g., via Bluetooth) and are housed in waterproof enclosures, enabling a cable-less configuration that simplifies deployment. However, wired connections are also supported when higher reliability is required, using waterproof connectors and marine-grade cabling to ensure robust operation in harsh environments.

Live video streams are transmitted using the Real-Time Streaming Protocol (RTSP) from both above-water and underwater cameras are used for real-time monitoring of the survey operations. The above-water camera supports collision avoidance with vessels and surface obstacles, while the underwater camera helps prevent grounding and collisions with submerged rocks, and allows real-time supervision of photogrammetric data acquisition. As shown in Fig. 1**,** the platform combines the georeferenced mapping subsystem with a navigation and communication subsystem, allowing remote supervision of the mission and online monitoring of both surface and underwater conditions.

The payload consists of a rigid frame supporting the positioning system and the optical sensors. The rigid structure includes a differential GNSS antenna mounted above the water surface and a multi-camera rig composed of two action cameras, one mounted above the water surface and one below the water surface, allowing simultaneous acquisition of images in air and underwater. The rigid connection between the GNSS antenna and the cameras ensures that the spatial offsets between sensors, referred to as lever-arm parameters, remain constant during acquisition, which is a fundamental requirement for direct georeferencing.

To achieve centimetric positioning accuracy, the system supports Real-Time Kinematic (RTK), Network RTK (NRTK), and Post-Processed Kinematic (PPK) positioning modes, with RTK and NRTK primarily used for navigation and real-time positioning, while PPK is used for trajectory refinement and photogrammetric georeferencing. The main GNSS receiver used on the platform and for ground truth measurements was a CHCNAV i89 GNSS, which supports RTK positioning via both radio link and NRTK corrections via internet connection, and allows raw GNSS data logging for post-processing. The receiver includes tilt compensation based on an integrated inertial measurement unit (IMU), which allows accurate point measurements even when the survey pole is not perfectly vertical. This feature is particularly important in shallow water environments, where wave motion and operator instability make it difficult to maintain the survey pole perfectly vertical during RTK measurements.

GNSS data are transmitted via Bluetooth to an ESP32 microcontroller that acts as a communication bridge between the GNSS receiver and the USV navigation system. The ESP32 parses the incoming NMEA messages and forwards positioning data both to the Pixhawk autopilot for navigation purposes and to the onboard logging system, where GNSS observations are stored for post-processing and trajectory reconstruction.

In addition to the CHCNAV i89 receiver, a low-cost alternative based on the u-blox ZED-F9P module was also tested. The receiver was integrated directly into the system of BlueBoat via a UART (Universal Asynchronous Receiver-Transmitter) connection, and a second identical unit was configured as a base station. This configuration enables RTK positioning with significantly reduced cost while maintaining compatibility with the overall system architecture.

The optical system is based on low-cost action cameras equipped with internal GNSS receivers. The use of lightweight action cameras was motivated by their low cost, small size, and ease of integration within a modular payload. In addition to standard action cameras (GoPro© v.10), a 360° camera (Insta360) was also tested to acquire panoramic imagery and extract both fisheye images and GNSS trajectory data from the video stream.

3.2 Differential GNSS integration for navigation and direct georeferencing

To enable both precise navigation and direct georeferencing of the acquired imagery, the USV was equipped with a differential GNSS positioning system capable of operating in Real-Time Kinematic (RTK), Network RTK (NRTK), and post-processed kinematic (PPK) modes. The use of centimetric positioning is particularly important in shallow-water photogrammetric surveys, where the camera operates close to the seabed and the swath width is therefore limited. In such conditions, survey lines must be planned with small spacing to ensure sufficient lateral overlap (sidelap) between image strips. With standard GNSS positioning based on code measurements, navigation errors could lead to insufficient overlap and gaps in photogrammetric coverage. The use of RTK positioning allows the platform to follow planned trajectories with sufficient accuracy to guarantee the required image overlap and complete coverage of the surveyed area.

In addition to navigation, the differential GNSS system provides the trajectory required for direct georeferencing of both above- and underwater imagery. The GNSS antenna is mounted above the water surface, while the cameras are mounted on a rigid frame connected to the antenna. At the same time, GNSS observations are logged for post-processing and trajectory reconstruction. The GNSS system also provides a common time reference used to synchronize all sensors, ensuring temporal consistency between positioning data and image acquisition..

(a) (b) (c)

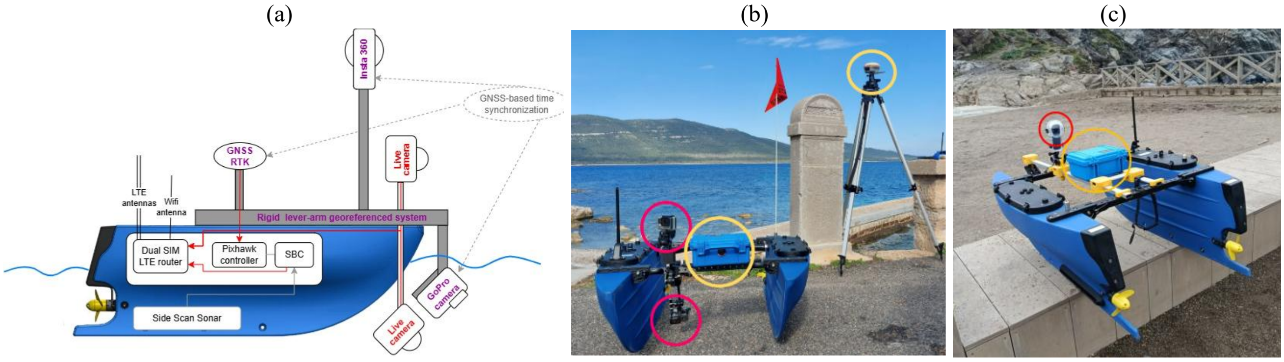

Figure 1. Conceptual architecture of the modular USV-based multi-sensor survey system. The platform integrates a rigid lever-arm georeferenced system with differential GNSS, synchronized above- and underwater optical sensors, navigation cameras for real-time mission supervision, onboard navigation and logging units, cellular communication, and optional geophysical payloads such as side-scan sonar and magnetometer. GNSS time is used as the common temporal reference for sensor synchronization and direct georeferencing (a). Two operational version of the developed modular, low-cost, cable-less payload: in (a) with two GPS enabled GoPro cameras, in (b) with an Insta360 camera above the water; in yellow RTK/PPK GNSS systems.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-955-2026

This integration allows the USV to operate both as a navigation platform following pre-planned trajectories and as a mobile mapping platform providing the trajectory required for direct georeferencing of multi-sensor data acquired at the air–water interface. While the differential GNSS system provides the platform trajectory and time reference, the geometric relationship between the positioning system and the optical sensors is defined by the rigid multi-camera system described in the following section.

3.3 Multi camera system for simultaneous under and above the water surveys

In previous works (Nocerino and Menna, 2020; Nocerino et al., 2018; Menna et al., 2018; Menna et al., 2013) the authors investigated a methodology to achieve accurate, unified 3D reconstructions of scenes that span both air and underwater environments, overcoming the geometric discontinuity introduced by refraction at the water interface. The approach is based on separating the photogrammetric processing of above-water and underwater image blocks—each treated with appropriate calibration models and camera parameters—and then rigorously linking them into a single reference system. This linkage can be achieved through:

1) the use of rigid cross-medium constraints, such as calibrated objects intersecting the water surface; or
2) synchronized imaging systems that enforce geometric relationships between at least two cameras operating in different media.

The second approach, based on synchronized cameras mounted on a rigid structure, is adopted in this work. The system uses a rigid multi-camera rig equipped with two action cameras, one operating above the water surface and one below the water surface, allowing simultaneous acquisition of images in air and underwater. The rigid connection between the cameras ensures that their relative orientation and position remain fixed during acquisition, thus providing a stable geometric constraint between the two photogrammetric blocks. This configuration allows the above-water and underwater models to be processed separately using appropriate camera functional models and then merged into a single reference system using the known rigid transformation between the two cameras.

In the proposed system, the rigid multi-camera rig is also physically connected to the differential GNSS antenna, forming a single rigid body that includes the positioning system and the optical sensors. This configuration allows the spatial relationship between the GNSS antenna and the camera optical centres to be determined through lever-arm calibration. Once this relationship is known, the position of each camera can be computed from the GNSS trajectory, enabling direct georeferencing of both the above- and underwater image blocks (provided that camera positions are not aligned along a straight line that would represent a degenerate case for the computation of the 3D transformation).

The choice of using cable-less lightweight action cameras was mainly motivated by their simple and modular integration, limited weight, and the possibility of synchronizing their internal clocks with GNSS time. The cameras can be easily mounted and removed together with the rigid rig, allowing calibration procedures to be performed outside the water in controlled conditions. This modular configuration also allows different sensors to be installed depending on the survey requirements, without modifying the main platform configuration.

This multi-camera rigid rig configuration therefore serves two main purposes: first, it enforces a stable geometric relationship between the above- and underwater cameras, enabling integrated photogrammetric reconstruction across the air–water interface; second, it provides a rigid geometric link between the cameras and the GNSS antenna, enabling direct georeferencing of the acquired imagery.

3.4 Stereo rig calibration and lever-arm estimation

To enable both the geometric integration of above- and underwater photogrammetric blocks into a single reference system and direct georeferencing, the relative geometry between the cameras and the GNSS antenna must be accurately determined. This requires the estimation of (i) the relative orientation between the two cameras mounted on the rigid rig (stereo rig calibration), and (ii) the lever-arm between the GNSS antenna phase centre and each camera.

The stereo rig calibration is performed following the approach described in Method 2 in Nocerino and Menna (2020), where two synchronized cameras mounted on a rigid support acquire images simultaneously in air and underwater. The method allows the estimation of the relative orientation between the two cameras without the need for objects intersecting the water surface, by enforcing rigid constraints between the two image blocks.
In addition to the relative orientation between the two cameras, the spatial relationship between the GNSS antenna phase centre and each camera projection centre must be determined. Let $\mathbf{X}_{GNSS}(t)$ be the GNSS antenna position at time t, and let $\mathbf{l}_c$ be the lever-arm vector between the GNSS antenna and the camera projection centre expressed in the body reference frame. The position of the camera projection centre can be computed as:

$$\mathbf{X}_c(t) = \mathbf{X}_{GNSS}(t) + \mathbf{R}_b^m(t)\,\mathbf{l}_c$$

where:

- $\mathbf{R}_b^m(t)$ is the rotation matrix from the body reference frame to the mapping reference frame,
- $\mathbf{l}_c$ is the lever-arm vector between the GNSS antenna phase centre and the camera projection centre.

In this study, both the stereo rig parameters and the lever-arm between the GNSS antenna phase centre and the cameras were estimated through a self-calibration procedure performed in air before installing the rig on the USV. During this calibration phase, images acquired simultaneously by the two cameras and GNSS observations are processed together in a bundle adjustment with additional parameters. The unknown parameters included:

- the relative orientation between the two cameras (baseline and boresight),
- the lever-arm between the GNSS antenna and each camera.

These parameters are estimated through a least-squares adjustment, allowing the entire system (GNSS antenna and cameras) to be treated as a single rigid body with known internal geometry. Once calibrated, the rigid transformation parameters were applied during the survey to compute the position of each camera projection centre from the GNSS trajectory through trajectory interpolation and lever-arm correction.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-955-2026

This calibration procedure represents an important step in the proposed workflow, as it allows accurate direct georeferencing of both above- and underwater imagery without the need for underwater ground control points, provided that the rigid geometry of the system is preserved during the survey. The system can therefore be interpreted as a rigid multi-sensor mobile mapping platform, where the GNSS antenna provides the platform trajectory, the stereo rig defines the geometric relationship between sensors operating in different media, and the lever-arm and boresight parameters allow the projection centres of each camera to be computed in a common reference system.

3.5 Camera Synchronization and Time Referencing (Geotagging)

In the proposed system, synchronization is achieved using GNSS time as a common temporal reference for all sensors. This ensures temporal consistency between GNSS positioning, optical image acquisition, and acoustic measurements, and allows all sensors to be integrated within a single georeferencing framework.

For the GoPro cameras, synchronization is performed using the experimental GoPro Labs firmware, which allows the internal camera clock to be synchronized with GNSS time whenever a GPS signal is available. The cameras are powered on and allowed to acquire GNSS time before being deployed underwater, and again after the survey when the cameras are brought back above the water surface. Although the GNSS signal is not available while the camera is underwater, the internal clock drift over the relatively short acquisition time can be considered negligible for the intended accuracy level. Therefore, image timestamps can be referenced to GNSS time and used for trajectory interpolation.

In addition to standard action cameras, a 360° camera (Insta360) was also tested. In this case, images were extracted from the recorded video stream rather than acquired as individual photographs. The video metadata include GNSS positioning information, which can be exported as a GPX trajectory. The workflow therefore consists of extracting individual frames from the video, extracting the corresponding GNSS trajectory in GPX format, and synchronizing frames and trajectory using their timestamps. The position of each extracted frame is then computed by interpolating the GNSS trajectory at the frame acquisition time, similarly to the workflow used for the GoPro images. Once synchronization is established, the coordinates of the camera projection centres are computed by interpolating the GNSS trajectory at the image acquisition time and applying the lever-arm offsets between the GNSS antenna and the camera optical centre.

This synchronization and interpolation workflow is implemented using Python scripts and metadata processing tools (e.g., ExifTool) to assign interpolated GNSS coordinates to each image before photogrammetric processing.

3.6 Direct georeferencing workflow

The direct georeferencing workflow adopted in this study is based on the interpolation of the GNSS trajectory at the image acquisition time and the application of lever-arm corrections to determine the coordinates of the camera projection centers. For each acquired image i, the GNSS epochs surrounding the image acquisition time t_i are first identified, such that:

```
FOR each image i:

1. Identify GNSS epochs surrounding image time ti:
   t0 ≤ ti ≤ t1

2. Interpolate GNSS position at image time ti:
   Xi = X0 + ((ti − t0) / (t1 − t0)) * (X1 − X0)
   Yi = Y0 + ((ti − t0) / (t1 − t0)) * (Y1 − Y0)
   Zi = Z0 + ((ti − t0) / (t1 − t0)) * (Z1 − Z0)

3. Apply lever-arm correction:
   Xcam = Xi + dx
   Ycam = Yi + dy
   Zcam = Zi + dz

4. Assign projection centre to image i:
   PCi = (Xcam, Ycam, Zcam)

OUTPUT:
Camera projection centres for all images
→ used in photogrammetric bundle adjustment
```

Figure 2. – Algorithm for direct georeferencing based on GNSS trajectory interpolation and lever-arm correction.

$$t_0 \leq t_i \leq t_1$$

where $t_0$and $t_1$are the GNSS epochs immediately before and after the image acquisition time. The GNSS position at the image acquisition time is then computed through linear interpolation of the GNSS trajectory using the as summarized by Algorithm 1 (Fig. 2). For each image, the GNSS trajectory is interpolated at the image acquisition time and the lever-arm offset between the GNSS antenna and the camera is applied to obtain the camera projection centre coordinates. This procedure allows the reconstruction of the camera trajectory even when the camera operates underwater and GNSS signal is not available, provided that clock drift remains negligible over the acquisition time and that the rigid geometric calibration between sensors is preserved. This assumption is reasonable considering the short acquisition time, the low platform dynamics, and the synchronization of the camera clocks with GNSS time before and after the survey.

4. Experiment

The experimental activity was designed to test the proposed modular USV-based multi-sensor system and to evaluate the performance of the direct georeferencing workflow for integrated above- and underwater photogrammetric surveys in shallow coastal environments. The main experimental site (Fig. 3) was located in the rocky coastal area of L'Argentiera, near Sassari in north-western Sardinia (Italy). L'Argentiera is a former mining site and is currently included among the sites of community importance and within the Historical and Environmental Geomining Park of Sardinia. The area is currently subject to environmental restoration and monitoring activities due to its geological, geomorphological, and ecological relevance. For these reasons, the site represents an ideal test bed for the developed USV-based survey system, as it requires periodic monitoring of both the emerged rocky cliffs and the shallow underwater environment.

The Argentiera site was selected for accuracy assessment and for integrated above- and underwater survey experiments due to the presence of shallow water, rocky seabed, and good water transparency conditions, which allow both underwater photogrammetry and aerial photogrammetry to be performed in the same area. The experimental setup included ground truth measurements acquired using a tilt-compensating RTK GNSS

(CHCNAV i89) receiver mounted on a survey pole to measure reference points on the seabed, and an UAV survey to provide a reference model of the coastal area. These datasets were used to evaluate the accuracy of the integrated photogrammetric survey performed with the USV platform.

The data were processed using the official Italian datum (ETRS89-ITA [RDN2008] – EPSG:6706), and the ellipsoidal heights were converted to orthometric heights using the GK2 grids, which implement the ITALGEO2005 geoid model.

4.1 Data acquisition

Ground truth data were acquired using a pair of CHCNAV i89 GNSS receivers operating in base and rover RTK configuration. One receiver was installed over a fixed reference point and used as a base station transmitting RTK corrections, while the second receiver was used as a rover mounted on a survey pole to measure reference points on both the seabed in shallow water and on the emerged coastal area. All RTK measurements were acquired with estimated centimetric accuracy (approximately 1–2 cm in planimetry and 2–3 cm in height). A set of 8 artificial targets was placed on the seabed in shallow water (depth between approximately 0.5 m and 2 m) and measured using RTK GNSS. These points were used as reference bathymetric points for accuracy assessment.

The underwater photogrammetric survey was performed using the USV platform in fully autonomous navigation mode. The acquisition leveraged RTK positioning integrated within the USV autopilot, enabling accurate trajectory control and repeatable survey patterns. The mission was planned to ensure sub-centimetric ground sample distance (GSD), with an average overlap of 80% and an average sidelap of 70% between adjacent survey lines. This resulted in a dense and regular image acquisition geometry, suitable for high-resolution photogrammetric reconstruction in shallow water, as illustrated in Fig. 4.

The UAV survey was conducted using a DJI Mini 2 to acquire images of the emerged coastal cliff and surrounding area. Ground targets placed on the emerged area were measured using RTK GNSS and used to georeference the UAV photogrammetric model in the same reference system (residuals on the measured reference points are equal to 1.2 cm in the North direction, 0.8 cm in the East direction and 1.8 cm in the vertical direction).

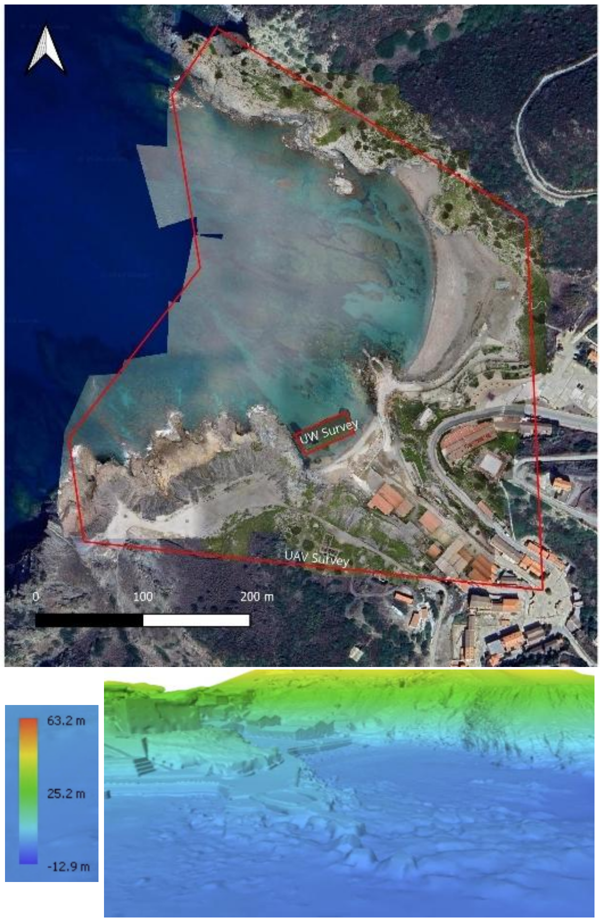

Figure 3. Left: UAV and USV survey coverage overlaid on a satellite image, with corresponding orthophotos (UAV: ~500 × 400 m; USV: 50 × 20 m). Right: height-colored mesh derived from UAV photogrammetry of the emerged area.

Due to the good water transparency conditions at the Argentiera site, the UAV photogrammetry also allowed partial reconstruction of the submerged seabed down to depths of approximately 4–5 m, providing an additional reference dataset for comparison with the underwater photogrammetric survey acquired from the USV platform.

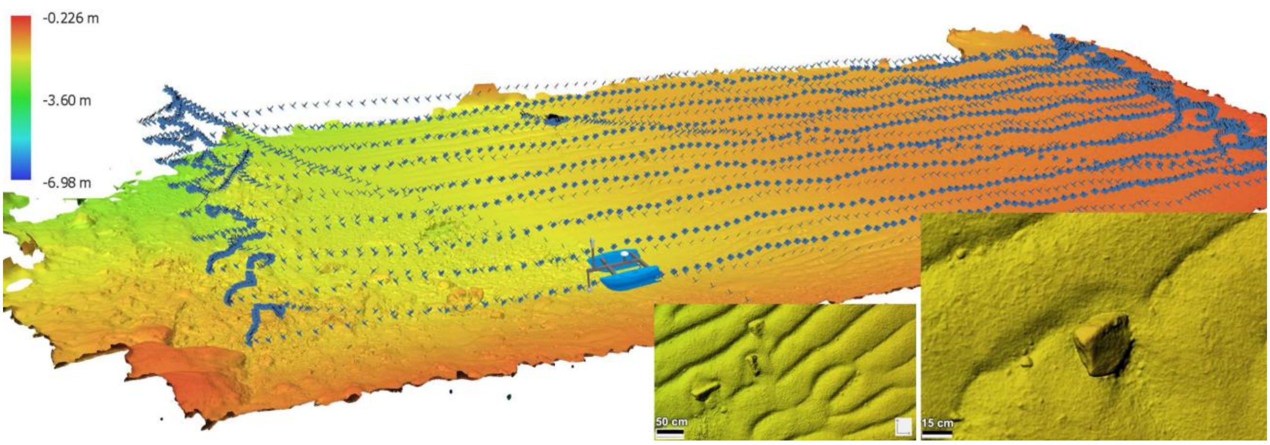

Figure 4. Depth-colored mesh derived from underwater GoPro imagery acquired from the USV, with camera positions shown along survey tracks. The regularity of the acquisition lines was ensured by RTK-based navigation implemented on the USV autopilot. A detail of the reconstructed seabed is shown in the bottom-right corner to highlight the spatial resolution achieved.

4.2 Accuracy assessment

The accuracy of the proposed direct georeferencing approach was evaluated using the RMSE, computed as the difference between the 3D coordinates of independent check points obtained from photogrammetry and RTK GNSS, respectively. These points were used to assess the absolute accuracy of the underwater photogrammetric model derived from the USV survey using a GoPro camera.

The results indicate that the model obtained through direct georeferencing achieves an absolute positioning accuracy on the order of a few centimetres (Table 1)

$RMSE_X$ $(East\ direction)$	$RMSE_Y$ $(North\ direction)$	$RMSE_Z$ $(Vertical\ direction)$
2.3 cm	1.5 cm	2.4 cm

Table 1. RMSE of underwater photogrammetry targets coordinates from USV compared with the RTK GNSS measured coordinates.

The residuals of the similarity transformation with scale factor between the GNSS geotagged camera positions and their respective coordinates from the bundle adjustment computed in free network in Agisoft Metashape resulted in a $RMSE_{X,Y,Z}$ for the GoPro of 2.9, 2.8, 4.8 cm respectively. These values are slightly higher than the values on the checkpoints. The greater errors for the camera positions are likely explained by the fact that the checkpoint distribution could be only carried out over a smaller area, e.g. the checkpoint could not be placed on the deepest part. While the horizontal error aligns with the expected accuracy of RTK GNSS positioning, the larger vertical error warrants further investigation. This discrepancy may be attributed to systematic uncompensated errors such as the absence of a rigorous correction for refraction effects at the air–water interface of the GoPro imaging system. Over an extended survey area, such effects can lead to the accumulation of vertical errors in the reconstructed model.

Moreover, a qualitative comparison was also performed between the underwater photogrammetric model acquired from the USV and the aerial photogrammetric model derived from the UAV survey in the emerged and shallow submerged areas by examining the elevation profiles extracted from the models obtained using the two approaches (Fig. 5): a good geometric consistency is shown between two submerged models, although some residual deviations might be attributed to incomplete compensation of refraction effects in the UAV-derived model.
The photogrammetric reconstruction of the emerged area shows a reasonable level of geometric consistency with the underwater model, suggesting that the adopted stereo-rig configuration and georeferencing framework can provide coherent results across the air–water interface. However, the above-water model exhibits a comparatively coarse level of geometric and radiometric detail.

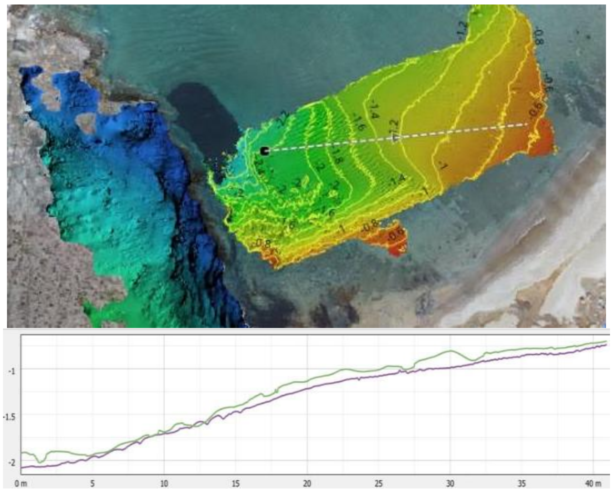

Figure 5. Top: depth-colored underwater and above the water model derived from USV-based photogrammetry overlaid on the orthophoto, with the location of the extracted underwater transect. Bottom: comparison between elevation profiles extracted along the same transect from the USV-derived model (purple) and the UAV-derived model (green), showing a generally consistent trend. Residual discrepancies are likely associated with incomplete compensation of refraction effects in the UAV-derived model.

This limitation is mainly attributed to the quality of the acquired images, as the action cameras used in the current configuration are not optimized for high-resolution terrestrial photogrammetry. A visual comparison between the reconstruction obtained from the Insta360 camera and the UAV-based photogrammetry is shown in Fig. 6. The reconstruction derived from the Insta360 images appears noticeably coarser, with reduced geometric detail and radiometric quality. This difference is mainly attributable to the characteristics of the imaging system and acquisition modality.

5. Conclusions

This study investigates a modular multi-sensor system for GNSS-supported photogrammetric surveying in shallow coastal environments, with particular focus on integrated surveys across the air–water interface. The experimental results indicate that the proposed workflow allows underwater photogrammetric data to be georeferenced with a few centimeters absolute accuracy using differential GNSS positioning and a calibrated geometric relationship between the GNSS antenna and the camera system, without the need for underwater ground control points.

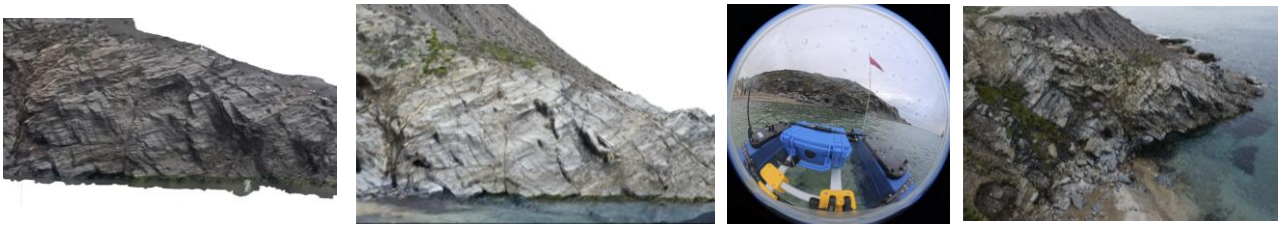

Figure 6. Side-by-side comparison between Insta360-based and UAV-based photogrammetric results. From left to right: (a) detail of the reconstruction derived from Insta360 images, (b) detail of the reconstruction derived from UAV images, (c) Insta360 input image, and (d) UAV input image.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-955-2026

A key aspect of the system is the use of a USV as the primary acquisition platform. The USV enables controlled and repeatable data acquisition in very shallow and nearshore environments that are difficult or unsafe to access with traditional vessels, while maintaining a stable geometric configuration between sensors. This makes it particularly effective for close-range underwater photogrammetry and for integrated surveys at the air–water interface. At the same time, the experimental results highlight some limitations of the current configuration, particularly with respect to the reconstruction of the emerged area. While the above- and underwater models show a reasonable geometric consistency, the photogrammetric reconstruction of the emerged part is comparatively coarse, mainly due to the quality and characteristics of the images acquired with action cameras. This aspect represents one of the main challenges identified in this first experimental campaign. This aspect will be further investigated in future work by evaluating the use of higher-resolution or larger-format cameras for the above-water acquisition. Moreover, this limitation can be effectively mitigated by integrating UAV-based photogrammetry, which provides higher-resolution and radiometrically richer data for the emerged coastal area, thus complementing the USV-based acquisition.

The modular architecture of the system allows the sensor configuration to be adapted to different survey requirements. In particular, the above-water camera can be replaced with higher-quality sensors to improve the reconstruction of the emerged part, while maintaining the same georeferencing framework. More broadly, although this study focuses on rocky coastal environments, the proposed system can be easily adapted to other shallow coastal settings, including low-relief environments and applications such as seagrass (e.g., *Posidonia oceanica*) monitoring, where high-resolution mapping of the seabed is required.

Future developments will focus on strengthening the integration between optical and acoustic sensors, in particular through the use of side-scan sonar (SSS), to support the generation of more complete and robust seabed models. Further work will also address the optimization of the acquisition workflow and the integration of multi-sensor data within a unified coastal mapping framework.

Acknowledgements

EN declares financial support was received from the National Biodiversity Future Center funded by the Italian Ministry of University and Research, PNRR, Missione 4, Componente 2, "Dalla ricerca all'impresa", Investimento 1.4 Project CN00000033. This work was supported by e.INS- Ecosystem of Innovation for Next Generation Sardinia (cod. ECS 00000038) funded by the Italian Ministry for Research and Education (MUR) under the National Recovery and Resilience Plan (NRRP) – MISSION 4 COMPONENT 2, "From research to business" INVESTMENT 1.5, "Creation and strengthening of Ecosystems of innovation" and construction of "Territorial R&D Leaders". The authors acknowledge the use of ChatGPT (OpenAI) for assistance with English language editing and improving the clarity of the manuscript.

References

Cao, B., Gao, Z. and Shang, W., 2023. Methodology for fine-scale seamless elevation model construction using unoccupied aerial and surface vehicles. *Estuarine, Coastal and Shelf Science*, *282*, p.108229.

Casella, E., Scicchitano, G. and Rovere, A., 2024. Accuracy and Precision of Shallow-Water Photogrammetry from the Sea Surface. Remote Sensing, 16(22), p.4321.

Furlani, S. and Antonioli, F., 2023. The swim-survey archive of the Mediterranean rocky coasts: Potentials and future perspectives. Geomorphology, 421, p.108529.

Hatcher, G. A., Warrick, J. A., Ritchie, A. C., Dailey, E. T., Zawada, D. G., Kranenburg, C., & Yates, K. K. (2020). Accurate bathymetric maps from underwater digital imagery without ground control. *Frontiers in Marine Science*, *7*, 525.

Hatcher, G.A., Warrick, J.A., Kranenburg, C.J. and Ritchie, A.C., 2023. Accurate maps of reef-scale bathymetry with synchronized underwater cameras and GNSS. *Remote Sensing*, *15*(15), p.3727.

Jaud, M., Delsol, S., Urbina-Barreto, I., Augereau, E., Cordier, E., Guilhaumon, F., Le Dantec, N., Floc'h, F. and Delacourt, C., 2024. Low-tech and low-cost system for high-resolution underwater RTK photogrammetry in coastal shallow waters. *Remote Sensing*, *16*(1), p.20.

Karaki, A.A., Ferrando, I., Federici, B. and Sguerso, D., 2024. Bridging Disciplines with Photogrammetry: A Coastal Exploration Approach for 3D Mapping and Underwater Positioning. Remote Sensing, 17(1), p.73.

Lo, E., Lozano Bravo, H., Hui, N., Nocerino, E., Menna, F., Rissolo, D., Kuester, F., 2024. Evaluation of the Accuracy of Photogrammetric Reconstruction of Bathymetry Using Differential GNSS Synchronized with an Underwater Camera. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-1/W1-2024, 211–218.

Menna, F., Nocerino, E., Remondino, F., 2018. Photogrammetric modelling of submerged structures: Influence of underwater environment and lens ports on 3D measurements. In: *Latest Developments in Reality-Based 3D Surveying and Modelling*, MDPI, 279–303.

Menna, F., Nocerino, E., Troisi, S., Remondino, F., 2013. A photogrammetric approach to survey floating and semi-submerged objects. In: *Videometrics, Range Imaging, and Applications XII*, Proc. SPIE 8791, 87910.

Moisan, E., Charbonnier, P., Foucher, P., Grussenmeyer, P., Guillemin, S., Koehl, M., 2015. Adjustment of Sonar and Laser Acquisition Data for Building the 3D Reference Model of a Canal Tunnel. *Sensors*, 15(12), 31180–31204.

Nocerino, E., Menna, F., 2020. Photogrammetry: Linking the world across the water surface. *J. Marine Sci. Eng.*, 8(2), 128.

Nocerino, E., Menna, F., Farella, E., Remondino, F., 2019. 3D Virtualization of an Underground Semi-Submerged Cave System. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W10, 857–864.

Van Der Lucht, J., Bleier, M., Leutert, F., Schilling, K., Nüchter, A., 2018. Structured-Light Based 3D Laser Scanning Of Semi-Submerged Structures. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2, 287–294.