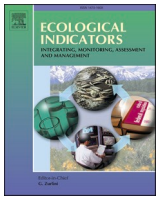
Ecological Indicators 190 (2026) 115335



Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind





Interchangeability of structure-from-motion platforms in marine monitoring

Jessica L.Rhodes [a,*], Russell B.Connelly [a], Boyd A.McKew [a], Jon Chamberlain [b]

[a] School of Life Sciences, University of Essex, CO4 3SQ, United Kingdom
[b] School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ, United Kingdom

ARTICLE INFO

Dataset link: Images and 3D Reconstructed Models (.OBJ & .MTL) of the Cromer Shoal Chalk Reef (Original data)

Keywords:
Structure-from-motion
Benthic monitoring
Chalk reefs
Software consistency
Computational efficiency
Habitat complexity

ABSTRACT

Structure-from-Motion (SfM) photogrammetry has revolutionised marine monitoring, however rapid software evolution challenges long-term data continuity. Across longitudinal datasets, changing reconstruction software may introduce artificial variability into ecologically relevant indicators. Establishing the consistency of 3D structural complexity indices is essential for detecting subtle changes preceding ecosystem collapse, especially in the globally rare chalk reefs of the North Sea (UK), where habitat loss is irreversible. This study evaluates the methodological interchangeability, indicator reproducibility, and computational efficiency of two leading SfM platforms, *Agisoft Metashape* and *3DF Zephyr*, using Remotely Operated Vehicle (ROV) imagery from the Cromer Shoal Chalk Beds.

Three complexity metrics (rugosity, relief, and vector dispersion) were compared across five software configurations representing a gradient of computational effort. Macro-topographical indicators such as rugosity remained statistically identical across platforms (GLMM, $p > 0.05$), though micro-textural outputs such as vector dispersion were significantly more sensitive to software-specific algorithms. High-resolution configurations converged on statistically similar ecological values at vastly different computational costs. *3DF Zephyr* exhibited greater mesh generation efficiency than *Agisoft Metashape* ($F_{4,50} = 4.83$, $p = 0.002$), whereas *Agisoft Metashape* at High setting generated significantly denser point-clouds than *Zephyr* Advanced ($F_{4,50} = 31.27$, $p < 0.001$) by an average of 3.8×10^7 points (Tukey HSD: $p < 0.001$).

These findings demonstrate that broad-scale structural indicators remain robust to software and pipeline changes, ensuring long-term data continuity. Nonetheless, the observed shifts in micro-textural metrics may reflect software artefacts rather than genuine ecological decline, highlighting the need for strict parameter standardisation across multi-year monitoring objectives. For conservation programmes, this methodological consistency enables practical transitions between software platforms to optimise field time and processing budgets.

1. Introduction

Topographical complexity is a key structural characteristic of benthic marine ecosystems that drives both biodiversity and function of marine communities. Therefore, it has become essential to be able to accurately quantify topographical complexity related metrics, to help understand micro- to ecosystem-scale processes in the marine environment, particularly within highly complex environments such as coral reefs (McCoy and Bell, 1991; Graham and Nash, 2013; Remmers et al., 2024). These metrics serve as a robust proxy for key ecological indicators, including habitat resilience, disturbance regimes, and biological community structure, such as species diversity and abundance (Bayley and Mogg, 2020). Technological advancements in remote sensing, specifically Structure-from-Motion (SfM) photogrammetry, have revolutionised our capacity to observe and quantify these environments, enabling detailed visualisation of previously inaccessible underwater habitats at unprecedented resolutions and spatial scales (Anderson et al., 2019).

Within this field, *Agisoft Metashape* has become a preferred commercial tool for benthic mapping due to its transparent parameter controls and consistent, detailed output (Conley and Hollander, 2021; Vincent et al., 2021). *3DF Zephyr* is a competing commercial option

* Corresponding author.
E-mail address: j.rhodes@essex.ac.uk (J. L.Rhodes).

https://doi.org/10.1016/j.ecolind.2026.115335
Received 19 June 2026; Received in revised form 11 August 2026; Accepted 14 August 2026
Available online 18 August 2026
1470-160X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

that generates reconstructions from both still imagery and video, using automated frame selection to optimise input data (Budău et al., 2023). While both platforms are widely used, there remains limited research evaluating their consistency and interchangeability in ecologically meaningful terms, particularly regarding topographic complexity when applied to non-tropical reef systems.

1.1. The geomorphology of chalk reef habitats

Software validation is essential in chalk reef habitats, which are among the rarest shallow marine environments globally. Unlike tropical coral reefs, where structural complexity is built through biological accretion of Scleractinia coral, the topography of marine chalk ecosystems is primarily erosional. Formed over millennia through the layered deposition of calcareous ooze (Savrda, 2012), these rocky reefs are shaped by continuous water movement that carves the softer substrate into a complex landscape of gullies, arches, and ledges (Buatois, 2003; Moffat et al., 2020). This eroded complexity provides functional niches for spawning and refuge, yet it is exceptionally fragile; as chalk cannot regenerate, any structural loss is permanent. This unique geomorphology poses a specific challenge for SfM monitoring: standard metrics, such as rugosity, and software parameters tailored for fine-scale branching structures of tropical corals may not accurately capture the broader topographical shifts characteristic of temperate rocky seabeds (Wright and Chamberlain, 2024). To address this, specialised computational metrics, such as relief, have been developed to quantify the depth and volume of these gully features (Wright and Chamberlain, 2024). As SfM validation efforts expand beyond tropical baselines (Lochhead and Hedley, 2022; Oliveira et al., 2022), it is imperative to include vulnerable habitats like chalk reefs, where structural degradation is irreversible.

A primary example of this unique and vulnerable geomorphology is the Cromer Shoal Chalk Beds (CSCB). This Marine Conservation Zone (MCZ), covering a total area of 320.5 km^2 along the North Norfolk coast, UK, is characterised by complex formations of flint, clay, carstone, and chalk outcrops that provide essential habitat for diverse benthic communities and commercially vital crustacean fisheries (Jackson et al., 2022). The gully features of this site drive local biodiversity and ecosystem resilience, yet they are difficult to quantify using traditional metrics (Wright and Chamberlain, 2024). Accurate monitoring is critical for assessing the conservation status of these protected areas; however, the subtidal nature and high-turbidity conditions of the North Sea limit diver-led surveys, necessitating the utilisation of Remotely Operated Vehicle (ROV)-based SfM photogrammetry.

While providing greater spatial coverage, as divers are limited by slack water windows (often only 30–60 min in the North Sea), the transition to ROV surveying introduces distinct photogrammetric challenges. Specifically, the increased velocity of ROV transit compared to SCUBA swimming speeds can produce image motion blur, which risks the smoothing of fine-scale topographic features during model reconstruction (Paterson et al., 2024). Furthermore, inconsistencies between filming speed and frame-extraction rates can lead to sub-optimal image overlap, a critical factor for the alignment accuracy of complex benthic surfaces (Marre et al., 2019). Consequently, the utility of ROV-derived models is inherently tied to the software's ability to navigate these platform-specific artefacts, highlighting the need for rigorous software validation.

1.2. Metric sensitivity

The reliability of 3D reconstructions is fundamentally contingent upon the mathematical sensitivity of the metrics applied to measure the digital surface (Young et al., 2017; Marre et al., 2019; Figueira et al., 2015). To ensure these tools are accessible for ecological monitoring, Young et al. (2017) expanded the application of traditional complexity indices, such as rugosity and vector dispersion (Risk, 1972; Luckhurst and Luckhurst, 1978; Carleton and Sammarco, 1987) into standardised digital workflows. However, the performance of these metrics is dictated by their spatial scaling properties (Friedman et al., 2012; Figueira et al., 2015; Marre et al., 2019).

Rugosity and relief function as area-based averages (Friedman et al., 2012; Lazarus and Belmaker, 2021). As these indices obtain topographical data over a broader spatial footprint, they tend to be resilient to algorithmic noise or minor point-cloud artefacts (Friedman et al., 2012). In these calculations, the averaging process effectively dampens the influence of platform-specific variations; a minor error in a single coordinate would be statistically diluted across the wider surface area. Consequently, these macro-topographic metrics are expected to remain robust across different SfM algorithms.

In contrast, vector dispersion, which utilises the unit-vector calculation of surface normals, operates at a finer, micro-textural scale (Young et al., 2017). This metric evaluates the angular uniformity of every individual face within a mesh, making it sensitive to the geometry created during the meshing process. While an area-based metric may overlook a slightly smoothed surface, a unit-vector metric might record that smoothing as a significant decrease in structural complexity. Understanding the interchangeability of SfM platforms, therefore, requires not only a comparison of absolute outputs but an evaluation of how different complexity indices scale with algorithmic variance.

1.3. Software evolution and monitoring consistency

As marine monitoring programs transition into decadal time-series (Edwards et al., 2010; Benway et al., 2019; Valdés et al., 2021), software evolution introduces a significant methodological challenge (Forsmoo et al., 2019; Smith et al., 2016). Long-term ecological studies rely on the fundamental assumption that observed changes in metrics reflect true biological or physical shifts rather than artefacts of software updates or platform changes. This presents a risk to study continuity: if platforms interpret the same imagery differently, the resulting algorithmic noise could mask subtle signs of reef degradation or inaccurately suggest recovery. Establishing the interchangeability of these tools is therefore a requirement for the integrity of global reef monitoring. Without validating this continuity, we risk misinterpreting biological change as a digital artefact.

We address this by evaluating the methodological consistency and computational trade-offs involved when transitioning between two leading SfM platforms, *Agisoft Metashape* and *3DF Zephyr*. Using high frame rate ROV imagery from the CSCB temperate rocky reef, we systematically test whether ecological metrics, specifically those sensitive to fine-scale topographic change, remain robust to algorithmic shifts or are significantly altered by software-specific processing. Furthermore, we quantify the cost of accuracy by assessing the relationship between computational effort (processing time) and the resulting structural output (point and mesh density). By establishing these benchmarks, this work provides a quantitative baseline to inform software selection and workflow optimisation, ensuring that digital reconstructions remain reliable methods to monitor ecological change.

We address the following research questions:

***RQ1*:** Do topographical metrics (rugosity, vector dispersion, and relief) remain consistent when transitioning between software platforms?

***RQ2*:** Does *3DF Zephyr* offer significant computational advantages over *Agisoft Metashape* when processing high frame rate ROV video imagery?

***RQ3*:** How do quality setting configurations impact the reconstructive reliability of reef models?

2. Materials and methods

To provide an overview of the experimental approach, Fig. 1 outlines the methodological workflow and processing pipeline, detailing the

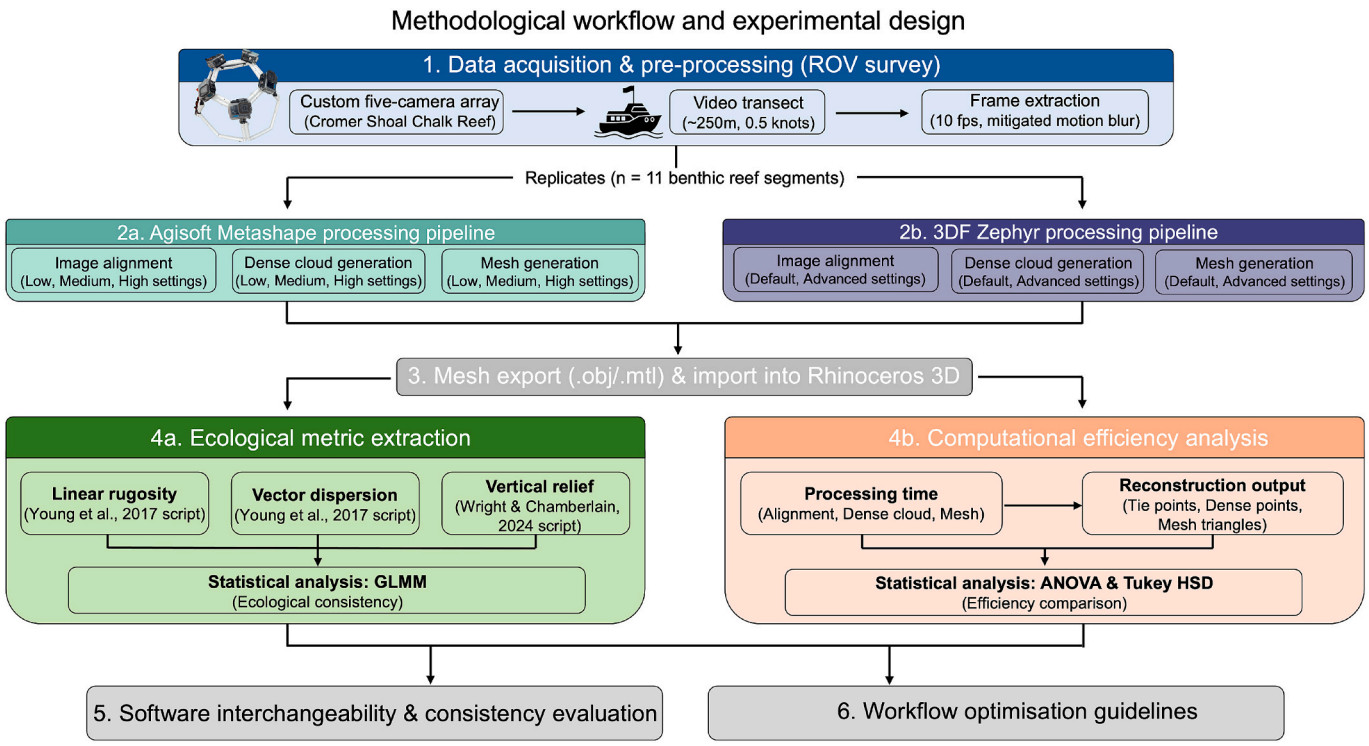


Fig. 1. Methodological workflow and experimental design illustrating the sequential stages of the study, from action camera video acquisition (using a custom five-camera array attached to an ROV) through parallel processing pipelines in *Agisoft Metashape* and *3DF Zephyr,* to mesh export, ecological metric extraction, and computational efficiency analyses.

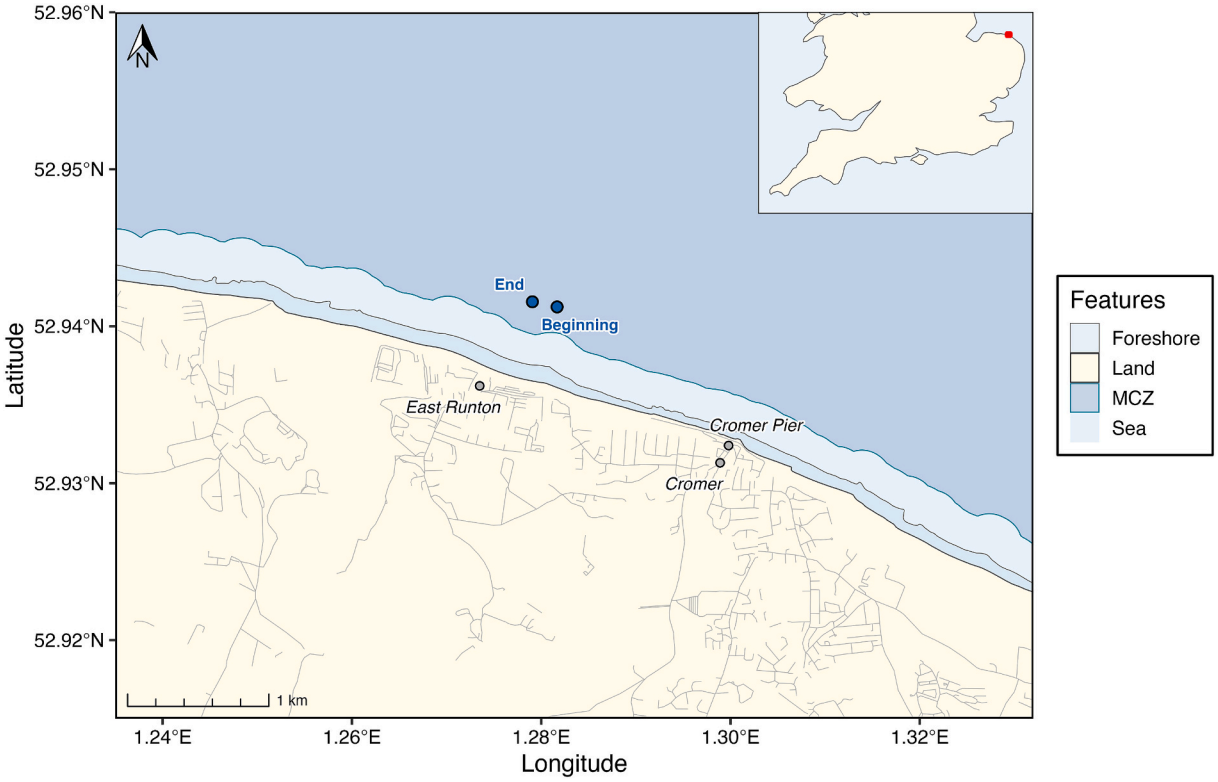


Fig. 2. Map of the start and end points of the 250 m benthic recording taken on the Cromer Shoal chalk reef, using a five-camera setup attached to a Remotely Operated Vehicle (ROV; BlueROV2) that was submerged from a moving boat. Points were recorded with an UnderWater Information System (UWIS) transponder attached to the ROV. Blue dots indicate the locations of the start (East) and end (West) points of the survey. The land, sea, foreshore, and Cromer Shoal MCZ are overlaid in yellow, light blue, outlined blue, and dark blue, respectively. The red circle within the inset map designates the wider sampling location of North Norfolk, East Anglia, UK. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

progression from action camera data acquisition and multi-software configuration through to structural complexity metric extraction and statistical evaluation.

2.1. Data acquisition and image pre-processing

Benthic imagery was acquired from the Cromer Shoal chalk reef (Fig. 2) on September 15th, 2023 (early autumn). Survey location was chosen in direct collaboration with the Eastern Inshore Fisheries and Conservation Authority (EIFCA) to align with their established regional monitoring framework within the MCZ. Data collection utilised a custom five-camera array (Wright and Chamberlain, 2024);Fig. S1) mounted to a Remotely Operated Vehicle (ROV; BlueROV2) and deployed from a surface vessel. The ROV maintained a consistent altitude of 1 m ± 10 cm above the seabed at a mean depth of 7 m ± 50 cm with a towing speed of approximately 0.5 knots throughout the ~250 m survey.

The multi-camera configuration was chosen to overcome inherent visual and geometric constraints associated with ROV-based video capture. Unlike stationary or diver-operated setups, vessel-towed ROVs operate at higher transit speeds, which restricts the field of view and limits sequential image overlap from a single downward-facing lens. The five-camera array was configured to capture the seabed simultaneously from sideways and downward-facing views (Fig. S2). This multi-directional image capture ensures the high degree of image overlap and intersection of angles required to accurately reconstruct complex, shadow-prone topographies (such as the gullies, overhangs, and ledges characteristic of temperate chalk reefs) while compensating for the rapid motion of the platform.

Temporal synchronisation across individual camera recordings was achieved prior to deployment by capturing a distinct, upward-facing motion signal at the start of the recording session before ROV submersion. This shared movement signal created a distinct temporal stamp in the video frames, enabling precise post-collection alignment and frame-level synchronisation across all recordings.

The video recordings captured by the five-camera configuration were subsequently converted into discrete frames at a rate of 10 frames per second (fps). This extraction frequency is significantly higher than the 1–5 fps extraction rates often used in comparable surveys, reflecting the increased velocity of the ROV and implemented to mitigate motion blur. Extracted frames from the ~250 m survey were sub-sectioned into 11 discrete replicates; each comprising 175 images, representing a ~ 10 × 4 m segment of the benthic environment.

2.2. Photogrammetric processing

To transition from raw visual imagery to structural complexity metric analysis, individual video frames were processed via Structure-from-Motion (SfM) photogrammetry, converting 2D pixel arrays into high-resolution 3D point clouds and meshes from which continuous numerical metrics were extracted.

All photogrammetric reconstructions were processed on a virtual machine provided by the School of Computer Science and Electronic Engineering (CSEE) at the University of Essex, ensuring consistent computational power across all setting parameters (Horizon computer specifications provided in Table S1). To evaluate the performance and potential interchangeability of *Agisoft Metashape* (v1.7.5.13) and *3DF Zephyr* (v7.51), 3D model replicates were generated using five configurations (Table 1).

These configurations, comprising three tiers in *Agisoft Metashape* and two in *3DF Zephyr*, were selected to represent a gradient of computational effort. This experimental design allowed for a direct comparison of whether lower-tier settings in one software could achieve statistical similarity with higher-tier settings in the other. For each of the 11 reef replicates, the standardised post-survey workflow described by Bayley and Mogg (2020) was applied across all five configurations, resulting in a total of 55 high-resolution 3D models for downstream structural analysis.

2.3. Quantification of structural complexity

To provide a comprehensive assessment of reef topography, three distinct structural complexity metrics were extracted from each 3D reconstruction. These metrics were selected to capture varying scales of topographic detail: Linear rugosity (R) as a standard measure of surface complexity; vector dispersion $\left(\frac{1}{k}\right)$ to quantify fine-scale angular uniformity of the 3D mesh; and relief to account for the vertical depth and gully structures characteristic of the North Norfolk chalk beds, which traditional complexity indices often underestimate (Wright and Chamberlain, 2024). Together, these metrics capture complementary, distinct dimensions covering broad macro-topography to fine-scale mesh orientation. Alternative descriptors, such as fractal dimension and surface roughness, were intentionally excluded to avoid metric redundancy and to prevent resolution-dependent software artefacts common in video-derived benthic reconstructions (Fukunaga and Burns, 2020; Fukunaga et al., 2020; Oh et al., 2025).

All measurements were conducted within the Rhinoceros 3D (v7; (McNeel, 2010)) environment using the 3D model replicates exported from *Agisoft Metashape* and *3DF Zephyr* in .obj / .mtl format. To ensure spatial consistency across replicates, measurements were aligned by sampling the central point of the 3D model and the helper-file of each complexity metric. The following commands were used for alignment: 'AreaCentroid', 'Move', 'Point', '0', to place both objects at 0,0,0 within the x,y,z coordinates of the *Rhino* workspace.

- **Linear rugosity** was determined by a virtual chain-and-tape method, where the custom Young et al. (2017) Rhino Python script returns R as the distance of the virtual chain along the mesh contour divided by the total length of the virtual chain.

Table 1
Processing parameters of the three Structure-from-Motion (SfM) workflow stages (alignment, dense cloud, and mesh generation) used to generate 3D reconstructions of 11 Cromer Shoal Chalk Reef (North Norfolk, UK) segments (10 × 4 m). The three software quality settings within *Agisoft Metashape* (v1.7.5.13) and two within *3DF* Zephyr (v7.51) represent a gradient of computational effort.

Software	Setting ID	Alignment parameters				Reconstruction parameters		Mesh creation parameters	
		Accuracy				Depth maps quality	Depth maps filtering mode	Reconstruction surface type	Reconstruction source data
Metashape	Low	Low				Low	Mild	Arbitrary	Dense cloud
Metashape	Medium	Medium				Medium	Moderate	Arbitrary	Dense cloud
Metashape	High	High				High	Aggressive	Arbitrary	Dense cloud
		Key points sensitivity	Matching type	Matching stage depth	Tie-points quantity	Image resolution	Noise filtering	Watertightness	Image resolution
Zephyr	Default	Medium	Default	Medium	High	50%	10%	10%	25%
Zephyr	Advanced	High	Accurate	High	Very high	70%	20%	50%	50%

- **Vector dispersion** was obtained with a second Young et al. (2017) Rhino Python script, which is calculated to the same resolution as *R* (1 cm), and has the same scaling system (0–1), where 1 indicates a highly complex surface, and 0 represents a flat plane.

$$Dense\ Cloud\ Efficiency = \frac{Number\ of\ Dense\ Cloud\ Points \div Number\ of\ Tie\ Points}{\mathrm{Td}\ (s)}$$

- **Vertical relief** was then calculated as the maximum vertical distance between the highest and lowest mesh vertices within the sample area, utilising the Wright and Chamberlain (2024) Python script.

$$Mesh\ Efficiency = \frac{Number\ of\ Mesh\ Triangles \div Number\ of\ Dense\ Cloud\ Points}{\mathrm{Tm}\ (s)}$$

Data collection yielded a nested sampling design per replicate, consisting of one relief measurement, six rugosity values, and 64 vector dispersion readings.

2.4. Statistical analysis

All analyses were performed in R (v4.5.2; (R Core Team, 2025)) with significance defined at $\alpha = 0.05$. To evaluate the impact of software selection and processing intensity on reef structural metrics, rugosity, vector dispersion, and relief were compared using a mixed-effects modelling framework. To accommodate the nested data structure and non-normal distributions, Generalised Linear Mixed Models (GLMMs) were implemented using the *lme4* (Bates et al., 2015) and *glmmTMB* (Brooks et al., 2017) packages. All raw data points were fed into the GLMM with replicate set as a random intercept. A Gamma distribution (log link) was applied to relief data, while rugosity and vector dispersion were modelled using a Beta distribution (logit link). For Beta models, data were transformed following Smithson and Verkuilen, 2006 to accommodate values bounded strictly between 0 and 1. For all models, the software setting configuration group was treated as a fixed effect, with reef replicate included as a random effect to account for inherent topographic variation between reef segments. Post-hoc pairwise comparisons were conducted using the *emmeans* package (Lenth et al., 2024) with Tukey-adjusted *p*-values and back-transformation of model estimates to the response scale.

Computational efficiency was evaluated by comparing processing durations and reconstruction outputs across the five software configurations. Differences in absolute processing time and total counts (tie points, dense points, and mesh triangles) were analysed using a one-way analysis of variance (ANOVA), followed by Tukey's Honest Significant Difference (HSD) post-hoc tests. Efficiency was further quantified as the rate of data production relative to the input available at the start of each stage, calculated as follows:

1. **Alignment Efficiency:** The number of tie points generated per second of alignment time (T_a):

$$Alignment\ Efficiency = \frac{Number\ of\ Tie\ Points}{Ta\ (s)}$$

2. **Dense Cloud Efficiency:** The ratio of dense cloud points produced per existing tie point, divided by the generation time (T_d):

3. **Mesh Efficiency:** Defined as the ratio of mesh triangles produced per existing dense point, divided by the generation time (T_m):

The relationship between processing effort and reconstruction output was subsequently evaluated using linear regression to calculate coefficients of determination (R^2), identifying the predictability of software performance across tiers.

3. Results

3.1. Ecological metric consistency and software interchangeability

The comparison of structural complexity metrics (rugosity, relief, and vector dispersion) revealed a high degree of consistency across software platforms, particularly at higher processing tiers. Rugosity was the most robust metric, with no significant differences observed across any software or quality settings (GLMM Tukey-adjusted post-hoc, $p > 0.05$ for all pairwise comparisons; Fig. 3). Similarly, relief demonstrated consistency between the highest tiers of both platforms. While *Metashape* High and *Zephyr* Advanced both significantly outperformed lower-tier settings ($p < 0.05$), they remained statistically interchangeable with one another (shared significance group 'ab'; Fig. 3).

Vector dispersion exhibited the highest sensitivity to processing parameters. Although lower-tier settings significantly underestimated surface complexity ($p < 0.01$), measurements converged at the highest resolution. *Metashape* High and *Zephyr* Advanced shared overlapping statistical groupings, suggesting that at a 10 fps ROV sampling rate, both software packages reconstruct fine-scale reef texture with comparable reliability despite utilising different underlying algorithms.

3.2. Computational efficiency and processing applications

While the topographical complexity outputs remained stable, the computational cost to reach these results varied by several orders of magnitude. *Metashape* Medium required the longest alignment time (1325 ± 475 s), approximately 4.4-fold longer than *Zephyr* Advanced ($F_{(4,50)} = 7.35$, $p < 0.05$) (Fig. 4A), yet it generated 25.7-fold more tie points (Fig. 4B). This trend continued into dense cloud generation,

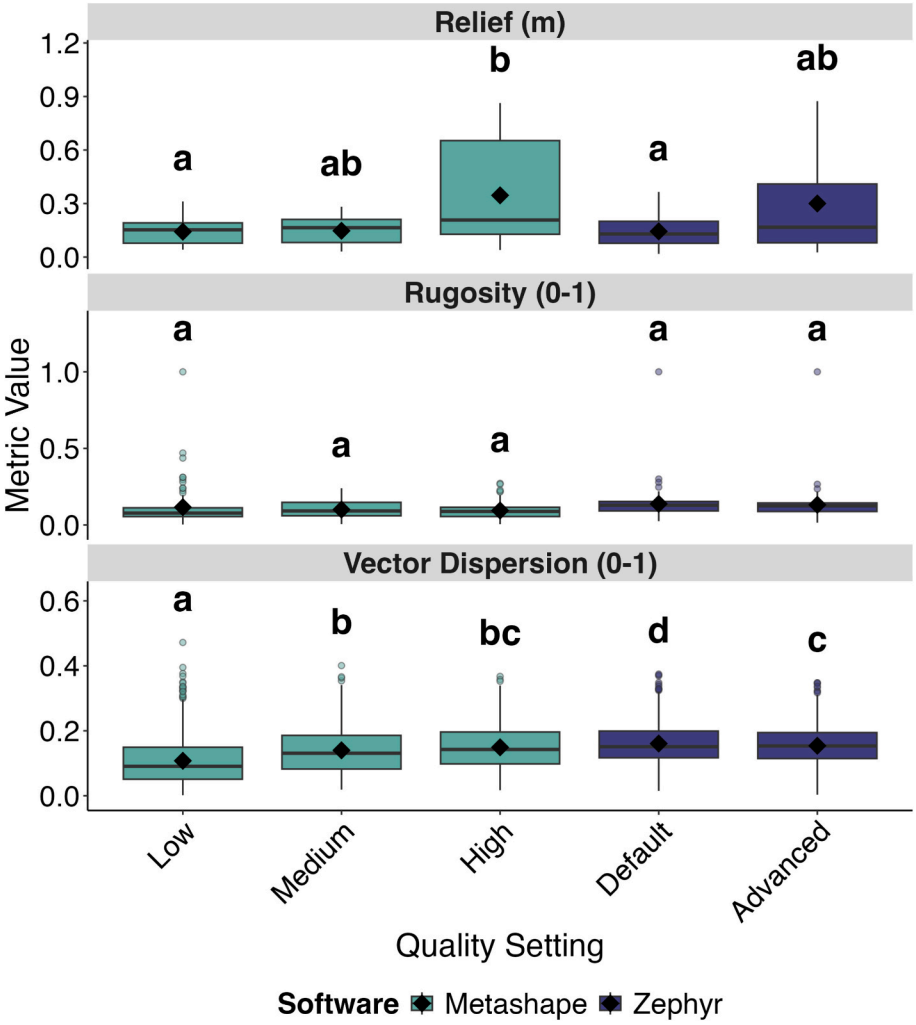


Fig. 3. The influence of Structure-from-Motion (SfM) software (*Agisoft Metashape* vs. *3DF Zephyr*) and processing quality settings on three key reef topographic metrics: Relief (m), Rugosity (0–1), and Vector Dispersion (0–1). Software settings represent progressive tiers of computational effort: Low, Medium, and High for Metashape, and Default and Advanced for Zephyr. Data was collected from the North Norfolk Cromer Shoal MCZ (UK) using a boat towed Remotely Operated Vehicle (ROV). Video imagery was extracted at 10 frames-per-second ($n = 11$ 3D model reef replicates per software setting configuration). Boxplots illustrate the median (centre line), the mean (black diamond), interquartile range (box), and 1.5 × IQR (whiskers), with open circles indicating individual outliers. Significance was determined using Generalised Linear Mixed Models (GLMM) with Beta (for Rugosity and Vector dispersion) and Gamma (for Relief) distributions to account for the nested structure and non-normality of the data. Letters above the bars denote statistically significant differences between groups based on Tukey's HSD post-hoc tests ($p < 0.05$). Groups sharing a common letter are not significantly different. Results indicate that while macro-metrics such as Rugosity remain stable across platforms, micro-metrics such as Vector Dispersion are highly sensitive to processing tiers, though they achieve statistical similarity at high-resolution settings (*Metashape* High and *Zephyr* Advanced).

where *Metashape* High produced the highest point density ($4.06 \times 10^7 \pm 9.95 \times 10^6$) (Fig. 4D), over 100 times more points than *Zephyr* Default, despite no significant increase in processing time over the *Metashape* Low tier (Fig. 4C).

Efficiency calculations (output per second) highlighted a distinct divergence in software strengths. *Metashape* High was significantly more efficient for point-based generation, yielding 315 ± 158 tie points per second (Fig. 5A) and over 41,000 dense points per second ($p < 0.05$) (Fig. 5B). Conversely, *Zephyr* Advanced demonstrated the highest relative mesh-processing efficiency, operating at 8570 ± 907 units per second (Fig. 5C). This performance was significantly more efficient than all other configurations (Tukey HSD, $p < 0.001$ for all pairwise comparisons), yielding a data-generation rate 1.8 times higher than the corresponding high-resolution configuration (*Metashape* High) and 339 times higher than the lowest-performing configuration (*Metashape* Low) (Fig. 5C).

3.3. Within-software variability

Regression analysis of the internal relationship between processing effort (time) and output (points/triangles) revealed high levels of stochasticity, particularly within *Metashape*. For the alignment stage, the processing time of *Metashape* High failed to explain any meaningful variation in tie-point yield ($R^2 = 0.001$), suggesting that longer alignment durations do not result in more points (Fig. S3). In contrast, *Zephyr* Advanced maintained a more predictable linear relationship for both alignment ($R^2 = 0.46$) (Fig. S3) and mesh generation ($R^2 = 0.92$) (Fig. S5). This suggests that while final

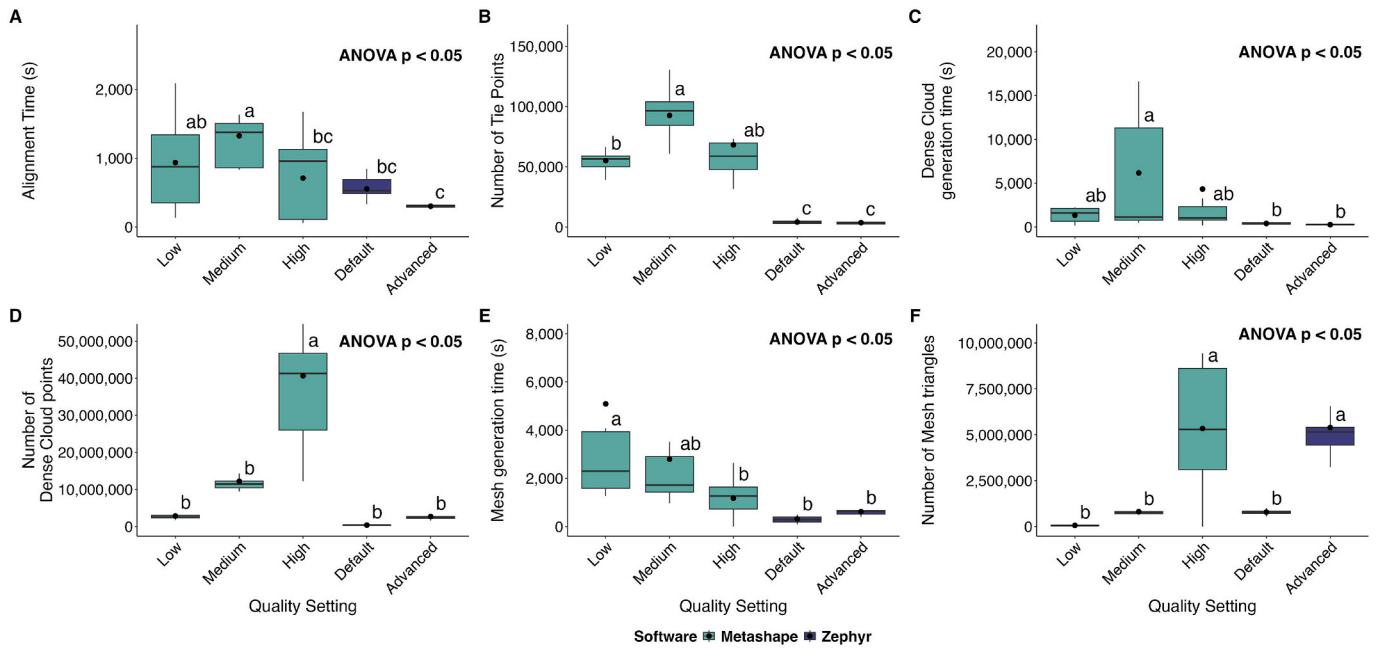


Fig. 4. Structure-from-Motion (SfM) software processing parameter comparisons between the different quality settings used in *Agisoft Metashape* (teal) and *3DF Zephyr* (dark blue) to generate 3D reconstructions of 11 Cromer Shoal Chalk Reef segments (North Norfolk, UK) (10 × 4 m). Software settings represent progressive tiers of computational effort: Low, Medium, and High for Metashape, and Default and Advanced for Zephyr. Boxplots display the processing time of each model generation stage of the workflow (A) Alignment, (C) Dense point cloud generation, (E) Mesh generation, alongside the corresponding output (B) number of tie points, (D) number of dense points, and (F) number of mesh triangles. Boxplots illustrate the median (centre line), the mean (black circle), interquartile range (box), and 1.5 × IQR (whiskers). The ANOVA displays the overall statistically significant difference between software settings. Letters above the interquartile range denote statistically significant differences between groups based on Tukey's HSD post-hoc tests ($p < 0.05$). Groups sharing a common letter are not significantly different. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ecological metrics are robust, the internal SfM software algorithms, specifically for *Metashape*, exhibit high replicate-to-replicate variance that does not scale predictably with computational effort for the generation of tie points (Fig. S3), dense cloud points (Fig. S4), and mesh triangles (Fig. S5).

4. Discussion

The rapid evolution of SfM platforms presents a dilemma for long-term monitoring: while newer algorithms offer increased efficiency for large-scale datasets, the impact of changing software mid-study on ecological time-series remains poorly understood. If algorithmic divergence leads to inconsistent complexity metrics, it risks misidentifying software-driven variance as true environmental change. A significant challenge in video-based monitoring is the inherent risk of structural underestimation. However, this methodology examines a potential mitigation strategy. By optimising the temporal resolution of towed action camera video frame-extraction (10 fps) and comparing the reconstructive consistency of *Agisoft Metashape* and *3DF Zephyr,* we demonstrate that software-parameter configuration can partially compensate for the limitations associated with non-still-image datasets.

4.1. Metric-dependent interchangeability and long-term data continuity

In addressing whether ecological metrics remain consistent across software platforms (RQ1), the statistical similarity observed in rugosity and relief across models generated by varying *Agisoft Metashape* and *3DF Zephyr* parameters, suggests that for broad-scale topographic monitoring, these platforms are functionally interchangeable. However, the significant divergence in vector dispersion, which captures fine-scale mesh surface orientation, indicates that software shifts pose a genuine risk to high-resolution surface analysis. These findings highlight that the risk of algorithmic noise masking biological change is highly dependent on the choice of ecological metric. These findings extend the work of Bayley and Mogg (2020) regarding the importance of standardised workflows for maintaining data quality in large-scale surveys by demonstrating that computational effort must also remain consistent. Essentially, transitioning from low to high-quality reconstruction settings on action camera data mid-survey could introduce ecological differences that are simply a result of digital artefacts. For long-term projects on vulnerable habitats like the Cromer Shoal chalk beds, transitioning between software mid-study may require tier-matching to ensure that the digital resolution of the eroded complexity remains consistent.

While standardising processing tiers and documenting software versions would help mitigate technical inconsistencies, ensuring long-term trust in automated 3D data interpretation also requires frameworks for identifying algorithm-induced artefacts. Recent advances in explainable artificial intelligence, such as generative AI-based counterfactual and geometric interpretation frameworks (e.g., (Holzinger et al., 2025; Rozajac et al., 2025)), emphasise that increasing algorithmic transparency is critical for distinguishing true environmental change from processing anomalies. Adopting these perspectives in marine photogrammetry bridges data-driven reproducibility toward deeper interpretability, ensuring that minor geometric variations driven by software-specific reconstruction pathways are not mistaken for genuine ecological degradation.

4.2. Computational optimisation

The convergence of vector dispersion measurements at the highest processing tiers suggests that differences between *Agisoft Metashape* and *3DF Zephyr* become insignificant when provided with sufficient computational resources and high-overlap imagery, directly answering RQ3. Consequently, the computational threshold (the point at which a setting is high enough to represent the habitat topography details) is

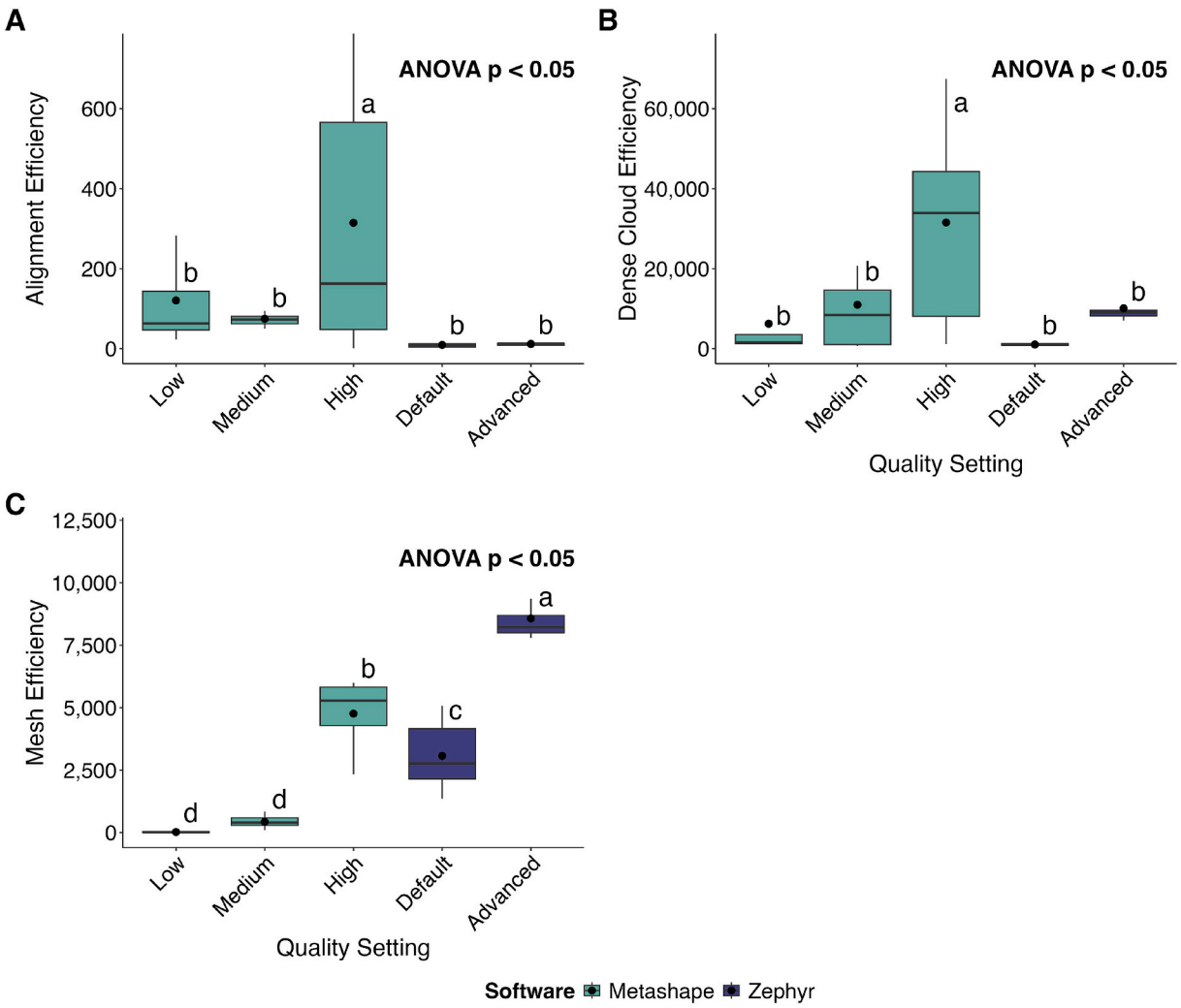


Fig. 5. Efficiency comparisons of the three main Structure-from-Motion (SfM) model generation stages (Alignment, Dense point cloud, and Mesh) between the different quality settings used in *Agisoft Metashape* (teal) and *3DF Zephyr* (dark blue) to generate 3D reconstructions of 11 Cromer Shoal Chalk Reef segments (North Norfolk, UK) (10 × 4 m). Software settings represent progressive tiers of computational effort: Low, Medium, and High for Metashape, and Default and Advanced for Zephyr. Efficiency is calculated as the rate of data production relative to the input at the start of each stage. Boxplots illustrate the median (centre line), the mean (black circle), interquartile range (box), and 1.5 × IQR (whiskers). The ANOVA displays the overall statistically significant difference between software settings. Letters above the interquartile range denote statistically significant differences between groups based on Tukey's HSD post-hoc tests ($p < 0.05$). Groups sharing a common letter are not significantly different. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

more critical for data continuity than the software differences themselves. However, for large-scale monitoring programs, the efficiency divergence noted in RQ2 becomes the deciding factor: if the goal is mesh-based metrics (rugosity/relief), *Zephyr's* nearly double processing efficiency during high-resolution reconstructions represent a substantial reduction in computational overhead for decadal studies.

The lack of relationship between processing time and tie-point yield in *Metashape* High ($R^2 = 0.001$) indicates that the software often experienced prolonged alignment phases that did not translate into increased structural data. This is likely a result of the algorithm searching for features in turbid imagery that ultimately failed to produce discrete tie-points. Conversely, *Zephyr's* more linear scaling ($R^2 = 0.46$) suggests a more consistent (albeit less dense), feature-matching strategy. The computational efficiency observed in *Zephyr's* mesh generation (8570 triangles/s) suggests a streamlined pathway for rapid benthic assessment. In the context of the UK's 2030 '30 × 30' conservation targets (The Department for Environment Food and Rural Affairs DEFRA, 2024), the ability to process large-scale ROV transects with minimal computational overhead is vital. These findings suggest that by prioritising mesh-optimised workflows, conservation bodies can maintain the high-frequency monitoring necessary to detect environmental changes in near real-time, rather than being delayed by the bottleneck of high-density point cloud alignment.

4.3. Methodology implications

The transition from traditional, diver-led SCUBA surveys to vessel-towed or ROV approaches inevitably introduces an underlying trade-off in data acquisition. While diving operations offer high-fidelity, with often motion-stabilised imagery (dependent on finning speed), they are restricted by spatial coverage and the temporal constraints of slack-water windows. Conversely, ROV surveys significantly scale up the area of monitored benthos, yet they introduce method-specific artefacts; critically, the risk of speed-induced motion blur. To compensate for these methodological limitations, our findings suggest that moderating ROV transit speed based on real-time environmental conditions (e.g., surge and turbidity) remains the primary field-based mitigation strategy to prevent significant image degradation.

However, even when platform speed is optimised, the deployment of video-extracted frames may yield a lower absolute estimate of structural

complexity compared to high-resolution still photography. Nonetheless, this systematic underestimation wouldn't invalidate the utility of ROV-based SfM for ecological monitoring purposes. Specifically, as the primary objective of decadal marine monitoring is the detection of relative environmental change rather than absolute geometric perfection, the trend is the critical focus (Henson et al., 2016; Båmstedt and Brugel, 2017). If a standardised 10 fps ROV workflow consistently under-represents rugosity by a stable margin (e.g., 5%), the data remains robust for identifying topographical degradation over a ten-year time series.

4.4. Study limitations

While these findings demonstrate robust macro-topographic consistency between *Agisoft Metashape* and *3DF Zephyr*, several methodological limitations must be considered when generalising these outcomes. Firstly, this evaluation is restricted to two commercial SfM software packages and does not encompass emerging open-source platforms or neural radiance field (NeRF)/Gaussian splatting workflows. Secondly, data collection was constrained to a single temperate marine habitat (the Cromer Shoal chalk reefs), characterised by erosional gully formations rather than the biological accretion typical of tropical coral reefs or cold-water coral systems. While macro-topographical indicators like rugosity are expected to remain stable across hard-substrate environments, micro-textural metrics such as vector dispersion may exhibit heightened algorithmic sensitivity in habitats with complex, fine-scale branching topography. Finally, the data acquisition protocol relied on a fixed-altitude, high-framerate (~250 m, 10 fps) ROV-towed video transect; variations in platform speed, turbidity, or the use of still photography rather than video-extracted frames can influence image alignment success and subsequent point density. Future research should test these software pipelines across diverse marine environments and data-collection configurations to establish a broader benchmark for marine photogrammetric standardisation.

4.5. Aligning SfM pipelines with ecological objectives

A common constraint to deep or highly turbid marine photogrammetry is the absence of an absolute, independent geometric benchmark (such as a LiDAR scan) to determine definitive accuracy of digital models (Maas, 2015; Nocerino et al., 2020; Istenič et al., 2019; Lechene et al., 2024). Whilst this remains true for our study, variations in point density, mesh triangles or structural complexity scores do not inherently equate to geometric truth. As these algorithm-specific shifts alter the geometric formulation of the digital surface, they could lead to vastly different outcomes rather than a singular correct measurement. Currently, the wider marine photogrammetry field lacks a globally standardised methodology (Rossi et al., 2021; Althaus et al., 2015), meaning there is no universally accepted way to measure benthic complexity across temporal and spatial scales (Remmers et al., 2024). Rather than viewing this lack of uniformity as a concern, our findings support the approach of workflows being tailored to individual project objectives. For instance, broad-scale management programs prioritising spatial coverage and rapid assessment may find area-averaged metrics utilising *3DF Zephyr's* faster processing time to be better suited to their logistical parameters. Conversely, where micro-textural erosion is the focus, denser, unit-vector processing may be required.

Ultimately, selecting a SfM pipeline should not be a search for an absolute universal standard, but a deliberate calibration of software strengths to specific ecological questions. By shifting the focus from absolute fidelity to methodological standardisation, these results offer a highly reassuring baseline for NGOs and conservation bodies. We demonstrate that high-coverage, cost-effective video imagery can reliably monitor fragile habitats such as the Cromer Shoal chalk reefs, even when operating within the boundaries of minor systematic underestimation.

5. Conclusion

To ensure data continuity and prevent software artefacts from masking true ecological change in long-term monitoring programmes, specific operational protocols are recommended when transitioning between SfM platforms or software versions. First, comprehensive version and parameter documentation must be maintained, recording not only the software platform and version (e.g., *Agisoft Metashape* v1.7.5 or *3DF Zephyr* v7.51) but also the exact configuration tier and parameter settings applied across all workflow stages (alignment, dense cloud, and mesh generation). Second, users must practice strict tier-matching across platforms mid-study by avoiding the mixing of resolution tiers, such as comparing low-tier models to high-tier models, to ensure corresponding mesh resolution with computation effort and prevent artificial shifts in micro-scale metrics such as vector dispersion. Finally, pre-transition validation should be conducted by processing a subset of archived raw data replicates through both established and proposed workflows, thereby quantifying platform-specific divergence to ensure relative trends remain ecologically valid and distinguishing software-induced noise from true topographical degradation.

CRediT authorship contribution statement

Jessica L.Rhodes: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Russell B.Connelly:** Writing – review & editing, Visualization, Validation. **Boyd A.McKew:** Writing – review & editing, Validation, Supervision. **Jon Chamberlain:** Writing – review & editing, Validation, Supervision, Resources, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the University of Essex Chancellor's PhD Scholarship. We are grateful to the School of Computer Science and Electronic Engineering (CSEE) at the University of Essex for providing the computational facilities and virtual machine infrastructure used in this study. We would also like to thank the Eastern Inshore Fisheries and Conservation Authority (EIFCA) for access to the data collection site.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2026.115335.

Data availability

Research Link Provided

Images and 3D Reconstructed Models (.OBJ & .MTL) of the Cromer Shoal Chalk Reef (Original data) (Zenodo)

References

Althaus, F., Hill, N., Ferrari, R., Edwards, L., Przeslawski, R., Schönberg, C.H., Stuart-Smith, R., Barrett, N., Edgar, G., Colquhoun, J., 2015. A standardised vocabulary for identifying benthic biota and substrata from underwater imagery: the CATAMI classification scheme. PloS One 10 (10), e0141039.

Anderson, K., Westoby, M.J., James, M.R., 2019. Low-Budget Topographic Surveying Comes of Age: Structure from Motion Photogrammetry in Geography and the Geosciences. SAGE Publications, Sage UK: London, England.

Båmstedt, U., Brugel, S., 2017. A cost-precision model for marine environmental monitoring, based on time-integrated averages. Environ. Monit. Assess. 189 (7), 354.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48.
Bayley, D.T., Mogg, A.O., 2020. A protocol for the large-scale analysis of reefs using structure from motion photogrammetry. Methods Ecol. Evol. 11 (11), 1410–1420.
Benway, H.M., Lorenzoni, L., White, A.E., Fiedler, B., Levine, N.M., Nicholson, D.P., DeGrandpre, M.D., Sosik, H.M., Church, M.J., O'Brien, T.D., Leinen, M., Weller, R.A., Karl, D.M., Henson, S.A., Letelier, R.M., 2019. Ocean time series observations of changing marine ecosystems: an era of integration, synthesis, and societal applications. Front. Mar. Sci. 6.
Brooks, M.E., Kristensen, K., Van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., Bolker, B.M., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal. 9 (2), 675.
Buatois, L.A., 2003. Ichnology of shallow marine deposits in the Miocene Chenque formation of Patagonia: complex ecologic structure and niche partitioning in Neogene ecosystems. Publicación Electrónica Asoc. Paleontol. Argent. 9 (1).
Budău, L., Pescari, S., Chendeş, R., Vîlceanu, C.-B., Herban, S., 2023. Comparative study on digital image processing for 3D modeling. Sci. Pap. Ser. E. Land Reclam. Earth Obs. Surv. Environ. Eng. 12.
Carleton, J., Sammarco, P.W., 1987. Effects of substratum irregularity on success of coral settlement: quantification by comparative geomorphological techniques. Bull. Mar. Sci. 40 (1), 85–98.
Conley, D.D., Hollander, E.N., 2021. A non-destructive method to create a time series of surface area for coral using 3D photogrammetry. Front. Mar. Sci. 8, 660846.
Edwards, M., Beaugrand, G., Hays, G.C., Koslow, J.A., Richardson, A.J., 2010. Multi-decadal oceanic ecological datasets and their application in marine policy and management. Trends Ecol. Evol. 25 (10), 602–610.
Figueira, W., Ferrari, R., Weatherby, E., Porter, A., Hawes, S., Byrne, M., 2015. Accuracy and precision of habitat structural complexity metrics derived from underwater photogrammetry. Remote Sens. 7 (12), 16883–16900.
Forsmoo, J., Anderson, K., Macleod, C.J.A., Wilkinson, M.E., DeBell, L., Brazier, R.E., 2019. Structure from motion photogrammetry in ecology: does the choice of software matter? Ecol. Evol. 9 (23), 12964–12979.
Friedman, A., Pizarro, O., Williams, S.B., Johnson-Roberson, M., 2012. Multi-scale measures of rugosity, slope and aspect from benthic stereo image reconstructions. PloS One 7 (12), e50440.
Fukunaga, A., Burns, J.H., 2020. Metrics of coral reef structural complexity extracted from 3D mesh models and digital elevation models. Remote Sens. 12 (17), 2676.
Fukunaga, A., Burns, J.H.R., Pascoe, K.H., Kosaki, R.K., 2020. Associations between benthic cover and habitat complexity metrics obtained from 3D reconstruction of coral reefs at different resolutions. Remote Sens. 12 (6), 1011.
Graham, N.A., Nash, K.L., 2013. The importance of structural complexity in coral reef ecosystems. Coral Reefs 32, 315–326.
Henson, S.A., Beaulieu, C., Lampitt, R., 2016. Observing climate change trends in ocean biogeochemistry: when and where. Glob. Chang. Biol. 22 (4), 1561–1571.
Holzinger, A., Lukač, N., Rozajac, D., Johnston, E., Kocic, V., Hoerl, B., Gollob, C., Nothdurft, A., Stampfer, K., Schweng, S., 2025. Enhancing trust in automated 3D point cloud data interpretation through explainable counterfactuals. Inf. Fusion 119, 103032.
Istenič, K., Gracias, N., Arnaubec, A., Escartín, J., Garcia, R., 2019. Scale accuracy evaluation of image-based 3D reconstruction strategies using laser photogrammetry. Remote Sens. 11 (18), 2093.
Jackson, A., Watson, D., Spray, R., 2022. Benthic assemblages in Cromer shoal chalk beds MCZ: comparisons among different types of terrain.
Lazarus, M., Belmaker, J., 2021. A review of seascape complexity indices and their performance in coral and rocky reefs. Methods Ecol. Evolution 12 (4), 681–695.
Lechene, M.A.A., Figueira, W.F., Murray, N.J., Aston, E.A., Gordon, S.E., Ferrari, R., 2024. Evaluating error sources to improve precision in the co-registration of underwater 3D models. Eco. Inform. 81, 102632.
Lenth, R.V., Piaskowski, J., Banfai, B., Bolker, B., Buerkner, P., Giné-Vázquez, I., Hervé, M., Jung, M., Love, J., Miguez, F., Riebl, H., Singmann, H., 2024. Emmeans: estimated marginal means, aka least-squares means.
Lochhead, I., Hedley, N., 2022. Evaluating the 3D integrity of underwater structure from motion workflows. Photogramm. Rec. 37 (177), 35–60.
Luckhurst, B.E., Luckhurst, K., 1978. Analysis of the influence of substrate variables on coral reef fish communities. Mar. Biol. 49 (4), 317–323.
Maas, H.-G., 2015. On the accuracy potential in underwater/multimedia photogrammetry. Sensors 15 (8), 18140–18152.
Marre, G., Holon, F., Luque, S., Boissery, P., Deter, J., 2019. Monitoring marine habitats with photogrammetry: a cost-effective, accurate, precise and high-resolution reconstruction method. Front. Mar. Sci. 6, 276.
McCoy, E.D., Bell, S.S., 1991. Habitat structure: the evolution and diversification of a complex topic. In: Habitat Structure: the Physical Arrangement of Objects in Space. Springer, pp. 3–27.
McNeel, R., 2010. Rhinoceros 3D, 7.0.
Moffat, C., Richardson, H., Roberts, G., 2020. Natural England marine chalk characterisation project. *Natural England* Research Reports, (080).
Nocerino, E., Menna, F., Gruen, A., Troyer, M., Capra, A., Castagnetti, C., Rossi, P., Brooks, A.J., Schmitt, R.J., Holbrook, S.J., 2020. Coral reef monitoring by Scuba divers using underwater photogrammetry and geodetic surveying. Remote Sens. 12 (18), 3036.
Oh, D., Cresswell, A.K., Thomson, D.P., Renton, M., 2025. Morphological composition influences redundancy, complementarity and ecological relevance of habitat complexity metrics in simulated coral communities. Ecol. Evol. 15 (9), e72077.
Oliveira, L.M.C.d., Lim, A., Conti, L.A., Wheeler, A.J., 2022. High-resolution 3D mapping of cold-water coral reefs using machine learning. Front. Environ. Sci. 10.
Paterson, Iona L.R., Dawson, Kathryn E., Mogg, Andrew O.M., Sayer, Martin D.J., Burdett, Heidi L., 2024. Quantitative comparison of ROV and diver-based photogrammetry to reconstruct Maerl bed ecosystems. Aquat. Conserv. Mar. Freshwat. Ecosyst. 34 (11), e70007.
R Core Team, 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Remmers, T., Grech, A., Roelfsema, C., Gordon, S., Lechene, M., Ferrari, R., 2024. Close-range underwater photogrammetry for coral reef ecology: a systematic literature review. Coral Reefs 43 (1), 35–52.
Risk, M.J., 1972. Fish diversity on a coral reef in the Virgin Islands. Atoll Research Bulletin. 153.
Rossi, P., Ponti, M., Righi, S., Castagnetti, C., Simonini, R., Mancini, F., Agrafiotis, P., Bassani, L., Bruno, F., Cerrano, C., Cignoni, P., Corsini, M., Drap, P., Dubbini, M., Garrabou, J., Gori, A., Gracias, N., Ledoux, J.-B., Linares, C., Mantas, T.P., Menna, F., Nocerino, E., Palma, M., Pavoni, G., Ridolfi, A., Rossi, S., Skarlatos, D., Treibitz, T., Turicchia, E., Yuval, M., Capra, A., 2021. Needs and gaps in optical underwater technologies and methods for the investigation of marine animal forest 3D-structural complexity. Front. Mar. Sci. 8 (2021).
Rozajac, D., Lukač, N., Schweng, S., Gollob, C., Nothdurft, A., Stampfer, K., Del Ser, J., Holzinger, A., 2025. Explaining 3D semantic segmentation through generative AI-based counterfactuals. Expert. Syst. 42 (12), e70163.
Savrda, C.E., 2012. Chalk and related deep-marine carbonates. Developments in Sedimentology: Elsevier, pp. 777–806.
Smith, M.W., Carrivick, J.L., Quincey, D.J., 2016. Structure from motion photogrammetry in physical geography. Prog. Phys. Geogr. 40 (2), 247–275.
Smithson, M., Verkuilen, J., 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods 11 (1), 54–71. https://doi.org/10.1037/1082-989X.11.1.54.
The Department for Environment Food & Rural Affairs (DEFRA), 2024. 30by30 on land in England: confirmed criteria and next steps. https://www.gov.uk/government/publications/criteria-for-30by30-on-land-in-england/30by30-on-land-in-england-confirmed-criteria-and-next-steps (Accessed: 03/06/2026).
Valdés, L., Bode, A., Latasa, M., Nogueira, E., Somavilla, R., Varela, M.M., González-Pola, C., Casas, G., 2021. Three decades of continuous ocean observations in North Atlantic Spanish waters: the RADIALES time series project, context, achievements and challenges. Prog. Oceanogr. 198, 102671.
Vincent, M., Isabel, U.B., Louis, F., Romain, P., Rémi, G., Kenneth, D., 2021. 'Underwater photogrammetry as an environmental assessment tool to monitor coral reefs and artificial structures. In: *Offshore Technology Conference*: OTC. D021S027R003.
Wright, J., Chamberlain, J., 2024. Investigating human impacts on rocky reefs using measures of complexity and relief from 3D photogrammetry. Ecosphere 15 (2), e4763.
Young, G., Dey, S., Rogers, A., Exton, D., 2017. Cost and time-effective method for multi-scale measures of rugosity, fractal dimension, and vector dispersion from coral reef 3D models. PloS One 12 (4), e0175341.