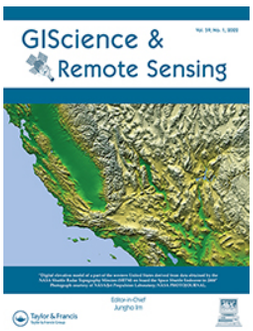

GIScience & Remote Sensing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgrs20

Satellite-derived bathymetry using machine learning and optimal Sentinel-2 imagery in South-West Florida coastal waters

S.S.J.D. Mudiyanselage, A. Abd-Elrahman, B. Wilkinson & V. Lecours

To cite this article: S.S.J.D. Mudiyanselage, A. Abd-Elrahman, B. Wilkinson & V. Lecours (2022) Satellite-derived bathymetry using machine learning and optimal Sentinel-2 imagery in South-West Florida coastal waters, GIScience & Remote Sensing, 59:1, 1143-1158, DOI: 10.1080/15481603.2022.2100597

To link to this article: https://doi.org/10.1080/15481603.2022.2100597

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 05 Aug 2022.

Submit your article to this journal

Article views: 1661

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgrs20

GISCIENCE & REMOTE SENSING
2022, VOL. 59, NO. 1, 1143–1158
https://doi.org/10.1080/15481603.2022.2100597

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Satellite-derived bathymetry using machine learning and optimal Sentinel-2 imagery in South-West Florida coastal waters

S.S.J.D. Mudiyanselage[a], A. Abd-Elrahman[a,b], B. Wilkinson[a] and V. Lecours[a]

[a]School Forest, Fisheries and Geomatics Sciences, University of Florida, Gainesville, FL, USA; [b]Gulf Coast Research Center, University of Florida, Plant City, FL, USA

ABSTRACT

This study examines the use of the Multi-Spectral Instrument (MSI) in Sentinel-2 satellite in combination with regression-based random forest models to estimate bathymetry along the extended southwestern Florida nearshore region. In this study, we focused on the development of a framework leading to a generalized Satellite-Derived Bathymetry (SDB) model applicable to an extensive and diversified coastal region (>200 km of coastline) utilizing multi-date images. The model calibration and validation were done using airborne lidar bathymetry (ALB). As ALB surveys are very expensive to conduct, the proposed model was trained with a limited and practically feasible ALB data sample to expand the model's practicality. Out of the three different sub-models introduced using varying combinations of historical satellite imagery, the combined-band model with the largest feature pool yielded the highest accuracy. The results showed root mean square error (RMSE) values of 8% and lower for the 0–13.5 m depth range (limit of the lidar surveys used) for all areas of interest, indicating the model efficiency and adaptability to varying coastal characteristics. The influence of training sample locations on model performance was evaluated using three distinct model configurations. The difference between these configurations was less than 5 cm, which highlights the robustness of the proposed SDB model. The quality of the satellite imagery is a significant factor that influences the accuracy of the bathymetry estimation. A preliminary methodology incorporating spectral data embedded in Sentinel-2 imagery to effectively select the most optimal satellite imagery was also proposed in this study.

ARTICLE HISTORY
Received 28 February 2022
Accepted 7 July 2022

KEYWORDS
Satellite-derived bathymetry; Sentinel-2; airborne lidar bathymetry; machine learning; random forest

1. Introduction

The world population in coastal regions has increased drastically over the last few decades, which emphasizes the significance of effective coastal zone management strategies. The dynamic and uncertain nature of coastal processes has enhanced the necessity to explore feasible techniques for retrieving data that enable rapid assessment of open coast topography and shallow bathymetry at an affordable cost. State-of-the-art techniques such as the use of echo-sounders or airborne lidar bathymetry (ALB) provide high accuracy and reliability. Single and multibeam echo-sounding missions often produce the most precise water depth information (Horta et al. 2014). ALB, which is unimpeded by maritime constraints, provides a rapid and efficient bathymetric data retrieval mechanism within the nearshore region up to depths of around 20 m, depending on turbidity and wave action. However, these technologies require expensive and sophisticated instrumentation, careful mission planning and sophisticated data acquisition and analysis expertise. In addition, frequent and large area coverage is specifically challenging with these methods, leading to paucities of bathymetric mapping with sufficient spatial extent. Moreover, echo-sounding is typically impractical in shallow waters. Developments in spaceborne remote sensing technology with improved data quality could complement or even replace these traditional methods. As a result, Satellite Derived Bathymetry (SDB) that primarily focuses on the nearshore region is rapidly gaining popularity owing to its cost-effectiveness, high-frequency mapping, and accessibility to remote coastal regions.

Sentinel-2A/B multispectral satellite provides such spaceborne remotely sensed data option, which is characterized by a relatively high spatial resolution of 10 m in the visible and near-infrared region of the electromagnetic spectrum, a nominal revisit time of 5 days, and open data access policy. As a result, many recent SDB studies have utilized Sentinel-2 images in

CONTACT S.S.J.D. Mudiyanselage sanduni.disanaya@ufl.edu

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

varying conditions and topographical regions [(Caballero and Stumpf 2020) (Page, Olmanson, and Mishra 2019) (Traganos et al. 2018) (Casal, Hedley, et al. 2020)]. Optical bathymetry inversion from multispectral imagery is built on the principle that the light penetration through the water column at different wavelengths is a function of sea-water properties. It can be hypothesized that the different combinations of these bands with their water penetration and absorption characteristics can effectively be used to derive shallow-water bathymetry.

The spaceborne remotely sensed data driven optical bathymetry inversion has been performed using various approaches, namely physics-based methods, empirical methods, and optimization-driven machine learning methods. A few studies have been conducted using physics-based bathymetry inversion algorithms similar to radiative transfer models primarily using hyperspectral imagery [(Brando et al. 2009) (Lee et al. 1999)]. Physics-based methods eliminate the need for a-priori known bathymetry data but are more computationally expensive (Casal et al. 2020). Lyzenga (1978) and Stumpf et al. (2003) provided widely used empirical algorithms that incorporate a multiple regression and a spectral log-ratio regression, respectively. Most algorithms consider the blue band the optimal reference band because of its shorter wavelength that enables a higher penetration. As a result, blue light is used either with green or red bands as predictor variables in empirically based SDB studies. The empirical coefficients are site-dependent and hence need to be tuned to meet varying coastal conditions such as the seabed type (Traganos et al. 2018). In addition, the Stumpf log-ratio method recommends specific band combinations based on the depth range of the region to which they are applied (Isabel and Stumpf 2020). Machine learning approaches such as random forest [(Sagawa et al. 2019) (Manessa et al. 2016)] and support vector machine (Misra et al. 2018) are used in a few studies that focus on transparent waters. Random forest is capable of producing more flexible and accurate SDB models when dealing with noisy images (Manessa et al. 2016). Using these established SDB algorithms, bathymetric maps are most often produced in mainly localized regions.

Coastal regions are unique and heterogeneous based on their inherent characteristics, namely the seabed albedo, turbidity, wave energy, and bottom topography. Some of these environmental attributes along a coastal strip pose many challenges in accurate satellite-based depth extraction. For instance, turbidity and chlorophyll content may impede SDB precision and enforce an SBD depth limit. Therefore, it is imperative to evaluate the robustness of the different SDB methodologies for reliable bathymetry estimation in distinctive conditions to enhance their operational validity. Most SDB studies focus on transparent or low-turbidity environments so that the impeding effect from turbidity in nearshore areas is minimized [(Isabel and Stumpf 2019) (Sagawa et al. 2019)]. Suspended solids and turbidity reduce the ability of bands to penetrate the water column, thus hampering the SDB accuracy. Coastal regions near tidal inlets, areas subjected to beach nourishment and extended shallower parts of the nearshore region are generally associated with elevated turbidity levels. It has been suggested that opting for the highest spectral ratio from a multi-scene composite image would improve the SBD results in turbid waters (Isabel and Stumpf 2020). The highest blue/green or blue/red ratio represents the maximum signal extracted from each pixel that could effectively deliver better depth agreement in varying water transparency conditions.

The use of image time series could be useful to address image incongruities in addition to combat varying water transparency conditions encountered in SDB. Multi-temporal image utilization in SDB analysis is approached by different ways in a few recent studies. In Chu et al. (2019), a time-domain noise removal was used to generate an optimal image from a Sentinel-2 time series. Nan et al. (2021) used 74 Sentinel-2 images and generated 74 bathymetric maps using the linear multiple regression with the ICESat-2 bathymetric data facilitating the model calibration. These 74 bathymetric maps were then used in a multi-temporal stacking method from which the median bathymetry value of 74 depth estimations for each pixel was selected as the final depth estimate. A similar utilization of multi-temporal images was carried out in Sagawa et al. (2019) where they used 135 Landsat-8 images over five areas to predict bathymetry using random forest regression. The median value of bathymetry estimates was selected as the output depth value. The median values were used over mean values to prevent the effect of outliers derived from poor-quality data from satellite images (Nan et al. 2021).

The main objective of this study is to propose an efficient and robust methodology to derive bathymetry in an extensive coastal nearshore region using multispectral imagery. We focused on the development of a framework leading to a generalized SDB model approach applicable to a larger diversified area using multi-date Sentinel-2 satellite imagery. As ALB surveys are costly operations, the proposed model was calibrated only with a limited and feasible ALB data sample size to expand the model practicality during its intended operational use. As a potential reference data alternative, few recent studies emphasize the utility of open-access ICESat-2 data to map shallow water bathymetry [(Parrish et al. 2019) (Yao et al. 2019)]. Because the quality of the satellite imagery is a vital factor influencing the accuracy of the bathymetry estimation, a methodology to effectively select the optimal satellite imagery from a large pool of images to be used as input to the bathymetry inversion model is also proposed.

The paper is structured as follows. *Section 2* presents the study region, materials, and methodologies, *Section 3* presents the selection of optimal Sentinel-2 imagery and the results of the SDB analysis, *Section 4* discusses the main findings, applicability, practical implications, and limitations of this study, and *Section 5* highlights the concluding remarks.

2. Materials and methodology

2.1 Study region

Florida's coastal waters are the most valuable (more than $30 billion in revenue per year), have the highest recreational use, and have one of the highest concentrations of coastal communities in the United States. With the likely increase in size and frequency of extreme weather events (Gao et al. 2012) that have the potential to significantly impact coastal systems (Ummenhofer and Meehl 2017), there is a critical need for seafloor data in Florida to produce maps to understand resource distribution, examine sea-level rise indicators, and inform decision-making in contexts such as coastal zone management, navigation, and coastal resilience planning. One area of particular relevance is the coastal strip from Tampa to Marco Island on Florida's west coast, which is significant for boosting the "Blue Economy" of the state. Since one of the main aims of this paper was to examine the validity of the proposed SDB model over a large and diversified coastal region, this area was particularly relevant such that a framework to derive bathymetry using consistently available satellite images in this extended region can be a great asset. In addition, this area is not characterized by the clear waters often used in other SDB work, such as those south of our study area, in the Florida Keys [(Isabel and Stumpf 2019) (McCarthy et al. 2022)].

The coastal strip from Tampa to Marco Island spans approximately 230 km and was divided into four main sub-regions (Figure 1). This region is characterized by a variety of coastal characteristics (e.g. presence of features like barrier islands and tidal inlets, variable turbidity, different water depth ranges) and has a rich archive of ALB data that can be used as ground-truth data. It exhibits a predominant southbound longshore sediment transport pattern with common local reversals occurring due to strong inlet dynamics (Van Gaalen, Tebbens, and Barton 2016), which leads to elevated turbidity levels. The diffuse attenuation coefficient K at 490 nm (K490) can be used as a proxy for water turbidity. K490 increases with increasing turbidity. K490 values within the range of 0.15–0.25 m^{-1} indicate moderate turbidity (van Woesik et al. 2020). A monthly distribution of K490 (4 km, SNPP satellite Ocean Color Level-3 VIIRS multi-sensor, SCI monthly) in February 2017 obtained from NOAA CoastWatch (coastwatch.noaa.gov) exceeds 0.5 m^{-1} within the region of interest in the present study. Water depth in this region is relatively shallow and the turbidity levels fluctuate between medium to high values. The turbidity levels can reach exceptionally higher levels from May to December during which a high frequency of storms occurs in this part of Florida.

2.2 Sentinel-2 satellite imagery

The Sentinel-2 Multi-Spectral Instrument (MSI) consists of twin satellites launched in 2015 and 2017 and facilitates a 290 km field of view, a 12-bit radiometric quantization, and a thirteen spectral band composition spanning from the visible to shortwave infrared (SWIR) region. It employs four bands (red, green, blue, and near-infrared) with a spatial resolution of 10 m, six bands of 20 m, and three bands of 60 m, including the highly penetrating coastal aerosol band. In comparison to all the open and free multispectral sensors available, Sentinel-2 enables one of the highest spatial

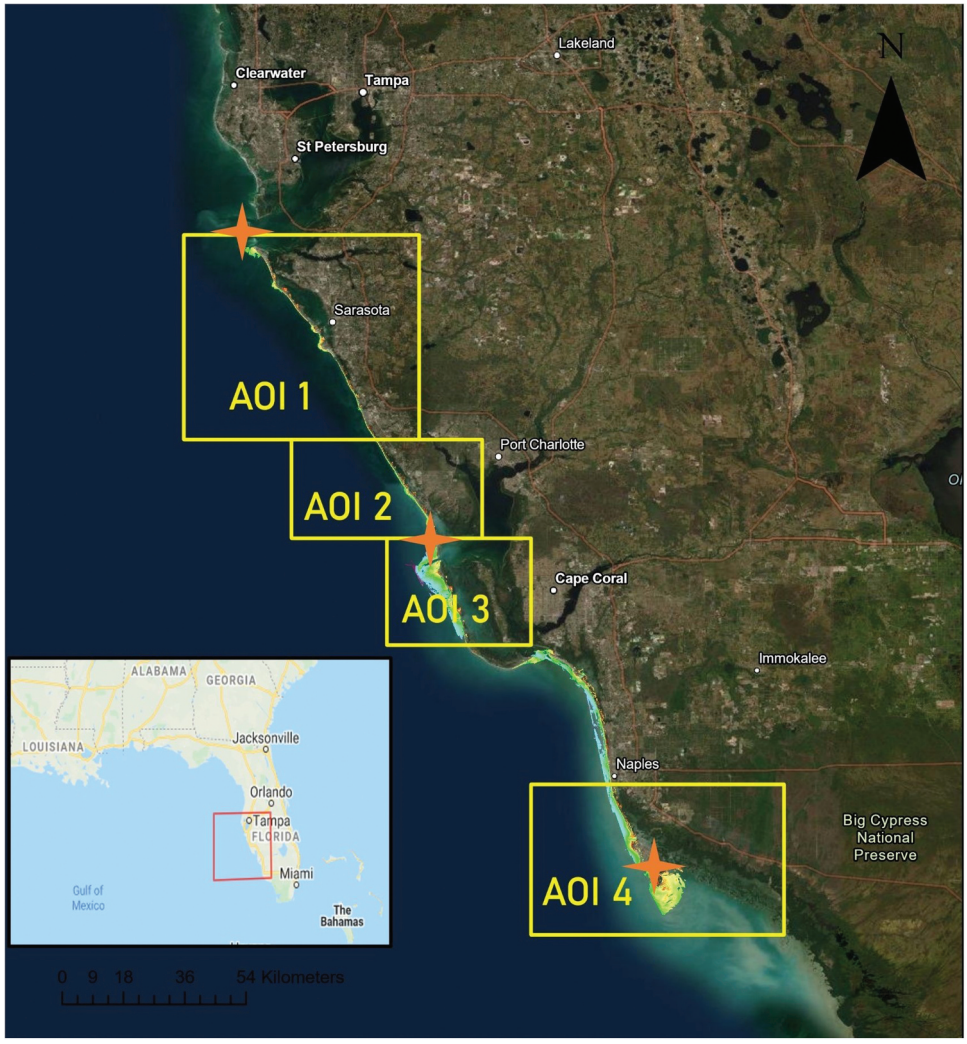

Figure 1. Study region is located on the South-West coast of Florida in the United States as indicated by the red rectangle in the inset map. It expands from Desoto (North Orange marker) to Marco Island (South Orange marker) with Boca Grande (Middle Orange marker) in the mid-point. Based on the availability of ground-truth data and the tile position of Sentinel-2 imagery, it was divided into four main areas of interest (i.e. AOI 1–4).

positioning accuracy products (Level-1C with <11 m accuracy at 95.5% confidence without ground control points) (Report, Data Quality. 2021).

Remote sensing images tend to contain various forms of noise such as clouds, cloud shadows, atmospheric turbulence, and sensor-related noise. The bathymetric inversions from multispectral sensors are even more prone to noise, as the solar radiation is subjected to several alterations when passing through both the atmosphere and water. Therefore, satellite images were filtered initially based on a number of criteria, namely date of image acquisition, degree of cloud coverage at the time of image acquisition, visual inspection based on water clarity and glinting effect. Apart from reducing the rigorous image pre-processing time, the initial filtering from a large image collection using the above criteria also enables more accurate and reliable bathymetry inversions. The filtering and image selection process was carried out on the Google Earth Engine (GEE) platform.

2.3 Airborne lidar bathymetry (ALB)

ALB provides a ground-truthing data with high accuracy in the nearshore region (International Hydrographic Organization 2008). The National Oceanic and Atmospheric Administration (NOAA)'s Digital Coast data repository provides open and free access to a number of ALB surveys performed from Tarpon Springs to Marco Island along the Gulf of

Table 1. Details of ALB and the selected optimal Sentinel-2 imagery used for each AOI.

AOI	ALB data acquisition	Sentinel-2 tile	Sentinel-2 image acquisition
AOI 1	October 2017	T17RLL tile	11/25/2017
			11/30/2017
			02/18/2018
AOI 2	October 2017	T17RLK tile	11/05/2017
			11/30/2017
			02/18/2018
AOI 3	May 2016	T17RLK tile	02/14/2016
			05/14/2016
AOI 4	May 2016	T17RMJ tile	10/21/2016
			11/10/2016

Mexico coast of Florida from 2015 to 2018. They have been collected mostly using a Riegl VQ-880-G topobathy lidar system which, in addition to using an infrared laser (1064 nm), uses a green laser (532 nm) that enables water penetration to the seabed in shallow waters. The data are provided already processed to an initial LAS format using Riegl RiProcess software. The processed lidar data are classified into classes including but not limited to ground, low point (noise), bathymetry point, water surface, and submerged object. These high-density lidar point clouds can be used to generate precise 3D elevation models incorporating accurate positional data (Isabel and Stumpf 2019). The Mean Lower Low Water (MLLW) level is used as the reference datum for the ALB data.

The most recent ALB dataset available for each part of the area of interest (AOI) was selected for the analysis (Table 1). The data cover different depth ranges in each of the surveys from 0 to 13.5 m depths. The variety of characteristics existing along the coastal strip spanning close to 230 km allowed us to evaluate the capacity of both the Sentinel-2 MSI and the proposed SDB model to predict nearshore bathymetry along such extended coastal stretches.

2.4 Methodology

The workflow of the methodology used in this study can be summarized using the following steps. The workflow of the methodology is graphically highlighted in Figure 2.

(1) ALB lidar data acquired along the southwestern coastal region of Florida after 2016 were filtered and downloaded. There were three different surveys done covering the coastal region from Desoto to Marco Island during this period:

i. 2017 NOAA NGS Desoto to Boca Grande
ii. 2016 NOAA NGS Boca Grande
iii. 2016 NOAA NGS Marco Island

The downloaded ALB data were filtered to retain only the bathymetric (seafloor) points, which represent points on the seafloor from which the lidar pulses are reflected in the ALB data. Using the filtered bathymetric points, 10 m spatial resolution Digital Bathymetric Models (DBMs) were created using arithmetic averaging to match the MSI's image resolution.

(2) Based on the ALB data acquisition date, we gathered Sentinel-2 imagery that was captured within a time window of 6 months (6 months before and 6 months after the ALB data acquisition). We used a cirrus and opaque cloud filter threshold of 1% in GEE to only include the images with minimal cloud coverage.

(3) We employed a terrestrial region mask in GEE using Landsat 8's SWIR band (B6). By experimenting with different threshold values, it was found that 0.08 performs best at capturing the optimum region most adjacent to the shoreline.

(4) We used a radiative transfer-based atmospheric correction algorithm developed by Page, Olmanson, and Mishra (2019) on the land masked images from step 3. Images were corrected for both Rayleigh scattering and ozone absorption. Further details on the image preprocessing can be found in Section 2.4.1.

(5) ArcGIS Pro V2.7.1 was used to extract, process, and analyze the spectral (pre-processed Sentinel-2 images) and DEM files of the ALB data. The training and testing data samples were generated via the "Create random points" tool.

(6) A random forest regressor model (Section 2.4.2) was used to derive bathymetry first by using a cross-validation scheme for each Sentinel-2 image collected in step 2. Using the results of cross-validation, the most optimal Sentinel-2 images were selected as described in Section 2.4.3.

(7) After selecting the Sentinel-2 images based on the cross-validation results, each DBM was divided into parallel and adjacent East-West strips of the size of approximately 0.5 square

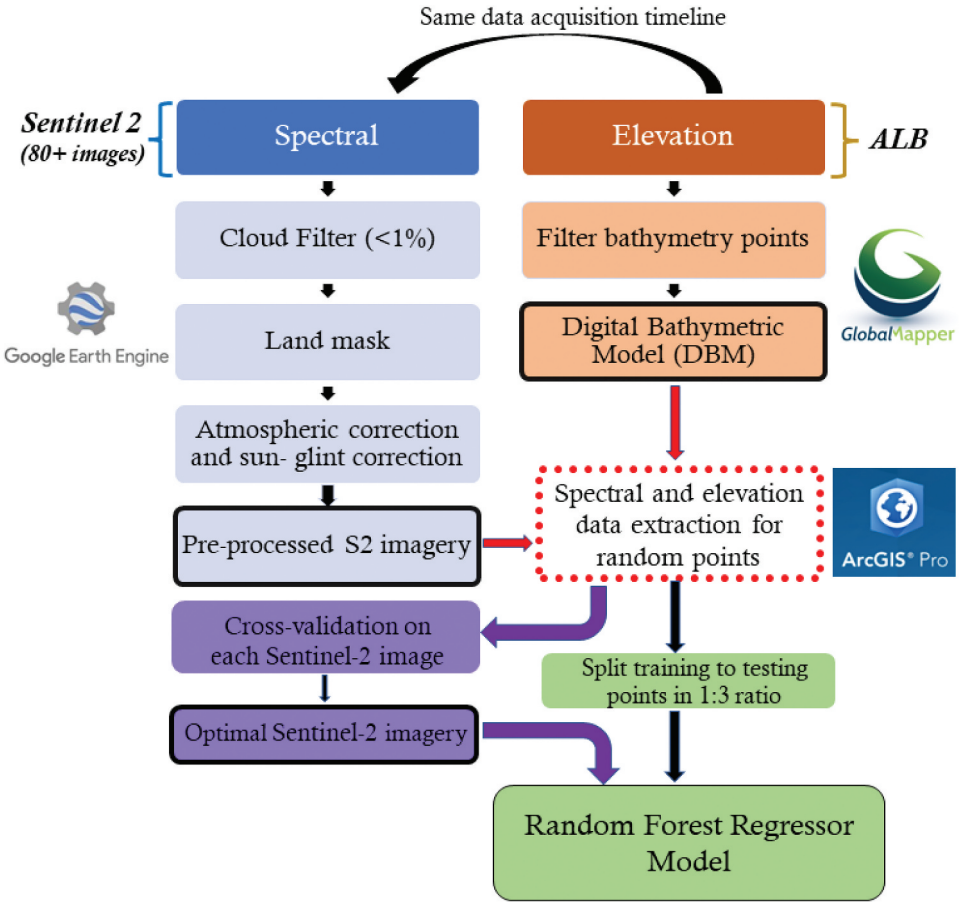

Figure 2. The workflow included processing of spectral and topography data, generation of random point samples, selection of optimal Sentinel-2 imagery using cross-validation, splitting to training and testing samples based on the point locations, and application of a random forest model

kilometers, and a specific number of random points (800 to 1000) were produced within each strip.

(8) A random forest regressor model was used for the analysis with a training-to-testing point sample ratio of 1 to 3. Three main sub-models were introduced within the random forest regressor configuration (Section 2.4.4). Training data were extracted in four different locations along each dataset to be representative of the area of interest. To assess the effect of training sample locations, three different geographical configurations of training data samples from non-overlapping regions were utilized for AOI 1–4 (Section 2.4.5).

(9) Several performance metrics were incorporated to evaluate the model performance in each of the AOIs.

2.4.1 Pre-processing of Sentinel-2 images

Several impeding factors can be identified that hinder the remotely sensed signals reflected from the shallow seabed. These originate from either atmospheric interference (e.g. aerosols and haze) or ocean surface (e.g. sun-glint, refraction, white caps). For instance, the top-of-atmosphere (TOA) signal for blue-to-red spectral bands could constitute up to 90% of scattering due to ozone and Rayleigh effects (Mishra et al. 2005). To correct for these undesirable yet natural intrusions, the selected imagery was pre-processed using atmospheric correction and sun-glint correction models. The radiative transfer-based atmospheric correction algorithm suggested by Page, Olmanson, and Mishra (2019) (Page, Olmanson, and Mishra 2019) was used to convert the top-of-atmosphere reflectance into surface reflectance. Sentinel 2 MSI Level 1C products assumed to be the sum of Rayleigh reflectance,

aerosol reflectance, and water-leaving reflectance were used as input to this algorithm (Liu et al. 2017). We used the GEE platform to pre-process the Sentinel-2 imagery. GEE maintains an open and peta-byte-scale repository from various satellite imaging systems, environmental elements, climate variables in a ready-to-use format that enables efficient data handling and processing (Gorelick et al. 2017). GEE environment also provides an internet-accessible application programming interface (API) integrated with cloud computing that enables rapid data processing and visualizing the outputs.

2.4.2 Random forest regressor model

Random forest regression (RFR) is a supervised learning algorithm that uses an ensemble learning regression method. This machine learning-based algorithm uses several decision trees for prediction, which leads to higher accuracy (Figure 3). The trees are trained by minimizing the sum of squared deviations about the mean (Criminisi, Shotton, and Konukoglu 2012). Random forest regressor models perform statistically better than other regression models (e.g. multiple linear regression and partial least square models) in terms of better tolerance to outlier data samples (de Santana, de Souza, and Jesus Poppi 2018). RFR has the potential to minimize model biases through the implementation of a bagging algorithm.

In the RFR model, we conducted several trials to test the model's performance and robustness across band combinations and image time series. An RMSE improvement in the range of 1 cm to 3 cm was achieved when adding the coastal aerosol band to the three visible range bands in the model. As a result, the most optimum results were obtained when using bands 1 to 4 (i.e. coastal aerosol, blue, green, red) as features in the RFR model. An optimization scheme was adopted for a number of decision trees in RFR model. Therefore, an optimized RFR configuration with different number of images/features based on the proposed sub-models (Section 2.4.4) was used in the present study.

2.4.3 Selection of optimal Sentinel-2 images

Even after filtering the multi-date satellite images based on their cloud coverage and correcting the effect of the atmospheric interference and sun glint, some images still suffered spectral distortions and produced high RMSE. As a secondary objective in this study, we studied the spectral characteristics of the images and modeled them against RFR RMSE. Different spectral combinations (independent variables) were linearly regressed against the RMSE of the SDB RFR models (dependent variable). We used the image collection obtained through the initial filtering described in Section 2.2 and corrected for atmospheric interference or sun glint. Data extraction (i.e. spectral and elevation information) was carried out for 10,000 random points distributed over each of the four AOIs. A cross-validation scheme was implemented using ten-fold cross-validation of each image at a time in the RFR model. Based on the cross-validation results, the most optimal Sentinel-2 images (Table 1) were identified for each AOI that were subsequently used in the main analysis described in Section 2.4.4.

2.4.4 RFR model implementation

Three main RFR sub-models were introduced and evaluated. The individual-image sub-model used spectral data from a single image at a time and the other two sub-models combined the spectral data

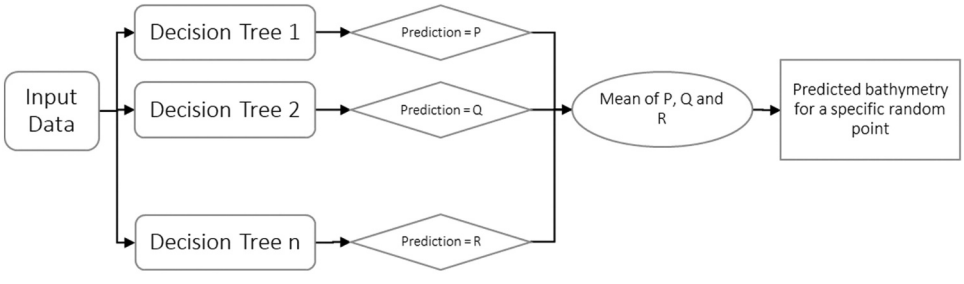

Figure 3. Random Forest model implementation for bathymetry estimation.

from multiple images. The three main sub-models assessed within each area of interest can be described as follows:

- Individual-image sub-model: Individual historical images were evaluated, taking one at a time.
- Mean-band sub-model: Mean values of the respective bands of multiple images were used as the predictor variables.
- Combined-band sub-model: The bands of historical images were combined to create a pool of bands to be used as RFR features. For example, three images with four bands each lead to twelve total features.

To capture the variability within the interest region, training data were selected at several non-adjacent locations along the coast. Model validation was carried out by testing on larger data samples covering the rest of the area. To maintain consistency, a training-to-testing point ratio of 1 to 3 was utilized across all sub-models. Root mean square error (RMSE), mean absolute error (MAE), and median error were used as the performance metrics.

2.4.5 Selection of training point samples

Special attention is needed to study the effect of sample size and samples' spatial distribution since relatively expensive ALB datasets are used for RFR model calibration. Better performance requires the training samples to be representative of different conditions prevailing within the coastal strip. In this study, training samples were gathered at four spatially distinctive and non-adjacent locations along each AOI. The effect on the positional locations of the training sample was evaluated using three non-overlapping training data collection configurations in AOI 1–4. For each configuration, the RFR RMSE, MAE, and median errors were calculated. These resulting error metrics were compared taking two configurations at a time.

3. Results

3.1 Optimal Sentinel-2 image selection based on spectral data

The box plots of RFR RMSE from the cross-validation scheme implemented on each image for each of the four areas of interest are shown in Figure 4. Based on the initial image selection criteria, six images were selected for AOI 1 and

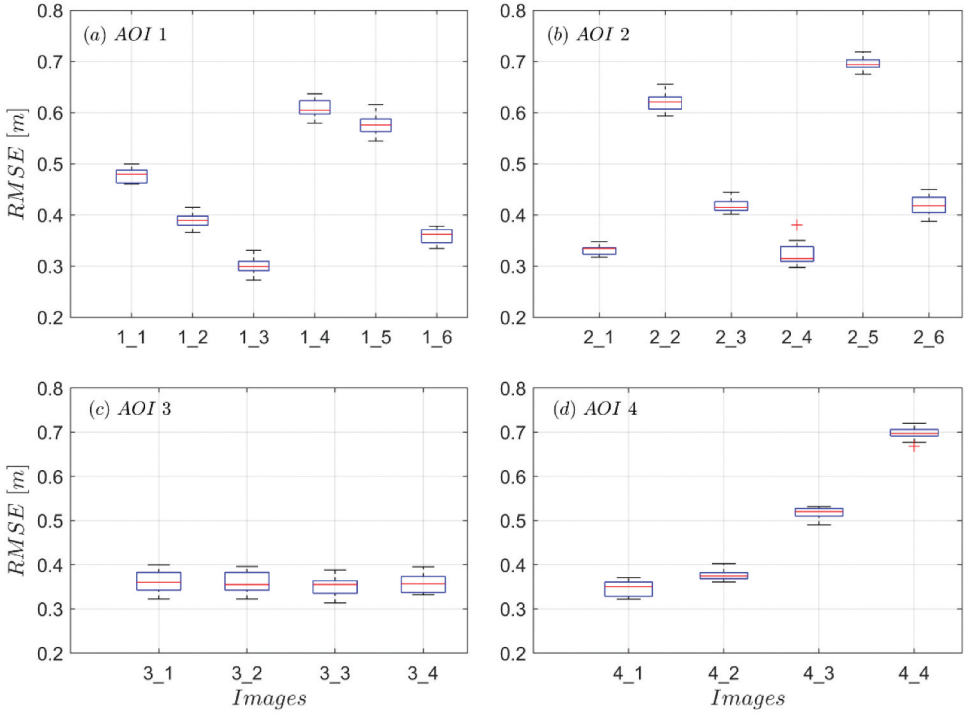

Figure 4. Box plot configurations of the tenfold cross-validation tests carried out using the initial image collection for AOI 1-4. The image collection after initial filtering based on the image acquisition, cloud coverage, visual sun-glint effect was used. The y-axis shows RMSE values while the x-axis indicates the individual images for each AOI. Images 1_2, 1_3, 1_6 (AOI 1), 2_1, 2_4, 2_6(AOI 2), 3_3, 3_4 (AOI 3), 4_1, 4_2 (AOI 4) were selected as the most optimal images to be used in the subsequent SDB analysis.

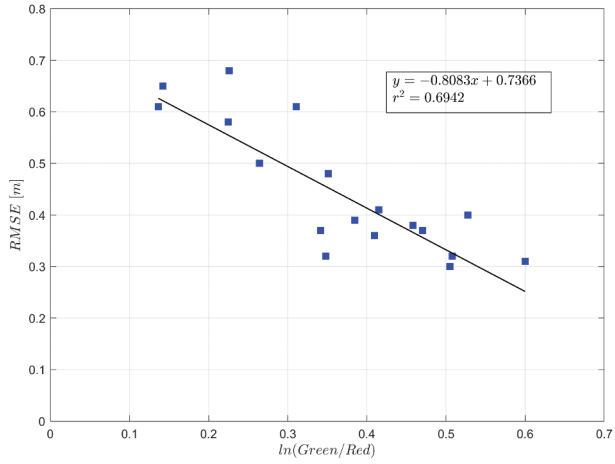

Figure 5. Cross-validated bathymetry RMSE vs. log ratio of mean green to mean red band values.

AOI 2, while four images were selected for AOI 3 and AOI 4. The interquartile range of the RFR RMSE values is relatively narrow (<8 cm) for all the images. The consistent small variation of RFR RMSE within the ten-fold structure for each image suggests high robustness and model stability of the RFR model. The log ratio of mean green band value to mean red band value as the independent variable produced a coefficient of determination (R^2) of 0.69, indicating a high goodness-of-fit to the proposed model (Figure 5). This result indicates the possibility of using the spectral values of the images to exclude the multi-date images that are expected to produce lower SBD RFR accuracy (i.e. high RMSE values). Based on the cross-validation RMSE results, the sub-optimal images were left out from the subsequent SDB analysis. As a result, three images from AOI 1 and 2 and two images from AOI 3 and 4 were selected.

3.2 Assessment of RFR models performance

The optimization of number of decision trees indicated that the accuracy did not improve beyond a threshold value of 30 decision trees. Consequently, 30 decision trees were adopted across all sub-models and all AOIs for a fair comparison. Table 2 reports the precision and accuracy of depth retrieval from all the sub-models. The depth ranges are 0–6.16 m, 0–8.20 m, 0–13.51 m, and 0–5.61 m for AOI 1, AOI 2, AOI 3, and AOI 4, respectively. The validation of the RFR models was carried out using testing point samples of 19,800, 20,400, 37,000, and 19,600 for AOI 1–4. RMSE values from the individual image models are quantitatively comparable with each other in each AOI. The combined-band model yields the most accurate bathymetry prediction in all AOIs, which is indicated by the lowest RMSE values compared to the individual-image and mean-band models. RMSE values are below 0.42 m in all the AOIs except for the case of AOI 3, which has the largest depth range. Based on these results, it is evident that RFR possesses the capability of using a relatively large pool of features without compromising its performance. It should be noted that based on

Table 2. Accuracy assessment results of SDB models for four AOIs. For each AOI, results of the individual image models, mean band model, and combined band model are listed. The performance metrics of RMSE = root mean square error, MAE = mean absolute error and median error are included.

Area of Interest	Used images	Number of features	Number of training points	Number of test points	RMSE (m)	RMSE/max depth %	MAE (m)	Median error (m)	Standard deviation (m)
AOI 1	1	4	6500	19,800	0.44	7.14	0.32	0.24	0.44
	2	4	6500	19,800	0.43	6.98	0.30	0.20	0.43
	3	4	6500	19,800	0.41	6.20	0.30	0.23	0.41
	Mean 1,2,3	4	6500	19,800	0.36	5.84	0.26	0.18	0.36
	Comb 1,2,3	12	6500	19,800	0.35	5.68	0.25	0.19	0.35
AOI 2	1	4	6500	20,400	0.40	4.87	0.30	0.24	0.39
	2	4	6500	20,400	0.42	5.12	0.29	0.20	0.42
	3	4	6500	20,400	0.55	6.70	0.36	0.24	0.54
	Mean 1,2,3	4	6500	20,400	0.46	5.61	0.31	0.22	0.45
	Comb 1,2,3	12	6500	20,400	0.34	4.14	0.23	0.17	0.34
AOI 3	1	4	8000	37,000	0.72	5.33	0.53	0.40	0.72
	2	4	8000	37,000	0.89	6.58	0.61	0.41	0.87
	Mean 1,2	4	8000	37,000	0.67	4.96	0.49	0.36	0.65
	Comb 1,2	8	8000	37,000	0.65	4.81	0.45	0.31	0.64
AOI 4	1	4	6000	19,600	0.47	8.37	0.33	0.22	0.45
	2	4	6000	19,600	0.45	8.02	0.33	0.24	0.45
	Mean 1,2	4	6000	19,600	0.50	8.91	0.34	0.22	0.49
	Comb 1,2	8	6000	19,600	0.41	7.31	0.29	0.19	0.41

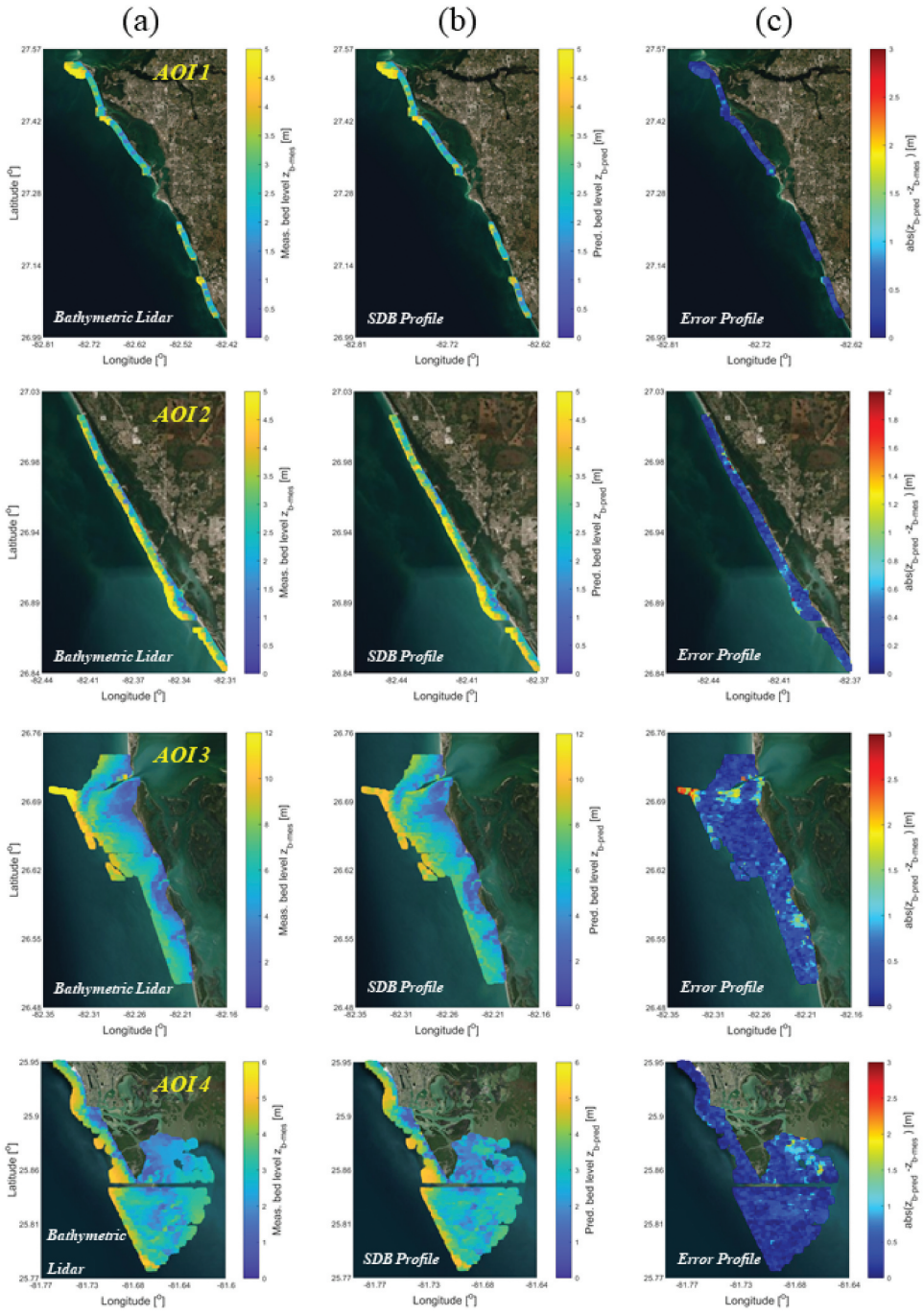

Figure 6. Bathymetric lidar (a), predicted SDB (b) and prediction error (c) maps for AOI 1-4 using combined-band model.

the F-tests carried out, the resulting variances for sub-models were all found to be significantly different from each other within each AOI ($p < 0.05$).

From hereafter only the combined-band image model results are further analyzed and discussed. Figure 6(a,b) show that ALB ground truth DBM and the predicted SDB for all four AOIs along the south-western coastal strip of Florida. Overall, the SDB profiles align with the bathymetric lidar measurements producing consistent shallower depths within the shallow and medium depths in the intermediate region. Figure 6(c) represents the difference between the ALB and predicted elevations, demonstrating a minimal number of extreme errors. The error map of AOI 3 in Figure 6(c) shows a cluster of high error points toward the deeper waters, adjacent to the tidal inlet and near the southern shoreline. Table 3 summarizes the number and proportions of high error points in each AOI.

Table 3. Number of high error points for each AOI. These values as a percentage of total number of testing points are indicated within brackets.

Area of interest	Number of testing points	Test points with error >1 m	Test points with error >2 m	Test points with error >3 m	Test points with error >4 m
AOI 1	19,800	292 (1.47%)	4 (0.02%)	0	0
AOI 2	20,400	254 (1.25%)	50 (0.24%)	8 (0.04%)	1 (0.01%)
AOI 3	37,000	3842 (10.38%)	510 (1.37%)	151 (0.41%)	31 (0.08%)
AOI 4	19,600	684 (3.49%)	35 (0.18%)	15 (0.07%)	1 (0.01%)

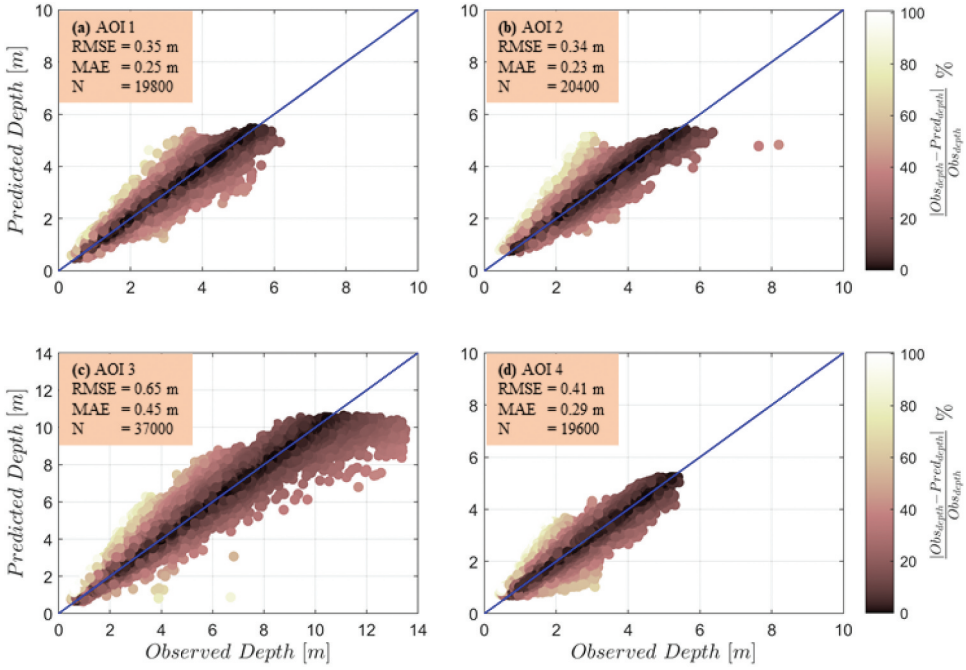

Figure 7. RFR predicted-observed plots for validation datasets (N) of 19800, 20400, 37000 and 19600 respectively for AOI 1-4. The blue line represents the predicted=observed graph. The error percentage is the absolute difference between the observed and measured depth divided by the measure depth. For illustration purposes, the limited number of outliers (test points with >3m and >4m errors) in Table 3 are removed.

AOI 3 with the largest water depth range produces the highest aggregation of points with errors exceeding 1 m.

Figure 7 illustrates the scatter plots of predicted and observed water depths for all the points in the validation datasets from AOIs 1–4. The points are color coded according to the relative error against the ground truth ALB data. The coefficient of determination R^2 is greater than 0.9 for all AOIs indicating high correlation. AOI 1, 2, and 4 feature a uniform under-prediction and over-prediction tendency across their entire depth range. This trend continues for AOI 3 until 10 m water depth, after which the SDB model depicts a distinctive under-prediction tendency (Figure 6(c)). A greater depth agreement in shallower waters (<10 m) is evident than in intermediate depths. This observation further gives an indication of the optical shallow water depth limit in this region. Figure 8 presents the distribution of the difference between the predicted depths and the ALB depths. It should be noted that the mean error is always close to zero for all the AOIs.

The results of the assessment of the influence of training sample locations using three trials are listed in Table 4. The RMSE difference for each sub-model was determined to be less than 5 cm highlighting the repeatability and the consistency of the proposed approach.

4. Discussion

This study examines the use of multi-spectral Sentinel-2 images in combination with random forest regression to estimate shallow water (0–13.5 m) bathymetry. All four AOIs produced the least RMSE values when

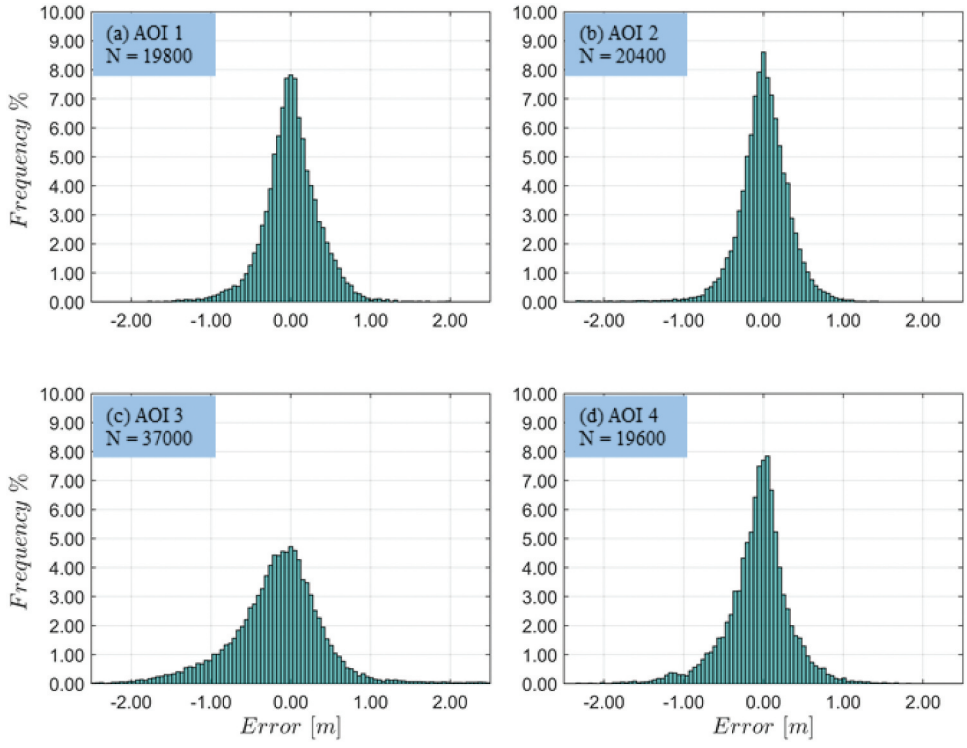

Figure 8. Histogram of the error in the depth estimated for AOI 1–4.

Table 4. Comparison of SDB predictions from three distinctive training sample configurations for AOI 1–4. The training samples were extracted from non-overlapping regions to evaluate the effect of positional location of data to train the proposed SDB model. The differences in RMSE, mean absolute error, and median error are listed from sixth column onwards taking two configurations out of three at a time.

					Difference for Training Sets 1 and 2			Difference for Training Sets 2 and 3			Difference for Training Sets 1 and 3		
Area of interest	Used images	Number of features	Number of training points	Number of test points	Δ RMSE (m)	Δ MAE (m)	Δ Median error (m)	Δ RMSE (m)	Δ MAE (m)	Δ Median error (m)	Δ RMSE (m)	Δ MAE (m)	Δ Median error (m)
AOI 1	1	4	6500	19,800	−0.04	−0.02	−0.02	0.03	0.04	0.03	−0.01	0.02	0.01
	2	4	6500	19,800	−0.01	0.00	0.01	0.09	0.03	0.01	0.08	0.03	0.02
	3	4	6500	19,800	−0.06	−0.04	−0.02	0.05	0.04	0.04	−0.01	0.00	0.02
	Mean 1,2,3	4	6500	19,800	−0.09	−0.04	−0.01	0.05	0.03	0.01	−0.04	−0.01	0.00
	Comb 1,2,3	12	6500	19,800	0.03	0.00	−0.01	0.03	0.02	0.02	0.06	0.02	0.01
AOI 2	1	4	6500	20,400	−0.01	−0.01	−0.01	0.03	0.03	0.03	0.02	0.02	0.02
	2	4	6500	20,400	−0.05	−0.03	−0.01	−0.05	−0.02	0.00	−0.1	−0.05	−0.01
	3	4	6500	20,400	−0.01	−0.03	−0.02	0.05	0.02	0.00	0.04	−0.01	−0.02
	Mean 1,2,3	4	6500	20,400	−0.01	−0.01	−0.01	0.00	0.00	0.00	−0.01	−0.01	−0.01
	Comb 1,2,3	12	6500	20,400	−0.06	−0.03	−0.02	0.06	0.01	0.00	0.00	−0.02	−0.02
AOI 3	1	4	8000	37,000	−0.02	−0.01	−0.02	0.02	0.01	0.01	0.00	0.00	−0.01
	2	4	8000	37,000	0.18	0.11	0.04	−0.03	0.00	0.03	0.15	0.11	0.07
	Mean 1,2	4	8000	37,000	0.04	0.02	0.00	0.01	0.01	0.01	0.05	0.03	0.01
	Comb 1,2	8	8000	37,000	0.08	0.03	0.00	0.02	0.03	0.02	0.1	0.06	0.02
AOI 4	1	4	6000	19,600	0.00	0.00	−0.01	0.01	0.00	0.01	0.01	0.00	0.00
	2	4	6000	19,600	0.03	0.04	0.04	−0.02	−0.02	−0.02	0.01	0.02	0.02
	Mean 1,2	4	6000	19,600	0.04	0.05	0.03	0.01	−0.02	−0.02	0.05	0.03	0.01
	Comb 1,2	8	6000	19,600	0.06	0.04	0.01	−0.03	−0.02	−0.01	0.03	0.02	0.00
	Mean abs error (m)				**0.05**	0.03	0.02	**0.03**	0.02	0.02	**0.04**	0.03	0.01

using the combined-band model with RMSE values of 0.35 m, 0.34 m, 0.65 m, and 0.41 m for AOI 1–4, respectively. These consistently low RMSE values are in line with or better than the accuracies attained by most of the past SDB studies, although due to different site conditions the results cannot directly be compared. A random forest and multi-temporal image approach-based study estimated SDB with Landsat-8

yielding RMSE of 1.41 m for depths of 0–20 m (Sagawa et al. 2019). A GEE-based SDB study by Traganos et al. (2018) (Traganos et al. 2018) focused on Eastern Mediterranean achieved RMSE of 1.5 m. A recent study conducted using Sentinel-2 images in two sites with turbid waters by Stumpf et al.. 2020 (Isabel and Stumpf 2020) yielded median errors of 0.5 m for depths of 0–13 m. To the best of our knowledge, no SDB study has focused on an extensive, and largely diversified coastal region using multi-image bands as individual variables with a generalized SDB approach. The high flexibility and adaptation embedded in the random forest algorithm can be identified as a key factor to obtain reliable and precise bathymetry predictions in long coastal stretches.

Random forest regression was successfully used in this study with promising results. This echoes other SDB studies (e.g. Sagawa et al. 2019), which used Landsat-8 image time-series data as input to a random forest model. The use of bands across multi-date images as individual features, employment of spatially independent training data, the division of training to testing samples (1 to 3) and limiting the Sentinel-2 image selection to shorter duration around the ground truth data acquisition date in our study provide more evidence and introduce implementation schemes when using RFR for SDB.

There could be a few other potential machine learning algorithms applicable to bathymetry inversion such as support vector machine (SVM) and artificial neural networks besides the RFR model. In nonlinear machine learning technique of SVM, the training vectors are mapped into a higher dimensional space incorporating a nonlinear kernel function to predict shallow water morphology. SVM has been used with multi-spectral imagery in a few SDB studies with reasonable bathymetry accuracy values in the nearshore region [(Colomina and Molina 2014) (Misra et al. 2018)]. In Misra et al. (2018), the clear water-dominated two study sites of depth ranges of 1–15 m and 1–3.5 m produced 8.26% and 14.43% of errors, respectively. The deep learning-based regression is another viable option for deriving bathymetry (Ceyhun and Yalçin 2010) that has the potential to further improve SDB results, especially when rich training datasets exist. Subsequently, further studies which compare different machine learning methods in relation to SDB are suggested.

The spatial distribution of the prediction errors needs to be studied further to better understand the factors including the inherent environmental conditions contributing to high or low errors. These environmental factors could be the benthic type, turbidity levels, and bottom texture. There have been few studies conducted on SDB where the over or under prediction of SDB is linked with the level of reflectance from the sea bottom (Traganos et al. 2018). The spatial distribution of residuals in AOI 3 depicts clusters of high error points in the northern deeper region, within the tidal inlet and near the southern shoreline. The accumulation of the largest residuals in the northern section of AOI 3 could be due to the dredging of the possible channel that runs from the tidal inlet toward the north-east direction along which the aggregation of high residuals is aligned. A further investigation using multi-temporal imagery in the region would provide a better explanation. Another potential approach to interpret the error histogram would be through the exploration of the spectral profiles of each band. The trends existing in the spectral information could be correlated to the error gradients. We hope to work on these different aspects to interpret the obtained error distribution maps as a follow-up study.

The proposed model is dependent on the quality of the Sentinel-2 imagery used. Our preliminary analysis to filter the optimal images addressed this element in the SDB model and provided an approach for image selection. However, it is possible to experience cases with limited availability of high-quality images, highlighting the need to develop new methods to improve image quality and accommodate lower-quality image in the SDB analysis. For example, clear pixels in cloudy images can effectively be filtered and used in SDB analysis in case of no or limited number of cloud-free satellite images. This also could allow for a tighter agreement of data acquisition periods between the Sentinel-2 images and ALB surveys to limit the errors due to temporal dataset acquisition discrepancies.

We tested an approach where only ALB data collected over limited areas were used for SDB RFR model training, which suggests the possibility of integrating ALB missions and SDB modeling to produce high frequency bathymetric model at reasonable cost. The proposed SDB model can also be implemented in

a new coastal region by acquiring a limited ALB data sample. The recurring potential of this SDB model to reproduce consistent bathymetry assessment in a challenging coastal environment would be a great testament. From an operational perspective, the mission planning of the ALB data acquisition would play a major role in retrieving a representative data sample. Failure to do so would adversely affect the training of the SDB model leading to inadequate learning for the RFR model to perform effectively.

Airborne lidar bathymetry provides an efficient alternative to vessel-based echo-sounding techniques, particularly in transparent shallow waters. The field tests conducted by the Integrated Mapping for the Sustainable Development of Ireland's Marine Resource (INFOMAR) program reveal that certain ALB sensors would not be efficient in highly turbid coastal regions as much as in transparent or less turbid waters (Coveney and Monteys 2011). However, technological advancements in modern topo-bathymetric lidar systems (e.g. CZMIL SuperNova) excel in facilitating maximum depth penetration, superior coverage even in turbid waters, and high-density point clouds, which is promising for training the prospective SDB models. The ALB point clouds used in this study were collected using Riegl VQ-880-G topo-bathymetric lidar sensor. The vertical accuracy of the post-processed ALB was 0.15 m. It should be noted that part of the uncertainties in the achieved RMSEs can be attributed to the accuracy of the ALB datasets used for training and validation.

ALB data can be expensive and thus limited in availability, so that alternative sources of reference data to calibrate the bathymetric models will be immensely useful. Nan et al. (2021), Hsu et al. (2021) and Yue et al. (2020) demonstrated examples of using ICESat-2 data to train the empirical SDB models. The fusion of ICESat-2 and repeating Sentinel-2 images in a multi-temporal stacking method enabled accurate bathymetry retrieval in areas where conventional ground-truth data are absent (Hsu et al. 2021). Consequently, with these emerging technologies SDB would no longer limited by local priori-measurements but could be upgraded as a tool to produce larger scale bathymetric maps of extensive nearshore regions at extremely low cost.

The temporal data acquisition discrepancies of Sentinel-2 images and ALB data constitute the major proportion of uncertainty in the obtained RMSEs of the present study. In addition, the vertical accuracy of ALB datasets, particularly within the turbid water segments, and the modest tendency of RFR model for overfitting biases can be identified as part of the uncertainties in the depth estimation accuracy levels reported.

The methodology of this study is constructed to generate a practically feasible and repeatable bathymetric inversion procedure. The routinely estimated bathymetry values would be useful for numerical models to predict sediment transport dynamics in both cross-shore (Hewageegana and Canestrelli 2021a), alongshore (Van Gaalen, Tebbens, and Barton 2016) directions and sandbar movements (Hewageegana and Canestrelli 2021b). This study can also be extended to areas with existing ALB data. As an initial step, it is possible to repeat the method within the nearshore regions around Florida as several ALB data missions have already been conducted between 2015 and 2021.

5. Conclusion

This paper presents a comprehensive framework to achieve accurate SDB in coastal regions with varying depths and characteristics using machine learning and multi-temporal images. An extensive coastal strip along southwestern Florida in the United States was explored that is representative of diversified nearshore regions. The SDB generation scheme developed in our study is expected to be applicable to coastal regions with low transparent waters that can be attributed to high energy wave environments and shallower depths. The depth estimation model was constructed using a smaller training sample and was validated by a much larger testing sample to expand its applicability.

Combining individual bands in multi-date images as the features in RFR model delivered highest accuracy values in each AOI. The RMSE of the derived coastal bathymetry in depths ranging 0–13.5 m was within 8% of ALB data. The consistently high accuracy values obtained in this study indicate the flexibility in RFR model approach in SDB analysis in broader areas. These findings are encouraging for future work that encompasses SDB analysis in long coastal stretches even with higher turbidity levels that deemed to be very challenging for accurate SDB predictions. We expect that the proposed SDB approach will contribute to many

of the applications that require shallow water bathymetry.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the United States Department of Commerce–National Oceanic and Atmospheric Administration (NOAA) through the University of Southern Mississippi under the terms of Agreement No. NA18NOS400198.

Data availability statement

The data that support the findings of this study are available in Mendeley Data repository at DOI: 10.17632/5m8bj4x7j2.1 (http://dx.doi.org/10.17632/5m8bj4x7j2.1).
Readers interested in the code developed in these study data will be shared upon request to the corresponding author. These data were derived from the following resources available in the public domain:

- Sentinel-2 data: From either EarthExplorer-USGS (https://earthexplorer.usgs.gov/) or data archive embedded in Google Earth Engine (https://earthengine.google.com/)
- ALB data: From https://www.ncei.noaa.gov/maps/bathymetry/

References

Brando, V. E., J. M. Anstee, M. Wettle, A. G. Dekker, S. R. Phinn, and C. Roelfsema. 2009. "A Physics Based Retrieval and Quality Assessment of Bathymetry from Suboptimal Hyperspectral Data." *Remote Sensing of Environment* 113 (4): 755–770. doi:10.1016/j.rse.2008.12.003.

Caballero, I., and R. P. Stumpf. 2020. "Towards Routine Mapping of Shallow Bathymetry in Environments with Variable Turbidity: Contribution of Sentinel-2A/B Satellites Mission." *Remote Sensing* 12: 3. doi:10.3390/rs12030451.

Casal, G., J. D. Hedley, X. Monteys, P. Harris, C. Cahalane, and T. McCarthy. 2020. "Satellite-Derived Bathymetry in Optically Complex Waters Using a Model Inversion Approach and Sentinel-2 Data." *Estuarine, Coastal and Shelf Science* 241 (March): 106814. doi:10.1016/j.ecss.2020.106814.

Casal, G., P. Harris, X. Monteys, J. Hedley, C. Cahalane, and T. McCarthy. 2020. "Understanding Satellite-Derived Bathymetry Using Sentinel 2 Imagery and Spatial Prediction Models." *GIScience and Remote Sensing* 57 (3): 271–286. doi:10.1080/15481603.2019.1685198.

Ceyhun, Ö., and A. Yalçin. 2010. "Remote Sensing of Water Depths in Shallow Waters via Artificial Neural Networks." *Estuarine, Coastal and Shelf Science* 89 (1): 89–96. doi:10.1016/j.ecss.2010.05.015.

Chu, S., L. Cheng, X. Ruan, Q. Zhuang, X. Zhou, L. Manchun, and Y. Shi. 2019. "Technical Framework for Shallow-Water Bathymetry with High Reliability and No Missing Data Based on Time-Series Sentinel-2 Images." *IEEE Transactions on Geoscience and Remote Sensing* 57 (11): 8745–8763. doi:10.1109/TGRS.2019.2922724.

Colomina, I., and P. Molina. 2014. "Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 92: 79–97. doi:10.1016/j.isprsjprs.2014.02.013.

Coveney, S., and X. Monteys. 2011. "Integration Potential of INFOMAR Airborne LIDAR Bathymetry with External Onshore LIDAR Data Sets." *Journal of Coastal Research* 62: 19–29. doi:10.2307/29783132.

Criminisi, A., J. Shotton, and E. Konukoglu. 2012. "Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning." *Foundations and Trends® in Computer Graphics and Vision* 7 (2–3): 81–227. doi:10.1561/0600000035.

de Santana, F. B., A. M. de Souza, and R. Jesus Poppi. 2018. "Visible and near Infrared Spectroscopy Coupled to Random Forest to Quantify Some Soil Quality Parameters." *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 191: 454–462. doi:10.1016/j.saa.2017.10.052.

Gao, Y., J. S. Fu, J. B. Drake, Y. Liu, and J. F. Lamarque. 2012. "Projected Changes of Extreme Weather Events in the Eastern United States Based on a High Resolution Climate Modeling System." *Environmental Research Letters* 7 (4): 044025. doi:10.1088/1748-9326/7/4/044025.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202: 18–27. doi:10.1016/j.rse.2017.06.031.

Hewageegana, V. H., and A. Canestrelli. 2021a. "Numerical Modeling of the Influence of Tides on Cross-Shore Sediment Dynamics of Dissipative Beaches under Moderate Wave Conditions." *Continental Shelf Research* 218 (November 2020): 104381. doi:10.1016/j.csr.2021.104381.

Hewageegana, V. H., and A. Canestrelli. 2021b. "On the Predictive Skill of Morphodynamic Models for Onshore Sandbar Migration." *Earth Surface Processes and Landforms* 46 (9): 1692–1712. doi:10.1002/esp.5079.

Horta, J., A. Pacheco, D. Moura, Ferreira, and Ó. Ferreira. 2014. "Can Recreational Echosounder-Chartplotter Systems Be Used to Perform Accurate Nearshore Bathymetric Surveys?" *Ocean Dynamics* 64 (11): 1555–1567. doi:10.1007/s10236-014-0773-y.

Hsu, H. J., C. Yuan Huang, M. Jasinski, L. Yao, H. Gao, T. Yamanokuchi, C. Gi Wang, et al. 2021. "A Semi-Empirical Scheme for Bathymetric Mapping in Shallow Water by ICESat-2 and Sentinel-2: A Case Study in the South China Sea." *ISPRS Journal of Photogrammetry and Remote Sensing* 178 (May): 1–19. doi:10.1016/j.isprsjprs.2021.05.012.

International Hydrographic Organization. 2008. *IHO Standards for Hydrographic Surveys*. MONACO: International Hydrographic Bureau.

Isabel, C., and R. P. Stumpf. 2019. "Retrieval of Nearshore Bathymetry from Sentinel-2A and 2B Satellites in South Florida Coastal Waters." *Estuarine, Coastal and Shelf Science* 226 (June): 106277. doi:10.1016/j.ecss.2019.106277.

Lee, Z., K. L. Carder, C. D. Mobley, R. G. Steward, and J. S. Patch. 1999. "Hyperspectral Remote Sensing for Shallow Waters: 2 Deriving Bottom Depths and Water Properties by Optimization." *Applied Optics* 38 (18): 3831. doi:10.1364/ao.38.003831.

Liu, H., L. Qingquan, T. Shi, H. Shuibo, W. Guofeng, and Q. Zhou. 2017. "Application of Sentinel 2 MSI Images to Retrieve Suspended Particulate Matter Concentrations in Poyang Lake." *Remote Sensing* 9 (7): 761. doi:10.3390/rs9070761.

Lyzenga, D. 1978. "Passive Remote-Sensing Techniques for Mapping Water Depth and Bottom Features." *Applied Optics* 17 (3): 379–383. doi:10.1364/AO.17.000379.

Manessa, M. D. M., A. Kanno, M. Sekine, M. Haidar, K. Yamamoto, T. Imai, and T. Higuchi. 2016. "Satellite-Derived Bathymetry Using Random Forest Algorithm and Worldview-2 Imagery." *Geoplanning: Journal of Geomatics and Planning* 3 (2): 117. doi:10.14710/geoplanning.3.2.117-126.

McCarthy, M. J., D. B. Otis, D. Hughes, and F. E. Muller-Karger. 2022. "Automated High-Resolution Satellite-Derived Coastal Bathymetry Mapping." *International Journal of Applied Earth Observation and Geoinformation* 107 (November 2021): 102693. doi:10.1016/j.jag.2022.102693.

Mishra, D. R., S. Narumalani, D. Rundquist, and M. Lawson. 2005. "Characterizing the Vertical Diffuse Attenuation Coefficient for Downwelling Irradiance in Coastal Waters: Implications for Water Penetration by High Resolution Satellite Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (1): 48–64. doi:10.1016/j.isprsjprs.2005.09.003.

Misra, A., Z. Vojinovic, B. Ramakrishnan, A. Luijendijk, and R. Ranasinghe. 2018. "Shallow Water Bathymetry Mapping Using Support Vector Machine (SVM) Technique and Multispectral Imagery." *International Journal of Remote Sensing* 39 (13): 4431–4450. doi:10.1080/01431161.2017.1421796.

Nan, X., M. Xin, M. Yue, P. Zhao, J. Yang, and X. Hua Wang. 2021. "Deriving Highly Accurate Shallow Water Bathymetry from Sentinel-2 and ICESat-2 Datasets by a Multitemporal Stacking Method." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (June): 6677–6685. doi:10.1109/JSTARS.2021.3090792.

Page, B. P., L. G. Olmanson, and D. R. Mishra. 2019. "A Harmonized Image Processing Workflow Using Sentinel-2/MSI and Landsat-8/OLI for Mapping Water Clarity in Optically Variable Lake Systems." *Remote Sensing of Environment* 231 (September 2018): 111284. doi:10.1016/j.rse.2019.111284.

Parrish, C. E., L. A. Magruder, A. L. Neuenschwander, N. Forfinski-Sarkozi, M. Alonzo, and M. Jasinski. 2019. "Validation of ICESat-2 ATLAS Bathymetry and Analysis of ATLAS's Bathymetric Mapping Performance." *Remote Sensing* 11 (14): 1634. doi:10.3390/rs11141634.

Report, Data Quality. 2021. "S2 MPC."

Sagawa, T., Y. Yamashita, T. Okumura, and T. Yamanokuchi. 2019 . "Satellite Derived Bathymetry Using Machine Learning and Multi-Temporal Satellite Images." *Remote Sensing* 10 11(10): 1155. doi:10.3390/rs11101155

Stumpf, R. P., K. Holderied, J. A. Robinson, G. Feldman, and N. Kuring. 2003. "Mapping Water Depths in Clear Water from Space." Proceedings of the 13th Biennial Coastal Zone Conference Baltimore, MD, 13-17 July, 5.

Traganos, D., D. Poursanidis, B. Aggarwal, N. Chrysoulakis, and P. Reinartz. 2018. "Estimating Satellite-Derived Bathymetry (SDB) with the Google Earth Engine and Sentinel-2." *Remote Sensing* 10 (6): 1–18. doi:10.3390/rs10060859.

Ummenhofer, C. C., and G. A. Meehl. 2017. "Extreme Weather and Climate Events with Ecological Relevance: A Review." *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1723): 20160135. doi:10.1098/rstb.2016.0135.

Van Gaalen, J. F., S. F. Tebbens, and C. C. Barton. 2016. "Longshore Sediment Transport Directions and Rates from Northern Maine to Tampa Bay, Florida: Literature Compilation and Interpretation." *Journal of Coastal Research* 32 (6): 1277–1301. doi:10.2112/JCOASTRES-D-15-00002.1.

van Woesik, R., L. M. Roth, E. J. Brown, K. R. McCaffrey, and J. R. Roth. 2020. "Niche Space of Corals along the Florida Reef Tract." *PLoS ONE* 15 (4): 1–16. doi:10.1371/journal.pone.0231104.

Yao, L., H. Gao, M. F. Jasinski, S. Zhang, and J. D. Stoll. 2019. "Deriving High-Resolution Reservoir Bathymetry from ICESat-2 Prototype Photon-Counting Lidar and Landsat Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 57 (10): 7883–7893. doi:10.1109/TGRS.2019.2917012.

Yue, M., X. Nan, Z. Liu, B. Yang, F. Yang, X. Hua Wang, and L. Song. 2020. "Satellite-Derived Bathymetry Using the ICESat-2 Lidar and Sentinel-2 Imagery Datasets." *Remote Sensing of Environment* 250 (July): 112047. doi:10.1016/j.rse.2020.112047.