Check for updates

Workflow for the Generation of Expert-Derived Training and Validation Data: A View to Global Scale Habitat Mapping

Chris M. Roelfsema[1]*, Mitchell Lyons[1], Nicholas Murray[1,2], Eva M. Kovacs[1], Emma Kennedy[1], Kathryn Markey[1], Rodney Borrego-Acevedo[1], Alexandra Ordoñez Alvarez[1], Chantel Say[1], Paul Tudman[1], Meredith Roe[1], Jeremy Wolff[1], Dimosthenis Traganos[3], Gregory P. Asner[4], Brianna Bambic[5], Brian Free[5], Helen E. Fox[5], Zoe Lieb[5] and Stuart R. Phinn[1]

[1] Remote Sensing Research Centre, School of Earth and Environmental Sciences, University of Queensland, Brisbane, QLD, Australia, [2] College of Science and Engineering, James Cook University, Townsville, QLD, Australia, [3] German Aerospace Center, Berlin, Germany, [4] Center for Global Discovery and Conservation Science, Arizona State University, Tempe, AZ, United States, [5] The National Geographic Society, Washington, DC, United States

Our ability to completely and repeatedly map natural environments at a global scale have increased significantly over the past decade. These advances are from delivery of a range of on-line global satellite image archives and global-scale processing capabilities, along with improved spatial and temporal resolution satellite imagery. The ability to accurately train and validate these global scale-mapping programs from what we will call "reference data sets" is challenging due to a lack of coordinated financial and personnel resourcing, and standardized methods to collate reference datasets at global spatial extents. Here, we present an expert-driven approach for generating training and validation data on a global scale, with the view to mapping the world's coral reefs. Global reefs were first stratified into approximate biogeographic regions, then per region reference data sets were compiled that include existing point data or maps at various levels of accuracy. These reference data sets were compiled from new field surveys, literature review of published surveys, and from individually sourced contributions from the coral reef monitoring and management agencies. Reference data were overlaid on high spatial resolution satellite image mosaics (3.7 m × 3.7 m pixels; Planet Dove) for each region. Additionally, thirty to forty satellite image tiles; 20 km × 20 km) were selected for which reference data and/or expert knowledge was available and which covered a representative range of habitats. The satellite image tiles were segmented into interpretable groups of pixels which were manually labeled with a mapping category via expert interpretation. The labeled segments were used to generate points to train the mapping models, and to validate or assess accuracy. The workflow for desktop reference data creation that we present expands and up-scales traditional approaches of expert-driven interpretation for both manual habitat mapping and map training/validation. We apply the reference data creation methods in the

context of global coral reef mapping, though our approach is broadly applicable to any environment. Transparent processes for training and validation are critical for usability as big data provide more opportunities for managers and scientists to use global mapping products for science and conservation of vulnerable and rapidly changing ecosystems.

Keywords: calibration, validation, coral reefs, training, habitat mapping, Allen Coral Atlas

INTRODUCTION

Our global environment is changing in response to various natural and anthropogenic processes, which is having a direct effect on living organisms (Tittensor et al., 2014). Understanding our environment and conserving natural resources requires consistent sources of relevant and up-to-date information. Information sources suitable for these purposes are becoming increasingly available and expanding in scope to cover continental to global domains. Mapping of the natural environment at a global scale has increased particularly in the last decade, with actionable datasets such as global forest watch (Hansen et al., 2013), ocean color (Groom et al., 2019), tidal flats (Murray et al., 2019), and ice and snow cover (Bormann et al., 2018) becoming widely available. Recent advances in global monitoring are primarily due to enhanced satellite image sensor capabilities, leading to greater availability of datasets with increased temporal and spatial resolution, often with improved accuracy (Murray et al., 2018). In addition, these advances have been catalyzed by faster access to and processing of large Earth observation datasets [e.g., Google Earth Engine; (Gorelick et al., 2017)]. These efforts have also been stimulated by a growing focus on global conservation targets and the status of ecosystems (Tittensor et al., 2014; Keith et al., 2015), open access to data, and growing cooperation between countries to manage the environment, such as the United Nation's Convention on Biological Diversity and the Sustainability Development Goals[1]. However, major challenges persist in global mapping efforts. These include a lack of standardized methods to collate reference datasets at regional to global scales, poor availability of open access data suitable for training and validating models, and that emergent mapping methods require large amounts of reference data to achieve desired map quality targets.

The majority of the world's local to global scale satellite image-based mapping programs train and validate their mapping algorithms using observations linked to a location in the field, and recording details of the feature of interest, which are either categorical (e.g., land cover type) or continuous (e.g., vegetation cover or canopy height). We refer to these as "reference data." The types of sampling units for reference data varies, ranging from point data to polygons, which are either directly annotated or labeled after some processing operations such as segmentation. Selecting an appropriate reference data form depends on the type of feature being mapped or the modeling strategy to be employed. Typically, validation is conducted using an independent process or sample (Congalton and Green, 2008), although increasingly

model-based estimates of map accuracy are being employed (Lyons et al., 2018). Map accuracy estimates are generally achieved within the mapping process by holding out a portion of training data to use for validation of model performance, or post-hoc with a completely independent reference dataset. Regardless of the methods employed for training maps and assessing accuracy, a fundamental assumption is that the reference data are a representative sample, that are accurate, and a confirmed record of the feature being mapped.

With a few notable exceptions, e.g., Millennium Reef Mapping, United States NOAA reef mapping, and Living Oceans Foundation, coral reef mapping efforts have been focused on relatively small reef areas (up to 300 km^2; Roelfsema and Phinn, 2013), making them suitable for detailed benthic habitat mapping with abundant and well distributed field-based benthic reference data (Andréfouët et al., 2003; Phinn et al., 2012). Geolocated photoquadrat surveys in combination with machine learning or video surveys have been shown to provide a valuable source for benthic reference data (Roelfsema and Phinn, 2010; Hamylton, 2011; González-Rivero et al., 2016; Li et al., 2019b). Geomorphic maps, on the other hand, have been developed mostly for larger reef systems and generally do not require field data. Geomorphic mapping methods are based on manual digitization or object-based analysis using expert knowledge of reef geomorphology and physical attributes combined with visual interpretation of satellite imagery (Andréfouët et al., 2006; Leon and Woodroffe, 2011; Roelfsema et al., 2020). Validation samples for geomorphic mapping are ideally provided by independent methods and/or analysts developing a reference data set manually (Andréfouët, 2008). A similar expert-led approach could be employed for benthic mapping, particularly over large and remote areas where field data may be absent.

Most reference field data collection methods are suitable for relatively small reef areas (less than 100 km^2), but they require intensive, detailed field surveys by trained experts. This typically prohibits detailed surveys from being employed to support mapping efforts over coral reef areas (Purkis et al., 2019) at regional to global scales (Andréfouët et al., 2006). Some programs have worked toward these goals; both Living Oceans Foundation Global Reef Expedition (LOF-GRE; Purkis et al., 2019) and United States NOAA (Monaco et al., 2012) benthic mapping efforts cover large extents of reef and include intensive field campaigns to collect reference data. At the global scale, a map of geomorphic zonation and reef extent, the Global Coral Reef Map UNEP 2018 (Spalding et al., 2017), was generated based on these approaches. However, this data set is a composite of the Millennium coral reef mapping project (Andréfouët et al., 2006), LOF-GRE (Purkis et al., 2019), NOAA (Monaco et al., 2012) and local data sets, and did not utilize a consistent reference

---

[1]https://www.cbd.int/gbo/gbo5/publication/gbo-5-en.pdf; https://sdgs.un.org/goals

data set for training and validation. This was due to typical challenges related to vast areas and remoteness, but also unique challenges in the marine environment where field data collection requires boating, underwater surveys, and alternate approaches to geolocate sample points as Global Navigation Satellite System (GNSS) signals do not penetrate the water column (Roelfsema and Phinn, 2010). Thus, existing global reef maps remain largely unvalidated.

This study presents an expert-driven workflow for generating training and validation data, for global scale mapping of coral reefs. We introduce the workflow which is being used to develop the first globally-consistent, fine-spatial scale maps of geomorphic zonation and benthic composition as part of the Allen Coral Atlas[2].

MATERIALS AND METHODS

General Overview
The core global coral reef mapping framework that underpins the Allen Coral Atlas has been published previously, providing a detailed description of the framework including the data inputs, classification approach and validation routine (Lyons et al., 2020). Here we discuss the implementation of the reference sample creation that underpins the mapping framework (**Figure 1**). For each region, our mapping process consists of a combination of machine learning and object-based analysis (Lyons et al., 2020), and produces two regional products: geomorphic zonation and benthic cover maps, following a well-defined classification scheme developed for the Allen Coral Atlas project (Kennedy et al., 2020). The mapping approach uses multiple input data sources including reference data sets for training and validation (the subject of this paper), and data layers derived from satellite imagery that represent physical attributes (depth, significant wave exposure, slope). Satellite imagery comprised a low tide mosaic of Planet Dove data, which was used to derive mosaics of sub-surface reflectance at a spatial resolution of ∼5 m × 5 m pixels, which along with water depth and benthic slope, is derived according to Li et al. (2019a). Finally, reference data (points) were randomly sampled from these layers (subsurface reflectance, depth) and this point dataset was evenly split into training and validation data sets.

Data
Here we focus on the reference data for the areas mapped by the Allen Coral Atlas project (www.allencoralatlas.org; **Figure 2**) up to December 2020, which includes Andaman Sea, East Africa, The Indonesian Archipelago, North Caribbean and Bahamas, Papua New Guinea and Solomon Islands, South Asia, South West Pacific, Timor Sea, West Indian Ocean Islands, and West Micronesia (**Figure 2A** and **Supplementary Table 1**; see www.allencoralatlas.org for the output map products). An example of the initialization of reference data creation is indicated for the PNG-Solomon Islands region (**Figures 2B,C**).

---

[2]www.allencoralatlas.org

This region represents a large region with a variety of reef types, geomorphic zones, and benthic composition.

Generating Reference Data for Training Coral Mapping Algorithms and Validating Output Geomorphic and Benthic Maps
For each of the mapping regions (**Figure 2A**), reference data were sourced through either new field data acquisitions, or, acquisitions via literature review and/or requests to the scientific community, government agencies and non-governmental organizations (NGOs), **Figure 2B** shows the distribution of reference data collated for the PNG-Solomon Islands region. The online search focused on peer-reviewed scientific papers and web-based data sets, or non-peer reviewed but well documented datasets. Data sets were both field data and/or maps. Targeted requests were conducted to the coral and seagrass email lists, and to global NGOs (e.g., The Nature Conservancy, Wildlife Conservation Society, and World Wildlife Fund, etc.). Prior to being included in the workflow, the data sourced through literature review and requests were only utilized if the observations were recent (last 10 years), had published (peer-reviewed) methods, were georeferenced, and included explicit information about the accuracy of the data. The publicly-available datasets were divided into two dependent upon the quality of georeferencing and the type of data: (1) accurately georeferenced benthic field data, and (2) benthic field data or a benthic/geomorphic habitat maps with approximate geolocations.

The benthic composition depicted in the publicly available datasets were cross-walked and relabeled to the general benthic cover classes used in the Allen Coral Atlas (Kennedy et al., 2020): coral/algae, seagrass, microalgal mats, sand, rubble, and rock.

Existing Coral Reef Reference Data Sets
Accurately georeferenced benthic field data
Photoquadrats were collected at various depths and locations for the purpose of ecological assessment (González-Rivero et al., 2014) and/or validation of satellite imagery-derived habitat maps (Roelfsema and Phinn, 2010). Photoquadrats were collected randomly or along transect surveys and accurately georeferenced using a surface GNSS device towed by diver or snorkeler, where the timestamp of the logged GNSS position was synchronized with the timestamp of photoquadrat capture. For most of these data sets, benthic composition was derived from georeferenced photoquadrats using manual interpretation using Coral Point Count Excel (Kohler and Gill, 2006), or the machine learning platform CoralNet (Beijbom et al., 2015).

Benthic field data or benthic/geomorphic habitat maps with approximate geolocations
These data sets included field data collected for ecological assessment purposes (English et al., 1997) or maps. With respect to the field data, accuracy information was often not provided and the positioning methods poorly described, and thus the approximate GNSS position of the benthic field data could vary between tens to hundreds of meters of the actual point in the field. Examples include data types such as (i) a single
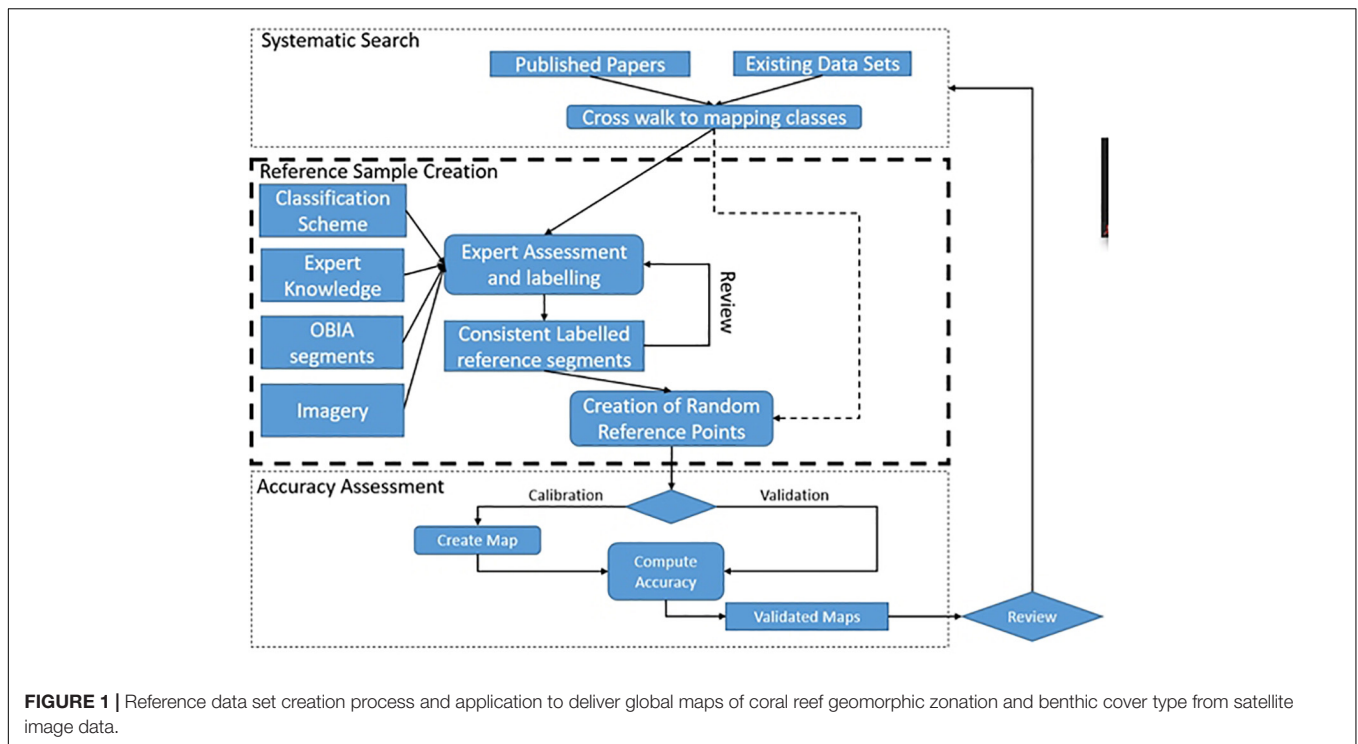
FIGURE 1 | Reference data set creation process and application to deliver global maps of coral reef geomorphic zonation and benthic cover type from satellite image data.

georeferenced point in the middle of a series of belt transects; (ii) a randomly placed series of photoquadrats; (iii) the survey vessel position; and (iv) a general description of the location. Maps of geomorphic zonation or benthic composition were generally georeferenced, but the thematic or spatial detail often varied and required expert interpretation.

Reference Data Segments From Satellite Imagery
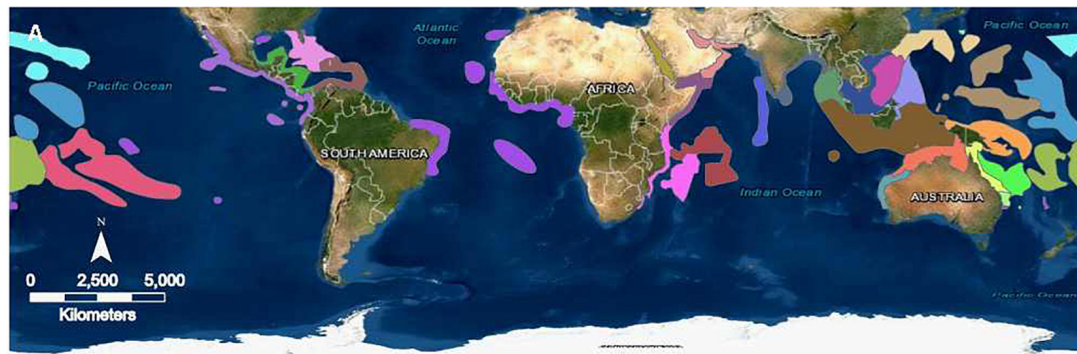
Reference data segment creation

Reference data segments were created from satellite image quadrat tiles (20 km × 20 km) for each mapping region. Reference image quadrat tiles (**Figure 2C**) were selected based on a combination of the presence of field data and/or maps (**Figure 2B**), expert knowledge of the site, quality of the Planet Dove satellite mosaic, and representation of the range of unique habitats within the region (e.g., fringing or atoll reefs, clear or less clear water, shallow or deep water (see example in **Figure 3A**). Image quadrat tiles and aggregated field data or maps were overlaid on the Planet Dove satellite image mosaic for the respective mapping region in a GIS environment (ArcPro).

Each Planet Dove reference image quadrat tile was segmented (via Trimble eCognition) into interpretable groups of pixels (segments) at geomorphic (**Figure 3B**) or benthic scale (**Figure 3C**). The aim of this stage was to generate a reference data set from these segments. To assign segments to mapping categories we used eight experts who had between two and 20 years of experience in field survey and remote sensing image analysis of coral reef and seagrass environments. The experts were trained and constantly reviewed in their ability to identify the different mapping categories. For each region, each expert was assigned a set of image quadrat tiles, and

for each the expert would manually assign mapping categories to segments. A maximum of 2 h was set per image quadrat tile, to assign segments with a geomorphic (**Figure 3B'**) or benthic label (**Figure 3C'**), however, in some cases this could be less dependent upon the extent of the reef surface area and complexity. Geomorphic classes followed the Reef Cover classification scheme (Kennedy et al., 2020) and included reef slope, reef crest, outer reef flat, inner reef flat, shallow lagoon, deep lagoon, back reef slope, sheltered slope, terrestrial reef flat, plateau, and patch reef. Assignment of these classes was based on the description of the individual geomorphic classes and expert visual interpretation of the imagery, water depth, slope, significant wave height, and existing geomorphic maps. Benthic class assignment was similar, however, it depended primarily on accurate georeferenced benthic field data in addition to interpretation of the less spatially-reliable benthic field data or maps.

To avoid introducing misclassified segments into the training set and to reduce the likelihood of error propagation in our mapping workflow, a protocol of quality assurance was developed. This included: (1) weekly review of examples of class assignments to reference segments by experts to fine-tune label assignment across experts; (2) all final reference segment assignment was reviewed by the most experienced expert for that region; (3) classification cues for geomorphic and benthic categories were created; and (4) confirmation of adherence to the classification scheme (Kennedy et al., 2020). Additionally, after reference data segments were created for a mapping region, each expert ranked the mapping categories from 1 to 10 where 1 represents a 51% confidence in labeling a segment with the specific class, and 10 represents 100% confidence. Based on
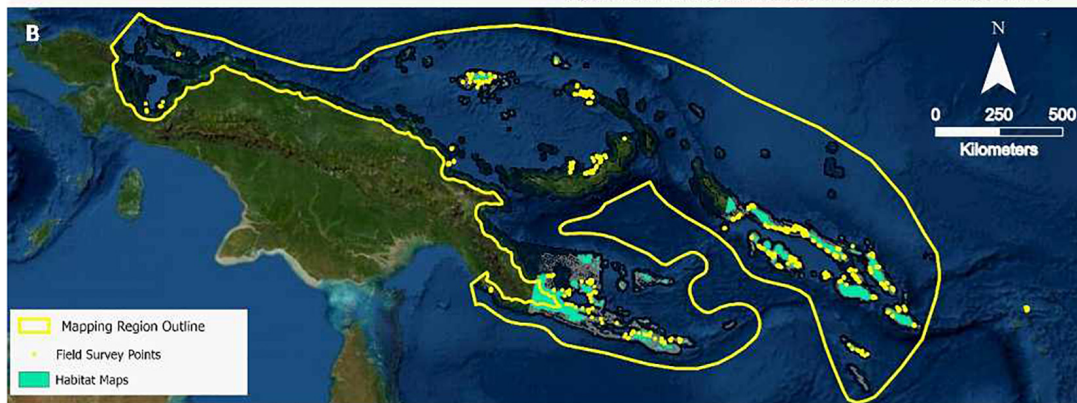
FIGURE 2 | Global reef mapping regions. **(A)** Division of the global reefs into similar biogeographic regions representing individual mapping regions. Regions completed as of December 2020 are indicated in bold type; **(B)** the distribution of existing field data and maps for the Papua New Guinea – Solomon Islands mapping region; and **(C)** Papua New Guinea – Solomon Islands region with the location of the reference data indicated (highlighted image quadrat tiles each of 20 km × 20 km).
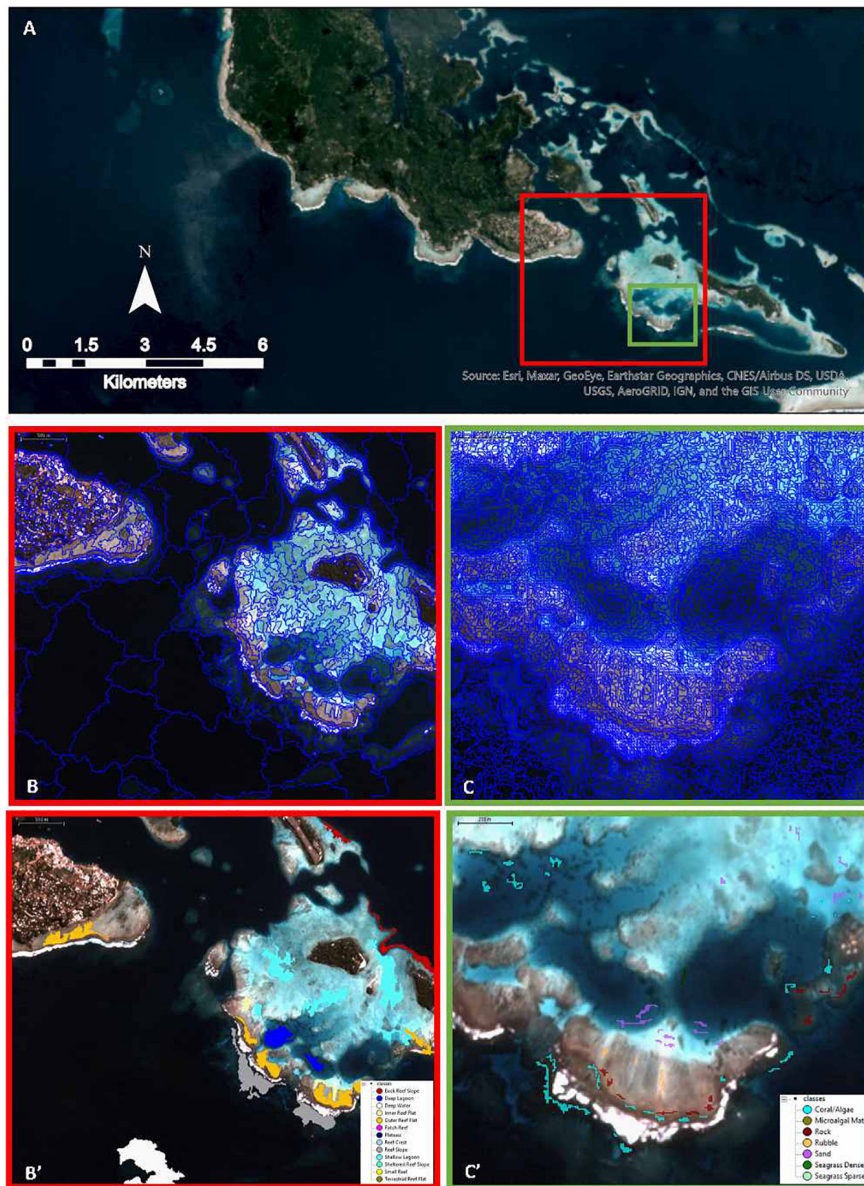
FIGURE 3 | The creation of reference data segments. **(A)** A representative Planet Dove image quadrat tile was segmented into interpretable groups of pixels (segments depicted by blue lines) at the **(B)** geomorphic and **(C)** benthic scale, using the object-based analysis software Trimble eCognition. Respective segments were assigned **(B')** geomorphic, or **(C')** benthic classes.

the review of this confidence ranking by the expert whom assigned the segments, further fine-tuning of the specific class was conducted for the label assignment across experts.

Sampling the reference data segments to create training and validation data sets

Point-based reference data samples were derived from each of the reference data segments. Subsequently, individually labeled reference data point samples were divided to create a training data set for the mapping process, and a validation data set to calculate the accuracy metrics. Training data were created via intersection with the mapping covariate data (e.g., satellite

imagery, depth, and slope, etc.), and validation data were simply the class labels.

Analysis of covariate data extracted for training data point samples

We explicitly compared the values of the covariate data from the mapping training samples to explore the variation among classes, regions and variables. To provide an overview of the type of information extracted for the training data and the variation encountered among both the mapping classes and mapping regions, we constructed box plots from the training points used for the mapping. Standard Tukey style box and whisker
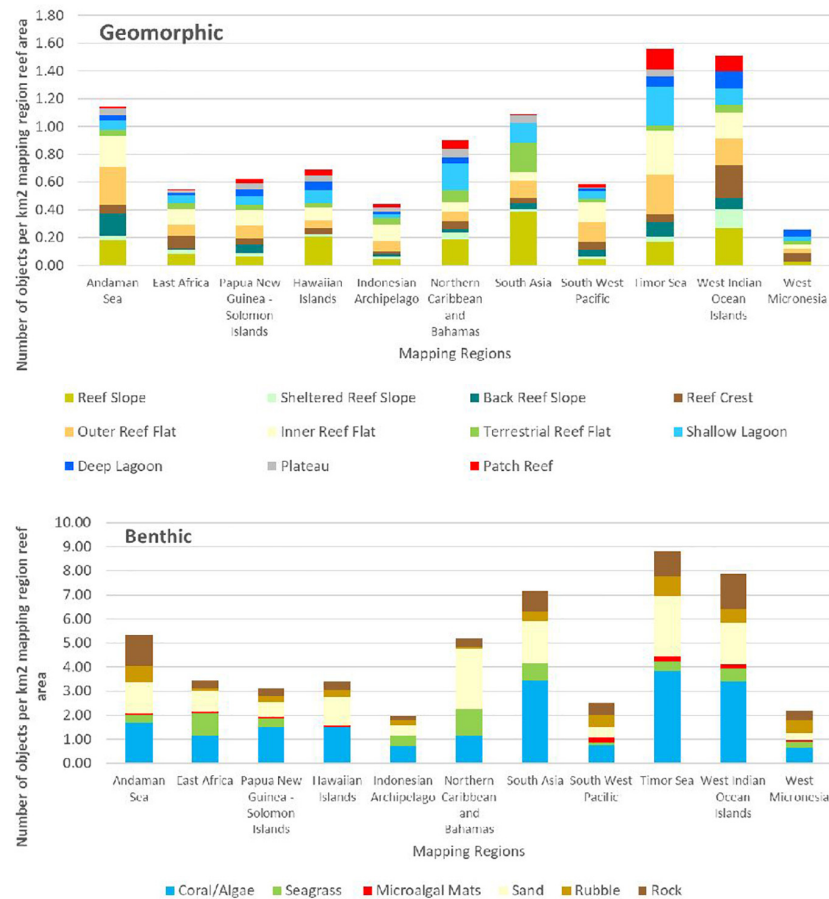
FIGURE 4 | Quantification of reference data segments. The number of geomorphic and benthic reference data segments per category, per region normalized for the extent of reef area within each mapping region.

plots were generated for each of the benthic and geomorphic mapping classes, for three key covariates; green reflectance image band as it demonstrated more variability than the blue or red bands, satellite-derived bathymetry and slope derived from the bathymetry. These plots were developed for all of the mapping regions combined, as well as separately for each region.

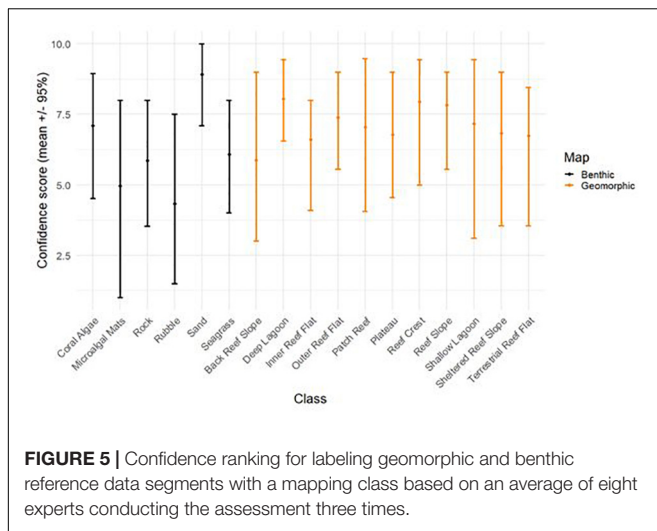## RESULTS AND DISCUSSION

Reference Data Segments

The number of reference segments varied depending on the size and complexity of the region (**Supplementary Table 1**). The minimum number of image quadrat tiles used for a region was 17 (Andaman Sea) and the maximum was 81 (Indonesian Archipelago, with an area 16 times larger). On average, fewer geomorphic than benthic segments were labeled per image quadrat tile (126 for geomorphic versus 593 for benthic), and on average, geomorphic segments were 166 times larger than benthic segments. This was expected as the geomorphic zone represents features that often cover hundreds to thousands of meters whilst benthic classes represent features that cover tens to hundreds

of meters. The reference data are accessible via an open source repository (Roelfsema et al., 2020).

Further investigation into the distribution of reference data segments per mapping class in each mapping region (**Figure 4**) indicated that for geomorphic segments, the most commonly sampled geomorphic classes were inner reef flat, outer reef flat and reef slope. For benthic segments, the most commonly sampled were sand and coral/algae. These classes usually constitute the largest areas on a reef and are, in general, identifiable with a higher confidence (**Figure 5**).

The West Indian Ocean Islands had the least number of reference data segments per reef surface area, even though the number of image quadrat tiles used was average, which is expected to be due to the similarity of reef types within that region.

Reference data segments created for the purpose of mapping global reefs represent a significant amount of data from which to extract training and validation points. Although confidence rankings are high for these reference data segments (**Figure 5**), the reliability is potentially low as they are based on expert interpretation of imagery and field data/maps. The estimated confidence that experts had in the ability to label a segment

FIGURE 5 | Confidence ranking for labeling geomorphic and benthic reference data segments with a mapping class based on an average of eight experts conducting the assessment three times.

with a certain mapping class was higher for geomorphic than for benthic segments (**Figure 5**). Confidence varied for geomorphic segments, on average ranking between 6.5 (back reef slope) and 8 (deep lagoon), and for benthic between 4 (rubble) and 9 (sand). It is known that manual interpretation by experts requires alignment of class descriptors with segments from an aerial photograph (Aswani and Lauer, 2006), or a photoquadrat (Beijbom et al., 2015). However, it has been demonstrated that increased cross-calibration between experts through training is known to improve class assignment (Done et al., 2017). Future work by the authors will focus on analysis of the variability between experts and how that could be tested for such large data sets as presented in this paper.

Manual interpretation was required for the assignment of geomorphic classes to segments, which was based upon: (1) distinguishing dark versus bright features (a surrogate for hard versus soft substrate, respectively), (2) the use of visual interpretation cues (e.g., color, texture, and brightness) in the satellite imagery, (3) physical attributes (e.g., depth, slope, and wave exposure), (4) neighborhood relationships (e.g., reef crest neighbors reef slope), and (5) detailed classification definitions (Kennedy et al., 2020). A similar process has been used previously for large scale global geomorphic mapping (Andréfouët et al., 2006). In that case, however, manual delineation was used to create the maps, rather than to create reference data. As geomorphic class assignment included examination of additional variables, this resulted in geomorphic classes being easier to determine via manual expert interpretation, as the position and physical environment provided extra clues for the interpreter to assign a class. Compared to the geomorphic classes, differentiation of benthic classes in satellite imagery was more dependent upon interpretation of satellite image color and texture than on physical attributes or neighborhood relationships. For the assignment of benthic classes to reference data segments, the mapping approach would ideally be driven by purposely-collected geolocated benthic field surveys that coincided with capture of the high-resolution satellite imagery such as Planet Dove (Andréfouët, 2008). As such, the presence

of benthic field data increased confidence in the assignment of benthic classes to segments, and in general, more field data was available for areas of the reef that were known to have higher coral cover. High confidence in sand, and low variability in that confidence, was likely the result of it being a distinctively bright feature on a coral reef in satellite imagery.

The Living Ocean Foundation (LOF) created detailed spatial and thematic benthic maps of 65,000 km[2] of coral reefs based on field data and coincident high spatial resolution satellite imagery over a 10 year period (Purkis et al., 2019). For most of these reef regions, detailed field campaigns focused on the collection of region-specific georeferenced training and validation data. This could be considered the "gold standard" to map reefs globally, however, it rapidly becomes infeasible with the current resources allocated to global reef mapping. Hence, there is a trade-off for generating reference data. We describe an approach for mapping the world's coral reefs (approx. 255,000 km[2]) in less time and at a lower cost per square kilometer (Spalding and Grenfell, 1997).
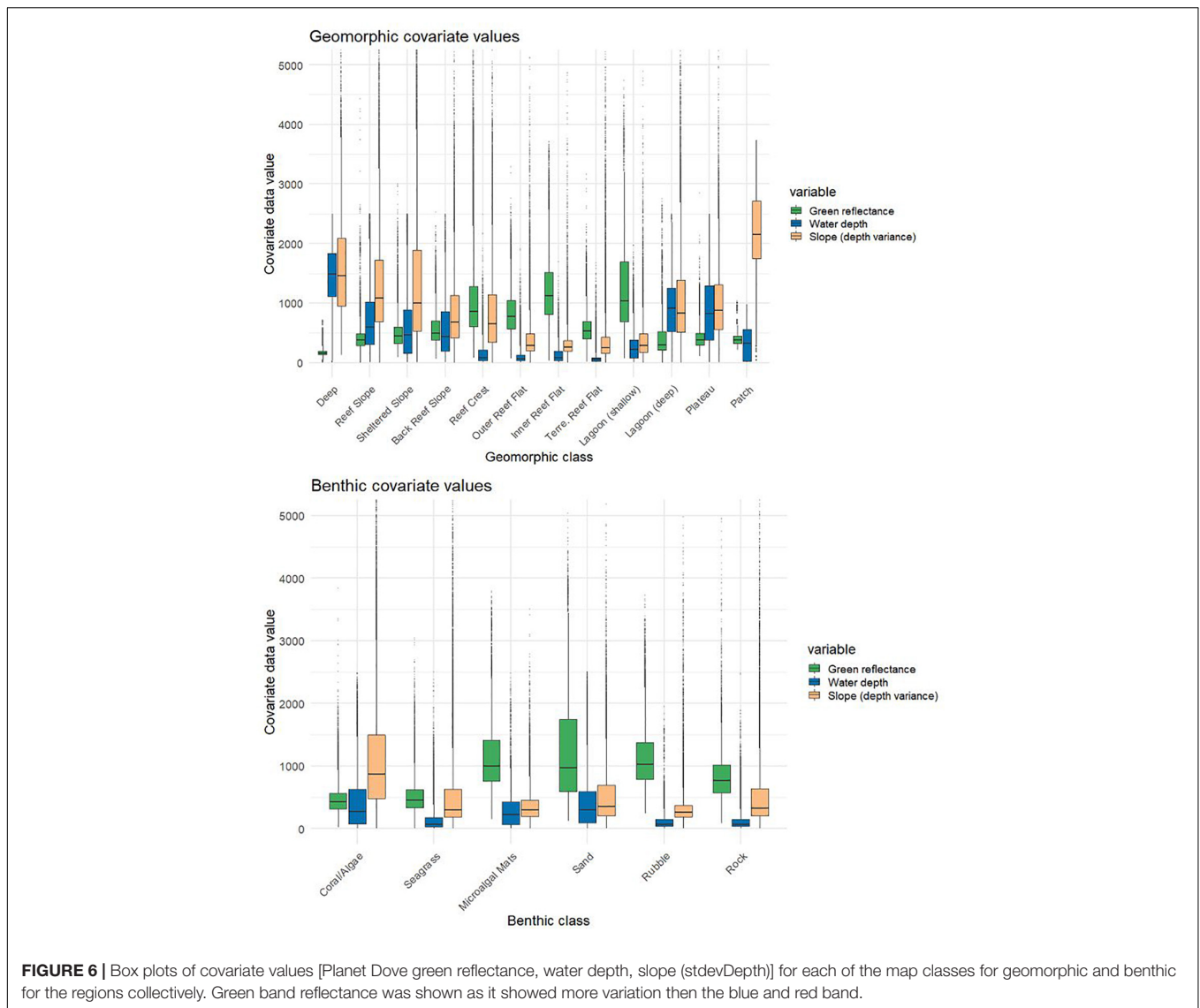
Characterization of the Training Data Point Samples for Individual Mapping Categories

Training points vary considerably among map classes, which is reflected in high variability of the Planet Dove green band reflectance values and the physical attributes of depth and slope (**Figure 6**).

Figure 6 shows that geomorphic and benthic classes are broadly differentiated by the green band reflectance of the Planet Dove imagery, water depth and slope. Some geomorphic classes such as reef slope and sheltered reef slope, are, however, similar, which is unsurprising given their main distinction is exposure to wave energy. Back reef slope exhibits a different pattern, with higher green band values in particular, which is due to the presence of bright sand and a low slope. Deep lagoon and plateau have similar covariate values, but are distinguished by their neighbors: plateau is predominantly surrounded by deep water while deep lagoon is typically surrounded by shallow water. Reef crest, outer reef flat and inner reef flat occur at similar depths but are differentiated in our data by relative brightness. Outer reef flat and inner reef flat tend to have different amounts of sand cover, whilst reef crest is brighter than both due to the presence of breaking waves that occur on shallow crest formations.

Variation in covariate values per region was also observed (PNG-Solomon Islands example in **Figure 7**), which indicates high variation in reef types and benthic composition around the world (see **Supplementary Figure 1** for plots of for the individual regions). For example, the PNG-Solomon Islands region's covariate values clearly reflect extremely diverse reef environments, which stretch across a very large area from a large landmass (PNG) to clear oceanic waters. In contrast, the Andaman Sea region is smaller and has less diversity of reef environments (**Supplementary Figure 1**).

All of the mapping regions have similar between-class variation, which is due to similarities in biophysical and geomorphological traits. However, we do note considerable variation among regions, much of which is idiosyncratic

FIGURE 6 | Box plots of covariate values [Planet Dove green reflectance, water depth, slope (stdevDepth)] for each of the map classes for geomorphic and benthic for the regions collectively. Green band reflectance was shown as it showed more variation then the blue and red band.

(**Supplementary Figure 1**). Given these variations in the reference data, an obvious avenue for future research is to assess the transferability of reference data sets from one mapping region to another and its downstream impact on the performance of classification models.

Considerations for Training and Validation Data Set Creation From a Global Mapping Perspective

There are several ways in which reference data sets could be used for training and validation samples. In the case of the global coral mapping described here, individual points were randomly sampled throughout the whole reference data set, and those points were split into training and validation data sets. This approach was preferable due to the already very sparse nature of the reference data sets. In cases where the reference data set provided a more thorough coverage of the mapping area,

the reference data itself could have been split into training and validation data sets, prior to sampling the training data from the covariate data. Another consideration is the proportion by which training and validation data are sampled – without strong *a-priori* information on the probability distribution of classes, in the case of the global mapping, we sampled an even number of points from each class. In other applications, either a weighting could be applied to the point sampling or the probability distribution of the reference data itself could be used.

Expert Interpretation Compared to New Field Surveys for Reference Data Creation

This study presents examples of reference data sets created via an expert knowledge-based workflow. Purposely planned field surveys would clearly provide a more reliable reference data source, yet these require a significant time and resource
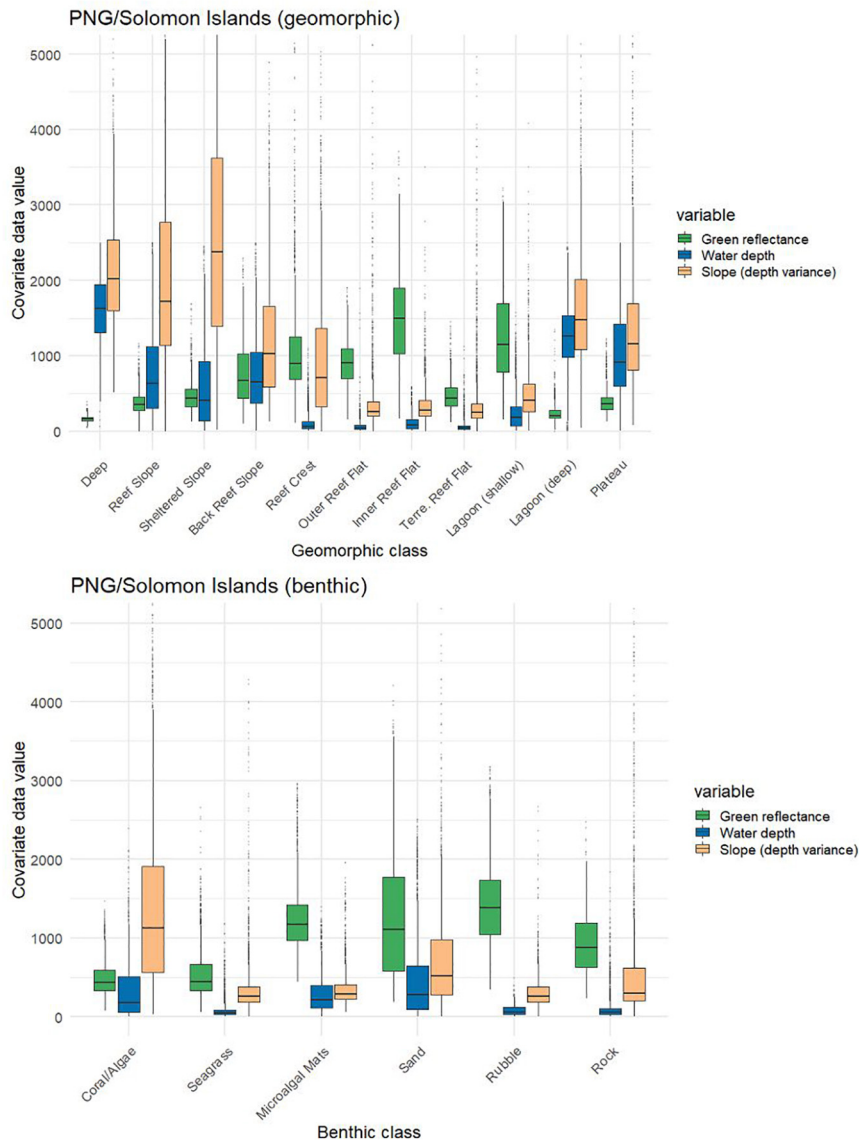
FIGURE 7 | Example of mean covariate values [green band reflectance, water depth, slope (stdevDepth)] for each of the mapping classes for geomorphic and benthic maps, highlighting the variability of the PNG-Solomon Islands Region.

investment, which is largely prohibitive for any large-scale mapping project. Nevertheless, our mapping workflow can effectively incorporate these data, should the opportunity to compile a globally extensive field data set arise. A simple estimate to achieve this for just one of our mapping regions, PNG-Solomon Islands, suggests it would take a 2,750% increase in effort to develop a reference data set based on new field surveys versus the approach we describe in this paper. In this region, sufficient field work would take an estimated 275 days for three people (770 work equivalent days), in stark contrast to the expert interpretation method described here, which totaled 30 days for one person. The PNG-Solomon Islands region extends 3,300 km west to east, and 1,000 km south to north, covering a reef area of 10,366 km^2, and represents around 6% of the world's

coral reefs (**Figure 2**). 55 image quadrat tiles were selected in this area (**Figure 2C**) and an average of 135 geomorphic and 593 benthic reference segments created per image quadrat tile (**Supplementary Table 1**). Each tile typically takes an expert about 2 h, with 55 image quadrat tiles taking an equivalent of 15 working days, with an additional 15 working days to search and gather existing data. A new field effort to gather benthic information by visiting the 593 reference segments within one reef area (image quadrat tile) would take 4.5 days with a three-person field team, including travel time and half a day for one person to analyze the field data, totaling 275 days (5 days × 55 image quadrat tiles). These field-based estimates are obviously conservative, as they do not consider travel time to the region, nor the added expense for accommodation and live-aboard staff.

TABLE 1 | Contrasting the desktop study versus field-based reference data set collection.

	Desktop study-based reference data creation	Field-based reference data from multiple sources
Consistency of creation	High	Low, unless a globally regulated method and program
Confidence in method	High	Low; unless a globally regulated method and program
Expertise level	High; expert in visual interpretation of benthic classes	High; expert in visual interpretation of benthic classes
Implementation globally	Possible with reasonable effort and resources to conduct and coordinate	Possible but significant effort and resources are required to conduct and coordinate
Thematic Accuracy	Low for many classes; High for easily defined classes (e.g., bright sand versus dark hard substrate)	High but variable, unless one person is responsible for the assignment
Positional accuracy	Variable	Variable
Repeatability over the global extent	High	Low

Table 1 compares the specific requirements deemed necessary for reference sample creation via either approach.

Thus, the expert interpretation desktop workflow presented here is a suitable and consistent approach for a large-scale effort with limited resources to support field data collection given that reefs require specialized field techniques unique for submerged environments, and/or there is limited access to mapping regions due to political unrest (e.g., South China Sea) extreme remoteness (e.g., Pacific reefs), or global crises such as the COVID-19 pandemic. The workflow is dependent upon a limited number of experts for interpretation, which have further limitations imposed due to the size of the data sets required. Trained citizen scientists could significantly enhance the capability for the creation of large region reference data sets, as is being done by NASA NeMO-Net (Van Den Bergh et al., unpublished[3]). Additionally the workflow presented could integrate a standard global monitoring protocol using photoquadrats such as the CATLIN seaview program (González-Rivero et al., 2014). Each of these options, however, requires a substantial investment of time and funding for development.

CONCLUSION

We presented a detailed desktop workflow to create reference data sets for the training and validation of high spatial and thematic resolution maps of coral reefs at a global scale, which was developed for the Allen Coral Atlas Global Coral Reef Mapping project.

The workflow presented here for the creation of reference data sets could be implemented for any environment. The minimum requirements for such a workflow would include access to expert knowledge, a detailed description of the classification scheme, and imagery from which the required mapping classes can be differentiated. The main advantages of the workflow are that it is applicable to any ecosystem anywhere, works across different spatial and thematic scales, provides a statistically sufficient sample set that is relevant for the proposed classification model, addresses class balance and minimum accuracy requirements,

and reference data sets from one area may inform reference data segment class assignment in another. However, the disadvantages are that an expert with knowledge or training in the area is required, consistency may vary between experts, no statistics are involved in selecting the image quadrat tiles for reference data segment creation, the quality and quantity of reference segments for each region is variable, and benthic or land cover class assignment would be based predominantly on color and/or texture, while geomorphic or topographic class assignment would be based on color, texture, and physical and/or environmental variables.

Our detailed description of reference data creation for global coral reef mapping through expert interpretation and quality control, provides the opportunity for others such as regional experts to participate in the creation of these data sets for their region. The consistency achieved by expert interpretation demonstrated by this study is unprecedented over the global extent, given the level of spatial (5 m pixels) and thematic detail mapped (more than 10 mapping classes), especially when compared to other dynamic global mapping efforts such as global forest gain/loss (Hansen et al., 2013), mud flats (Murray et al., 2019), or mangroves (Bunting et al., 2018).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Figshare.com: https://doi.org/10.6084/m9.figshare.c.5233847.

AUTHOR CONTRIBUTIONS

CR, ML, NM, EMK, SP, and EK designed the study. CR, EMK, EMK, KM, AO, CS, PT, JW, DT, BB, BF, and ZL collected data. CR, ML, EMK, EK, RB, PT, MR, and JW created reference samples. CR, ML, RB-A, EMK, EK, and PT analyzed data. CR wrote the first draft of the manuscript. CR, ML, NM, EK, SP, HF, and GA contributed to the writing and editing of the manuscript. All authors contributed to the article and approved the submitted version.

---

[3]Van Den Bergh, J., Chirayath, A. L., Torres-Pérez, J., and Segal-Rozenhaime, M. (2021). NASA NeMO-Net's citizen science video game – 3D classification of coral reefs from airborne fluid lensing and satellite remote sensing. *Front. Mar. Sci.* this issue.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2021.643381/full#supplementary-material

REFERENCES

Andréfouët, S. (2008). Coral reef habitat mapping using remote sensing: a user vs producer perspective. Implications for research, management and capacity building. *J. Spatial Sci.* 53, 113–129. doi: 10.1080/14498596.2008.9635140

Andréfouët, S., Hochberg, E. J., Payri, C., Atkinson, M. J., Muller-Karger, F. E., and Ripley, H. (2003). Multi-scale remote sensing of microbial mats in an atoll environment. *Int. J. Remote Sens.* 24, 2661–2682. doi: 10.1080/0143116031000066909

Andréfouët, S., Muller-Karger, F. E., Robinson, J. A., Kranenburg, C. J., Torres-Pulliza, D., Spraggins, S., et al. (2006). "Global assessment of modern coral reef extent and diversity for regional science and management applications: a view from space," in *Proceedings of the 10th International Coral Reef Symposium*, (Okinawa: International Coral Reef Society).

Aswani, S., and Lauer, M. (2006). Benthic mapping using local aerial photo interpretation and resident taxa inventories for designing marine protected areas. *Environ. Conserv.* 33, 263–273. doi: 10.1017/S0376892906003183

Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P., et al. (2015). Towards automated annotation of benthic survey images: variability of human experts and operational modes of automation. *PLoS One* 10:e0130312. doi: 10.1371/journal.pone.0130312

Bormann, K. J., Brown, R. D., Derksen, C., and Painter, T. H. (2018). Estimating snow-cover trends from space. *Nat. Clim. Chang.* 8, 924–928. doi: 10.1038/s41558-018-0318-3

Bunting, P., Rosenqvist, A., Lucas, R. M., Rebelo, L.-M., Hilarides, L., Thomas, N., et al. (2018). The global mangrove watch—a new 2010 global baseline of mangrove extent. *Remote Sens.* 10:1669. doi: 10.3390/rs10101669

Congalton, R. G., and Green, K. (2008). *Assessing the Accuracy of Remotely Sensed Data: Principles and practices. Mapping Science*. Boca Rotan FL: CRC Press. doi: 10.1201/9781420055139

Done, T., Roelfsema, C., Harvey, A., Schuller, L., Hill, J., Schläppy, M.-L., et al. (2017). Reliability and utility of citizen science reef monitoring data collected by Reef Check Australia, 2002–2015. *Mar. Pollut. Bull.* 117, 148–155. doi: 10.1016/j.marpolbul.2017.01.054

English, S., Wilkinson, C. R., and Baker, V. (1997). *Survey Manual for Tropical Marine Resources*. Townsville TSV: Australian Institute of Marine Science.

González-Rivero, M., Beijbom, O., Rodriguez-Ramirez, A., Holtrop, T., González-Marrero, Y., Ganase, A., et al. (2016). Scaling up ecological measurements of coral reefs using semi-automated field image collection and analysis. *Remote Sens.* 8:30. doi: 10.3390/rs8010030

González-Rivero, M., Bongaerts, P., Beijbom, O., Pizarro, O., Friedman, A., Rodriguez-Ramirez, A., et al. (2014). The Catlin Seaview Survey – kilometre-scale seascape assessment, and monitoring of coral reef ecosystems. *Aquatic Conserv. Mar. Freshw. Ecosyst.* 24, 184–198. doi: 10.1002/aqc.2505

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. doi: 10.1016/j.rse.2017.06.031

Groom, S., Sathyendranath, S., Ban, Y., Bernard, S., Brewin, R., Brotas, V., et al. (2019). Satellite Ocean Colour: current status and future perspective. *Front. Mar. Sci.* 6:485. doi: 10.3389/fmars.2019.00485

Hamylton, S. (2011). Estimating the coverage of coral reef benthic communities from airborne hyperspectral remote sensing data: multiple discriminant function analysis and linear spectral unmixing. *Int. J. Remote Sens.* 32, 9673–9690. doi: 10.1080/01431161.2011.574162

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science* 342:850. doi: 10.1126/science.1244693

Keith, D. A., Rodríguez, J. P., Brooks, T. M., Burgman, M. A., Barrow, E. G., Bland, L., et al. (2015). The IUCN red list of ecosystems: motivations, challenges, and applications. *Conserv. Lett.* 8, 214–226. doi: 10.1111/conl.12167

Kennedy, E., Roelfsema, C., Lyons, M., Kovacs, E., Borrego-Acevedo, R., Roe, M., et al. (2020). Reef Cover: a coral reef classification to guide global habitat mapping from remote sensing. *bioRxiv [Preprint]* doi: 10.1101/2020.09.10.292243

Kohler, K. E., and Gill, S. M. (2006). Coral Point Count with Excel extensions (CPCe): a visual basic program for the determination of coral and substrate coverage using random point count methodology. *Comput. Geosci.* 32, 1259–1269. doi: 10.1016/j.cageo.2005.11.009

Leon, J., and Woodroffe, C. D. (2011). Improving the synoptic mapping of coral reef geomorphology using object-based image analysis. *Int. J. Geogr. Inf. Sci.* 25, 949–969. doi: 10.1080/13658816.2010.513980

Li, J., Knapp, D. E., Schill, S. R., Roelfsema, C., Phinn, S., Silman, M., et al. (2019a). Adaptive bathymetry estimation for shallow coastal waters using Planet Dove satellites. *Remote Sens. Environ.* 232:111302. doi: 10.1016/j.rse.2019.111302

Li, J., Schill, S. R., Knapp, D. E., and Asner, G. P. (2019b). Object-based mapping of coral reef habitats using planet dove satellites. *Remote Sens.* 11:1445. doi: 10.3390/rs11121445

Lyons, M., Roelfsema, C., Kennedy, E., Kovacs, E., Borrego, R., Markey, K., et al. (2020). Mapping the world's coral reefs using a global multiscale earth observation framework. *Remote Sens. Ecol. Conserv.* 6, 557–568. doi: 10.1002/rse2.157

Lyons, M. B., Keith, D. A., Phinn, S. R., Mason, T. J., and Elith, J. (2018). A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.* 208, 145–153. doi: 10.1016/j.rse.2018.02.026

Monaco, M. E., Andersen, S. M., Battista, T. A., Kendall, M. S., Rohmann, S. O., Wedding, L. M., et al. (2012). *National Summary of NOAA's Shallow-Water Benthic Habitat Mapping of U.S. Coral Reef Ecosystems. NOAA Technical. Memorandum*. Silver Spring, MD: NOAA. Centers. Coastal. Monit. Assess. Biogeogr.

Murray, N. J., Keith, D. A., Bland, L. M., Ferrari, R., Lyons, M. B., Lucas, R., et al. (2018). The role of satellite remote sensing in structured ecosystem risk assessments. *Sci. Total Environ.* 61, 249–257. doi: 10.1016/j.scitotenv.2017.11.034

Murray, N. J., Phinn, S. R., DeWitt, M., Ferrari, R., Johnston, R., Lyons, M. B., et al. (2019). The global distribution and trajectory of tidal flats. *Nature* 565, 222–225. doi: 10.1038/s41586-018-0805-8

Phinn, S. R., Roelfsema, C. M., and Mumby, P. J. (2012). Multi-scale, object-based image analysis for mapping geomorphic and ecological zones on coral reefs. *Int. J. Remote Sens.* 33, 3768–3797. doi: 10.1080/01431161.2011.63 3122

Purkis, S. J., Gleason, A. C. R., Purkis, C. R., Dempsey, A. C., Renaud, P. G., Faisal, M., et al. (2019). High-resolution habitat and bathymetry maps for 65,000 sq. km of Earth's remotest coral reefs. *Coral Reefs* 38, 467–488. doi: 10.1007/s00338-019-01802-y

Roelfsema, C., and Phinn, S. (2010). Integrating field data with high spatial resolution multispectral satellite imagery for calibration and validation of coral reef benthic community maps. *J. Appl. Remote Sens.* 4:043527. doi: 10.1117/1. 3430107

Roelfsema, C. M., Kovacs, E. M., Ortiz, J. C., Callaghan, D. P., Hock, K., Mongin, M., et al. (2020). Habitat maps to enhance monitoring and management of the Great Barrier Reef. *Coral Reefs* 39, 1039–1054. doi: 10.1007/s00338-020-01 929-3

Roelfsema, C. M., and Phinn, S. R. (2013). *Validation.* Coral Reef Remote Sensing. Berlin: Springer, 375–401. doi: 10.1007/978-90-481-9292-2_14

Spalding, M., Burke, L., Wood, S. A., Ashpole, J., Hutchison, J., and zu Ermgassen, P. (2017). Mapping the global value and distribution of coral reef tourism. *Mar. Policy* 82, 104–113. doi: 10.1016/j.marpol.2017.05.014

Spalding, M. D., and Grenfell, A. M. (1997). New estimates of global and regional coral reef areas. *Coral Reefs* 16, 225–230. doi: 10.1007/s003380050078

Tittensor, D. P., Walpole, M., Hill, S. L. L., Boyce, D. G., Britten, G. L., Burgess, N. D., et al. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science* 346:241. doi: 10.1126/science.1257484