Journal of *Marine Science and Engineering*

MDPI

Article

Observed Spectral-Reflectance Zoning for Model Selection in Satellite-Derived Bathymetry of Heterogeneous Coral Reef Waters Using ICESat-2 and Sentinel-2

Yuan Le [1], Bin Xu [1], Yifu Chen [2,*], Lin Wu [3], Lin Yi [1] and Xiaoyu Sun [1,*]

[1] School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China; leyuan@cug.edu.cn (Y.L.); 1202521272@cug.edu.cn (B.X.); 17639050930@163.com (L.Y.)
[2] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[3] Key Laboratory of Geological Survey and Evaluation of Ministry of Education, China University of Geosciences, Wuhan 430074, China; wulin@cug.edu.cn
* Correspondence: yifuchen.cug@foxmail.com (Y.C.); 1202321125@cug.edu.cn (X.S.)

Abstract

Accurate and spatially continuous shallow water bathymetry supports coral reef geomorphological investigation, ecosystem assessment, and marine resource management. However, spectral heterogeneity arising from substrate properties, water depth, and water-column conditions complicates model selection. This study evaluates whether bathymetric model performance differs among observed spectral-reflectance zones and whether zonal model selection improves overall retrieval. ICESat-2 bathymetric points, Sentinel-2 imagery, and Allen Coral Atlas data were analyzed at Lingyang Reef, Huaguang Reef, and East Cay Reef. Representative substrate classes identified using Allen Coral Atlas data were used to derive Sentinel-2 green-band reflectance thresholds. Water pixels were then classified into high-, medium-, and low-reflectance zones. The band-ratio, multiple linear regression (MLR), random forest (RF), and extreme gradient boosting (XGBoost) models were constructed and validated within each zone. When averaged across the three study areas, MLR achieved the best performance in the high reflectance zones, with mean R^2 and RMSE values of 0.71 and 0.49 m, respectively. RF achieved the best mean performance in the medium reflectance zones, with mean R^2 and RMSE values of 0.89 and 0.62 m. In the low reflectance zones, RF and XGBoost outperformed Stumpf and MLR, with XGBoost yielding mean R^2 and RMSE values of 0.85 and 1.20 m, respectively. The optimal outputs from the different reflectance zones were integrated and smoothed to generate continuous bathymetric maps. The integrated zonal method achieved RMSE values of 0.76, 0.61, and 0.84 m at the three reefs, respectively, representing numerical reductions of approximately 3.8%, 4.7%, and 12.5% relative to the best unified model. These findings demonstrate that observed spectral-reflectance zoning provides a practical basis for matching bathymetric models to different reflectance conditions, supporting bathymetric mapping in heterogeneous coral reef waters.

Keywords: shallow water bathymetry; coral reef; substrate reflectance; model applicability; ICESat-2; Sentinel-2; observed spectral-reflectance zoning

Check for updates

Received: 19 August 2026
Revised: 8 October 2026
Accepted: 20 October 2026
Published: 26 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Accurate and spatially continuous shallow-water bathymetry is fundamental to understanding coral reef geomorphological evolution, assessing ecosystem conditions, and

supporting the sustainable use and management of marine resources. It also provides essential geospatial information for marine environmental monitoring, coastal engineering, navigation safety, and island and reef spatial planning [1,2]. Conventional shallow water bathymetric surveys rely primarily on shipborne sonar and airborne LiDAR bathymetry. Although shipborne sonar provides highly accurate depth measurements, its application in shallow and topographically complex waters is constrained by vessel accessibility, survey efficiency, and operational cost. Airborne LiDAR bathymetry enables the rapid acquisition of high-accuracy topographic information but requires costly flight operations and is affected by weather, airspace restrictions, and flight-line configuration [3–6]. These constraints limit the repeated acquisition of spatially continuous bathymetric data over extensive coral reef environments. Satellite-derived bathymetry (SDB), with its broad coverage, repeat observations, and high acquisition efficiency, therefore provides an important complement to conventional bathymetric surveys [7,8].

Optical SDB estimates water depth from light propagation and attenuation, providing spatially continuous estimates where sufficient seafloor signal remains detectable [9]. Existing methods include empirical, semi-analytical, and machine learning approaches. Empirical models, such as the Lyzenga multispectral and Stumpf log-band-ratio models, offer simplicity and efficiency [10–14]. The former estimates depth from linear combinations of transformed spectral variables, whereas the latter uses differential spectral attenuation to reduce sensitivity to bottom-albedo variations. However, their performance depends on the representativeness of calibration samples and the consistency of local optical conditions. Semi-analytical models explicitly relate reflectance to water-column properties, bottom reflectance, and depth [15–19], whereas machine learning models capture nonlinear spectral–depth relationships [1,20]. Multi-temporal observations, feature extraction and optimization, and neural-network modeling have further expanded the capacity of data-driven SDB to represent complex spectral–depth relationships [21–23]. These advances also place greater emphasis on training-data representativeness and model generalization across acquisition dates, geographic areas, and water optical conditions. High accuracy within a calibration area does not necessarily establish reliable performance under different environmental conditions. Recent deep-learning studies have incorporated multiscale feature extraction, with MuSRFM integrating features across spatial scales and DA-SDB combining multiscale feature learning with physically assisted representations and domain alignment to improve spatial and temporal transferability [24,25]. Reliable bathymetric control data are also essential. After photon extraction and refraction correction, ICESat-2 provides sparse along-track depth observations in clear shallow waters [26–30]. Their integration with Sentinel-2 imagery supports spatially continuous bathymetric mapping [31,32]. However, selecting suitable models remains challenging in heterogeneous coral reef environments.

In optically shallow waters, part of the solar radiation entering the water column reaches the seafloor and is reflected upward before being detected by a satellite sensor. The resulting water-leaving spectral signal therefore contains contributions from both the water column and the seafloor. These contributions follow the same radiative transfer principles, but different seafloor substrates have distinct inherent reflectance characteristics. Consequently, sand, rubble, rock, seagrass, microalgal mats, and coral may produce different spectral responses at comparable water depths, while similar spectral observations may correspond to different depths under different bottom reflectance backgrounds [12,17,33]. This spectral ambiguity increases the complexity of applying a unified bathymetric model across heterogeneous reef environments. Its influence may also differ among model types. Empirical models depend on relatively consistent spectral patterns, semi-analytical models require reliable parameterization of bottom and water-column optical properties, and ma-

chine learning models are sensitive to the representativeness and distribution of training data [34–36]. The relative performance of different bathymetric models may therefore vary with substrate reflectance and sampling conditions.

Previous studies have incorporated substrate information into bathymetry through bottom-reflectance retrieval [18], sediment-informed neural networks [37], bottom-type-adaptive empirical models [38], and hyperspectral methods considering spatial and spectral substrate characteristics [39]. Zoning and substrate-informed approaches have also improved retrieval in complex reef environments [31,40]. Mixed reef pixels often exhibit transitional spectral characteristics, while detailed substrate classification requires substantial field data or reliable prior information. Grouping reef waters according to observed spectral reflectance provides a practical representation of spectral heterogeneity without requiring detailed substrate labels for every water pixel. However, the present strategy still relies on prior substrate information to select representative samples for threshold derivation. Nonlinear machine-learning and deep-learning models may already capture part of the substrate- and water-column-related variability when training data are representative. Explicit zoning should therefore be evaluated for its additional value in model selection. Grouping reef waters according to observed spectral reflectance provides a practical representation of spectral heterogeneity without requiring detailed substrate labels for every water pixel. However, the present strategy still relies on prior substrate information to select representative samples for threshold derivation. Its contribution lies in comparing candidate bathymetric models under different observed reflectance conditions and integrating locally selected models, rather than eliminating substrate-data requirements or establishing universal zoning thresholds. Whether similar model-performance patterns occur across different reefs is examined through site-specific calibration and evaluation; transfer without recalibration is not assessed.

To address this gap, this study investigated the applicability of four bathymetric models under different substrate reflectance conditions at Lingyang Reef, Huaguang Reef, and East Cay Reef. ICESat-2 along-track bathymetric points were combined with Sentinel-2 multispectral imagery, and Allen Coral Atlas substrate-type data were used to characterize the green-band reflectance distributions associated with six representative coral reef substrate types. The study areas were divided into high-, medium-, and low-reflectance zones, within which the Stumpf band-ratio model, multiple linear regression, random forest, and extreme gradient boosting (XGBoost) were independently constructed and validated. Their bathymetric accuracy and applicability across the three study areas were compared, and the optimal outputs from the individual reflectance zones were combined through spatial harmonization and boundary smoothing to produce continuous bathymetric maps. This study aims to clarify how the applicability of the selected bathymetric models varies with substrate reflectance conditions and to determine whether reflectance zoning can provide a practical basis for model selection in heterogeneous coral reef waters.

2. Study Areas and Materials

This study was conducted at three shallow coral reef sites with contrasting geographic and geomorphic settings: Lingyang Reef and Huaguang Reef in the Xisha Islands of the South China Sea, and East Cay Reef in the southern Great Barrier Reef, Australia. The analysis integrated ICESat-2 ATL03 photon data, Sentinel-2 multispectral imagery, Allen Coral Atlas substrate-type data, tidal information, and airborne LiDAR bathymetry data available for Lingyang Reef. ICESat-2 provided along-track bathymetric control points, while Sentinel-2 imagery supplied spatially continuous spectral observations. The Allen Coral Atlas data were used to characterize the reflectance backgrounds associated with different substrate types, tidal information was used for vertical datum correction, and

the airborne LiDAR bathymetry data provided independent validation at Lingyang Reef and Huaguang Reef. The datasets used in this study and their respective purposes are summarized in Table 1.

Table 1. Datasets used in this study and their purposes.

Dataset	Purpose in This Study
ICESat-2 ATL03 photon data	Provides along-track water depths following seafloor photon extraction and correction, with separate datasets used for model fitting and accuracy validation.
Sentinel-2 MSI multispectral imagery	Provides spectral data for water-depth retrieval and green-band reflectance for zoning.
Allen Coral Atlas benthic habitat maps	Provides the spatial distribution of six substrate categories for calculating category-specific green-band reflectance statistics and informing reflectance-zoning thresholds.
Airborne laser bathymetry (ALB) data	Provides independent validation of bathymetric estimates at Lingyang Reef and Huaguang Reef, and supports error assessment across water-depth intervals.
Tide-gauge observations and NAO.99b ocean tide model	Provides tidal heights for correcting acquisition-time water-level differences and harmonizing bathymetric data to a consistent water-level reference.

2.1. Study Areas

Lingyang Reef and Huaguang Reef are located in the Xisha Islands of the South China Sea, whereas East Cay Reef is located in the Swain sector of the southern Great Barrier Reef, Australia (Figure 1). Lingyang Reef, located at 16°26′–16°29′ N and 111◦34′–111°36′ E, has a compact, approximately triangular shape, with a broad shallow reef flat surrounding a relatively enclosed lagoon. Huaguang Reef, located at 16°10′–16°15′ N and 111°36′–111°46′ E, is elongated and contains an extensive central lagoon and heterogeneous shallow reef surfaces. East Cay Reef, located at 21°28′–21°29′ S and 152°33′–152°35′ E, is also elongated, with broad shallow reef flats and an open lagoon directly connected to the surrounding sea. These differences in reef configuration and lagoon openness are associated with contrasting shallow water substrate distributions and reflectance backgrounds, providing suitable conditions for comparing bathymetric model applicability across the three sites. The geographic locations of the study sites and the distributions of ICESat-2 ground tracks are shown in Figure 1. Lingyang Reef and Huaguang Reef are remote offshore reefs with relatively limited direct terrestrial influence. However, their shallow-water environments exhibit spatial variations in geomorphic setting and benthic composition, which can affect the spectral characteristics observed by optical remote sensing [41,42]. East Cay Reef is also an offshore reef, where depth, wave exposure, geomorphic setting, and benthic composition contribute to spatial variations in reef habitats [43]. These environmental differences may alter water-column optical properties and benthic reflectance, thereby affecting the remotely sensed spectral signal among the study sites [44].

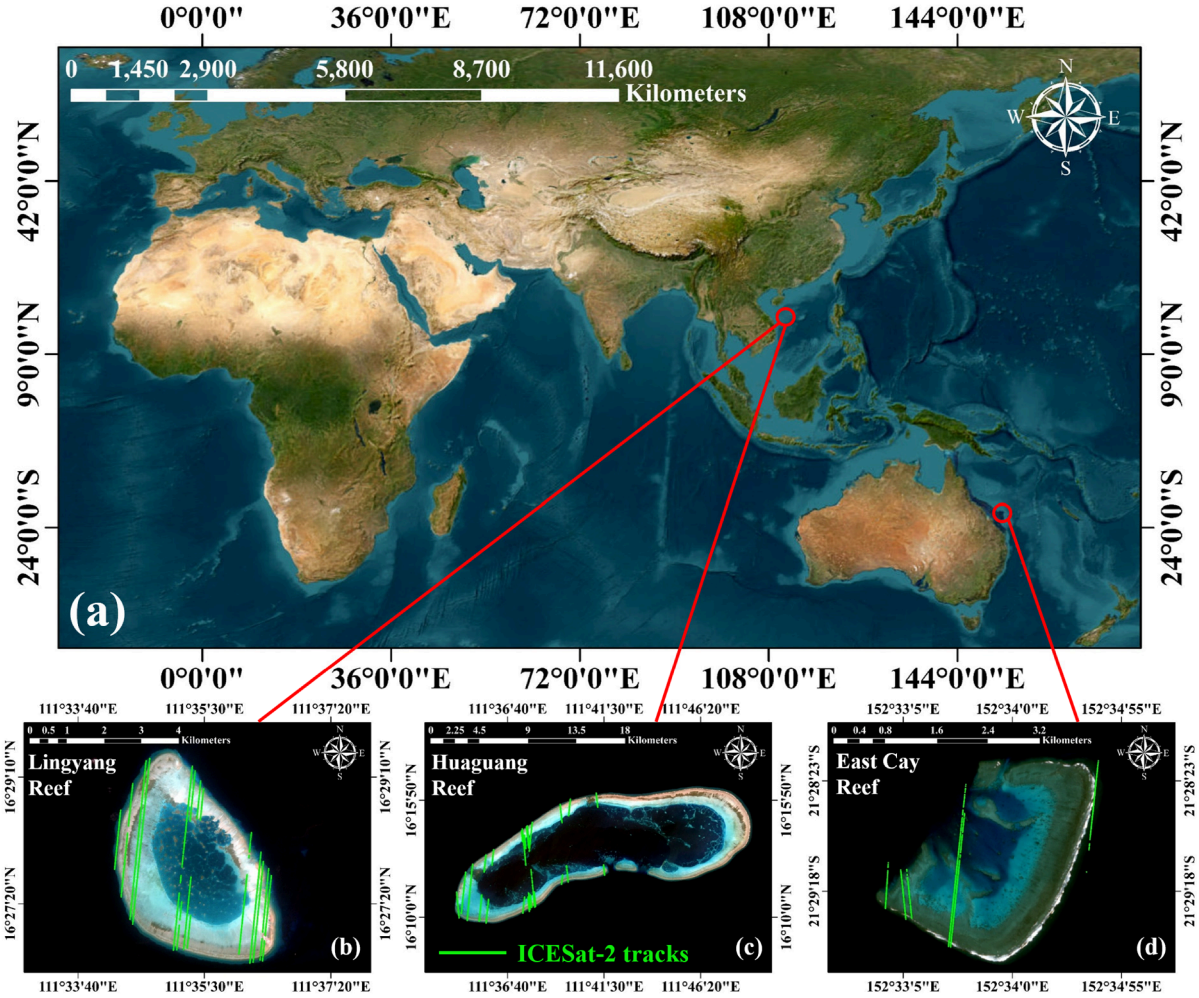

Figure 1. Geographic locations of the three study sites and spatial distribution of ICESat-2 ground tracks. (**a**) Regional locations of the study sites; (**b**) Lingyang Reef; (**c**) Huaguang Reef; and (**d**) East Cay Reef. Yellow circles and lines in (**a**) indicate the locations of the study sites, and green lines in (**b**–**d**) represent the ICESat-2 ground tracks.

2.2. ICESat-2 ATLAS Data

NASA ICESat-2 ATL03 photon data were used as the active remote sensing data source in this study. The ATL03 product, namely the global geolocated photon product, is stored in HDF5 format and records the time, longitude, latitude, and ellipsoidal height of each photon event [45]. The photon data are organized by six beams, including gt1l, gt1r, gt2l, gt2r, gt3l, and gt3r. ATL03 preserves the photon-level observations needed to inspect sea-surface and seafloor returns and to implement the extraction and correction of bathymetric reference samples. It was therefore selected to generate reference depths through a consistent processing workflow across the three study sites, rather than directly adopting the classified and refraction-corrected bathymetric estimates provided by the higher-level ATL24 product [46]. The ATL03 data used in this study were obtained from the NASA Earthdata Search platform (https://search.earthdata.nasa.gov/search, accessed on 19 October 2026). After preprocessing, the extracted along-track bathymetric points were used as control samples for model development and accuracy assessment. Across the three study sites, 17 ATL03 products comprising 30 beam tracks were used for model training, and 11 products comprising 20 beam tracks were used for validation. The acquisition dates, numbers of products and beam tracks, and training or validation roles at each study site are summarized in Table 2.

Table 2. Acquisition dates and usage of ICESat-2 ATL03 data at the three study sites.

Study Site	Purpose	Acquisition Dates	Number of Products	Number of Beam Tracks
Lingyang Reef	Training	18 August 2021; 5 November 2022; 15 August 2023; 14 November 2023; 12 February 2024; 12 August 2024	6	13
	Validation	11 April 2024; 11 July 2024; 8 January 2025	3	7
Huaguang Reef	Training	19 January 2020; 19 April 2020; 17 March 2022; 15 April 2022; 15 July 2022; 11 July 2024	6	11
	Validation	19 May 2021; 16 May 2023; 15 August 2023; 14 November 2023; 12 August 2024	5	9
East Cay Reef	Training	4 November 2019; 28 January 2023; 4 June 2023; 23 June 2024; 23 November 2024	5	6
	Validation	30 November 2024; 31 May 2025; 30 August 2025	3	4

2.3. Sentinel-2 Multispectral Imagery

Sentinel-2 multispectral imagery was used as the passive optical remote sensing data in this study. The data were downloaded from the Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu/, accessed on 19 October 2026). The imagery used in this study was acquired by Sentinel-2A, Sentinel-2B, and Sentinel-2C, each equipped with a Multispectral Instrument (MSI) providing 13 spectral bands at spatial resolutions of 10 m, 20 m, and 60 m [47]. The Sentinel-2 images used in this study were Level-2A products. One Level-2A image was selected for each study site based on spatial coverage and image quality. Detailed information on the Sentinel-2 images used in this study is provided in Table 3. These products have been orthorectified and atmospherically corrected, and the pixel values are provided as bottom-of-atmosphere reflectance. Considering the water penetration capability and spatial resolution requirements in shallow water areas, the 10 m blue, green, red, and near-infrared bands were selected for subsequent processing.

Table 3. Sentinel-2 imagery used for the three study sites.

Study Site	Acquisition Date	Product ID
Lingyang Reef	4 March 2025	S2C_MSIL2A_20250304T025641_N0511_R032_T49QEU_20250304T065712.SAFE
Huaguang Reef	31 December 2025	S2A_MSIL2A_20251231T030201_N0511_R032_T49PET_20251231T085713.SAFE
East Cay Reef	18 April 2025	S2B_MSIL2A_20250418T000229_N0511_R030_T56KMB_20250418T010528.SAFE

2.4. Allen Coral Atlas Substrate-Type Data

Allen Coral Atlas substrate-type data were used as auxiliary reference data for selecting representative substrate samples. The dataset is derived from high-spatial-resolution Planet Dove satellite imagery, together with physical wave modeling and field ecological monitoring information, and provides spatially continuous information on coral reef substrate distributions [48]. The original substrate classifications of the three study sites

are shown in Figure 2. In this study, the Allen Coral Atlas was used to identify typical substrate-dominated areas at Lingyang Reef, Huaguang Reef, and East Cay Reef. Six representative substrate types were considered, including sand, rubble, rock, seagrass, microalgal mats, and coral/algae habitat. The coral/algae habitat class is hereafter referred to as coral habitat for brevity. These data were used only for substrate sample selection and were not used as input variables in the bathymetric inversion models.

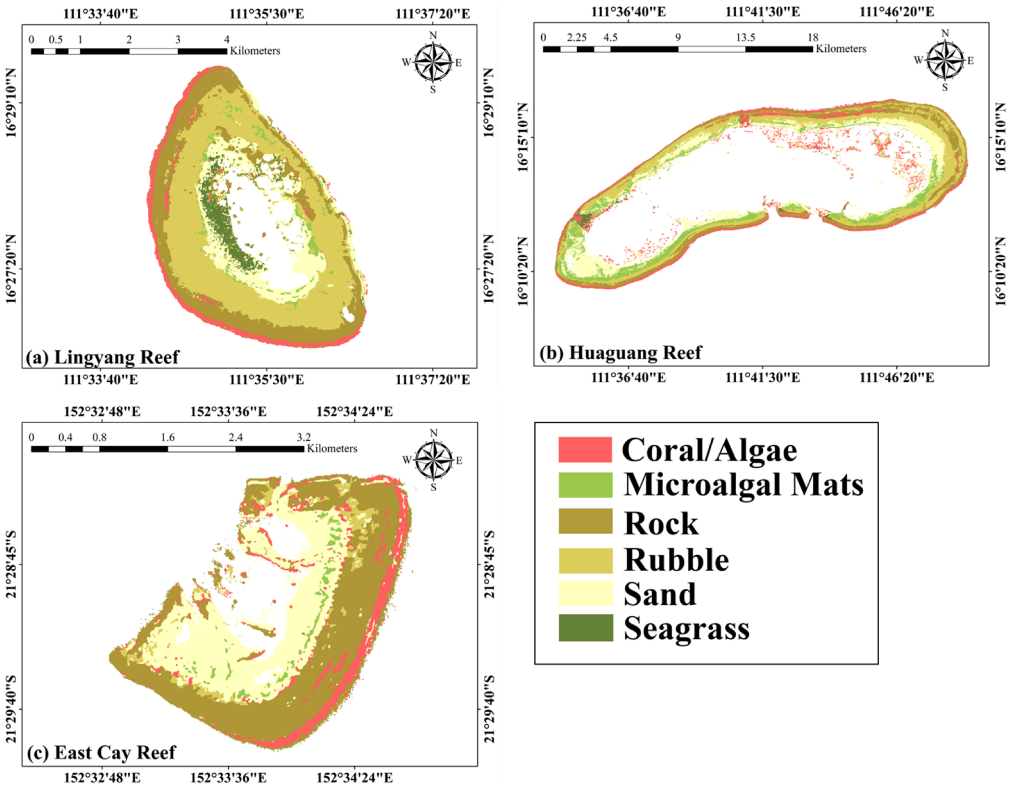

Figure 2. Original substrate classifications derived from the Allen Coral Atlas for the three study areas: (**a**) Lingyang Reef, (**b**) Huaguang Reef, and (**c**) East Cay Reef.

2.5. Tidal Data and Validation Data

Airborne LiDAR bathymetry data were available for both Lingyang Reef and Huaguang Reef and were used for external accuracy assessment. The datasets were acquired on 16 July 2018 and 21 May 2023, respectively, using the Mapper5000 airborne full-waveform LiDAR system developed by the Shanghai Institute of Optics and Fine Mechanics (Shanghai, China). The system employs a 532 nm green laser for bathymetric measurements and has a reported depth accuracy of 0.23 m. For East Cay Reef, one ICESat-2 ground track was reserved as an independent test dataset and was excluded from both model training and optimal model selection. This track was used exclusively for the final accuracy assessment.

The original ICESat-2 and Mapper5000 elevations were referenced to the WGS84 ellipsoid. ICESat-2 water depths were calculated from the elevation difference between the detected sea surface and the refraction-corrected seafloor, whereas Mapper5000 water depths were derived from the temporal separation between surface and bottom waveform returns. Both datasets therefore provided depths relative to the water surface at their respective acquisition times, without requiring an ellipsoid-to-geoid conversion for depth calculation. Hourly tide-gauge observations for Sansha City were obtained from the China Maritime Service Network and interpolated to 1 min intervals using B-splines to estimate tidal heights at the bathymetric and Sentinel-2 acquisition times. These tide-gauge

observations were used for Lingyang Reef and Huaguang Reef, whereas tidal heights for East Cay Reef were predicted using the NAO.99b model at the reef location [49]. The tidal heights corresponding to the ICESat-2 modeling tracks and independent validation data, and the Sentinel-2 imagery are summarized in Table 4. Depths were adjusted according to

$$d_{image} = d_{obs} + \eta_{image} - \eta_{obs} \quad (1)$$

where d_{obs} is the depth at the bathymetric acquisition time and η_{image} and η_{obs} are tidal heights at the Sentinel-2 and bathymetric acquisition times, respectively, expressed relative to the same tidal reference within each site. The tidal height difference was applied uniformly within each reef; spatial tidal variations were not explicitly modeled. Water depth was defined as positive downward, and the final water-level reference was the image-time water surface rather than a fixed mean-sea-level or chart datum.

Table 4. Tidal heights corresponding to the modeling tracks and validation data.

Study Site	Training Data	Tidal Height (m)	Validation Data	Tidal Height (m)	Tidal Heights Sentinel-2 Imagery
Lingyang Reef	18 August 2021	0.2154	11 April 2024	0.3766	0.1187
	5 November 2022	0.0845	11 July 2024	−0.2145	
	15 August 2023	0.3256	8 January 2025	0.1456	
	14 November 2023	−0.0458			
	12 February 2024	0.1124			
	12 August 2024	−0.2215			
Huaguang Reef	19 January 2020	0.0541	19 May 2021	0.2465	−0.3462
	19 April 2020	0.1985	16 May 2023	−0.1117	
	17 March 2022	−0.3124	15 August 2023	0.1785	
	15 April 2022	0.1479	14 November 2023	−0.2541	
	15 July 2022	−0.0789	12 August 2024	−0.2907	
	11 July 2024	−0.1408			
East Cay Reef	4 November 2019	0.3273	30 November 2024	0.2992	−0.1975
	28 January 2023	−0.1286	31 May 2025	−0.4206	
	4 June 2023	−0.0909	30 August 2025	0.0774	
	23 June 2024	0.4401			
	23 November 2024	0.2922			

3. Methods

An overall experimental framework was established to evaluate the applicability of bathymetric models under different substrate reflectance conditions, as shown in Figure 3. The primary data sources included Sentinel-2 multispectral imagery, ICESat-2 ATL03 photon data, Allen Coral Atlas substrate-type data, ALB bathymetric data, and tidal data. Optical imagery and ICESat-2 ATL03 data were preprocessed separately. Sentinel-2 preprocessing included basic image processing and glint correction, whereas ICESat-2 ATL03 processing included sea-surface detection, photon denoising, seafloor photon extraction, refraction correction, and tidal correction. Zoning thresholds were then calculated from the spectral reflectance statistics of representative substrate samples to generate three substrate reflectance zones. Band selection was conducted within each zone, and Stumpf, MLR, RF, and XGBoost models were constructed. Independent validation samples and accuracy metrics were used to compare the performance of the four models under different substrate reflectance conditions. Finally, the zonal inversion results were mosaicked and smoothed to generate shallow water bathymetric results, which were evaluated using validation data and compared with unzoned unified modeling.

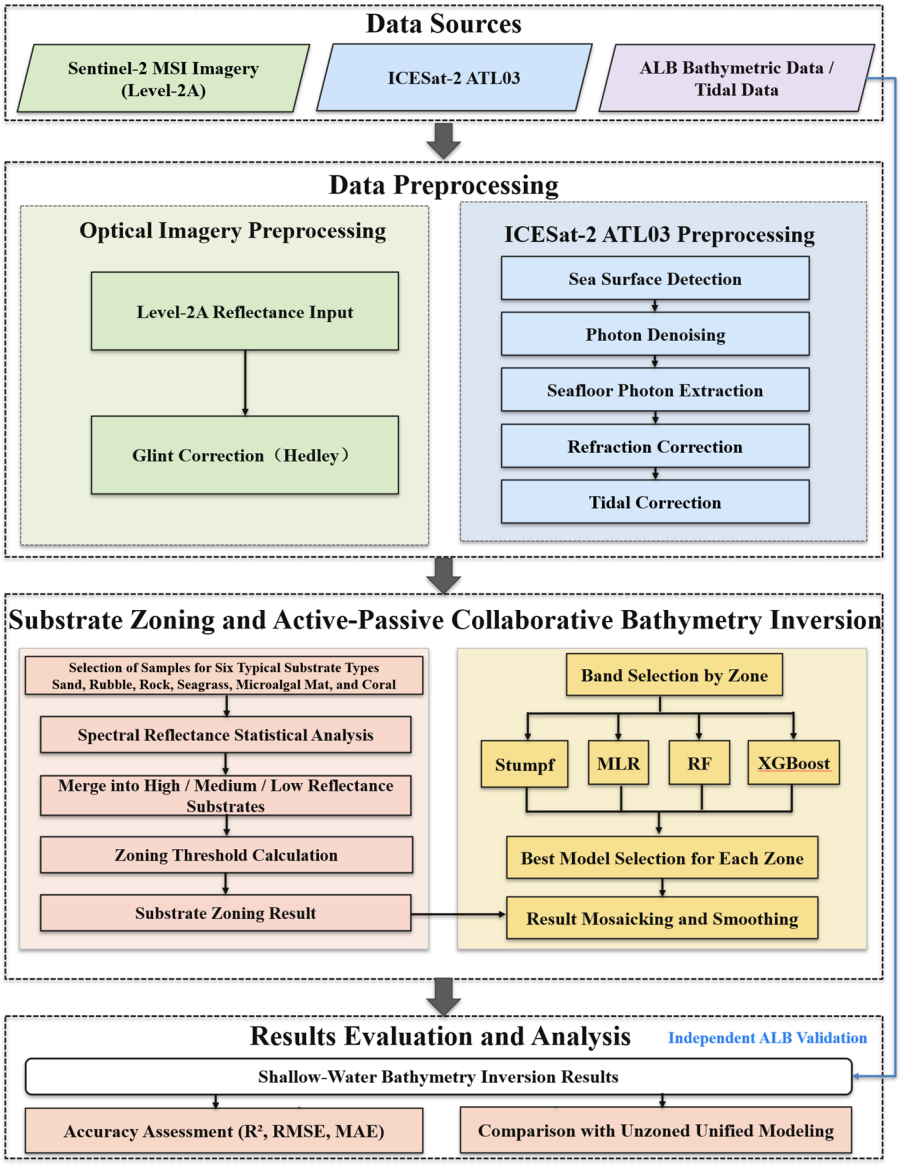

Figure 3. Workflow for bathymetry inversion using observed spectral-reflectance zoning and combined ICESat-2 and Sentinel-2 data.

The Sentinel-2 Level-2A products had already undergone radiometric calibration, orthorectification, and atmospheric correction through the standard processing chain. The scaled digital values of the selected 10 m bands were converted to bottom-of-atmosphere reflectance using the product scale factor. The images were then clipped to the respective study areas, and non-water areas were excluded from the subsequent analysis. Because only the native 10 m blue, green, red, and near-infrared bands were used, no inter-band resampling was required. Sunglint was corrected using the Hedley method [50], which estimates the glint contribution to each visible band from its linear relationship with the near-infrared band over selected deep-water samples. No additional spatial smoothing or adjacency-effect correction was applied before substrate zoning and bathymetric modeling.

3.1. Bathymetric Control Point Extraction from ICESat-2 ATL03 Data

ICESat-2 ATL03 photon data contain sea-surface photons, seafloor photons, and a large amount of background noise. In addition, the positions of underwater photons are affected by refraction and tidal variations. Therefore, signal extraction and geometric correction were required before the ATL03 data could be used for shallow water bathymetry mod-

eling. ATL03 was selected to allow direct control over photon extraction and subsequent corrections within a consistent processing workflow across the three study sites. Although ATL24 provides derived bathymetric information, direct processing of ATL03 enabled inspection of the photon distributions and intermediate processing results used to generate the bathymetric samples.

First, statistical histograms were constructed from along-track photon elevations to identify the instantaneous sea surface and initially separate surface photons from underwater photons. Photon processing was performed in along-track segments containing 10,000 raw photons, with confidence levels from 0 to 4 retained as input. DBSCAN was then applied to the underwater photons. The neighborhood radius was set to 1.5 m for daytime observations and 2.5 m for nighttime observations. MinPts was determined adaptively from the estimated signal-plus-noise and background-noise densities within each segment, with a lower bound of 3. Background-noise density was estimated from photons within the lowest 5 m elevation interval of each raw-photon segment. This initial filtering removed spatially isolated background noise while retaining candidate seafloor photons. The retained underwater photons were subsequently processed using the random sample consensus (RANSAC) algorithm with piecewise polynomial fitting along the track. Segment-wise fitting was used to accommodate local variations in seafloor topography and reduce the influence of residual outliers on seafloor signal extraction.

Because laser pulses are refracted at the air–water interface, the original horizontal positions and vertical elevations of seafloor photons deviate from their actual locations. Refraction correction was performed under a locally planar sea-surface assumption, using the sea-surface normal and laser incidence geometry to correct the horizontal positions and vertical elevations of the seafloor photons [51]. Along-track water depths were calculated from the corrected seafloor photon elevations and the corresponding instantaneous sea-surface elevations. Finally, tide-gauge observations and the NAO.99b tidal model were used, as described in Section 2.5, to harmonize the water-level references associated with the different acquisition times of ICESat-2, Sentinel-2, and airborne LiDAR data. The resulting along-track bathymetric points were used for subsequent zonal modeling and accuracy validation.

3.2. Substrate Zoning Method

An observed spectral-reflectance zoning strategy was adopted to compare bathymetric model performance under contrasting spectral conditions. The zones were defined using Sentinel-2 observed green-band reflectance, with Allen Coral Atlas data guiding representative substrate sample selection. Because observed reflectance is influenced by substrate properties, water depth, water-column optical properties, and residual surface and atmospheric effects, these zones characterize observed spectral conditions rather than independently retrieved intrinsic substrate reflectance.

First, based on the Allen Coral Atlas substrate-type data, six representative substrate types, including sand, rubble, rock, seagrass, microalgal mats, and coral/algae, were identified in the three study sites. The selected substrate-dominated regions were then overlaid with the preprocessed Sentinel-2 multispectral imagery to extract corresponding spectral reflectance values. The mean spectral reflectance of each substrate type was then calculated for subsequent spectral analysis. The number of valid pixels and the mean and standard deviation of Sentinel-2 green-band reflectance for each substrate type at the three study sites are summarized in Table 5. These statistics describe the observed green-band reflectance within the mapped substrate classes, including variations associated with water depth, water-column conditions, and mixed pixels.

Table 5. Valid pixel counts and Sentinel-2 green-band reflectance statistics for six substrate types at the three study sites.

Study Site	Substrate Type	Number of Valid Pixels	Green-Band Mean ± SD
Lingyang Reef	Sand	23,634	0.2258 ± 0.0393
	Rubble	63,853	0.2799 ± 0.0396
	Rock	36,068	0.2133 ± 0.0289
	Seagrass	7208	0.1951 ± 0.0127
	Microalgal mats	3366	0.2837 ± 0.0346
	Coral/Algae	9498	0.1687 ± 0.0103
Huaguang Reef	Sand	236,335	0.2486 ± 0.0567
	Rubble	224,389	0.2355 ± 0.0432
	Rock	143,142	0.1853 ± 0.0346
	Seagrass	12,168	0.1684 ± 0.0227
	Microalgal mats	118,237	0.2700 ± 0.0358
	Coral/Algae	110,288	0.1547 ± 0.0185
East Cay Reef	Sand	19,843	0.2295 ± 0.0354
	Rubble	2968	0.1828 ± 0.0184
	Rock	24,612	0.1785 ± 0.0520
	Seagrass	—	—
	Microalgal mats	1044	0.2504 ± 0.0149
	Coral/Algae	5966	0.1968 ± 0.1051

As shown in Figure 4, different substrates exhibit spectral differences in the visible range, although some substrate types remain spectrally similar. Based on the relative positions of their mean reflectance curves, the representative substrate samples were grouped into high-, medium-, and low-reflectance categories. The mean spectral curve of each group was calculated by averaging the reflectance values of all sample pixels within that group at each band. The green-band reflectance was selected as the zoning variable because of its water penetration capability and sensitivity to spectral differences in shallow reef waters. At each study site, the mean green-band reflectance was calculated for each group, and the midpoint between the means of adjacent groups was used as a boundary threshold. These two thresholds were then applied to classify water pixels into low-, medium-, and high-reflectance zones, representing observed spectral-reflectance conditions, for subsequent band selection and bathymetric model construction within each zone.

3.3. Bathymetric Inversion Algorithms

After observed spectral-reflectance zoning, bathymetric inversion models were constructed separately within each zone. Each ICESat-2-derived bathymetric point was matched to the Sentinel-2 pixel containing its geographic location, and the spectral values of that pixel were paired with the point's depth. When multiple bathymetric points fell within the same 10 m pixel, all points were retained as separate samples, sharing the same spectral values while preserving their individual depths. The Stumpf band-ratio and multiple linear regression (MLR) models used the blue (B2) and green (B3) bands, whereas RF and XGBoost used all four native 10 m bands: blue (B2), green (B3), red (B4), and near-infrared (B8). These band configurations were applied consistently across all reflectance zones. The four models were independently constructed within each observed spectral-reflectance zone, and the optimal model in each zone was selected according to the accuracy assessment described in Section 3.4.

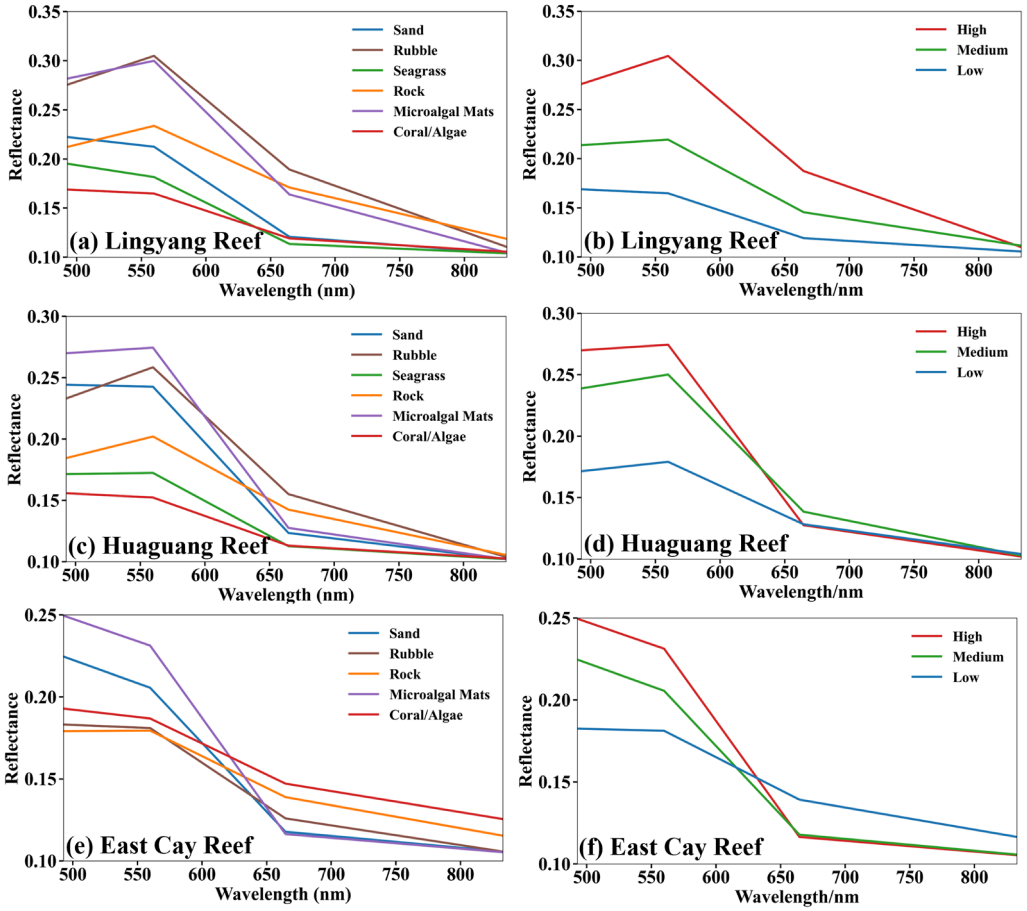

Figure 4. Typical substrate spectra from Sentinel-2. (**a**,**c**,**e**) Mean reflectance spectra associated with individual substrate classes at Lingyang Reef, Huaguang Reef, and East Cay Reef, respectively; (**b**,**d**,**f**) corresponding mean spectra of the high-, medium-, and low-reflectance groups used to derive the green-band zoning thresholds at the three reefs, respectively.

3.3.1. Band Ratio Model

The Stumpf band-ratio model exploits the differential attenuation of light at different wavelengths in the water column by constructing a logarithmic ratio between two spectral bands. The logarithmic band ratio can reduce part of the influence of variations in bottom reflectance and illumination while retaining sensitivity to water-depth variations [41]. In this study, candidate band combinations were evaluated separately within each substrate reflectance zone, and the selected combination was used to calibrate the Stumpf model:

$$Z = m_0 + m_1 \frac{\ln[n\rho(\lambda_i)]}{\ln[n\rho(\lambda_j)]} \quad (2)$$

where Z denotes the retrieved water depth; $\rho(\lambda_i)$ and $\rho(\lambda_j)$ are the sunglint-corrected Sentinel-2 bottom-of-atmosphere reflectance values of the two selected spectral bands; n is a constant used to ensure positive logarithmic values and was set to 1000 in this study; and m_0 and m_1 are regression coefficients obtained through model calibration.

3.3.2. Multiband Linear Regression Model

The multiband linear regression model (MLR) is a linear empirical method developed from classical bathymetric models [10,13]. This model assumes that shallow water depth can be expressed as a linear combination of sunglint-corrected Sentinel-2 bottom-of-atmosphere reflectance values from the selected spectral bands. By integrating spectral responses from different wavelengths, the MLR model can provide richer information for

bathymetric inversion, especially in areas with complex substrates or unstable single-band depth responses. In this study, bands with high correlations with water depth were selected within each substrate reflectance zone, and the model was expressed as

$$Z = a_0 + \sum_{i=1}^{k} a_i X_i \tag{3}$$

where Z denotes the retrieved water depth; k is the number of bands used in the model; X_i represents the reflectance variable of the i-th band; a_i is the corresponding regression coefficient; and a_0 is the constant term. The coefficients were estimated using the least squares method.

3.3.3. Random Forest Model

Random forest (RF) is a machine learning algorithm based on the bagging ensemble strategy and consists of multiple independent regression trees [52,53]. During model training, random sampling of training samples and random selection of input features are introduced to improve model robustness and reduce overfitting. RF can effectively handle high-dimensional spectral features and complex nonlinear relationships. In shallow water bathymetry, the relationship between spectral reflectance and water depth is often affected by substrate variability, and thus may not be strictly linear. In this study, multiband spectral variables within each substrate reflectance zone were used as input features, and ICESat-2-derived water depths were used as target values to construct the RF regression model. The final predicted depth is expressed as

$$Z = \frac{1}{N} \sum_{i=1}^{N} T_i(x) \tag{4}$$

where Z denotes the retrieved water depth; N is the number of regression trees; and $T_i(x)$ represents the prediction of the i-th regression tree for the input sample x. The RF model was implemented using scikit-learn with 100 trees and a random seed of 42. The remaining hyperparameters were left at their default values: no maximum tree-depth limit, all input features considered at each split, a minimum of one sample per leaf, and a minimum of two samples required to split an internal node. Bootstrap sampling was enabled, with each tree trained on a sample drawn with replacement whose size equaled the number of training samples.

3.3.4. Extreme Gradient Boosting Model

XGBoost is an ensemble learning algorithm that constructs regression trees sequentially to improve prediction accuracy and computational efficiency [54]. The XGBoost models were implemented using the Python XGBoost package, version 3.1. Unlike random forest, which builds multiple regression trees in parallel, XGBoost constructs trees sequentially, with each newly added tree fitting the residuals of the previous iteration. The algorithm introduces regularization terms into the objective function to control model complexity, thereby enhancing fitting ability while reducing overfitting. In shallow water bathymetry, the relationship between spectral reflectance and water depth can become nonlinear under complex substrate conditions. In this study, multiband spectral variables within each substrate reflectance zone and ICESat-2-derived depths were used to construct the XGBoost regression model. The objective function is expressed as

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{5}$$

where l is the loss function; y_i is the observed water depth; $\hat{y}_i^{(t-1)}$ is the prediction from the previous iteration; $f_t(x_i)$ is the prediction of the newly added tree at iteration t; and $\Omega(f_t)$ is the regularization term used to constrain model complexity. The XGBoost model used a squared-error loss for regression, with 100 boosting trees, a learning rate of 0.1, and a random seed of 42. The regularization parameters were retained at their default values: $\alpha = 0$ for L1 regularization, $\lambda = 1$ for L2 regularization, and $\gamma = 0$ for the minimum loss reduction required for a split.

3.4. Accuracy Assessment

The accuracy of the bathymetric inversion results was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). R^2 indicates the explanatory ability of the model for water depth variations. RMSE reflects the overall deviation between the retrieved and measured depths, and is sensitive to large errors. MAE represents the mean absolute difference between the retrieved and measured depths. These metrics were used to compare the performance of different models and substrate reflectance zones. The formulas are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y_i})^2} \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{7}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right| \tag{8}$$

where n is the number of validation samples; y_i is the measured depth of the i-th validation point; $\hat{y}_i$ is the retrieved depth; and $\overline{y_i}$ is the mean value of the measured depths.

For each image dataset and each substrate reflectance zone, the four candidate models were evaluated using validation samples that were not involved in model training. The model with the highest validation R^2 was selected for zonal integration based on its ability to explain the observed depth variation. RMSE and MAE were used to characterize the magnitude of bathymetric errors in meters.

4. Results

4.1. ICESat-2 ATLAS Bathymetric Extraction Results

After sea-surface identification, photon denoising, refraction correction, and tidal normalization, ICESat-2 along-track bathymetric control points were obtained for subsequent model construction, as shown in Figure 5. Taking the GT2R track on 15 August 2023 and the GT3R track on 18 August 2021 as examples, the extraction results are shown in Figure 5a,c, respectively. The light gray points represent the original photons, the blue points indicate the extracted sea-surface photons, the orange points denote the extracted seafloor photons, and the red points represent the seafloor photons after refraction correction. Photon filtering removed most spatially isolated background photons while retaining continuous sea-surface and seafloor photon sequences. After refraction correction, the seafloor photon profiles showed a more coherent bathymetric pattern and were generally consistent with reef-topographic variations.

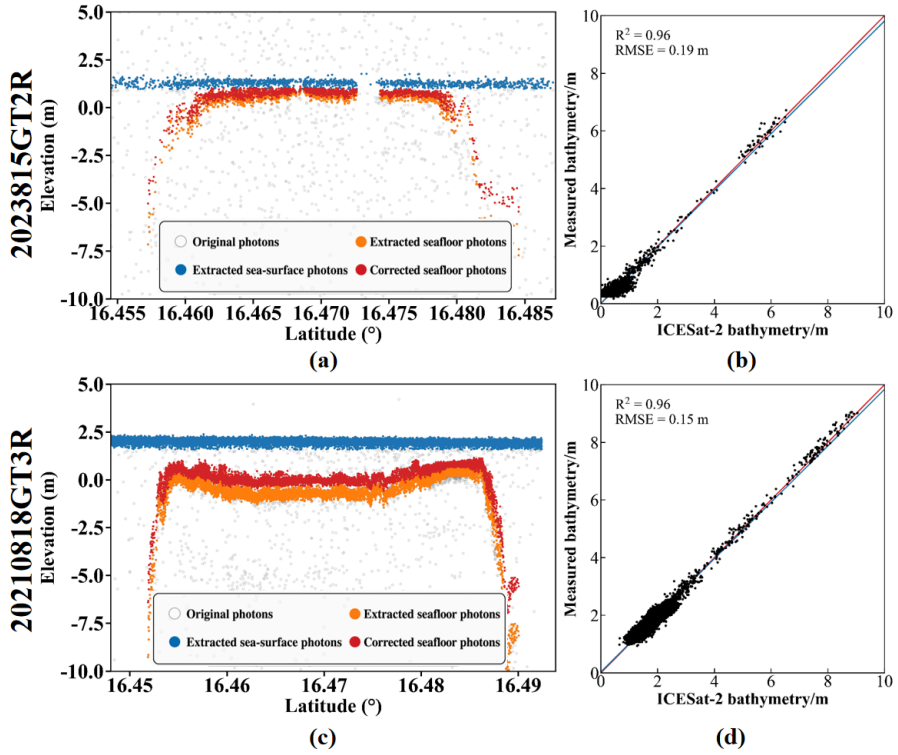

Figure 5. ICESat-2 seafloor photon extraction, refraction correction, and validation against ALB bathymetry for two representative tracks at Lingyang Reef. (**a**,**b**) GT2R acquired on 15 August 2023; (**c**,**d**) GT3R acquired on 18 August 2021. In (**a**,**c**), gray points represent the original ATL03 photons, blue points the extracted sea-surface photons, orange points the extracted seafloor photons, and red points the refraction-corrected seafloor photons. In (**b**,**d**), black points show ICESat-2-derived depths versus measured bathymetry; the red line denotes the 1:1 relationship and the blue line denotes the fitted regression.

The validation results are shown in Figure 5b,d. The validation points were concentrated around the 1:1 line, and both datasets achieved an R^2 of 0.96. The two validation groups both achieved an R^2 of 0.96, with RMSE values of 0.19 m and 0.15 m, respectively. These values describe agreement between ICESat-2-derived and ALB depths along the two selected tracks. The reported ALB depth accuracy of 0.23 m is a reference performance indicator rather than a lower bound on differences within every matched subset. Therefore, the observed RMSE values should not be interpreted as independently established absolute ICESat-2 accuracy or evidence of superior accuracy relative to ALB. The agreement supports the use of the extracted depths as bathymetric control data, subject to reference data uncertainty and temporal mismatch.

4.2. Substrate Zoning Results

For each reef, the mean Sentinel-2 green-band reflectance was calculated for the representative samples in each of the low-, medium-, and high-reflectance groups. The two zoning thresholds were defined as the arithmetic midpoints between adjacent group means. Group means and thresholds were calculated independently for Lingyang Reef, Huaguang Reef, and East Cay Reef using the same procedure. The threshold intervals for observed spectral-reflectance zoning are listed in Table 6. The medium-to-high reflectance thresholds were similar at Lingyang Reef and Huaguang Reef (0.2620 and 0.2623, respectively), whereas a lower threshold of 0.2184 was obtained for East Cay Reef.

Table 6. Threshold intervals for observed spectral-reflectance zoning in different study areas.

Study Area	Low-Reflectance Zone	Medium-Reflectance Zone	High-Reflectance Zone
Lingyang Reef	[0,0.1921)	[0.1921,0.2620)	[0.2620,0.5520]
Huaguang Reef	[0,0.2147)	[0.2147,0.2623)	[0.2623,0.6116]
East Cay Reef	[0,0.1934)	[0.1934,0.2184)	[0.2184,0.7696]

Using the thresholds in Table 6, the Sentinel-2 imagery was reclassified into high-, medium-, and low-reflectance zones, as shown in Figure 6. The three reflectance zones exhibited distinct spatial distributions within each reef. High- and medium-reflectance zones were mainly distributed over shallow reef-flat and marginal areas, whereas low-reflectance zones occupied larger portions of lagoonal and relatively deeper-water regions. Although the proportions and spatial configurations of the three zones differed among the study areas, the zoning results clearly separated areas with contrasting reflectance backgrounds for subsequent model comparison.

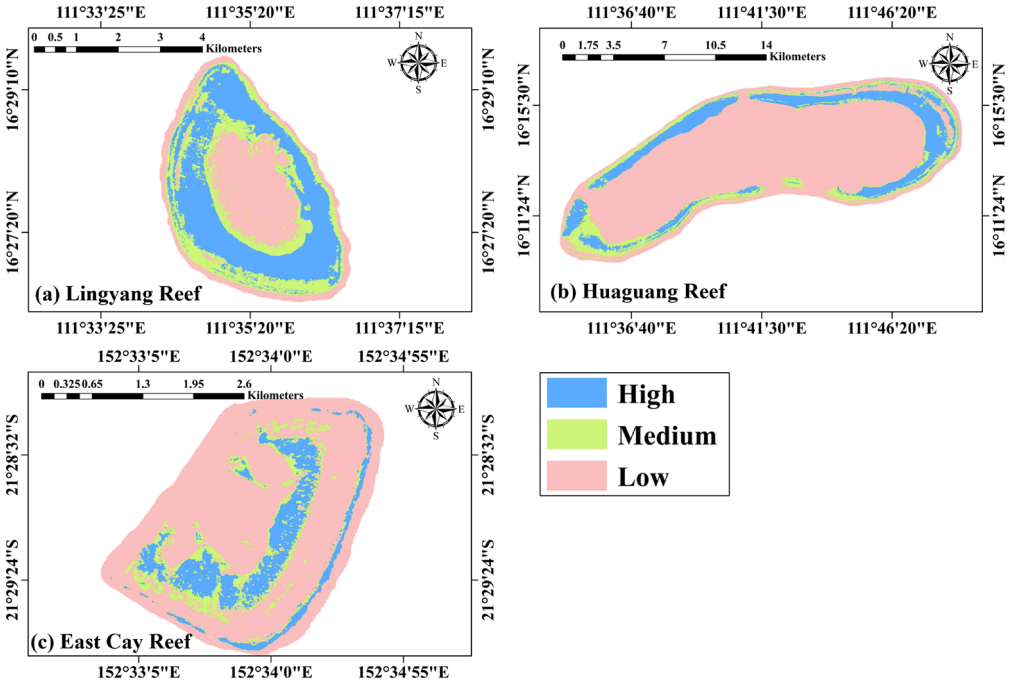

Figure 6. Spatial distributions of the Sentinel-2 green-band reflectance zones at (**a**) Lingyang Reef, (**b**) Huaguang Reef, and (**c**) East Cay Reef. Blue, green, and pink areas denote the high-, medium-, and low-reflectance zones, respectively.

4.3. Accuracy Comparison of Different Bathymetric Models

Based on the substrate zoning results, the Stumpf band-ratio, MLR, RF, and XGBoost models were constructed separately within the high-, medium-, and low-reflectance zones. For each substrate reflectance zone, the optimal model was selected according to the accuracy assessment strategy described in Section 3.4. The accuracy assessment results for Sentinel-2 are summarized in Table 7. Overall, the preferred model varied among the three reflectance conditions, with linear models showing stronger overall performance in the high-reflectance zones and RF and XGBoost becoming more competitive in the medium and low reflectance zones. The standard deviations of the validation depths were 0.63–1.27 m in the high-reflectance zones, 1.43–2.55 m in the medium-reflectance zones, and 2.62–3.64 m in the low-reflectance zones. Greater depth variance can yield a higher R^2 for a similar

RMSE. Comparisons across zones therefore considered the validation-depth distributions together with R^2, RMSE, and MAE.

Table 7. Model accuracy assessment results for different substrate zones using Sentinel-2 images.

Study Area	Zone	Validation n	Depth Range (m)	Depth SD (m)	Model	R^2	RMSE (m)	MAE (m)
Lingyang Reef	High	12297	0.26–5.03	0.63	Stumpf	0.66	0.55	0.46
					MLR	0.62	0.56	0.47
					RF	0.65	0.55	0.44
					XGBoost	0.70	0.53	0.43
	Medium	5819	0.25–13.07	1.87	Stumpf	0.81	0.82	0.64
					MLR	0.78	0.87	0.68
					RF	0.90	0.60	0.48
					XGBoost	0.90	0.59	0.48
	Low	1394	0.30–26.32	3.46	Stumpf	0.70	1.89	1.18
					MLR	0.69	1.92	1.19
					RF	0.87	1.24	0.85
					XGBoost	0.87	1.24	0.88
Huaguang Reef	High	9874	0.49–5.18	0.81	Stumpf	0.58	0.52	0.41
					MLR	0.63	0.49	0.39
					RF	0.53	0.55	0.46
					XGBoost	0.57	0.53	0.44
	Medium	7676	0.57–12.72	1.43	Stumpf	0.82	0.61	0.46
					MLR	0.82	0.60	0.44
					RF	0.86	0.54	0.41
					XGBoost	0.87	0.52	0.41
	Low	4865	0.44–18.69	3.64	Stumpf	0.80	1.63	1.13
					MLR	0.75	1.82	1.33
					RF	0.89	1.20	0.82
					XGBoost	0.90	1.17	0.79
East Cay Reef	High	2652	0.46–7.65	1.27	Stumpf	0.81	0.51	0.36
					MLR	0.87	0.43	0.33
					RF	0.76	0.58	0.44
					XGBoost	0.78	0.55	0.43
	Medium	2242	0.72–10.78	2.55	Stumpf	0.92	0.75	0.47
					MLR	0.90	0.80	0.49
					RF	0.92	0.72	0.49
					XGBoost	0.83	1.06	0.64
	Low	3900	0.21–27.07	2.62	Stumpf	0.73	1.28	0.81
					MLR	0.67	1.41	0.99
					RF	0.77	1.18	0.66
					XGBoost	0.77	1.18	0.63

In the high-reflectance zones, the best-performing model differed among the three study areas. XGBoost achieved the highest R^2 and the lowest RMSE at Lingyang Reef, whereas MLR performed best at Huaguang Reef and East Cay Reef. Averaged across the three study areas, MLR achieved the highest mean R^2 of 0.71 and the lowest mean RMSE and MAE of 0.49 m and 0.40 m, respectively. Stumpf showed comparable but slightly lower overall performance, with a mean RMSE of 0.53 m. On average, MLR performed best across the three study areas under high reflectance conditions, although the model with the highest local accuracy differed among the reefs.

In the medium-reflectance zones, RF and XGBoost generally achieved higher accuracies than Stumpf and MLR. XGBoost produced the best local results at Lingyang Reef and Huaguang Reef, whereas RF achieved the lowest RMSE at East Cay Reef. Across the three study areas, RF yielded the highest mean R^2 of 0.89 and the lowest mean RMSE and MAE of 0.62 m and 0.46 m, respectively. XGBoost showed similar performance at Lingyang Reef and Huaguang Reef, but its RMSE increased to 1.06 m at East Cay Reef, whereas RF produced more consistent results among the three study areas.

In the low-reflectance zones, the performance differences between the machine learning and empirical/linear models became more pronounced. RF and XGBoost outperformed Stumpf and MLR in all three study areas. XGBoost showed the best or tied-best performance across the three reefs, with mean R^2, RMSE, and MAE values of 0.85, 1.20 m, and 0.77 m, respectively. RF produced very similar results, with corresponding mean values of 0.84, 1.21 m, and 0.78 m. In comparison, the mean RMSE values of Stumpf and MLR increased to 1.60 m and 1.72 m, respectively. Thus, the performance advantage of RF and XGBoost over the empirical and linear models was most evident in the low-reflectance zones.

4.4. SDB Results

Based on the optimal models selected for the different reflectance zones in Section 4.3, the corresponding bathymetric results were integrated and smoothed to generate the final SDB maps. The bathymetric results for Lingyang Reef, Huaguang Reef, and East Cay Reef are shown in Figures 7–9, respectively. In each figure, the results from Stumpf, MLR, RF, and XGBoost are compared with the integrated zonal result and the corresponding substrate zoning map.

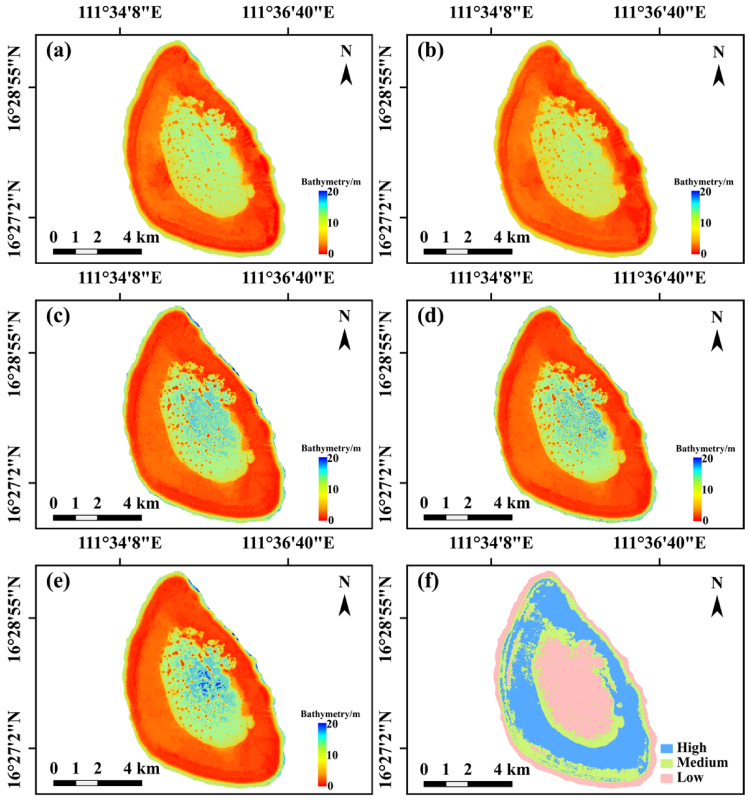

Figure 7. Comparison of bathymetric results derived from Sentinel-2 images over Lingyang Reef: (**a**) band ratio model; (**b**) multiband linear regression model; (**c**) random forest model; (**d**) extreme gradient boosting model; (**e**) integrated zonal result based on the optimal model selected for each reflectance zone; and (**f**) substrate zoning result.

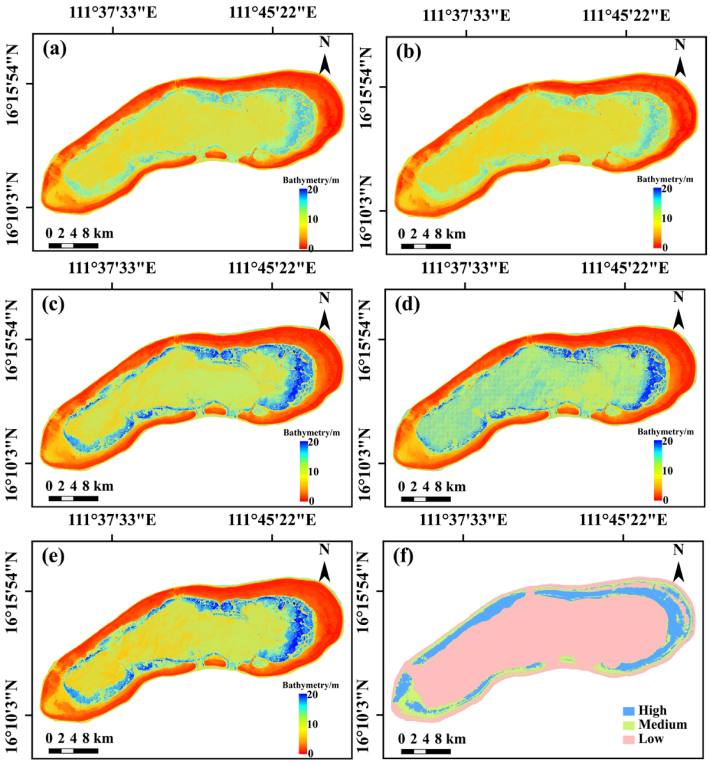

Figure 8. Comparison of bathymetric results derived from Sentinel-2 images over Huaguang Reef: (**a**) band ratio model; (**b**) multiband linear regression model; (**c**) random forest model; (**d**) extreme gradient boosting model; (**e**) integrated zonal result based on the optimal model selected for each reflectance zone; and (**f**) substrate zoning result.

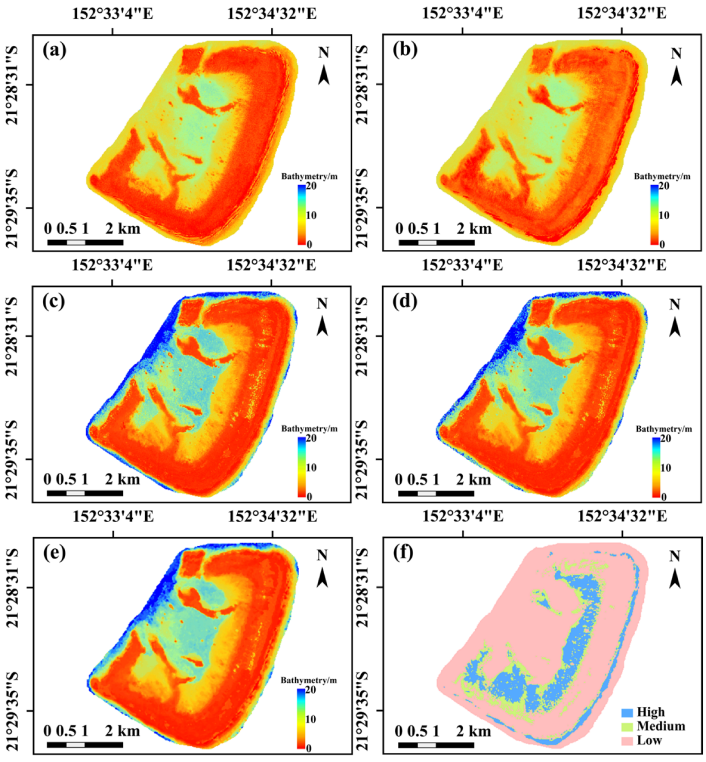

Figure 9. Comparison of bathymetric results derived from Sentinel-2 images over East Cay Reef: (**a**) band ratio model; (**b**) multiband linear regression model; (**c**) random forest model; (**d**) extreme gradient boosting model; (**e**) integrated zonal result based on the optimal model selected for each reflectance zone; and (**f**) substrate zoning result.

Clear differences in the spatial distribution of retrieved water depths were observed among the four models. The Stumpf and MLR results were generally smoother, but they tended to retrieve shallower depths in deeper lagoonal areas and along the reef margins. In contrast, RF and XGBoost retained greater depth variation in these areas and more clearly reproduced the transition from shallow to deep water. However, local fragmentation and small spatial fluctuations were also more evident in some parts of the RF and XGBoost results. These differences were particularly apparent in areas with large depth gradients and heterogeneous reflectance.

The integrated zonal results combined the models that achieved the highest validation accuracy in the high-, medium-, and low-reflectance zones. Compared with single-model results, the integrated zonal maps retained the stronger response of RF or XGBoost in deeper areas, where these models performed better, while preserving relatively continuous transitions between adjacent reflectance zones. The spatial differences shown in Figures 7–9 are consistent with the quantitative comparison presented in Section 4.3.

4.5. Bathymetric Profile Curves

To further compare the ability of the different models to reproduce local depth variations, representative profile lines were selected at Lingyang Reef, Huaguang Reef, and East Cay Reef. The retrieved depths from the four individual models and the integrated zonal result were compared with ICESat-2 bathymetric points, as shown in Figure 10. Figure 10a,c and e show the locations of the selected profiles, whereas Figure 10b,d and f show the corresponding bathymetric profile curves. The light blue shading shows a fixed ± 2 m reference band around the ICESat-2 reference profile, selected to facilitate visual comparison of local departures from the reference depths.

Across the three profiles, the differences among the models were relatively small along sections with gradual depth variation but became more pronounced near slope breaks and transitions from shallow to deep water. At Lingyang Reef and Huaguang Reef, the depths retrieved by the models generally followed the ICESat-2 depth variations, with the largest differences occurring locally near sections of rapid depth change.

At East Cay Reef, the differences among the models were more pronounced in the deeper part of the profile. MLR increasingly underestimated the ICESat-2 depths after the profile entered the deeper section, whereas RF and XGBoost followed the local deepening trend more closely, although some local fluctuations remained. The integrated zonal result showed relatively close agreement with the ICESat-2 profile across the three study areas and maintained continuous depth transitions between shallow reef-flat sections and deeper lagoonal areas or waters along the reef margins. These profile comparisons are consistent with the quantitative results reported in Section 4.3.

4.6. Overall Accuracy Assessment of the Integrated Zonal Model

The overall accuracy of the integrated zonal model was compared with that of the four unzoned baseline models at Lingyang Reef, Huaguang Reef, and East Cay Reef. The accuracy statistics are summarized in Table 8. The integrated zonal model achieved R^2 values of 0.93, 0.91, and 0.87, RMSE values of 0.76, 0.61, and 0.84 m, and MAE values of 0.54, 0.42, and 0.48 m at the three reefs, respectively. Compared with RF, which achieved the lowest RMSE among the four unzoned baseline models at each reef, the integrated zonal model reduced RMSE by approximately 3.8%, 4.7%, and 12.5%, respectively. The largest reduction was observed at East Cay Reef. At Huaguang Reef, however, RF achieved a slightly lower MAE than the integrated zonal model (0.41 versus 0.42 m).

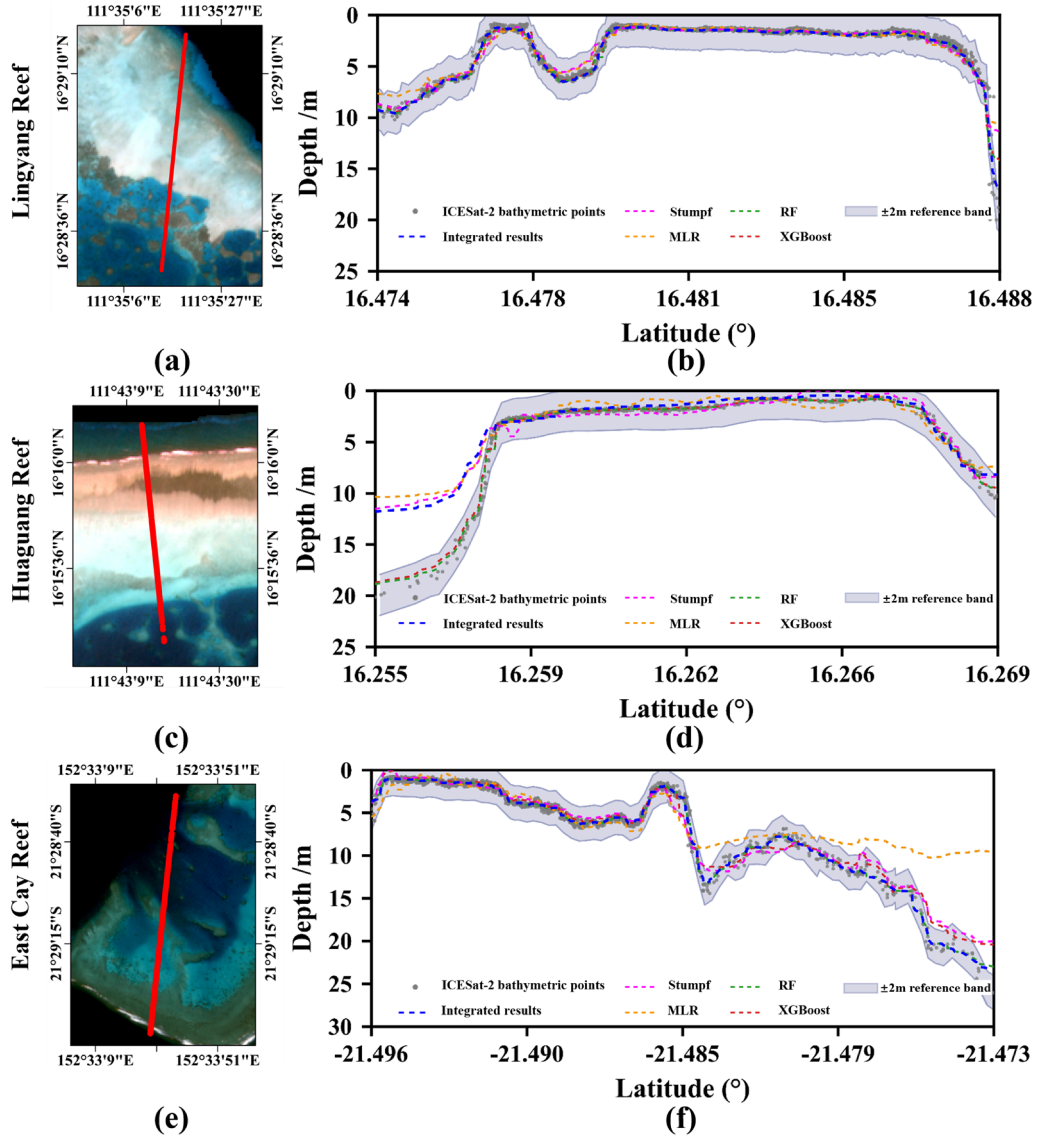

Figure 10. Comparison of bathymetric profile curves from different models. (**a**,**c**,**e**) show the profile line locations in Lingyang Reef, Huaguang Reef, and East Cay Reef, respectively; (**b**,**d**,**f**) show the corresponding comparisons between depths retrieved by the models and ICESat-2 bathymetric points along the profiles. The light blue shading indicates the ±2 m reference band used for visual comparison.

Table 8. Accuracy statistics of different bathymetric inversion models using Sentinel-2 images.

Study Area	Model	R^2	RMSE (m)	MAE (m)
Lingyang Reef	Integrated method	0.93	0.76	0.54
	Stumpf	0.87	1.10	0.77
	MLR	0.87	1.29	0.86
	RF	0.92	0.79	0.56
	XGBoost	0.91	0.85	0.60
Huaguang Reef	Integrated method	0.91	0.61	0.42
	Stumpf	0.88	0.72	0.46
	MLR	0.85	0.83	0.52
	RF	0.91	0.64	0.41
	XGBoost	0.88	0.73	0.52

Table 8. *Cont.*

Study Area	Model	R^2	RMSE (m)	MAE (m)
East Cay Reef	Integrated method	0.87	0.84	0.48
	Stumpf	0.78	1.30	1.03
	MLR	0.77	1.16	0.87
	RF	0.83	0.96	0.56
	XGBoost	0.85	1.07	0.63

Figure 11 presents the relationships between estimated and reference water depths for the five approaches at each reef. The integrated zonal model achieved higher R^2 values than Stumpf and MLR at all three reefs. Compared with RF and XGBoost, it achieved the highest R^2 at Lingyang Reef and East Cay Reef and the same R^2 as RF at Huaguang Reef. The corresponding residual distributions are presented in Figure 12, complementing the overall accuracy statistics in Table 8.

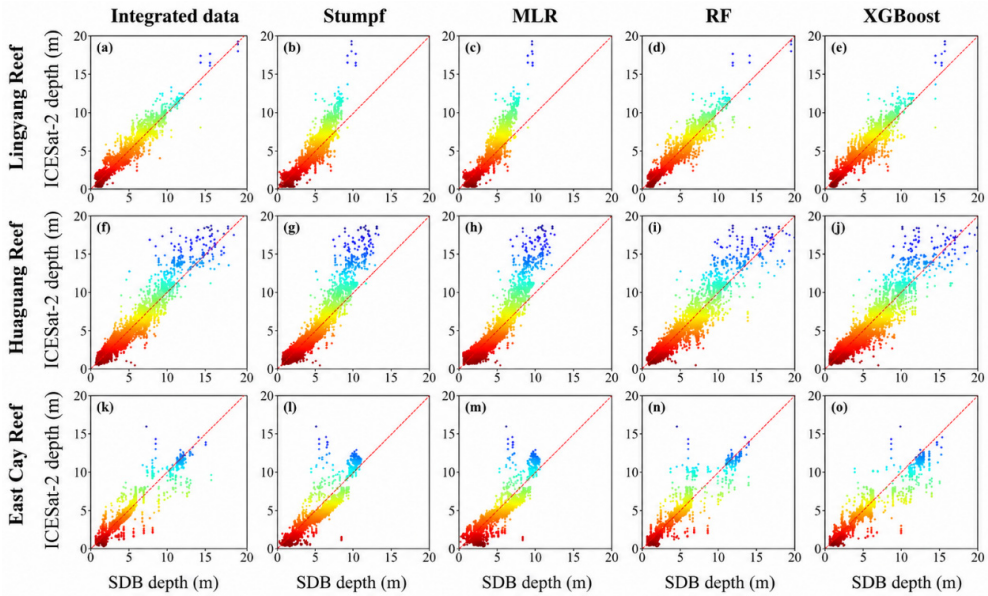

Figure 11. Scatterplots of SDB estimates against independent validation depths: (**a**) integrated zonal result at Lingyang Reef; (**b**) Stumpf at Lingyang Reef; (**c**) MLR at Lingyang Reef; (**d**) RF at Lingyang Reef; (**e**) XGBoost at Lingyang Reef; (**f**) integrated zonal result at Huaguang Reef; (**g**) Stumpf at Huaguang Reef; (**h**) MLR at Huaguang Reef; (**i**) RF at Huaguang Reef; (**j**) XGBoost at Huaguang Reef; (**k**) integrated zonal result at East Cay Reef; (**l**) Stumpf at East Cay Reef; (**m**) MLR at East Cay Reef; (**n**) RF at East Cay Reef; and (**o**) XGBoost at East Cay Reef. The red line in each panel represents the 1:1 relationship.

To assess variations in bathymetric accuracy with water depth, the independent ALB validation data at Lingyang Reef were divided into four depth intervals: 0–5, 5–10, 10–15, and 15–20 m. MAE and RMSE were calculated separately for the integrated zonal results and the four unzoned baseline models within each interval, as shown in Figure 13.

The errors of all models generally increased with water depth. Within the 0–10 m range, differences in MAE and RMSE among the models were relatively small. Beyond 10 m, the errors of Stumpf and MLR increased more rapidly, whereas RF and XGBoost showed slower error growth. Nevertheless, the errors of RF and XGBoost also increased in the 15–20 m interval. The integrated zonal results maintained relatively low errors across the four depth intervals. This depth-dependent assessment was restricted to Lingyang Reef, where independent ALB validation data were available.

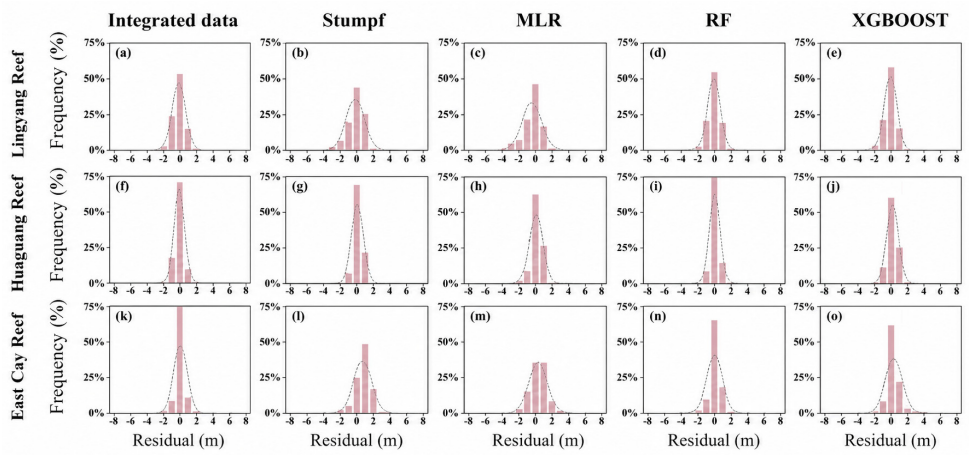

Figure 12. Residual distributions of SDB estimates derived from Sentinel-2 imagery: (**a**) integrated zonal result at Lingyang Reef; (**b**) Stumpf at Lingyang Reef; (**c**) MLR at Lingyang Reef; (**d**) RF at Lingyang Reef; (**e**) XGBoost at Lingyang Reef; (**f**) integrated zonal result at Huaguang Reef; (**g**) Stumpf at Huaguang Reef; (**h**) MLR at Huaguang Reef; (**i**) RF at Huaguang Reef; (**j**) XGBoost at Huaguang Reef; (**k**) integrated zonal result at East Cay Reef; (**l**) Stumpf at East Cay Reef; (**m**) MLR at East Cay Reef; (**n**) RF at East Cay Reef; and (**o**) XGBoost at East Cay Reef.

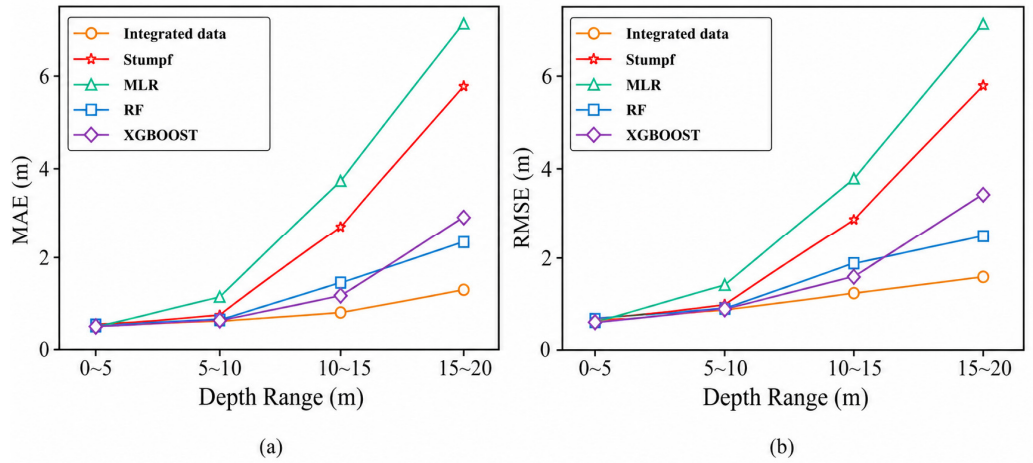

Figure 13. (**a**) MAE and (**b**) RMSE of the integrated zonal result and four bathymetric models for different water depth intervals at Lingyang Reef, evaluated using independent ALB validation data.

5. Discussion

5.1. Model Suitability Under Different Substrate Reflectance Conditions

The relative performance of the four models varied among the reflectance zones. In the high-reflectance zones, MLR achieved the lowest RMSE at Huaguang Reef and East Cay Reef, with values of 0.49 and 0.43 m, respectively. At Lingyang Reef, XGBoost performed best, although its RMSE of 0.53 m was only slightly lower than the 0.56 m obtained by MLR. These results show that empirical and linear models remained competitive under high-reflectance conditions, and that greater model complexity did not consistently improve accuracy.

The advantage of RF and XGBoost became more evident in the medium- and low-reflectance zones. At Lingyang Reef, their RMSE values in the medium-reflectance zone were 0.60 and 0.59 m, compared with 0.82 m for Stumpf and 0.87 m for MLR. In the low-reflectance zone, both nonlinear models achieved an RMSE of 1.24 m, whereas Stumpf and MLR yielded 1.89 and 1.92 m. Huaguang Reef showed a similar pattern. However, XGBoost was less effective in the medium-reflectance zone at East Cay Reef, where its RMSE reached 1.06 m compared with 0.72 m for RF. Thus, nonlinear models were generally

advantageous under lower-reflectance conditions, but their relative performance remained site-dependent.

Substrate reflectance provides a physical context for these differences. At comparable depths and under similar water-column conditions, a brighter substrate produces a stronger bottom-reflected signal. Such a signal can make spectral variations associated with water depth more readily detectable. Conversely, when the bottom contribution is weaker, water-column contributions and residual spectral variability become more influential in the observed signal. The resulting spectral–depth relationship may be less adequately represented by a fixed band ratio or a linear combination of bands. The flexibility of RF and XGBoost in representing nonlinear responses offers a plausible explanation for their stronger performance in the lower-reflectance zones.

Nevertheless, the reflectance categories describe the signal observed by Sentinel-2 rather than intrinsic substrate reflectance alone. Differences within substrate classes and mixtures of substrates within individual pixels contribute to spectral variability, while water depth and attenuation modify the signal reaching the sensor. Consequently, the correspondence between the six substrate types and the three reflectance categories represents a statistical association rather than a fixed optical classification.

The integrated zonal result accommodated these local differences in model suitability. Its largest improvement occurred at East Cay Reef, where MLR performed best in the high-reflectance zone and RF or XGBoost performed better in the lower-reflectance zones. The overall RMSE decreased from 0.96 m for the best unzoned model to 0.84 m for the integrated zonal result. Smaller reductions occurred at Lingyang Reef and Huaguang Reef, where the unzoned RF model already achieved relatively high accuracy. The benefit of integration therefore varied with the degree of complementarity among the models within each reef.

5.2. Influence of Water Depth on Inversion Accuracy

Water depth affects both the strength and spectral composition of the bottom contribution to satellite observations. Light traveling through the water column undergoes wavelength-dependent absorption and scattering before reaching the seabed and during its return toward the surface. As depth increases, the bottom contribution generally decreases, while the relative importance of the water-column signal changes. Bottom-reflectance differences that are distinguishable in shallow water can therefore become less apparent at greater depths.

The depth-dependent results at Lingyang Reef were consistent with this weakening of bottom information. Within 0–10 m, differences in MAE and RMSE among the models were relatively small. Beyond 10 m, Stumpf and MLR errors increased more rapidly, whereas RF and XGBoost showed slower error growth. This contrast suggests that the fixed spectral relationships used by the empirical and linear models became less effective in the deeper intervals. RF and XGBoost were better able to represent the relationships captured by the calibration samples, although their errors also increased in the 15–20 m interval. The integrated zonal result maintained relatively low errors across the four intervals, but likewise remained subject to the depth-dependent limitations of optical bathymetry. Accordingly, similar substrates at similar depths need not produce identical spectral responses under different water conditions. These processes provide a possible explanation for part of the variation in model performance among reefs, alongside differences in substrate composition and sample distributions. The model rankings themselves, however, do not establish differences in water clarity among the study areas.

The spatial overlap between low-reflectance zones and deeper lagoonal areas connects these optical effects to the zoning results. A low observed reflectance can arise from a

relatively dark substrate, attenuation over a longer water path, or their combined influence. The stronger performance of RF and XGBoost in these zones is therefore consistent with a more complex spectral–depth relationship, rather than an effect attributable exclusively to substrate brightness. Taken together, the results show that observed reflectance provides a useful basis for local model comparison, while water depth and water-column processes remain essential to interpreting the resulting performance differences.

5.3. Limitations

Several limitations should be considered when interpreting the model performance observed in this study. First, the three study areas are coral reef environments dominated by relatively clear shallow waters. Patterns observed here may therefore differ in turbid coastal waters, estuaries, or regions with high concentrations of suspended matter, where the optical properties of the water column exert a stronger influence on the observed spectral signal. Water-column optical variability was not independently quantified in this study. Its contribution to the observed reflectance zones and differences in model performance therefore remains unresolved. Consequently, the results characterize model performance under the observed spectral conditions and should not be interpreted as isolating the effects of substrate reflectance, water depth, or water-column variability.

Second, the distribution of ICESat-2 samples along satellite tracks limits the spatial density and coverage of available bathymetric data. This limitation is more pronounced in less extensive reflectance zones and deeper water intervals, where relatively small sample sizes may affect model calibration and the stability of accuracy estimates. In addition, ALB reference data were available for Lingyang Reef and Huaguang Reef, whereas East Cay Reef was evaluated using an independent ICESat-2 track excluded from training and model selection. Differences in reference coverage limit cross-site comparability. Cross-temporal comparisons assume broadly stable reef morphology under ordinary conditions, although local changes caused by waves, currents, ecological processes, or storms cannot be excluded. Tidal correction cannot remove these changes. Without higher-accuracy multibeam reference data, measurement uncertainty and temporal changes could not be separately quantified; validation statistics therefore represent cross-dataset agreement rather than independently established absolute accuracy. Confidence intervals for the RMSE differences between the integrated zonal model and the best unified model were not estimated, and their statistical significance remains untested. Future assessment should use track- or spatial-block-based resampling to account for spatial dependence.

Third, the classification of reef areas into zones with high, medium, and low reflectance provides a simplified representation of continuous substrate and spectral heterogeneity. Mixed pixels may contain different proportions of coral, sand, rubble, vegetation, and other substrate components, but these subpixel contributions were not explicitly quantified in the present analysis. The reflectance zones may also overlap with differences in water depth, as indicated by the analysis of different depth intervals at Lingyang Reef.

The operational transferability of the zoning strategy remains unverified. Each reef was evaluated using locally derived thresholds and calibrated models, and no leave-one-reef-out experiment was conducted. Where Allen Coral Atlas or equivalent substrate information is unavailable, the present threshold derivation procedure requires an alternative source of representative samples. Image-based sample selection or unsupervised reflectance grouping may offer alternatives, but their effectiveness requires separate validation. Future work should evaluate the observed patterns of model performance using more independent bathymetric datasets and a wider range of water conditions. More balanced sampling across reflectance and depth intervals, together with in situ substrate spectra, quantitative information on mixed-substrate composition, and independent measurements

of water-column optical properties, would help distinguish the respective influences of substrate reflectance, water depth, and water-column optical variability on bathymetric model performance.

6. Conclusions

This study combined ICESat-2 bathymetric control points with Sentinel-2 multispectral imagery to evaluate four bathymetric models within high-, medium-, and low-observed green-band reflectance zones at Lingyang Reef, Huaguang Reef, and East Cay Reef. The main conclusions are as follows:

(1) Model performance varied among the observed reflectance zones. MLR achieved the best mean performance in the high-reflectance zones, with an R^2 of 0.71 and an RMSE of 0.49 m. RF achieved the highest mean accuracy in the medium-reflectance zones, with an R^2 of 0.89 and an RMSE of 0.62 m.

(2) RF and XGBoost consistently outperformed Stumpf and MLR in the low-reflectance zones across the three reefs. XGBoost achieved a mean R^2 of 0.85 and an RMSE of 1.20 m, while RF produced very similar results.

(3) The depth-stratified analysis using ALB data at Lingyang Reef showed more pronounced differences among models beyond 10 m, where RF and XGBoost exhibited slower increases in error than Stumpf and MLR. However, the separate contributions of observed reflectance and water depth remain unresolved.

(4) Integrating the locally selected models yielded RMSE values of 0.76, 0.61, and 0.84 m at Lingyang Reef, Huaguang Reef, and East Cay Reef, respectively, representing reductions of approximately 3.8%, 4.7%, and 12.5% relative to the best single model applied across each study area. These improvements were observed on the evaluated datasets; their statistical significance remains untested.

Observed green-band reflectance zoning provided a practical basis for local model comparison and selection. However, observed reflectance incorporates benthic properties, water depth, and water-column optical conditions, whose independent effects were not quantified. Application to other reefs requires local calibration and validation.

Author Contributions: Conceptualization, Y.L., Y.C., L.W. and L.Y.; Methodology, Y.L., Y.C. and L.W.; Software, Y.L. and Y.C.; Validation, Y.L., Y.C. and L.W.; Formal analysis, Y.L., Y.C. and L.W.; Investigation, Y.L., B.X., Y.C., L.W. and L.Y.; Resources, Y.L., B.X., Y.C., L.Y. and X.S.; Data curation, Y.L., B.X., Y.C., L.W. and X.S.; Writing—original draft, B.X., L.W., L.Y. and X.S.; Writing—review & editing, B.X., L.W., L.Y. and X.S.; Visualization, B.X., L.Y. and X.S.; Supervision, Y.L., B.X., L.W. and L.Y.; Project administration, Y.L., B.X., Y.C., L.Y. and X.S.; Funding acquisition, Y.L., B.X., Y.C., L.Y. and X.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program of China (Grant No. 2023YFB3907705-3), the National Natural Science Foundation of China (Grant No. 42671454) and the Innovation Research Group Project of Hubei Provincial Natural Science Foundation (Grant No. 2026AFA005).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sun, M.; Yu, L.; Zhang, P.; Sun, Q.; Jiao, X.; Sun, D.; Lun, F. Coastal water bathymetry for critical zone management using regression tree models from Gaofen-6 imagery. *Ocean Coast. Manag.* **2021**, *204*, 105522. [CrossRef]
2. Hedley, J.D.; Roelfsema, C.M.; Chollett, I.; Harborne, A.R.; Heron, S.F.; Weeks, S.J.; Skirving, W.J.; Strong, A.E.; Eakin, C.M.; Christensen, T.R.L.; et al. Remote sensing of coral reefs for monitoring and management: A review. *Remote Sens.* **2016**, *8*, 118. [CrossRef]
3. Hickman, G.D.; Hogg, J.E. Application of an airborne pulsed laser for near shore bathymetric measurements. *Remote Sens. Environ.* **1969**, *1*, 47–58. [CrossRef]
4. Klemas, V. Beach profiling and lidar bathymetry: An overview with case studies. *J. Coast. Res.* **2011**, *27*, 1019–1028. [CrossRef]
5. Kachelreiss, D.; Wegmann, M.; Gollock, M.; Pettorelli, N. The application of remote sensing for marine protected area management. *Ecol. Indic.* **2014**, *36*, 169–177. [CrossRef]
6. Yeu, Y.; Yee, J.J.; Yun, H.S.; Kim, K.B. Evaluation of the accuracy of bathymetry on the nearshore coastlines of western Korea from satellite altimetry, multi-beam, and airborne bathymetric lidar. *Sensors* **2018**, *18*, 2926. [CrossRef] [PubMed]
7. Jawak, S.D.; Vadlamani, S.S.; Luis, A.J. A synoptic review on deriving bathymetry information using remote sensing technologies: Models, methods, comparisons. *Adv. Remote Sens.* **2015**, *4*, 147–162. [CrossRef]
8. Ashphaq, M.; Srivastava, P.K.; Mitra, D. Review of near-shore satellite derived bathymetry: Classification and account of five decades of coastal bathymetry research. *J. Ocean Eng. Sci.* **2021**, *6*, 340–359. [CrossRef]
9. Chu, S.; Cheng, L.; Ruan, X.; Zhuang, Q.; Zhou, X.; Li, M.; Shi, Y. Technical framework for shallow water bathymetry with high reliability and no missing data based on time-series Sentinel-2 images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8745–8763. [CrossRef]
10. Lyzenga, D.R. Passive remote sensing techniques for mapping water depth and bottom features. *Appl. Opt.* **1978**, *17*, 379–383. [CrossRef] [PubMed]
11. Lyzenga, D.R. Shallow water bathymetry using combined lidar and passive multispectral scanner data. *Int. J. Remote Sens.* **1985**, *6*, 115–125. [CrossRef]
12. Stumpf, R.P.; Holderied, K.; Sinclair, M. Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnol. Oceanogr.* **2003**, *48*, 547–556. [CrossRef]
13. Spitzer, D.; Dirks, R.W.J. Bottom influence on the reflectance of the sea. *Int. J. Remote Sens.* **1987**, *8*, 279–308. [CrossRef]
14. Wang, J.; Chen, Z.; Yu, K.; Huang, R.; Wang, Y. Water depth information extraction with multispectral remote sensing in coral reefs region. *Remote Sens. Technol. Appl.* **2018**, *33*, 61–67. (In Chinese) [CrossRef]
15. Lee, Z.; Carder, K.L.; Mobley, C.D.; Steward, R.G.; Patch, J.S. Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Appl. Opt.* **1999**, *38*, 3831–3843. [CrossRef] [PubMed]
16. Lee, Z.; Carder, K.L.; Arnone, R.A. Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* **2002**, *41*, 5755–5772. [CrossRef] [PubMed]
17. McKinna, L.I.W.; Fearns, P.R.C.S.; Weeks, S.J.; Werdell, P.J.; Reichstetter, M.; Franz, B.A.; Shea, D.M.; Feldman, G.C. A semianalytical ocean color inversion algorithm with explicit water column depth and substrate reflectance parameterization. *J. Geophys. Res. Ocean.* **2015**, *120*, 1741–1770. [CrossRef]
18. Ma, Y.; Zhang, H.; Li, X.; Wang, J.; Cao, W.; Li, D.; Lou, X.; Fan, K. An exponential algorithm for bottom reflectance retrieval in clear optically shallow waters from multispectral imagery without ground data. *Remote Sens.* **2021**, *13*, 1169. [CrossRef]
19. Ma, Y.; Han, T.; Wang, E.; Lee, Z.; Prasad, S.; Vosaki, G.W.; Cao, W.; Li, D.; Wang, J.; Lou, X.; et al. A practical and efficient model for benthic habitat parameters retrieval in optically shallow waters from four-band multispectral imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *139*, 104475. [CrossRef]
20. Cheng, J.; Cheng, L.; Chu, S.; Li, J.; Hu, Q.; Ye, L.; Wang, Z.; Chen, H. A comprehensive evaluation of machine learning and classical approaches for spaceborne active-passive fusion bathymetry of coral reefs. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 381. [CrossRef]
21. Santos, R.; Quartau, R.; Lamas, L.; Cândido, R.; Pinto, J.P. Application of Neural Networks for Bathymetric Estimation Using a Short Time-Series of Sentinel-2. A Case Study: Setúbal-Troia, Portugal. *Remote Sens. Appl. Soc. Environ.* **2025**, *39*, 101683. [CrossRef]
22. Sagawa, T.; Yamashita, Y.; Okumura, T.; Yamanokuchi, T. Satellite Derived Bathymetry Using Machine Learning and Multi-Temporal Satellite Images. *Remote Sens.* **2019**, *11*, 1155. [CrossRef]
23. Tran, H.T.T.; Nguyen, Q.H.; Pham, T.H.; Ngo, G.T.H.; Pham, N.T.D.; Pham, T.G.; Tran, C.T.M.; Ha, T.N. Novel Learning of Bathymetry from Landsat 9 Imagery Using Machine Learning, Feature Extraction and Meta-Heuristic Optimization in a Shallow Turbid Lagoon. *Geosciences* **2024**, *14*, 130. [CrossRef]
24. Qin, X.; Wu, Z.; Luo, X.; Shang, J.; Zhao, D.; Zhou, J.; Cui, J.; Wan, H.; Xu, G. MuSRFM: Multiple scale resolution fusion based precise and robust satellite derived bathymetry model for island nearshore shallow water regions using sentinel-2 multi-spectral imagery. *ISPRS J. Photogramm. Remote Sens.* **2024**, *218*, 150–169. [CrossRef]

25. Liu, C.; Xie, H.; Luan, K.; Xu, Q.; Sun, Y.; Ji, M.; Tong, X. Generalized satellite-derived bathymetry across spatial and temporal domains: A domain-adaptive deep learning approach with multi-source remote sensing data. *ISPRS J. Photogramm. Remote Sens.* **2025**, *230*, 452–468. [CrossRef]
26. Parrish, C.E.; Magruder, L.A.; Neuenschwander, A.L.; Forfinski-Sarkozi, N.; Alonzo, M.; Jasinski, M. Validation of ICESat-2 ATLAS bathymetry and analysis of ATLAS's bathymetric mapping performance. *Remote Sens.* **2019**, *11*, 1634. [CrossRef]
27. Ma, Y.; Xu, N.; Liu, Z.; Yang, B.; Yang, F.; Wang, X.H.; Li, S. Satellite-derived bathymetry using the ICESat-2 lidar and Sentinel-2 imagery datasets. *Remote Sens. Environ.* **2020**, *250*, 112047. [CrossRef]
28. Ranndal, H.; Sigaard, C.P.; Kliving, P.; Baltazar, A.O.; Nielsen, K. Evaluation of a statistical approach for extracting shallow water bathymetry signals from ICESat-2 ATL03 photon data. *Remote Sens.* **2021**, *13*, 3548. [CrossRef]
29. Zhang, X.; Ma, Y.; Li, Z.; Zhang, J. Satellite derived bathymetry based on ICESat-2 diffuse attenuation signal without prior information. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *113*, 102993. [CrossRef]
30. Xi, X.; Wang, Z.; Wang, C. Bathymetric extraction method of nearshore based on ICESat-2/ATLAS data. *J. Tongji Univ. Nat. Sci.* **2022**, *50*, 940–946. (In Chinese) [CrossRef]
31. Li, S.; Wang, X.H.; Ma, Y.; Yang, F. Satellite-derived bathymetry with sediment classification using ICESat-2 and multispectral imagery: Case studies in the South China Sea and Australia. *Remote Sens.* **2023**, *15*, 1026. [CrossRef]
32. Li, J.; Dong, Z.; Chen, L.; Tang, Q.; Hao, J.; Zhang, Y. Multi-temporal image fusion-based shallow water bathymetry inversion method using active and passive satellite remote sensing data. *Remote Sens.* **2025**, *17*, 65. [CrossRef]
33. Lesser, M.P.; Mobley, C.D. Bathymetry, water optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. *Coral Reefs* **2007**, *26*, 819–829. [CrossRef]
34. Cao, B.; Liu, H.; Cao, B. Making satellite-derived empirical bathymetry independent of high-quality in-situ depth data: An assessment of four possible model calibration data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *213*, 332–348. [CrossRef]
35. Chen, B.; Yang, Y.; Lin, M.; Zou, B.; Chen, S.; Huang, E.; Xu, W.; Tian, Y. Satellite retrieval of bottom reflectance from high-spatial-resolution multispectral imagery in shallow coral reef waters. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *139*, 104483. [CrossRef]
36. Xu, C.; Ruan, X.; Shen, C.; Tao, Z.; Xu, X.; Zheng, J.; Wu, W. A subregional shallow water bathymetry derivation method for coral reef using ICESat-2 and Sentinel-2 combined with sediment information. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *142*, 104698. [CrossRef]
37. Yao, C.J.; Yu, Z.; Wang, J.; Qian, C.; Xu, J.H. Shallow sea water depth inversion from WorldView-3 multispectral images based on seabed sediment classification. *Bull. Surv. Mapp.* **2023**, *7*, 25–31. (In Chinese) [CrossRef]
38. Ji, X.; Ma, Y.; Zhang, J.; Xu, W.; Wang, Y. A sub-bottom type adaption-based empirical approach for coastal bathymetry mapping using multispectral satellite imagery. *Remote Sens.* **2023**, *15*, 3570. [CrossRef]
39. Ye, L.; Chu, S.; Chen, H.; Cheng, J.; Xu, R.; Qu, Z. Hyperspectral satellite-derived bathymetry considering substrate characteristics. *Mar. Geod.* **2025**, *48*, 533–557. [CrossRef]
40. Chen, A.; Ma, Y.; Zhang, J. Partition satellite derived bathymetry for coral reefs based on spatial residual information. *Int. J. Remote Sens.* **2021**, *42*, 2807–2826. [CrossRef]
41. Zhang, X.; Li, Y.; Du, J.; Qiu, S.; Xie, B.; Chen, W.; Wang, J.; Hu, W.; Wu, Z.; Chen, B. Effects of ocean warming and fishing on the coral reef ecosystem: A case study of Xisha Islands, South China Sea. *Front. Mar. Sci.* **2022**, *9*, 1046106. [CrossRef]
42. He, M.; He, J.; Zhou, Y.; Sun, L.; He, S.; Liu, C.; Gu, Y.; Li, P. Coral reef applications of Landsat-8: Geomorphic zonation and benthic habitat mapping of Xisha Islands, China. *GISci. Remote Sens.* **2023**, *60*, 2261213. [CrossRef]
43. Roelfsema, C.; Kovacs, E.; Ortiz, J.C.; Wolff, N.H.; Callaghan, D.; Wettle, M.; Ronan, M.; Hamylton, S.M.; Mumby, P.J.; Phinn, S. Coral reef habitat mapping: A combination of object-based image analysis and ecological modelling. *Remote Sens. Environ.* **2018**, *208*, 27–41. [CrossRef]
44. Hedley, J.D.; Roelfsema, C.M.; Phinn, S.R.; Mumby, P.J. Environmental and sensor limitations in optical remote sensing of coral reefs: Implications for monitoring and sensor design. *Remote Sens.* **2012**, *4*, 271–302. [CrossRef]
45. Neumann, T.A.; Martino, A.J.; Markus, T.; Bae, S.; Bock, M.R.; Brenner, A.C.; Brunt, K.M.; Cavanaugh, J.; Fernandes, S.T.; Hancock, D.W.; et al. The Ice, Cloud, and Land Elevation Satellite-2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. *Remote Sens. Environ.* **2019**, *233*, 111325. [CrossRef] [PubMed]
46. Magruder, L.; Parrish, C.; Perry, J.; Holwill, M.; Swinski, J.P.; Kief, K. ICESat-2 coastal and nearshore bathymetry product algorithm development. *Earth Space Sci.* **2025**, *12*, e2025EA004390. [CrossRef]
47. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]
48. Lyons, M.B.; Murray, N.J.; Kennedy, E.V.; Kovacs, E.M.; Castro-Sanguino, C.; Phinn, S.R.; Borrego Acevedo, R.; Ordoñez Alvarez, A.; Say, C.; Tudman, P.; et al. New global area estimates for coral reefs from high-resolution mapping. *Cell Rep. Sustain.* **2024**, *1*, 100015. [CrossRef]

49. Matsumoto, K.; Takanezawa, T.; Ooe, M. Ocean tide models developed by assimilating TOPEX/POSEIDON altimeter data into hydrodynamical model: A global model and a regional model around Japan. *J. Oceanogr.* **2000**, *56*, 567–581. [CrossRef]
50. Hedley, J.D.; Harborne, A.R.; Mumby, P.J. Technical note: Simple and robust removal of sun glint for mapping shallow-water benthos. *Int. J. Remote Sens.* **2005**, *26*, 2107–2112. [CrossRef]
51. Chen, Y.; Zhu, Z.; Le, Y.; Qiu, Z.; Chen, G.; Wang, L. Refraction correction and coordinate displacement compensation in nearshore bathymetry using ICESat-2 lidar data and remote-sensing images. *Opt. Express* **2021**, *29*, 2411–2430. [CrossRef] [PubMed]
52. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
53. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
54. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.