Seafloor Substrate Characterization from Shallow Reefs to the Abyss:
Spatially-continuous seafloor mapping using multispectral satellite imagery, and multibeam bathymetry and backscatter data within the Pacific Remote Islands Marine National Monument and the Main Hawaiian Islands.[1]

Dayton Dove, Tomoko Acoba, and Annette DesRochers
June 2017

Summary

The report discusses the development of classification approaches for predicting the presence of hardbottom seafloor substrates in shallow (0–30 m) and mesophotic to deeper-water environments (> 30 m). In both shallow and deeper-water environments, an unsupervised machine-learning approach (clustering) was utilized to classify the seafloor, employing available remote-sensing data for each environment. In shallow-water environments, high-resolution multispectral satellite imagery (WorldView-2) was used while ship-based multibeam sonar data were used to characterize deeper-water environments. The final classification approach for the deeper water environments incorporated backscatter data and three bathymetric derivatives, including slope, standard deviation, and relative distance to the mean in the clustering analysis. Backscatter is widely accepted as an important metric in detecting benthic habitat variation, and the addition of the combined bathymetric derivatives provided unique indicators of hardness that improved the accuracy of the substrate predictions. These methods are intended to be applied where suitable ground-truth data are not available (i.e., most of the U.S. Pacific Islands Region). Hard substrates serve as predictive surrogates for benthic biodiversity and abundance, and these map products provide environmental baseline information useful to further understanding of ecosystem

---

[1] PIFSC Internal Report IR-18-018.
Issued 18 October 2018.

processes and interactions (e.g., benthic habitats and bottom fish), as well as supply habitat metrics for marine spatial planners and policy makers.

The approach was first developed and tested extensively for the west coast of Hawaii Island, and it was immediately employed in both shallow-reef and deep-water environments to characterize substrates within the Pacific Remote Islands Marine National Monument (PRIMNM). The map products will inform the development of the first Monument Management Plan for the PRIMNM. The shallow substrate maps also inform considerations regarding critical habitat designation for listed corals within the PRIMNM under the U.S. Endangered Species Act. The approach is also being applied in the main Hawaiian Islands (MHI) to examine how seafloor substrate constrains the distribution of bottomfish (Richards et al., 2016), as well as within the Atlantis ecosystem model to investigate the relative influence of physical substrate on ecosystem responses under changing climate and oceanographic conditions (Weijerman et al., 2015).

This report describes the overall background, objectives, and results on the development of this modelling approach. The results provided here are for illustrative purposes only and are not the final products. The objective of this report is to demonstrate the pros and cons of the mapping approach in both shallow and deep-water environments, primarily organized around figure examples. A separate protocol document has also been produced, which provides step-by-step procedures for implementing this approach in both shallow and deep-water environments.

Background

As NOAA and the international community have been transitioning toward implementation of Ecosystem-Based Fisheries Management, there has been significant global effort in research and data acquisition directed towards better understanding and characterization of benthic habitats (e.g., Harris and Baker, 2012). Equipped with extensive, high-quality datasets, marine spatial analysts have sought to develop methods that more accurately classify seafloor substrates, which underpin many habitat models. Providing continuous spatial information on the morphology, composition, and dynamics of the seafloor is useful for fundamental ecosystem baseline studies and provides spatial metrics necessary for marine spatial planning and policy making. Hard substrates have been observed across many environments to support greater biodiversity and abundance (Knowlton, 2001; Taylor and Wilson, 2003; McArthur et al., 2010; Davies and Guinotte, 2011; Pyle et al., 2016).

Multispectral satellite imagery, airborne hyperspectral imagery, and Lidar data have been commonly used to characterize the seafloor in shallow reef-dominated environments. Due to the reduction of the image acquisition cost, much progress has involved the application of satellite imagery to extract useful information about coral habitats (e.g., geomorphic zones; Hochberg et al., 2003; Phinn et al., 2012; Asner et al., 2017). Where the multispectral satellite data (e.g., WorldView-2, IKONOS, Quickbird) are augmented with georeferenced ground-truth data, spectral matching enables classification of biological cover and benthic communities. Further processing of the imagery can extract derivatives like bathymetry, which are further employed to improve seafloor classifications (Mumby et al., 2004). However, the use of multispectral satellite imagery is limited within the penetration depth of the optical range of the electromagnetic spectrum.

For deeper waters beyond the penetration depth of satellite imagery, progress has focused on the treatment of multibeam echosounder bathymetry and backscatter data. With the bathymetry data, researchers have shown

that including terrain variables/derivatives (e.g., slope, rugosity, northness, eastness) within classification models improves the ability to predict benthic biodiversity (Wilson et al., 2007; Dunn and Halpin, 2009; Lecours et al., 2016a, b). In effect, the morphology derivatives (e.g., slope, rugosity) serve as useful indicators for potential presence of hard substrates, and the aspect-related variables (e.g., northness, eastness) provide a proxy on whether the distribution of benthos is constrained according to local/regional hydrodynamic forces and related biogeochemical cycles. Specifically, slope informs the likelihood of sediment accretion, rugosity indicates seafloor roughness or complexity, and northness and eastness are derivatives of aspect (cosine and sine of aspect, respectively), which aid in describing the seabed in terms of exposure to wave and current energy (Calvert et al., 2015). Backscatter data, which express the intensity of the acoustic return, provide a useful proxy for the texture and hardness of the seafloor, independently providing information on the distribution of sediments and potential presence of hard substrates (Brown et al., 2011). However, it should be noted that empirical studies have suggested that when coarse sediment fractions are greater than approximately 10%, it is not possible to infer further physical characteristics from the data. That is, a relatively low percentage of coarse sediment can swamp the acoustic signal, thereby making predominantly soft substrates with some coarse material indistinguishable from hard substrates (Goff et al., 2004). This suggests that backscatter data alone should not be used to predict the presence of hard substrates. Moreover, while geometric corrections are made to the backscatter data to account for seafloor shape (co-registered bathymetry), loss of signal still occurs along steep slopes, which can result in backscatter-only based classifications under-predicting the presence of hard substrates.

With both the satellite imagery and multibeam data, efforts have focused on classification methodologies that typically involved introducing more quantitative and automated techniques (rather than manual editing/digitizing) to remove user bias, increase efficiency, and ideally generate more statistically-meaningful results (Brown et al., 2011; Phinn et al., 2012; Calvert et al., 2015; Diesing et al., 2014; Porskamp et al., 2018).

Objectives

For at least 10 years, scientists within the Pacific Islands Fisheries Science Center (PIFSC) have been working to better characterize the seafloor (Miller et al., 2008; Weiss et al., 2008; Ehses and Rooney, 2015; Suka and Rooney, 2017). In addition to partnering with scientists from the University of Hawaii (UH), PIFSC has collaborated with NOAA's National Centers for Coastal and Ocean Science (NCCOS; Lundblad et al., 2006), who have also undertaken mapping efforts in the Pacific (Battista et al., 2007). The work presented here, which commenced in Sept. 2017, represents a continuation of these efforts to more accurately characterize the seafloor substrate in order to better understand the nature and distribution of the benthos (and the pelagic organisms that depend on these habitats), and to support other ecosystem sciences that utilize substrate information to guide their work. More specifically, the objective of this project was to develop and test seafloor mapping and classification approaches for the data-rich area off the west coast of Hawaii Island and to apply the method in data-poor environments (both shallow and deep water) around the U.S. Pacific Islands without the benefit of ground-truth data. This paper discusses both the development of the method in West Hawaii and the application of the method in the PRIMNM and MHI. Suitable ground-truth data are typically sparse-to-absent across these areas (i.e., the vast majority of the U.S. Pacific Islands Region), indicating that map predictions will be based on remotely-sensed data alone.

Method Development

The method described herein discriminates between hard and soft seafloor substrates only rather than more detailed classifications (e.g., sediment composition or biological cover) due to the absence of ground-truth data in most areas, and in consideration of the level of information/detail that can be consistently and confidently extracted from remote-sensing data alone. As described above, characterizing the distribution of hard substrates is useful for characterizing and understanding a range of ecosystem processes.

Hard substrate is defined as areas where lithified outcrop/pavement (e.g., limestone pavement or basalt), hard biogenic reef (e.g., coral reef or rhodolith beds), or cobble-sized (> 64 mm) or larger unconsolidated particles (e.g., reef rubble) comprise the majority component at the seafloor. Soft substrate indicates the presence of unconsolidated sediment with particles predominantly smaller than cobble size (e.g., carbonate sand).

Prior to applying these methods in the PRIMNM and regionally across the MHI, the initial test area was off the west coast of Hawaii Island. This area was chosen based on data availability; there were multibeam, satellite (WorldView-2), and bathymetric Lidar data available within the area to develop the methodology and good, accurately geolocated ground-truth data were available for validation.. The ground-truth data supported a case study in mapping seafloor substrate within mesophotic waters using a supervised classification approach (Suka and Rooney, 2017), and these case-study results proved useful as comparison to the final results from this analysis.

Spatial scale

Spatial scale can be a confusing term and is inconsistently used to describe both the spatial resolution of underlying data and/or the final thematic scale of a map (e.g., 1:100,000). When predicting seafloor substrates, it is important to address the spatial resolution of the input data carefully as it has been frequently shown that benthic and pelagic organisms respond to abiotic (non-living) covariates at different spatial scales (Costa et al., 2015; Pittman and Brown, 2011; Sekund and Pittman, 2017; Wilson et al., 2007).

For the substrate mapping described in this report, the highest spatial-resolution available was used for the underlying data (e.g., 5-m gridded bathymetry preferred over 20-m gridded bathymetry). The spatial resolution refers to the size of the smallest possible feature that can be detected. With the multibeam bathymetry and backscatter data in particular, it is important to use the highest spatial resolution data because the analysis assumes that hard substrates are associated with more rugose surfaces. Because the scale of this roughness in reef and rock tends to occur over scales ranging from centimeters to 10s of meters, higher spatial-resolution data enabled better discrimination between rough vs. smooth terrains (i.e., hard vs. soft). Seamounts and coral reefs are ideal environments in which to apply this assumption. However, this assumption would not be so valid in continental shelf settings or any seafloor environment where significant morphology is associated with mobile sediment bedforms.

Substrate mapping was conducted with multibeam data at one spatial scale in both the PRIMNM and MHI (specific to the spatial resolution of the underlying data in each area). In the MHI, bathymetry and backscatter data and bathymetric derivatives were produced at multiple spatial resolutions to be included within species distribution models (e.g., GAMS, MaxENT) in support of other PIFSC studies, such as Atlantis ecosystem modelling and bottomfish stock assessment. Within the Atlantis project, model runs will examine whether different functional groups are more/less dependent on seafloor metrics at different spatial scales (i.e., data at 5-m, 25-m, 125-m, and 625-m spatial resolutions). While extracting the bathymetric derivatives at different spatial

resolutions may seem a trivial step, care should be taken as there can be significant variation in the resulting multi-scaled surfaces depending on how the data are processed (e.g., resampling then calculating derivatives vs. calculating derivatives with different neighborhood windows; Dolan, 2012).

Classification method

Increasingly large data volumes and an increasing requirement to produce non-biased, reproducible results are moving mapping away from manual interpretation (hand digitization; Battista et al., 2007). More recently, there has been significant effort in developing robust, quantitative methods to automate (or semi-automate) the seafloor mapping process (Diesing et al., 2014). Many of these efforts have the benefit of field-acquired validation (ground-truth) data with which to train and assess the accuracy of the classification model, which is broadly known as supervised classification (Hasan et al., 2012; Suka and Rooney, 2017). This approach, however, is not applicable to mapping in the PRIMNM or the MHI due to the remoteness and extensiveness of these areas, respectively. While there were spatially disparate survey data available in some locations (typically diver-based surveys in depths from 0–30 m), the spatial accuracy (e.g., ±50 m) and limited geographic extent of these data rendered them unsuitable for classifying high-resolution remote-sensing data (i.e., 5–10 m resolution multibeam data and 2-m resolution satellite imagery).

Therefore, it was determined this project must rely on unsupervised classification approaches (Calvert et al., 2015; Asner et al., 2017). Two machine-learning approaches that are broadly used—clustering and Object-Based Image Analysis (OBIA) segmentation—were considered for this analysis. In both circumstances, computational power is relied upon to identify boundaries and segment classes, at which point the interpreter attributes the predicted classes according to their expert opinion (e.g., boundary encompassing rugged bathymetry and high backscatter is associated with hard substrate). A **clustering-based approach** was employed for the following reasons: (1) clustering on both multibeam and WorldView-2 seafloor metrics proved consistently effective (Figure 1, Figure 2), (2) the process was straightforward (via Esri ArcGIS) and could be implemented by trained non-specialists, and (3) reasons (1) and (2) justified not purchasing OBIA software (Trimble eCognition). Trimble's eCognition is an expensive solution, and attempts to access a trial version of RSOBIA (OceanWise; Le Bas, 2016) for testing/comparison purposes were unsuccessful. Since this project aimed to adapt methods that could be *readily* employed by other scientists within PIFSC, the clustering-based approach in ArcGIS was deemed the best, most cost-effective approach currently available. Should these software packages become more accessible to PIFSC scientists in the future, it would be informative to compare results from an OBIA approach using eCognition or RSOBIA with results from the cluster-based approach using ArcGIS presented herein.

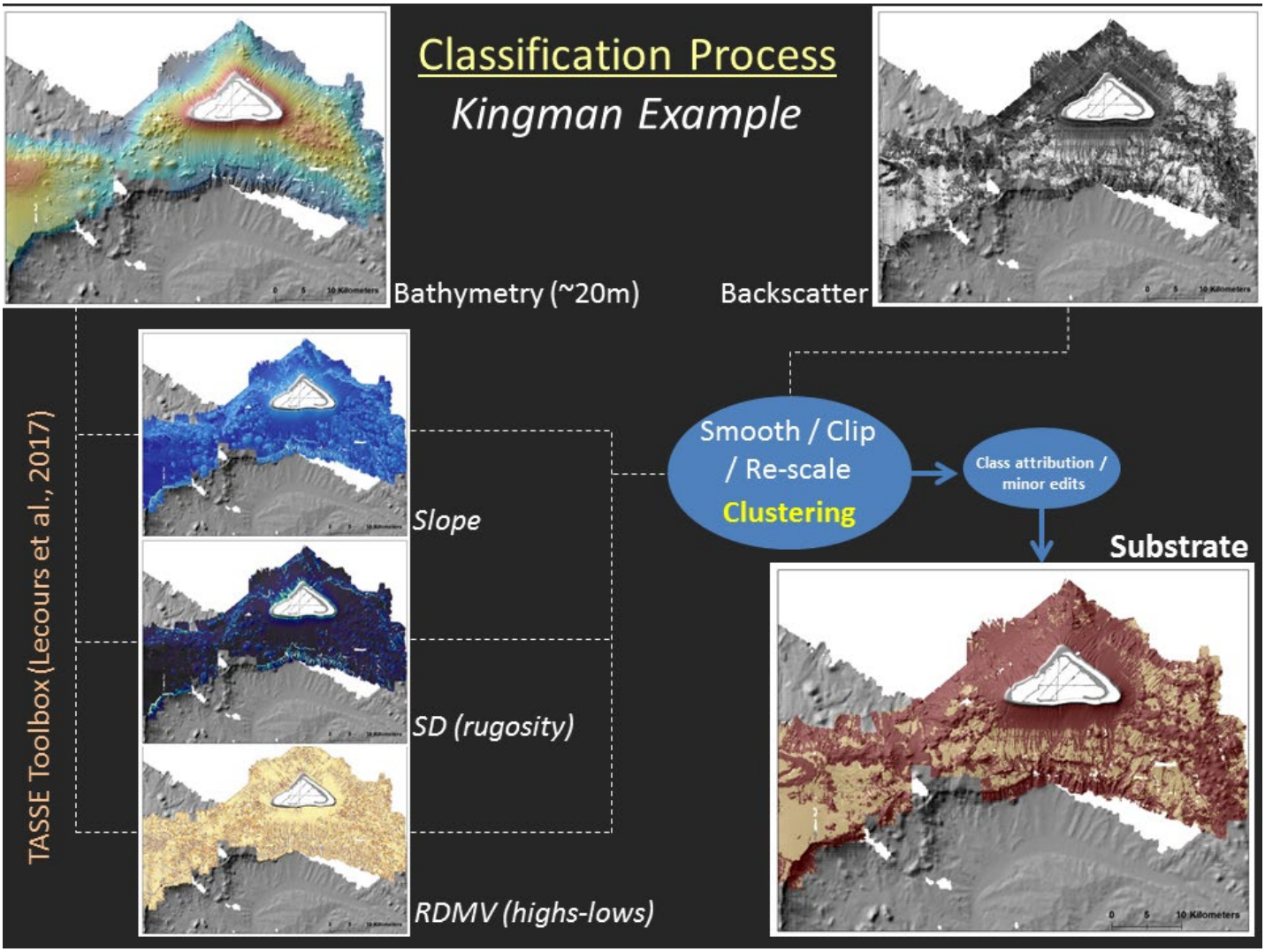

Figure 1: Diagram of simplified classification approach at Kingman Reef with multibeam data in deep water. Bathymetry derivatives and backscatter data were pre-processed and run through cluster analysis to identify groups (classes) of shared attributes. To produce the binary substrate map, the mapping practitioner then made an informed judgement in attributing the predicted substrate classes (e.g., classes 1–3 = hard, shown in dark red; 4–5 = soft, shown in tan). Where required, minor manual editing was undertaken to remove spurious predictions (e.g., nadir noise).

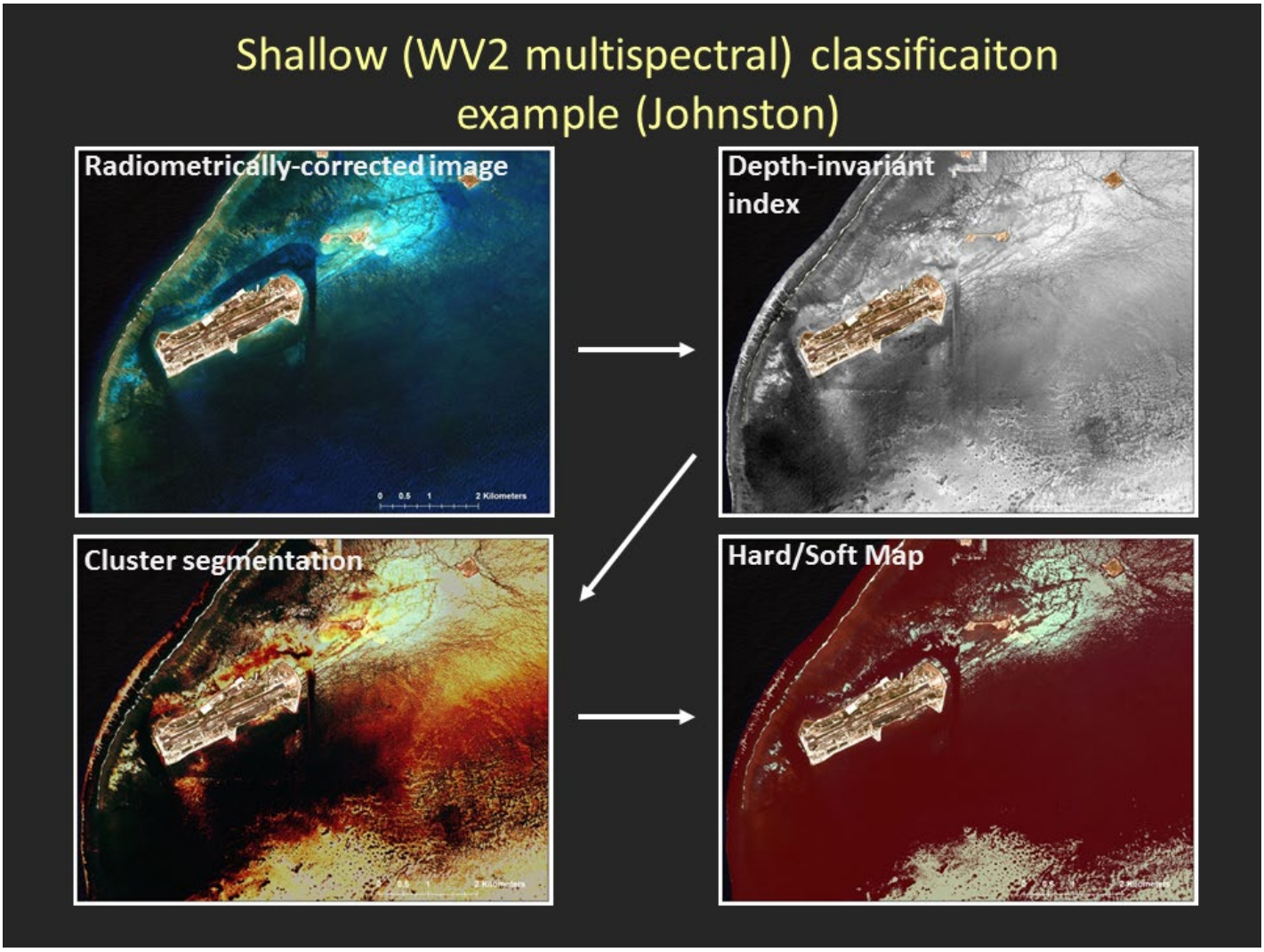

Figure 2: Diagram of simplified classification approach using WorldView-2 imagery for shallow-water areas at Johnston Atoll. Depth-invariant indices were prepared for the 5 spectral bands (10 band pairs) and run through cluster analysis to identify groups (classes) of shared attributes. To produce the binary substrate map, the mapping practitioner then made an informed judgement in attributing the predicted classes (e.g., classes 1–3 = hard, shown in dark red; 4–5 = soft, shown in tan). Where required, area-specific class attribution (typically because the effect of depth was not fully removed) and minor manual editing were undertaken to remove inaccurate predictions.

OBIA methods differ from clustering in that pixels are assessed at multiple resolutions and organized into homogeneous objects (Blaschke, 2010). Segmentation, therefore, is carried out on these objects rather than the underlying pixels. OBIA is not strictly an unsupervised approach, but the segmentation of image objects can be run in a supervised or unsupervised way (Dragut et al., 2010). The unsupervised approach is similar to clustering of raster pixels in that it segments the data into features of shared properties. There are many examples of researchers effectively applying OBIA methods with similar datasets and incorporating training data to supervise the approach (Phinn et al., 2012; Lecours et al., 2016b), but there is no consensus on whether OBIA or clustering provides a superior approach. The most commonly used OBIA algorithms are implemented through Trimble's eCognition, but there are increasing implementations (and approximations) available via Matlab, R, and Grass GIS. Two approximations of OBIA were tested within ArcGIS—RSOBIA and the Segment Mean Shift tool—and in both cases somewhat inferior results were achieved compared with clustering, and parameterization was less flexible.

Clustering is an unsupervised classification of patterns, which organizes data into sensible groupings (clusters; Jain, 2010). There are many clustering algorithms, of which K-means clustering is probably the most common. For this project, the ISO Cluster Unsupervised Classification tool within ESRI ArcGIS 10.X was employed. This tool employs a combination of ISODATA clustering (segregating groups of shared attributes) and maximum likelihood classification. This tool proved very useful as it can incorporate multiple covariates, therefore running the cluster analysis over multiple dimensions, and it runs very efficiently. ISODATA clustering is similar to K-means but differs in that the number of clusters can be adjusted automatically during model iterations. There is some indication that new clustering approaches may be more effective (e.g., DBscan and HDBSCAN), and exploring these algorithms (new options through ArcGIS Pro) could provide improvement to the method in the future.

The ISO cluster tool can be run in ArcGIS without specifying the number of clusters (predicted classes). However, it was determined that running clustering on WorldView-2 imagery at 15 classes and on multibeam data at 5 classes consistently gave good results based on visual interpretation and comparison of results from several different iterations that required minimal-to-no manipulation. Working with the multibeam data showed more potential for varying numbers of classes, particularly in the MHI where clustering at 2, 4, and 7 classes all gave valid results (predicting more or less hard areas) and showed variable sensitivities to bathymetry derivatives vs. backscatter. After running the ISO cluster tool, which assigns an integer to each raster class (e.g., 1–15), it was up to the interpreter to attribute these classes, choosing the threshold between hard and soft. This was the first subjective step in the mapping process and involved visually assessing the cluster results to ascertain if the boundaries matched the underlying data based on the practitioner's understanding of the seafloor environment (e.g., coral reef vs. unconsolidated sediment in shallow waters; or rugose, high-slope bedrock vs. mostly flat, low backscatter, sediment-filled depressions in deeper waters). This subjective step can be justified for two main reasons. First, the human eye is effective at discriminating between natural variations vs. noise. The value in this method was using the machine-learning and computational power to prepare accurate line-work, then the mapping practitioner made informed decisions on map attributions. Second, unless conducting supervised classification as was the case for the habitat maps previously prepared for West Hawaii by Suka and Rooney (2017), there were not many other options. Most remote-sensing scientists still use human perception to determine classification thresholds with computer-generated unattributed classes. This approach—including the subjective steps in the process—was initially developed for the West Hawaii area, the results of which were compared with the Suka and Rooney (2017) analysis before the method was applied elsewhere.

The protocol document provides further details on how the model covariates (e.g., bathymetric derivatives, backscatter, or depth-invariant indices) should be prepared to ensure satisfactory clustering results. This may include smoothing the covariates to remove spurious data/noise and normalizing the covariates to a common scale, so each covariate is equally weighted in the classification. The clustering process may also incorrectly classify noise in model covariates data (e.g., nadir noise in multibeam backscatter), and the protocol document describes several ways in which such errors are corrected.

Selection of seafloor covariates

This section describes the abiotic variables chosen for the semi-automated mapping of seafloor substrate, enabling clear discrimination between hard and soft classes. These covariates are different for deep vs. shallow-water environments.

Deep water environments

Data Sources

PIFSC's Ecosystem Sciences Division (ESD) previously conducted multibeam mapping surveys across the Pacific Islands Region using NOAA Ship *Hi`ialakai* and the R/V *AHI*. The *AHI* was deployed from the *Hi`ialakai* to perform multibeam surveys in shallow, nearshore waters with depths usually within ~10–300 m. Multibeam operations on the *Hi`ialakai* were conducted in deeper waters to complement and overlap the *AHI* shallow-water data, with depths exceeding 4500 m in some areas. Ship-based multibeam data were also acquired in specific areas around the Pacific by the Center for Coastal and Ocean Mapping/Joint Hydrographic Center (CCOM/JHC) at the University of New Hampshire and by UH's R/V *Kilo Moana* and following the expansion of the Pacific Marine National Monuments by Schmidt Ocean Institute's R/V *Falkor* and NOAA Ship *Okeanos Explorer*. While other surveys were completed across the region, these were the primary high-resolution multibeam data sources available for the PRIMNM. In the MHI, synthesized multibeam bathymetry and multibeam backscatter products, compiled from multiple multibeam surveys around the islands including data acquired by ESD, were prepared by scientists at UH. Bathymetric LiDAR data were also available.

Multibeam bathymetry

Bathymetry data provide depth and important information about the structure of the seafloor. Numerous derivatives, such as slope, rugosity, curvature, and aspect can be calculated from a continuous bathymetric surface. However, not all derivatives are useful for predicting hard vs. soft substrates (e.g., aspect), and many are spatially correlated and, therefore, do not provide unique information. Lecours et al. (2016a) provides a useful comparison of bathymetric derivatives (i.e., terrain attributes), and together with many of the publications previously listed, the efficacy of incorporating different bathymetric derivatives was investigated. Within the West Hawaii method development area (Figure 3), a range of derivatives were tested, including slope, rugosity (including standard deviation [SD], a measure of rugosity), curvature (plan, profile, and tangential), aspect (eastness and northness), bathymetric position index (BPI), vector ruggedness, relative distance to the mean (RDMV; measure of relative highs and lows), and bedform fill approach (relative highs). Derivatives that require qualitative decision making to be constructed were avoided so that the method could be applied consistently and efficiently. Several tools were used to calculate these metrics, including ESRI ArcGIS, ENVI, and the Orfeo toolbox. Individual and groups of derivatives were tested in a bootstrap fashion on both 2-m resolution Lidar data and 5-m resolution multibeam bathymetry data. To test the effectiveness of the variables, the clustering approach was run with individual and combinations of variables, then qualitatively (visually) assessed to determine how the classification results compared with the bathymetric data and true-color satellite imagery, specifically examining whether the classification accurately delimited the boundary between upstanding reef and sediment. A quantitative accuracy assessment of the cluster results was run with respect to available validation data previously compiled for the supervised map of the West Hawaii (Table 1; Suka and Rooney, 2017). The test results were also visually compared to the map for West Hawaii and to the previously compiled (and manually drawn from satellite data) shallow habitat maps produced by Battista et al. (2007).

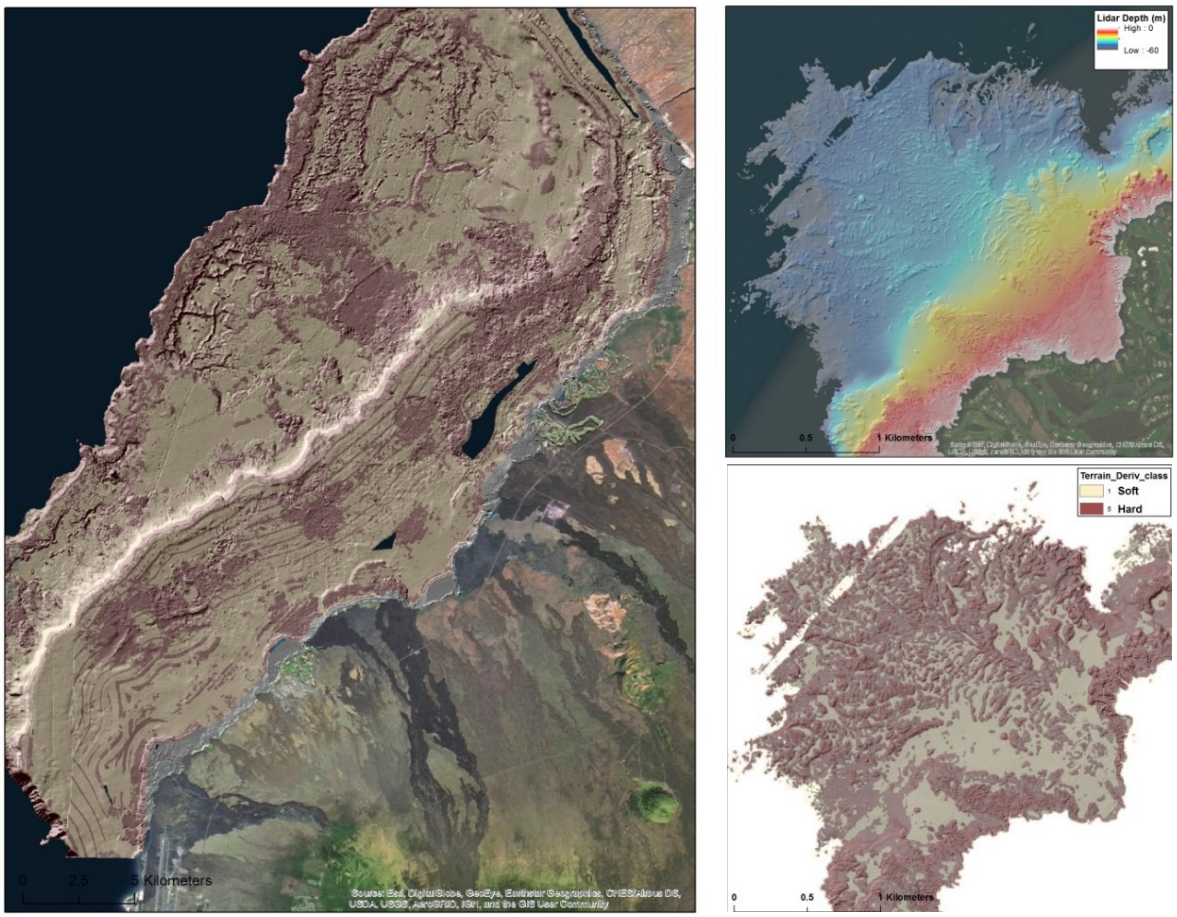

Figure 3: West Hawaii method development area. Left panel shows draft substrate map based on multibeam data (5-m gridded resolution), with narrow ribbon of Lidar hillshade-bathymetry data shown along the coast in grayscale. Top right panel shows higher resolution Lidar bathymetry (2-m gridded resolution), and the bottom right panel shows the test substrate map (based on bathymetric derivatives alone).

Ideally, one would conduct a sensitivity analysis using boosted regression trees to examine which derivative combinations best predict classes observed in validation data. Though the data here were significantly better with respect to geographic accuracy and video interpretation than elsewhere in the MHI, the ground-truth data were sampled at a limited depth range (< 100 m), and therefore covered only a fraction of the areas which had both multibeam in deeper water and Lidar in shallower waters. The accuracy assessment was informative and showed reasonable results, but should not be viewed as the absolute accuracy of each method (Table 1). The supervised method of Suka and Rooney (2017) gave highest (87%) accuracy because it used validation data in its predictions. The multibeam results were interpreted to be artificially low because the backscatter data were noisy in that region and were compiled from multiple surveys. By normalizing to common scale and removing spuriously small/noisy predicted polygons, the accuracy could be improved.

Ultimately, and with the observation that fewer derivatives produced the most consistently accurate results, the following were selected as the bathymetric derivatives to be incorporated as covariates in the clustering: slope, SD, and RDMV. These three combined provided several unique indicators of hardness, including: (1) higher slopes associated with hard substrates (the angle of repose was at 15–25° in wet unconsolidated sands, so greater slopes were likely to comprise outcrops of consolidated material), (2) more rugose seafloor was associated with hard substrates (e.g., complex reef topography vs. smooth sediment-covered lows, and (3) relative bathymetric highs were more likely hardbottom and relative lows were more likely soft (e.g.,

upstanding lava flow or reef platform). This final assumption was not valid in some scenarios, such as gullies, where coarser material was found due to high-energy gravity currents. In such cases, predictions from backscatter data were relied upon to supersede the morphology. The fill metric was calculated from multibeam data to capture relative highs and provided marginal improvements in predicting hard substrates. The assumption was that the reduced spatial resolution of the data inhibited the clear delimiting discrete features like reef spurs, but also prediction accuracy was reduced due to different morphologies (e.g., eroded fossil reef terraces).

Based on visual comparison to the underlying data, it was found that employing these three aforementioned derivatives alone, calculated from 2-m Lidar data, did a better, more consistent job of discriminating between hard and soft substrates than the benthic habitat product produced by NCCOS based on manual interpretation of satellite data (Battista et al., 2007).

Table 1. Accuracy assessment comparing the results from five different approaches used to predict seafloor substrates for West Hawaii.

Layer	Total Samples	Correct Prediction	Overall Accuracy
Suka and Rooney (2017): supervised classification	1802	1570	0.871
Lidar Bathymetry: Tier 4 (Slope, SD, RDMV, Fill)	1802	1384	0.768
Lidar Bathymetry: Tier 3 (Slope, SD, RDMV)	1802	1375	0.763
Multibeam Bathymetry (Slope, SD, RDMV) and Backscatter	799	590	0.738
Satellite-derived bathymetry	1140	732	0.642

Multibeam backscatter

Backscatter data provide a useful proxy for the texture and composition of the seafloor, as high backscatter intensities likely reflect harder substrate. Backscatter frequently scores as one of the most important seafloor metrics in benthic habitat variation. Therefore, the multibeam backscatter data were used together with the three bathymetric derivatives in the clustering analysis. It was determined that there was sometimes value in rescaling the backscatter data to exert either more or less influence on the final map product. However, when applying this approach across the PRIMNM and the MHI, there was only one instance where it seemed required to increase the relative importance of the backscatter data in the clustering (lower-resolution data used in the analysis for the Line Islands). This was encouraging in that it simplified the modelling approach, not requiring practitioners to tinker with too many parameters. Backscatter values, however, are not linearly related to seafloor hardness (i.e., backscatter is insensitive to changes in the seafloor where content of coarser materials is greater than ~10–20%). It is therefore recommended that backscatter data alone should not be used to predict presence of hard substrate.

Another way to include and potentially amplify the characteristics of backscatter data is to also incorporate backscatter derivatives in the modelling (similar to bathymetry). One such approach that has had some success is calculating textural metrics from the backscatter data, and the gray-level co-occurrence matrix (GLCM) is probably the most common implementation (Samsudin and Hasan, 2017). A range of GLCM textures were explored, such as entropy, correlation with backscatter data (as well as bathymetry) using a number of software implementations (Orfeo toolbox, ENVI, Grass GIS), but ultimately these metrics were not included in the final

modelling approach. While GLCM seemed to provide minor improvement within some individual survey areas, the results were inconsistent between datasets and different survey areas. This instability with GLCH textures seemed to result from inconsistencies in software implementation and excessive sensitivity to noise in the data. This required too much area-specific tuning of the method, which contradicted one of the primary objectives to apply mapping methodologies consistently across multiple environments.

Shallow water environments

The seafloor metrics derived from WorldView-2 imagery in shallow waters were classified in a way similar to the multibeam covariates used in deeper waters (clustering), but the covariates themselves were of a different type. Coral reef mapping with satellite data is a relatively established practice (Mumby et al., 2004); however, most studies rely on field-acquired ground-truth data (e.g., Phinn et al., 2012), which were not available for remote or large spatial areas. Even with validation data though, these studies struggled with the effect of the water column, and final classifications included prefixes such as deep or shallow for sand, lagoon, and reef. There have been efforts to map coral extent over multiple atolls using unsupervised approaches and simple true-color (RGB) images from multispectral data (e.g., Asner et al., 2017), and while results were promising, they too found that water depth inhibits the accuracy of predictions.

After literature review and experimentation in West Hawaii and PRIMNM, it was determined that using depth-invariant seafloor indices (Mumby and Edwards, 2000) provided the best available approach to discriminate hard vs. soft substrates without validation data. This was due to the success of using only the seafloor indices to resolve hard vs. soft areas, the lack validation data, and the relatively poor performance when using satellite-derived bathymetry (attempt to emulate the process with deeper multibeam data) as described below in the *Satellite-derived bathymetry* section.

Depth-invariant seafloor indices

In attempting to discriminate seafloor character using satellite imagery, the ultimate objective was to resolve bottom reflectance. However, while various radiometric corrections improve the measure of reflectance at the sea-surface, resolving bottom reflectance was not feasible due to the number of unknown variables (e.g., water depth at each pixel). Instead, calculating depth-invariant seafloor indices provided an effective proxy for bottom reflectance by comparing the attenuation coefficients among different spectral bands (Mumby and Edwards, 2000). With WorldView-2 imagery for example, five spectral bands in the visible wavelength were used to derive ten potential depth-invariant indices or band pairs (e.g., coastal-blue, coastal-green, coastal-yellow).

These depth-invariant indices could then be visually assessed and compared with true color imagery together with any available ground-truth data to examine which band pairs best reflect the natural variation at seafloor. Preliminary clustering on these band pairs was run to confirm that the depth-invariant indices were resolving the visually detected substrate boundaries. Incorporating multiple band pairs into the clustering was tested, but more often than not, clustering on one band pair (e.g., blue-green) performed at least as well as combining multiple band pairs; hence, the simpler option was chosen.

This analysis (clustering on WorldView-2 depth-invariant band pairs) has only been applied across the PRIMNM to date. There was not one band pair that worked best everywhere, but there were some consistencies. Band pairs 2 (coastal-green) and 5 (blue-green) were commonly incorporated into the final clustering. Also, the depth-invariant indices did not always entirely remove the effect of water depth. For this reason, establishing

different rule sets with the clustered classes to attribute hard and soft substrates in different parts of the study area was required. This was accomplished using a mask and was typically a fairly quick process.

The visual results of this classification appear effective, and when compared against available validation data within the PRIMNM, the accuracy produced 95–99% agreement. However, this analysis should not be treated as conclusive, because the validation data were derived from dive-samples with poorly constrained geographic certainty.

Table 2: Preliminary accuracy assessment comparing classification results to validation data based on survey locations in the PRIMNM.

Island	Total Samples	Correct Prediction	Overall Accuracy
Kingman	173	168	0.970
Jarvis	308	293	0.950
Palmyra	216	215	0.995
Howland	83	81	0.970

Atolls within the PRIMNM are ideal environments in which to conduct this unsupervised classification due to the shallower depths, clarity of water, and consistency of the water column across the study area. Despite this, the classification approach struggled in two settings, intertidal areas where limestone pavement was strongly bleached (spectrally bright areas misclassified as soft) and within lagoon settings (i.e., Palmyra) where sediments comprised significant silt fraction (spectrally dark areas misclassified as hard substrate). Turbidity in the water column or significant variation in the spectral character of sediment resulted in incorrect classifications.

Satellite-derived bathymetry

An initial objective in shallow settings was to employ a similar set of model covariates within the cluster analysis as were being developed in deeper waters using multibeam data. This would entail using satellite-derived bathymetry to extract derivatives (e.g., slope, rugosity), then using the satellite imagery as ~equivalent to the multibeam backscatter. This was not possible since the satellite-derived bathymetry consistently suffered from a random speckle-noise artifact. This artifact was observed in all data processed by ESD, and also by other groups (e.g., NCCOS). The satellite-derived bathymetry also had inversion issues. These issues were more commonly observed in shallow areas (e.g., bathymetric lows with high reflectance were incorrectly presented as bathymetric highs).

Results

The final shallow and deep-water substrate maps for the PRIMNM are being included in the *Coral Reef Ecosystem Monitoring Report for the PRIMNM* (in development at the time this report was published), and the shallow substrate maps were provided to the Pacific Islands Regional Office to support management of ESA-listed corals and related critical habitat activities. In the MHI, substrate maps, together with multi-scale bathymetric derivatives, are being used as environmental covariates within the Atlantis ecosystem model (i.e., GAMS) and to inform efforts to improve bottomfish stock assessments.

Pacific Remote Islands Marine National Monument

Deep water environments

Within the PRIMNM deep-water areas, there was variable accuracy in the predicted substrates that was largely dependent on the quality of the input data. Backscatter data quality varied most significantly.

Substrate mapping using multibeam data for Palmyra, Kingman, and Jarvis was based on relatively adequate quality bathymetry and backscatter data (20–60 m resolution bathymetry) acquired by the *Hi'ialakai,* with minimal noise and effective normalization of backscatter intensities between lines and across sites (Figure 1, Figure 4). This improved the quality of predictions, and substrate classes were appropriately associated with both key bathymetric (e.g., slope/rugosity) and backscatter features (e.g., gullies comprising relatively coarse material, and high amplitude returns associated with ferromanganese crusts; Figure 5).

Hard substrates at these three locations were found to be linked to high slope, high complexity areas associated with fossil reefs, ridges, and escarpments, gullies, volcanic cones, and ferromanganese crusts. Fossil coral reef terraces observed down to ~300 m are associated with previously lower relative sea levels. There are also multiple deep terraces (and bounding escarpments) associated with the phased development of the oceanic plateau and the complex of seamounts in the Line Islands (Figure 4, Figure 6). There are commonly submarine channels or gullies radiating out from the seamounts. These gullies host relatively strong density/gravity currents which either expose underlying bedrock through erosion (or preventing deposition), or deposit relatively coarse sediments compared to areas outside of gullies. The flat terraces commonly support deposition and preservation of unconsolidated sediments, whereas the escarpments and ridges are associated with hard substrate. However, this region is also known to host deep-sea minerals, and indeed ferromanganese crusts associated with a complex of volcanic cones were observed on flanks of Kingman Reef (Figure 5). Hard substrates associated with these crusts were found on steep slopes. Hard substrates became dominant above ~15–20° within the PRIMNM (and in the MHI) and above ~30° in most other places (e.g., Friedman and Robinson, 2002). This is consistent with the angle of repose for saturated sands observed in laboratory experiments, as well as empirically. However, hard substrate was also observed on relatively low-slope areas, thereby highlighting the importance of backscatter data to distinguish these hard areas from low-slope areas with unconsolidated sediments.

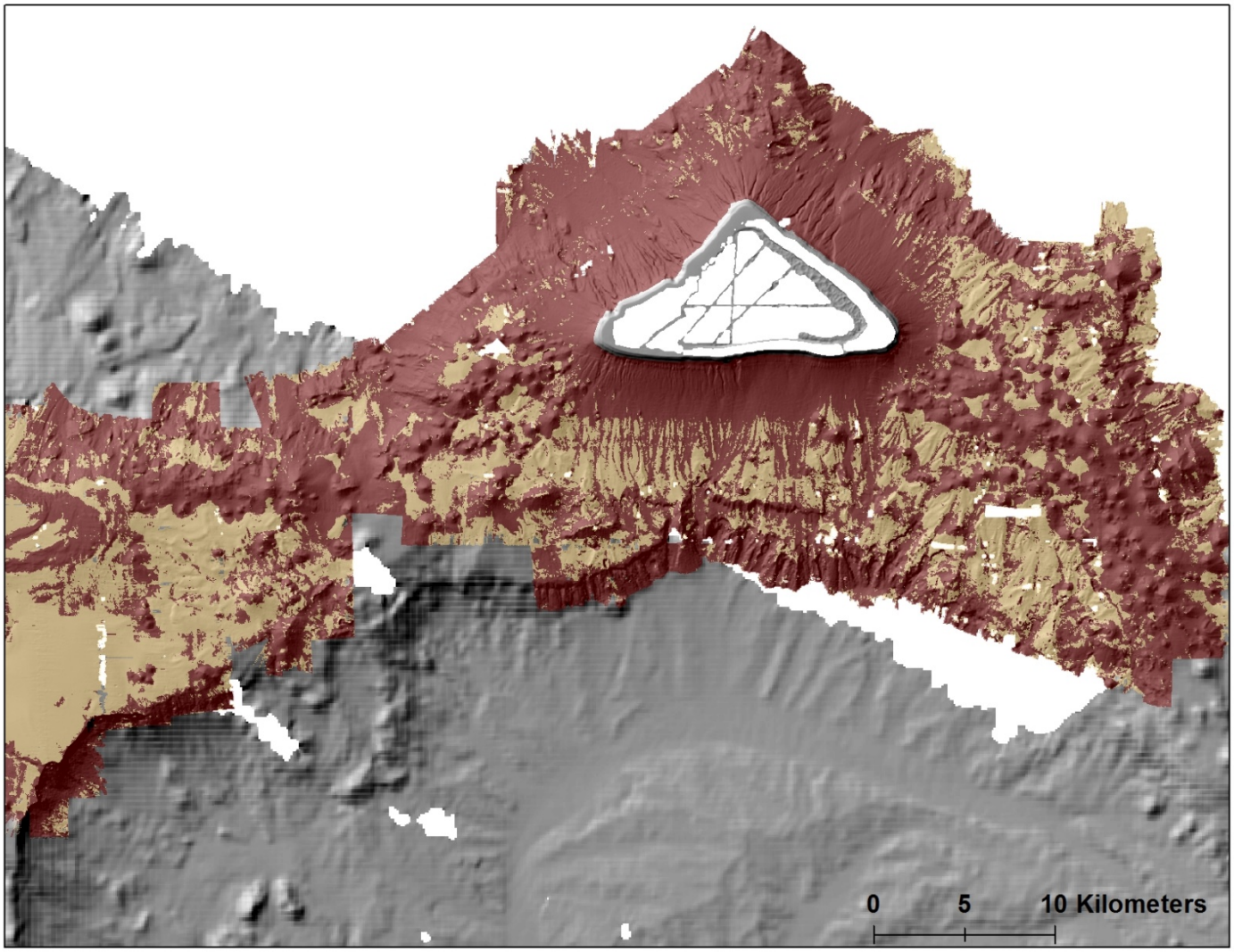

Figure 4: Substrate map for Kingman Reef based on sufficient quality bathymetry and backscatter data. Hard substrate is shown in dark red and soft substrate is shown in tan. The grayscale hillshade is based on CCOM/JHC bathymetry (see Figure 10).

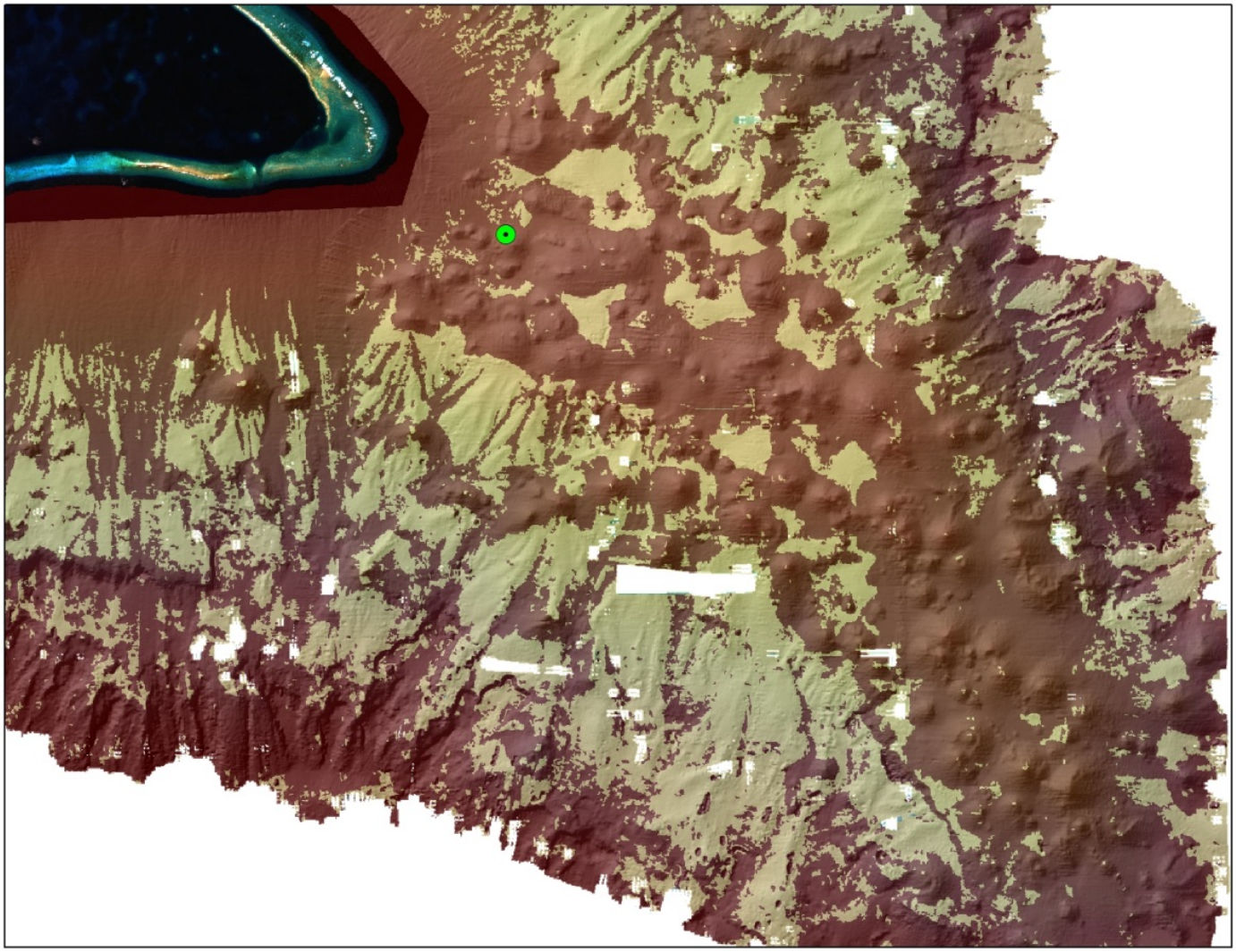

Figure 5: Zoomed-in view of the southeast flank of Kingman Reef showing seafloor substrates overlaid on hillshade bathymetry. Hard substrates are shown in dark red/brown and soft substrates are shown in tan. Predicted hard substrates were associated with steep slopes and escarpments, gullies (where gravity currents expose bedrock), and with apparent ferromanganese crusts associated with field of volcanic cones. Sample data acquired by the R/V *Okeanos Explorer* (green point) confirm mineral crust. Predicted hard substrates associated with crusts was based on high backscatter values rather than morphological character.

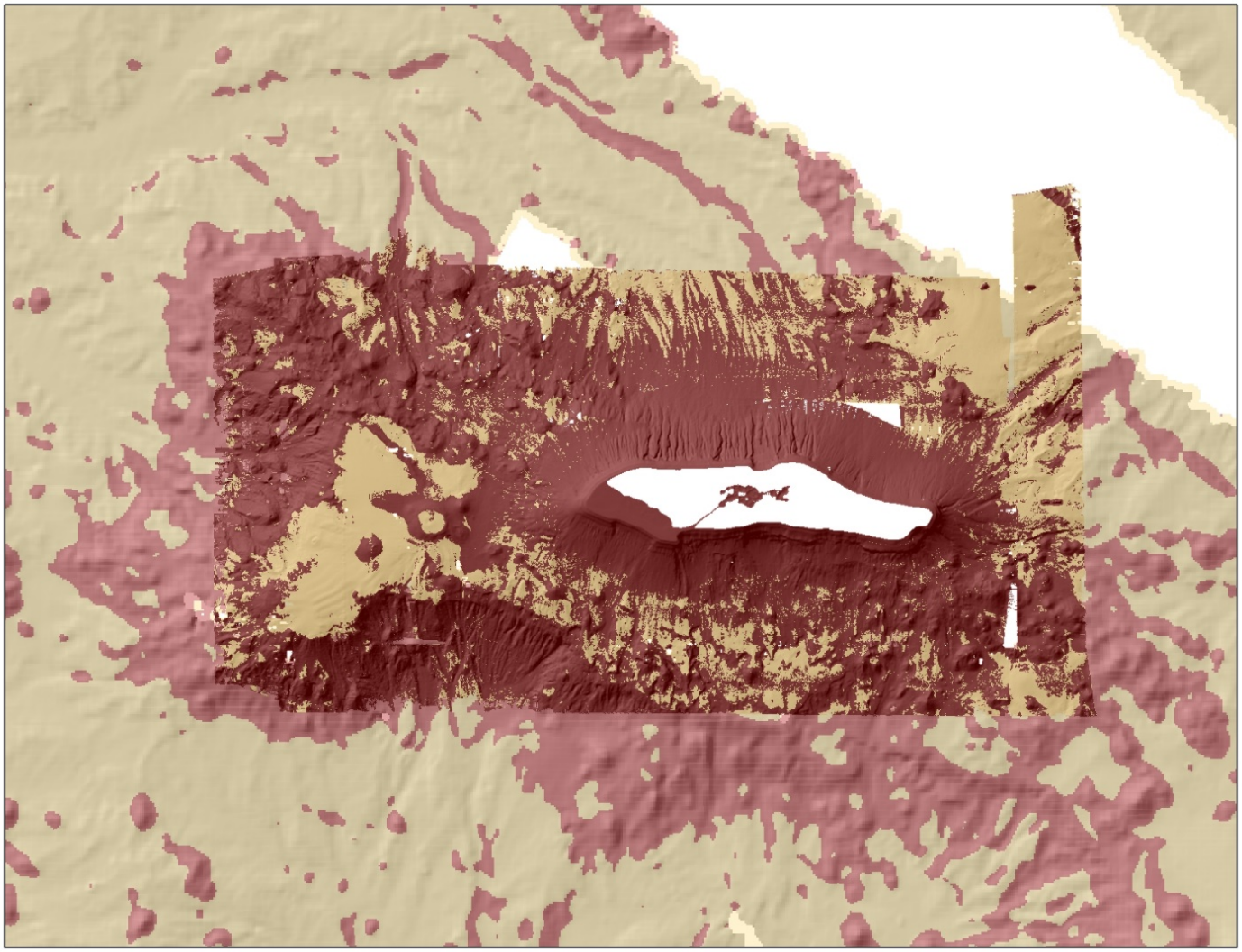

Figure 6: Substrate maps for Palmyra Atoll derived from multibeam data. Hard substrate is shown in dark red and soft substrate is shown in tan. Substrates derived from ESD-acquired multibeam data shown in darker shades (40-m resolution grid) are overlaid on substrates derived from coarser-resolution multibeam data acquired by CCOM/JHC shown in lighter shades (100-m resolution grid).

Relatively poor quality multibeam data were acquired at Howland, Baker, Wake, and to a lesser extent, Johnston (Figure 7). The backscatter data in particular suffered from excessive noise from heave, nadir, along-track gain errors, and poor normalization between survey lines. For substrate mapping, noisy variations in the backscatter data had relatively little influence on the final substrate predictions, as the cluster analysis found fewer natural groupings between the bathymetry and backscatter data resulting in under-predicted hard substrates (Figure 7).

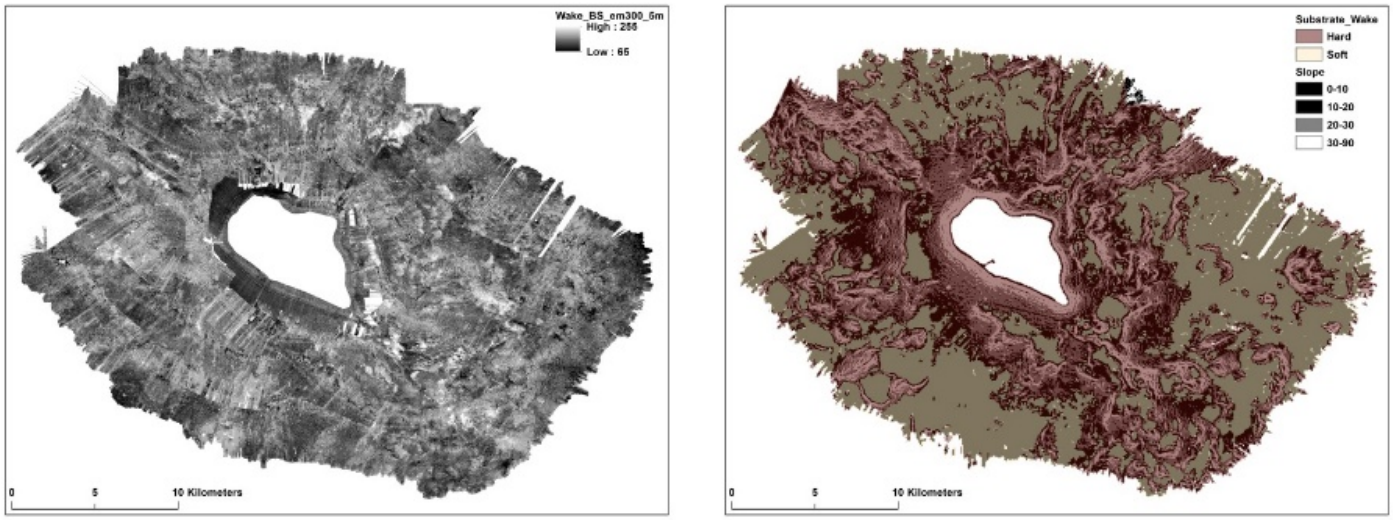

Figure 7: The final predicted substrate map for Wake (*right*) was based primarily on bathymetric derivatives because of the issues with the multibeam backscatter data available for Wake (*left*). The substrate map is overlaid on the contoured slope map, demonstrating that hard substrates were primarily predicted in high slope/high rugosity locations.

Mid-water data gaps

Within the PRIMNM areas, analyses were conducted using all *Hi'ialakai*-acquired multibeam datasets. However, the *AHI*-acquired multibeam data were not used. While the *AHI*-acquired bathymetry data were of sufficient quality, the backscatter data consistently were not, with variation in backscatter intensity primarily due to ship track and across-track gain errors (Figure 8). In other words, the backscatter data poorly reflect substrate variation, which inhibits the clustering process. It was, therefore, decided substrate maps would not be produced on bathymetric derivatives alone if backscatter data could not be incorporated in the analysis. One exception to this occurred at Palmyra where the shallow *AHI* bathymetry was used to predict substrates within the nearshore deeper lagoons (beyond the limit of the satellite imagery that was used to predict substrates for the shallower areas surrounding the lagoons).

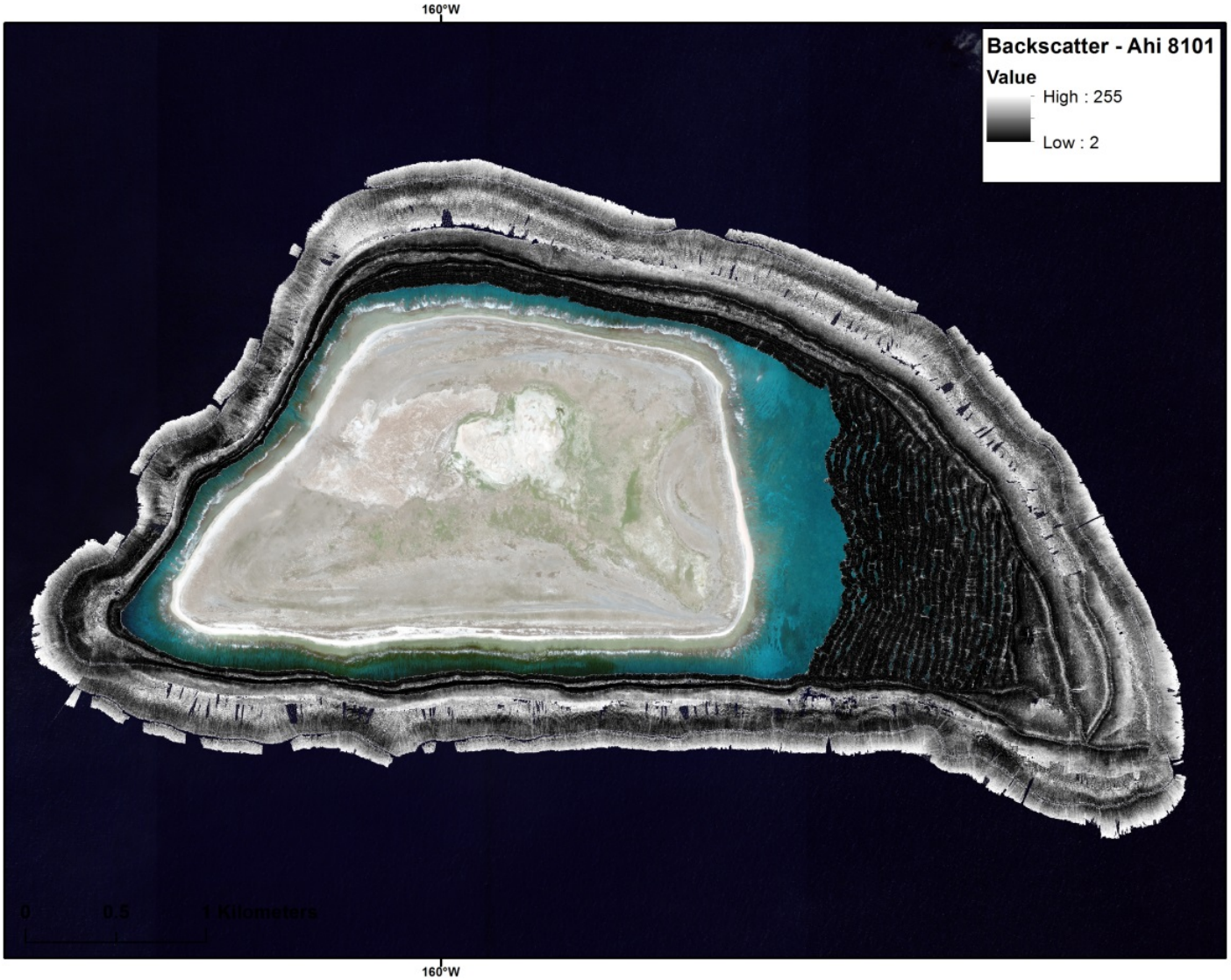

Figure 8: Map of Jarvis showing *AHI*-acquired backscatter data of insufficient quality for cluster analysis overlaid on a WorldView-2 image.

Because the *AHI* data were not incorporated in the substrate mapping for any of the PRIMNM islands, there is commonly a gap between the extent of the shallow satellite-based classification and the deeper multibeam-based classification (Figure 9). This gap extends between ~30 m and 250 m depths and a covers a geographic width (plan view) of 100–800 m. This is unfortunate, as this is a key depth range relevant to potential mesophotic corals (Rooney et al., 2010).

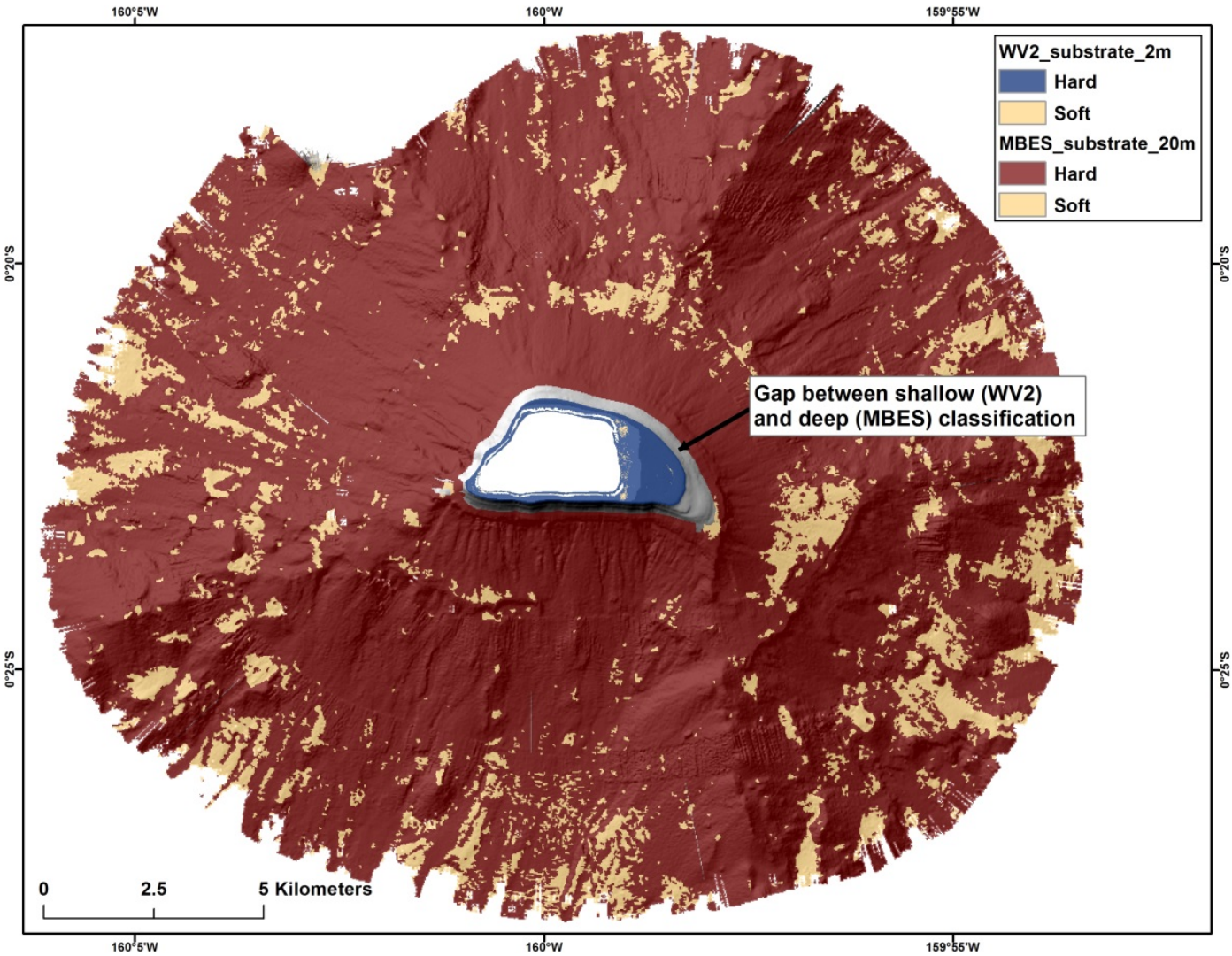

Figure 9: An example of the gap between shallow and deep-water substrate maps for Johnston Atoll, where insufficient quality shallow multibeam backscatter was excluded from the analysis.

Third-party data

Third-party multibeam data acquired by CCOM/JHC were used to cover broader areas within the PRIMNM boundary around the Line Islands (Kingman and Palmyra). These data, though gridded to 100 m, were of similar quality to the *Hi'ialakai* data for these areas and produced acceptable predicted substrates (Figure 10). There were further third-party data available, acquired by the *Okeanos Explorer*, *Falkor,* and the *Kilo Moana*, though it was determined that only data available as gridded surfaces (i.e., data that were already processed and ready for analysis) would be incorporated in the analysis, as long as no further processing or preparation of the data was required.

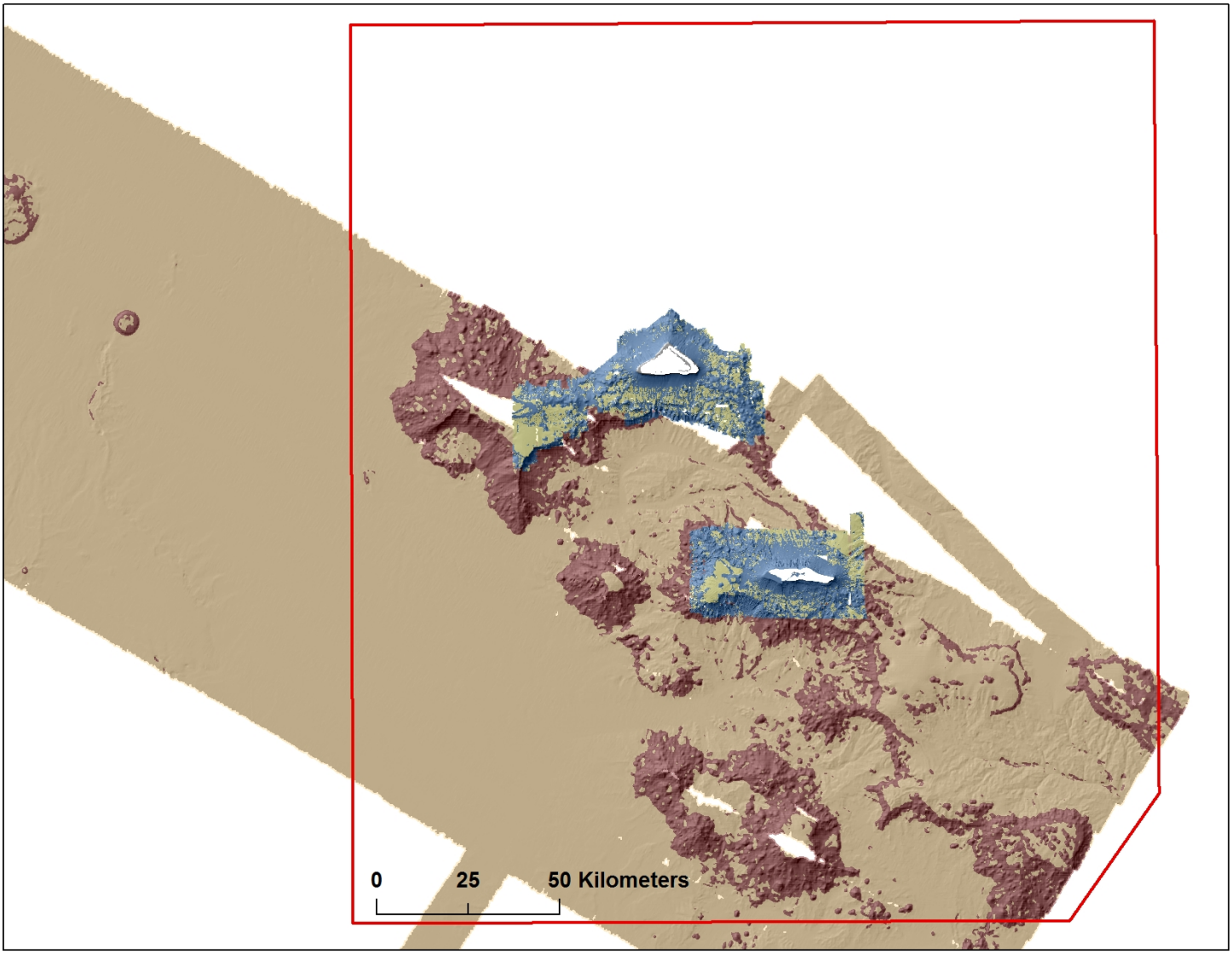

Figure 10: Substrate map based on multibeam data for the Kingman Reef and Palmyra Atoll unit of the PRIMNM (boundary shown in red). Substrates derived from ESD-acquired multibeam data (40-m resolution grid) shown in blue/tan (hard/soft, respectively) are overlaid on substrates derived from coarser-resolution (100-m grid) data acquired by CCOM/JHC (dark red = hard, tan = soft).

Shallow water environments

Shallow-water substrate mapping in the PRIMNM with multispectral satellite imagery using cluster analysis on depth-invariant seafloor indices proved very successful (Table 2, Figure 11) as detailed delineations were captured between hard and soft substrates (Figure 12). The clustering-based method was applied consistently across the islands (Figure 2), though there was some variation in the level of post-clustering cleaning and area-specific class attribution. The latter aspect was time consuming as it entailed assessing the accuracy of the initial cluster predictions, establishing area-specific rule sets, producing a mask to apply the different rule sets, and lastly compiling the final substrate map. Setting different rule sets was also necessary if multiple WorldView-2 image tiles were required to cover the entire reef area at an island (Figure 11).

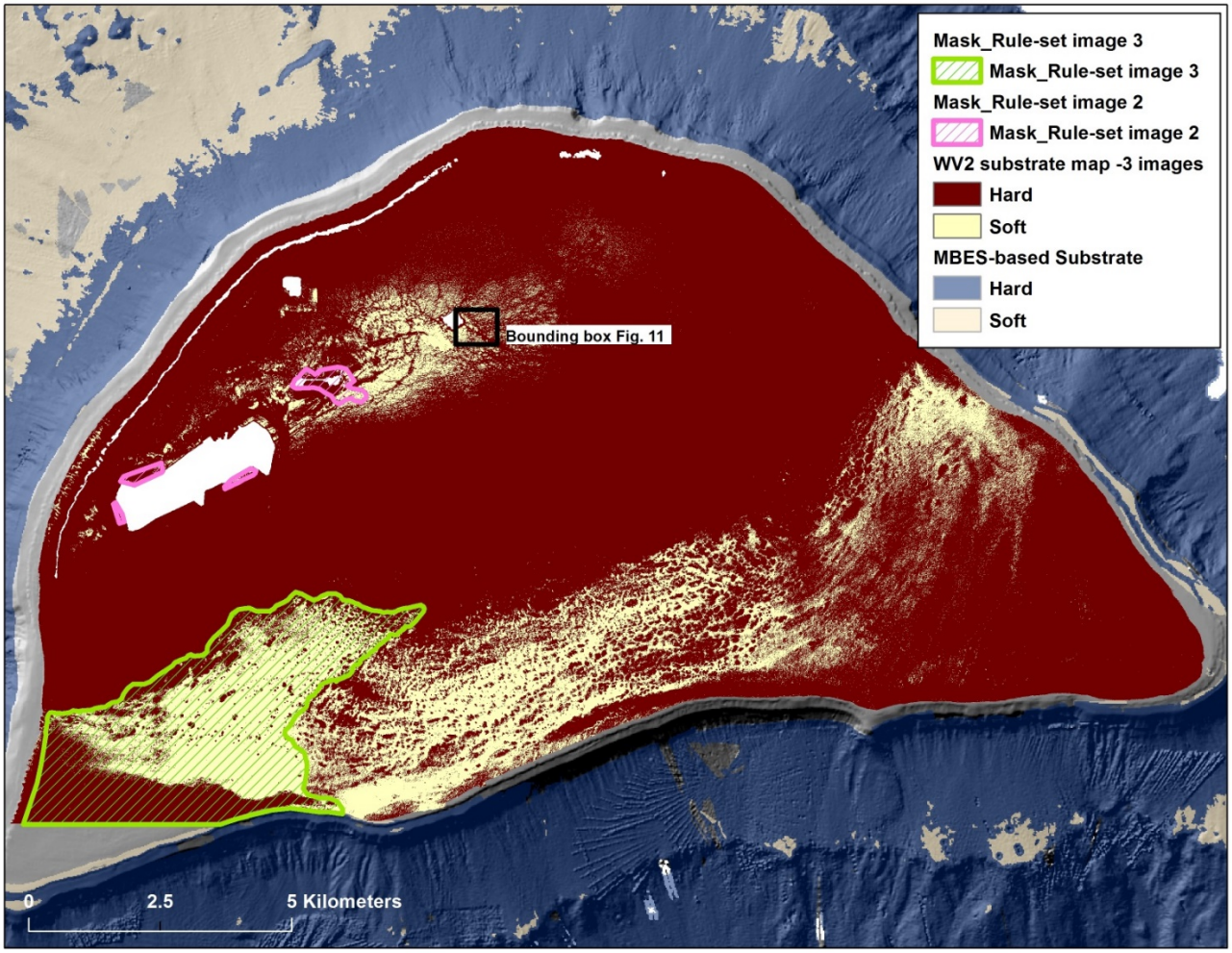

Figure 11: Shallow-water substrate map for Johnston Atoll based on three WorldView-2 images. Hard and soft class attributes were determined from 15 classes for each image. Green and pink polygons show areas where different class attribute rule sets were applied. Once hard and soft areas were correctly attributed from the cluster results for each of three images, the results were combined to form the final substrate map. The deep-water substrate map is shown in blue/tan, and the gap between the shallow and deep substrate maps is shown in grey.

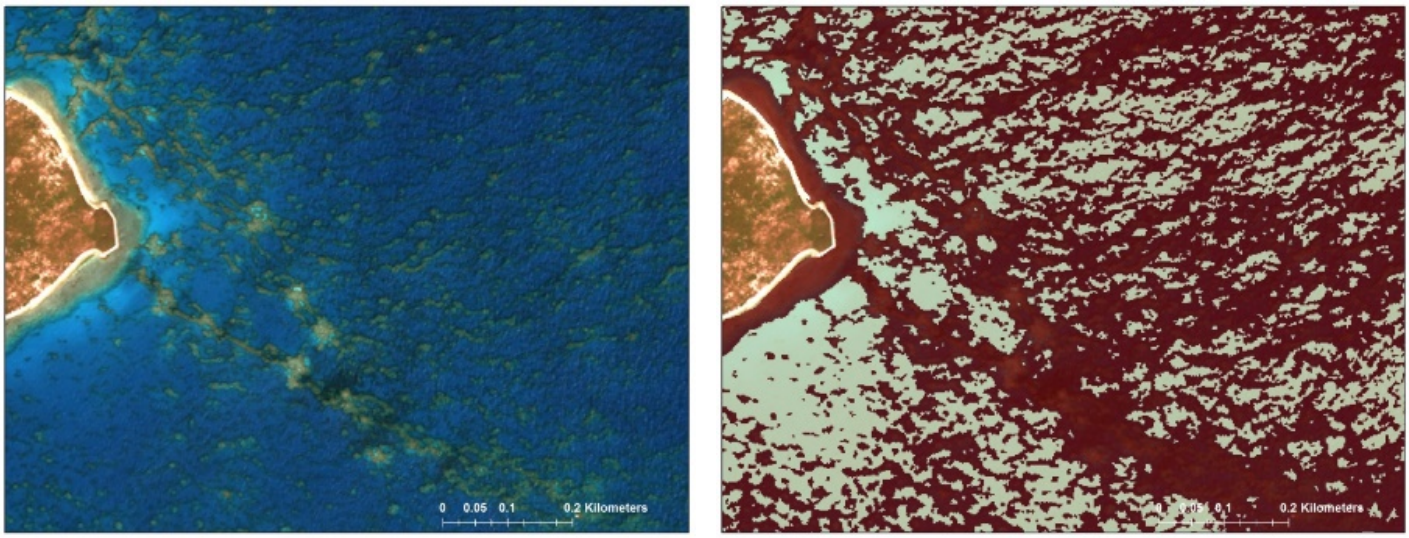

Figure 12: Example of detailed, accurate seafloor characterization at Johnston Atoll. Left panel shows the WorldView-2 image, with the final substrate map shown in the right panel (dark red = hard, tan = soft).

There were two circumstances where cluster-based delineations were not usable. In the first case, silty (and dark) sediments, such as those found in the shallow lagoon at Palmyra, confounded the classification and led to soft sediments being misclassified as hard substrate. Manual edits were made to correct the substrate map (Figure 13). In the second case, intertidal areas where limestone pavement is strongly bleached and spectrally bright led to the misclassification of these areas as soft substrate. This circumstance was most pronounced at Jarvis. The intertidal areas were not of strong interest to PIFSC, and thus removed from the analysis. These cases are indicative of limitations to the cluster-based delineation method associated with mapping submarine environments where there is turbidity, variation in the chemical and biological properties of the water column, or diverse sediment types with different spectral qualities.

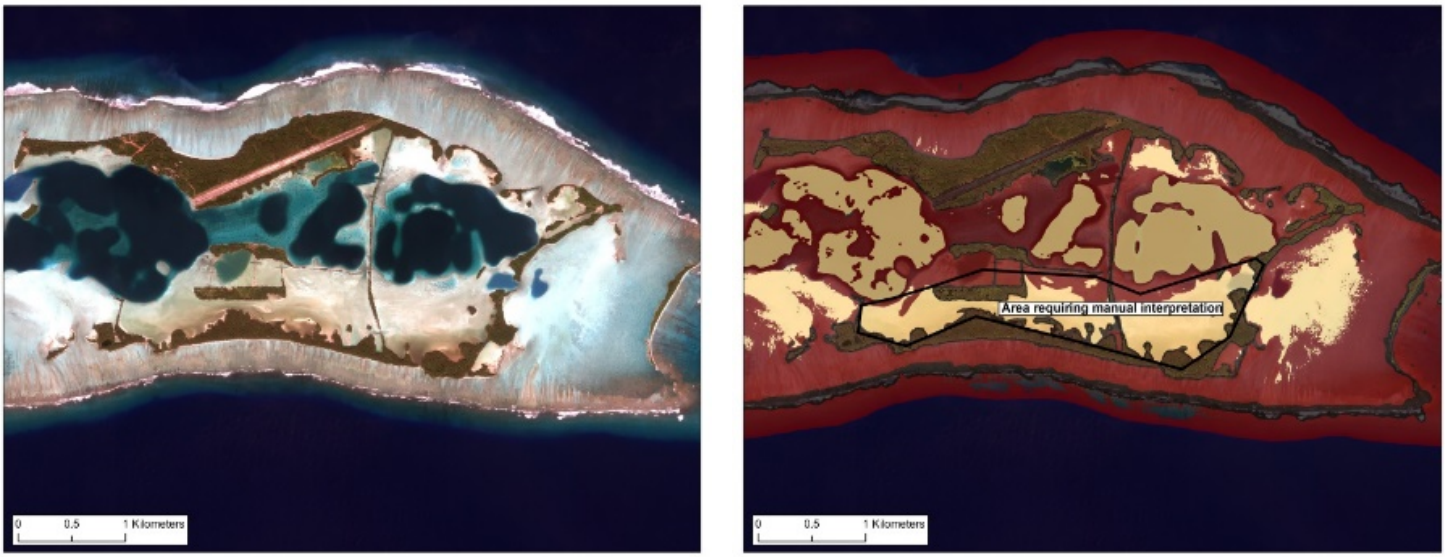

Figure 13: Silty sediment areas confound clustering of WorldView-2 depth-invariant indices at Palmyra, requiring manual interpretation. The WorldView-2 image for Palmyra is shown in the left panel, and the satellite-based substrates indicating problematic silty sediment areas are shown in the right panel.

Preserve initial clustering result

At times it can be useful to preserve the initial clustering result, at it includes a gradient of classes from hard to soft. While this was true to an extent for the multibeam-based classification where multiple covariates were used in the cluster analysis, it was especially useful to preserve the initial result from clustering of WorldView-2 depth-invariant indices. The clustering on just one covariate (depth-invariant index), typically to 15 classes, consistently revealed the gradient between hard and soft end-members (Figure 14). On forereef areas, for example, there is commonly variable amounts/thicknesses of unconsolidated sediment over reef or limestone pavement. To produce a substrate map, a threshold must be determined to delineate hard and soft areas. However, viewing the initial cluster result may be independently useful for coral reef scientists to examine where there is relatively more or less sediment in the system (keeping in mind sediment distribution may change over tidal, storm, and seasonal cycles).

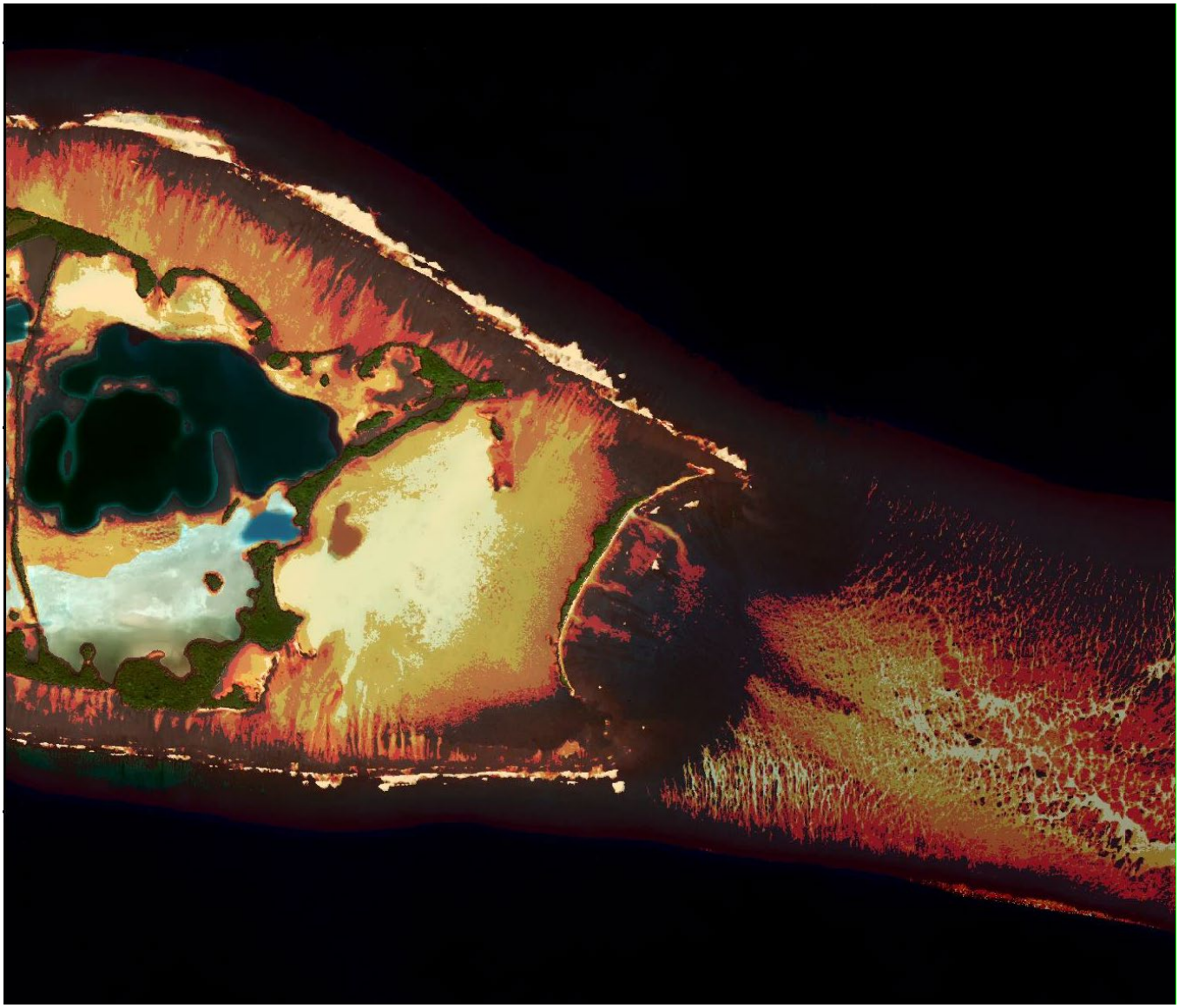

Figure 14: Initial clustering results overlaid on WorldView-2 image at Palmyra Atoll, demonstrating how multiple classes preserve useful information about the gradient between hard and soft end-members. Here, lighter classes indicate greater proportion of soft sediments at seafloor. It is therefore important to preserve this result prior to combining classes to form the final binary substrate map. Areas in the lagoon where manual edits were made were clipped from the map (*center left*).

Main Hawaiian Islands

The impetus for substrate mapping in the MHI was twofold: to support improved bottomfish assessments and to provide environmental seafloor covariates for integrated ecosystem modeling. In both cases, researchers required substrate information for the extent of the MHI offshore areas between 0 m and 500 m depths. The multibeam-based method used to predict substrates in the PRIMNM was also applied in the MHI. All classifications for this analysis were based on synthesized multibeam bathymetry (Figure 15) and multibeam backscatter data (Figure 16) at 5-m spatial resolution; however, these preliminary results could be significantly improved by, for example, using Lidar data to predict substrates in shallow waters. As such, frequently there was a gap in data coverage in the MHI substrate maps between ~0 m and 30 m in areas. Despite the limitations described below, these island wide, high-resolution multibeam datasets were the best available and were successfully used to produce a substrate map for the MHI (Figure 17).

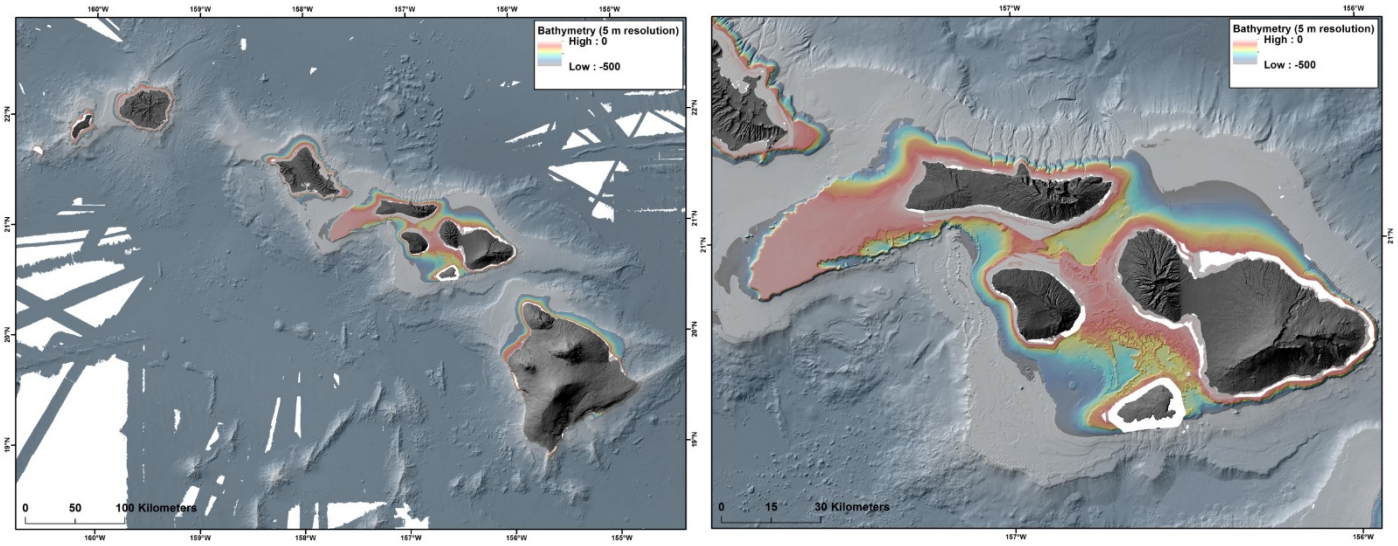

Figure 15: Bathymetry synthesis data compiled by UH for the MHI (*left*) and zoomed in to the Maui Nui region of the MHI (*right*); 5-m resolution synthesis (0–500 m depths) shown in rainbow color-scale and 50-m resolution synthesis (~0–6,000 m) shown in blue-scale (with hillshade).

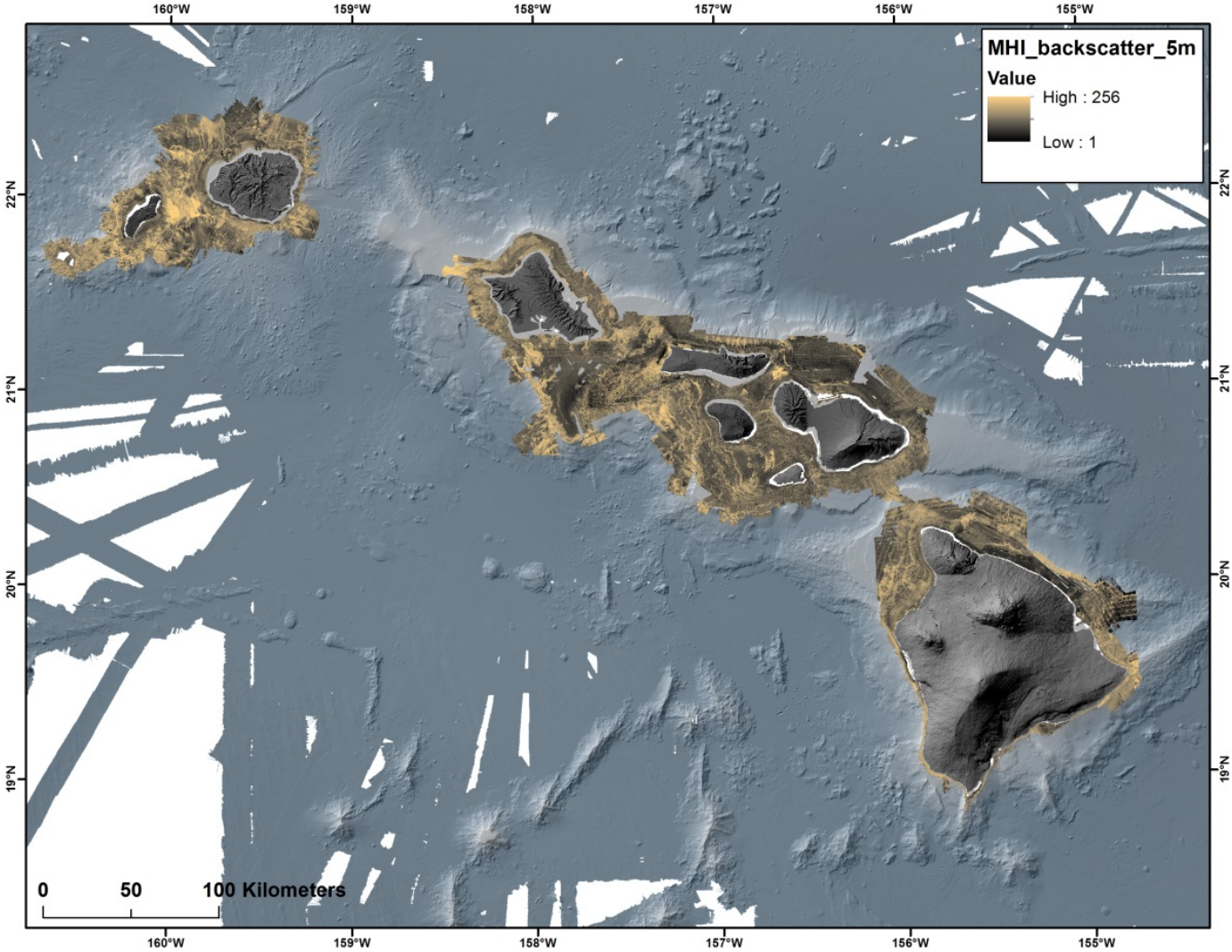

Figure 16: Multibeam backscatter synthesis data for the MHI compiled by UH.

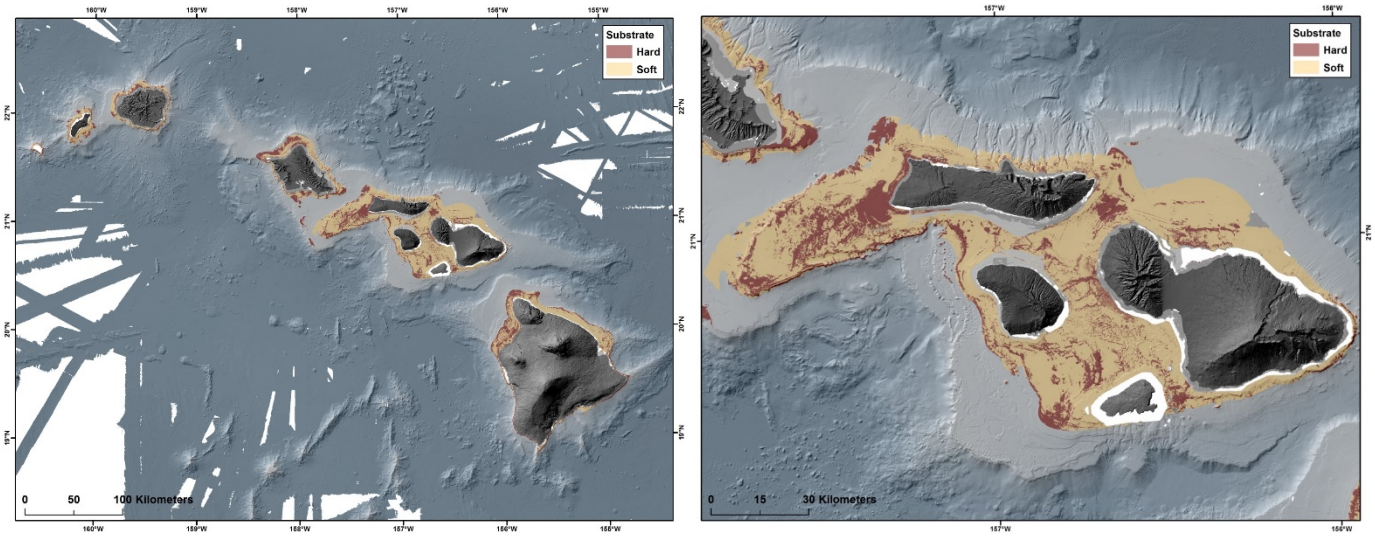

Figure 17: Updated substrate map (0–500 m depths) for the MHI (*left*) and zoomed in to the Maui Nui region of the MHI (*right*). Predicted substrates were based on the classification approach using both bathymetric derivatives and backscatter data (Figure 1). Hard substrates shown in dark red were commonly associated with high-slope areas (e.g., slopes, ridges, and lava flows), and with ~low-slope, high-backscatter areas (e.g., the area between Maui and Molokai).

The 5-m backscatter synthesis was prepared in support of PIFSC bottomfish work and is described in a NOAA Technical Memorandum (Richards et al., In Review). The backscatter synthesis has several issues characteristic of a compilation product consisting of multiple surveys acquired from different systems over multiple years. Nevertheless, the synthesis, which aimed to normalize all survey backscatter intensities to a common scale, is as good as can be expected with the available data. Significant differences among surveys exist in many areas, which ultimately influenced the accuracy of the substrate classification and further highlighted the importance of incorporating bathymetric derivatives in the classification approach. Many recommend that backscatter data alone should not be used to conduct substrate mapping because of the tendency to underestimate hard areas along slopes. The inconsistencies across this dataset only amplify this recommendation (e.g., Brown et al., 2011; Figure 18). There were also concerns about geographic (horizontal) errors in the synthesis data (~5–10 m), though this was not confirmed.

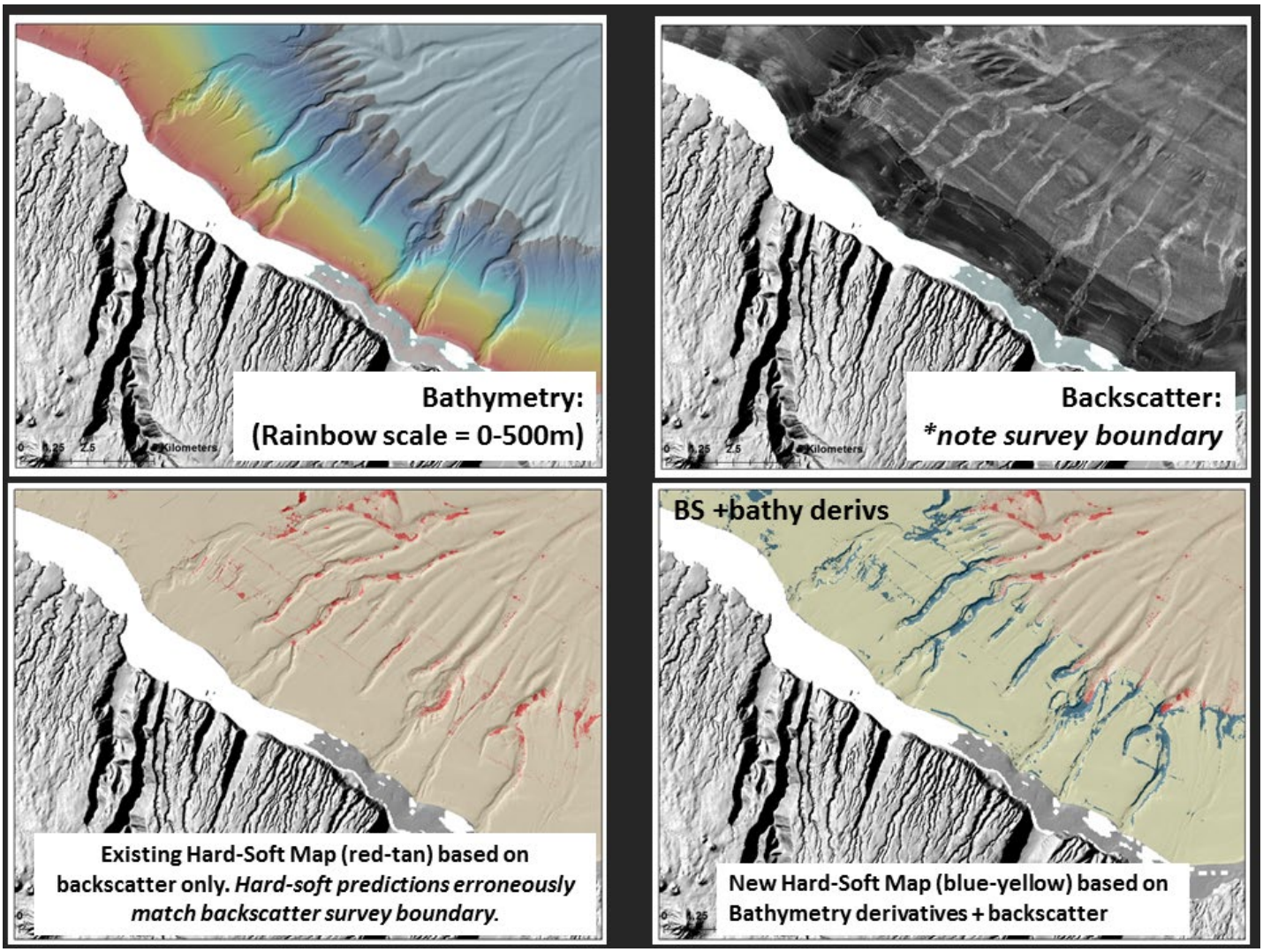

Figure 18: Comparison between the existing backscatter-based substrate map (*bottom left*) and the new substrate map (*bottom right*) for an area offshore of northern Hawaii. The existing substrate map incorrectly indicates the substrate boundary along the multibeam survey boundary in UH's backscatter synthesis (*top right*). The new substrate map more accurately predicts the continuation of hard substrates (shown in blue) in high slope/high rugosity areas within the submarine gullies and canyons. Incorporating bathymetry derivatives (derived from the bathymetry data shown in the top left panel) in the classification mitigates incorrect predictions based on backscatter alone and also reduces nadir noise apparent in the backscatter data (though does not remove it). The terrestrial DEM hillshade is shown in each panel.

Concurrent with the backscatter synthesis production, a bathymetry synthesis was also created using the same survey data. Prior to limiting the depth range of interest to 0–500 m, a number of errors related to re-gridding and geographic transformations were observed in deeper waters, which UH attempted to resolve. There remained, however, issues with the data in the 0–500 m depth range, namely excessive smoothing of the seafloor surface (Figure 19). This was particularly significant to the analysis, as the ability to detect surface variation and roughness had a significant effect on the classification of hard vs. soft. Efforts to improve the 5-m bathymetry product by reducing the smoothing consequently introduced noise and spikes in the data elsewhere. An earlier version of the bathymetry synthesis (same extent, same input data) previously compiled by UH was also available and was found to be of superior quality compared with the recent synthesis (less noise, while preserving high-resolution variation). Though it was preferred to use both of the recently prepared synthesis datasets for the analysis, the recent backscatter synthesis and the previously prepared bathymetry synthesis were used instead for the formal substrate mapping conducted around the MHI.

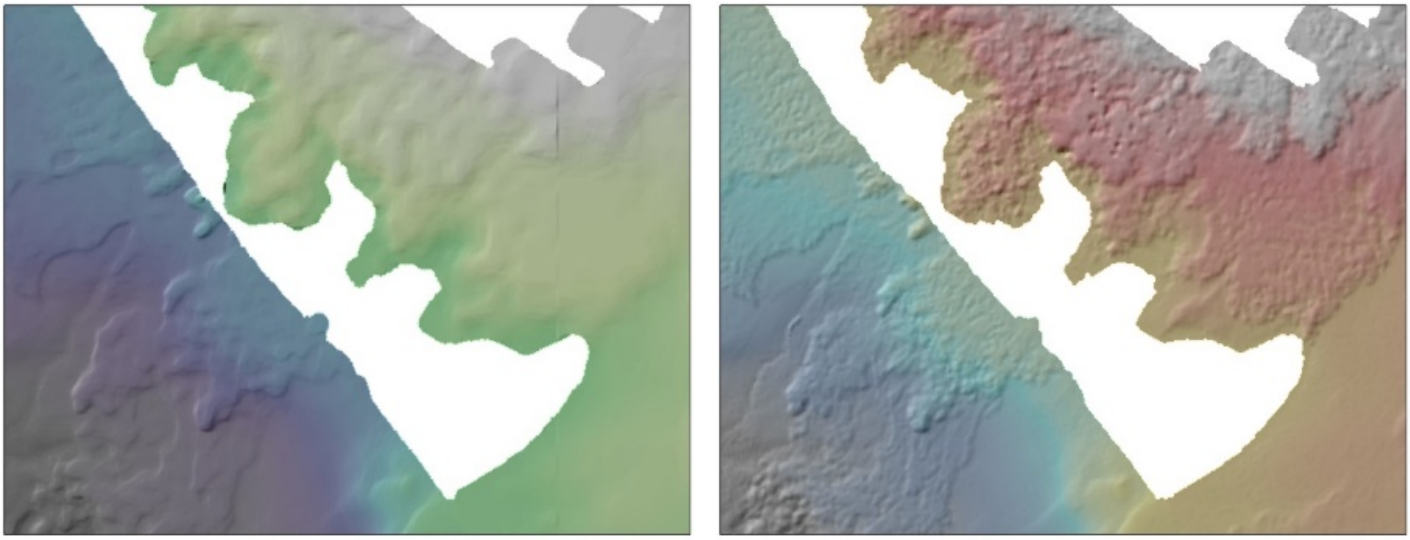

Figure 19: Comparing the MHI bathymetry synthesis-derived datasets for a lava flow area south of Maui. Left panel shows initial recent 5-m bathymetry synthesis prepared by UH, where smoothing removed real bathymetric variation. Right panel shows previously compiled bathymetry synthesis that preserved high-resolution variation without introducing noise and data spikes.

Comparing cluster results

Within the MHI, three versions of the substrate maps were produced, differing according to the number of classes included in the cluster analysis (2, 4, and 7 classes), with each judged to be valid according to visual assessment of the underlying data and in comparison with available validation data (Figure 20, Figure 21). Based on subjective assessment alone, there was no way of determining which of these maps was superior. The final binary threshold between hard and soft was then determined between the classes (e.g., for the 7-class version, clusters 1–4 = soft and 5–7 = hard). The three versions varied in relative sensitivity to morphology vs. backscatter covariates (e.g., 2-class and 4-class versions predicted more hard areas along slopes). All three versions are being incorporated within the Atlantis ecosystem model to determine which version best predicts the abundance and distribution of biomass of marine organisms.

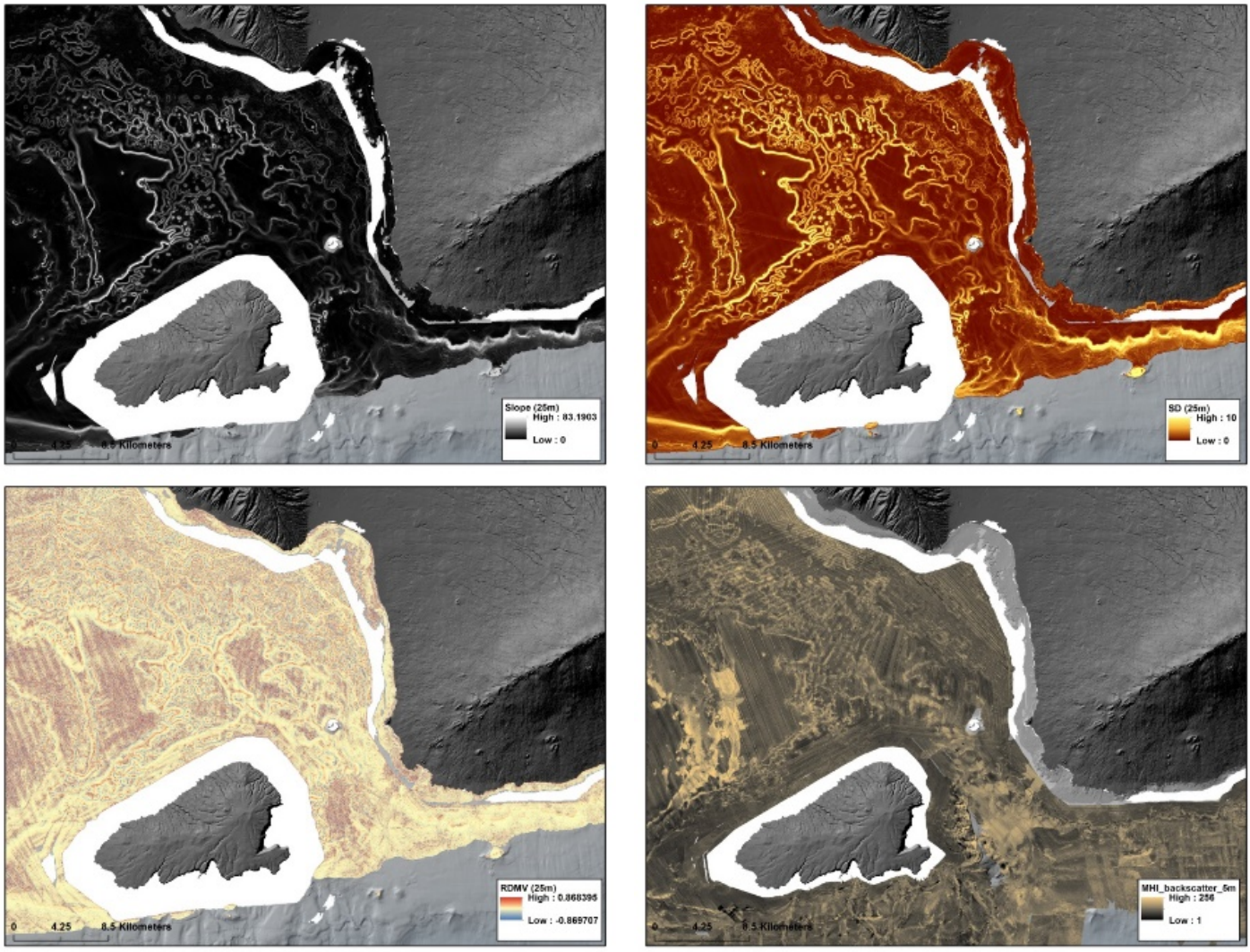

Figure 20: Bathymetric derivatives (slope, *top left*; SD, *top right*; RDMV, *bottom left*) and backscatter data (*bottom right*) incorporated as clustering covariates in the substrate analysis for the MHI, shown for the area between Maui and Kaho'olawe.

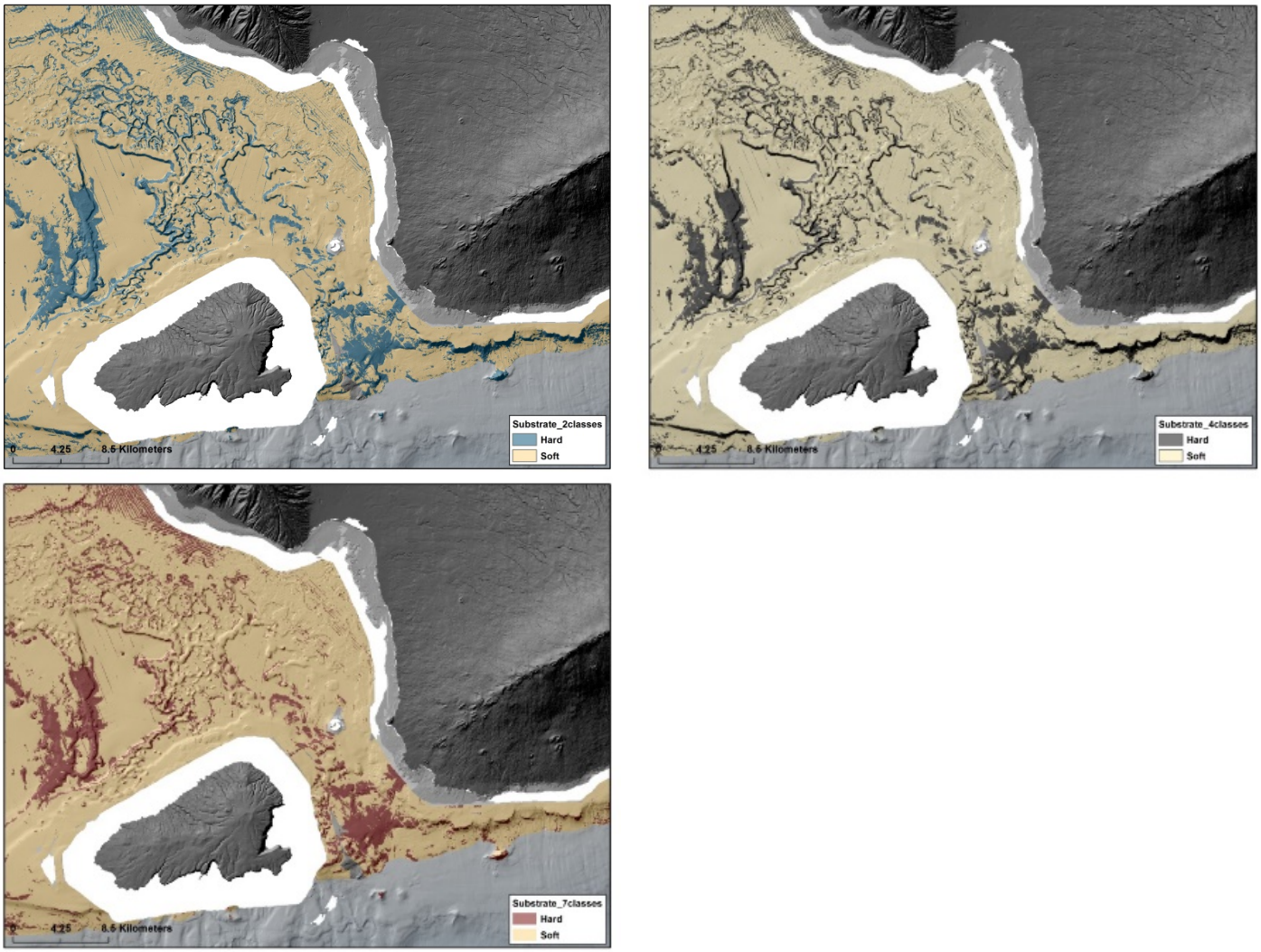

Figure 21: Three versions of the substrate map for the MHI, shown for the area between Maui and Kaho'olawe, based on the number of classes specified in the cluster analyses: 2 classes (*top left*), 4 classes (*top right*), and 7 classes (*bottom*). The 2-class and 4-class versions predicted more hard substrates along slopes, though there were subtle differences between the two.

Multi-scale seafloor metrics

In addition to the substrate maps that were provided to species distribution researchers, the bathymetry, bathymetric derivatives, and backscatter data were also provided at a number of different spatial resolutions (25 m, 125 m, and 625 m), as numerous studies indicate that marine organisms respond to these seafloor characteristics at varying spatial scales (Figure 22). The bathymetric derivatives provided were slope, SD, RDMV, northness, and eastness. While the northness and eastness were not used to predict substrates, these aspect-related variables serve as useful proxies for local hydrodynamic conditions, which cannot be measured adequately.

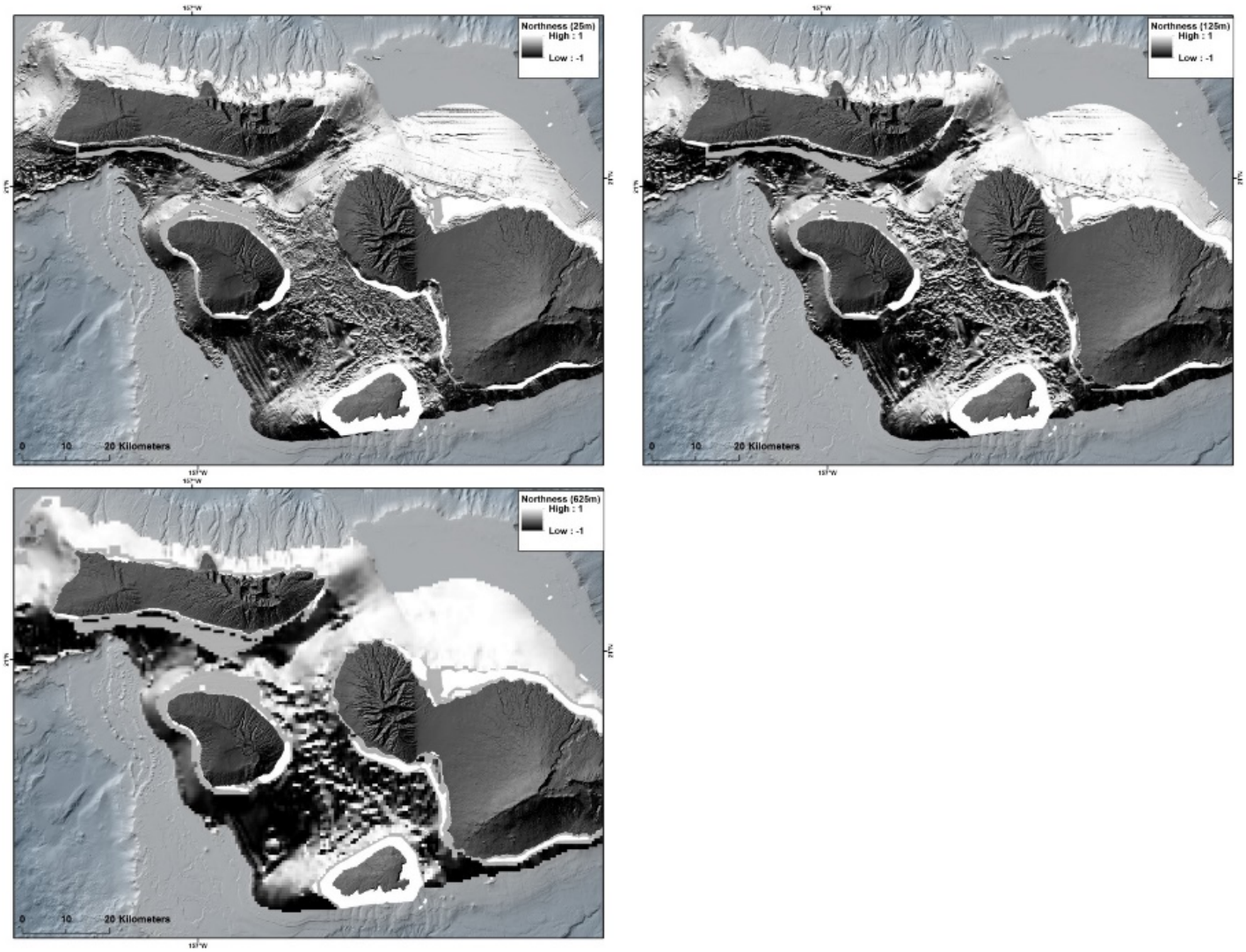

Figure 22: Multi-scale bathymetric derivatives (northness shown as an example) prepared at 25 m (*top left*), 125 m (*top right*), and 625 m (*bottom*) spatial resolutions, which demonstrate that measures of northness show varying levels of detail. While some organisms may be constrained by local-scale heterogeneity that is depicted by the 25-m grid, some may instead be more constrained by regional-scale patterns depicted by the 625-m grid.

Conclusions

Clustering is an effective, largely repeatable classification approach for predicting hard- and soft-bottom seafloor substrates in both shallow and deeper-water environments, specifically for places that lack adequate ground validation data where supervised approaches are otherwise not possible. The approach requires subjective input in the mapping process to determine classification thresholds; however, there is still value in this approach as the majority of the classification is automated and the final steps allow for input where the human eye can detect characteristics that the model cannot.

In deeper-water environments where multibeam sonar data are available, both backscatter and bathymetry data—in the form of bathymetric derivatives—are essential in accurately discriminating between hard and soft substrates. In shallow-water environments, satellite-derived bathymetric derivatives cannot be used in the same

way as multibeam-based bathymetric derivatives, because current approaches to estimate depths from satellite imagery create artifacts that impede this approach.

Alternatively, for shallow-water areas where high-resolution satellite imagery is available, depth-invariant indices are useful in similarly discerning hard and soft substrates using a clustering-based approach. This approach is most effective in study areas with high water clarity and consistent water column properties and is generally easier to automate in forereef habitats. As these circumstances are not always the case (areas that typically confound satellite-based mapping efforts include lagoons, intertidal zones, and turbid waters), the approach allows for minor manual manipulation to fine-tune the initial substrate predictions to correct areas where there are obvious variations, natural or otherwise, that are misinterpreted during the automated classification.

Recommendations for improved seafloor substrate maps

Several suggested improvements are provided here for consideration should further substrate mapping efforts within the U.S. Pacific Islands Region continue at PIFSC or elsewhere.

Improve remote-sensing data

Multibeam data, and backscatter data could be improved as acquisition guidelines/technology have vastly advanced over the last several years. Such data could be acquired from survey vessel platform, or from an AUV. AUVs provide an ideal platform for coral reef studies in shallow water and in mesophotic depths where high-resolution data can be recovered by surveying close to the seafloor. AUVs can be equipped with standard multibeam equipment and chemical and oceanographic sensors, including a sub-bottom profiler that provides 3-D information on geological character (e.g., thickness of mobile sediments). New multispectral multibeam backscatter would improve the ability to resolve detailed characteristics of seafloor composition. These data are particularly important for populated places, such as the MHI, where numerous research projects are taking place to support Ecosystem Based Management or Ecosystem Based Fisheries Management. John Smith at the University of Hawaii is updating the backscatter synthesis for Johnston in the PRIMNM. The analysis can be rerun once the updated synthesis is available to improve the deeper-water substrate map for Johnston.

Lidar bathymetry in the West Hawaii method development area provided accurate substrate maps using the bathymetric derivatives alone. These maps were at least as accurate as the existing statewide NCCOS habitat maps, with far more consistency and details. It would be straightforward to apply this method using Lidar data across the Hawaiian Islands.

Further improvements regarding technologies and techniques towards the applications of multispectral satellite and airborne hyperspectral imagery on the marine sciences are expected. There is perpetual sensor development coupled with improved classification routines for handling multispectral satellite data, and automated techniques for extracting benthic cover from satellite constellations are anticipated (e.g., Dove satellites). There is also promise with airborne hyperspectral data research in higher-resolution studies happening in Hawaii (G. Asner of the Carnegie Airborne Observatory). Collaboration with these groups would be beneficial, as there is interest in linking their local-scale mapping with regional-scale and deeper-water substrate maps.

Improved satellite-derived bathymetry data (without random speckle noise) may allow satellite-derived bathymetric variables (as initially planned) to be incorporated in substrate characterization.

Improve ground-truth data

Acquiring high-quality and sufficient ground-truth data is often considered a requirement to ensure high-accuracy benthic habitat maps, especially if there is a need to provide a quantitative assessment of uncertainty with the predictions. Most of PIFSC's currently available ground-truth data were not optimized for substrate classification (e.g., inaccurate or inconsistent video interpretation) and are of insufficient spatial accuracy to be used in quantitative mapping of high-resolution remote-sensing data. Ground-truth data will enable higher-precision when making predictions using supervised mapping approaches.

Incorporate additional geological information as environmental covariates

The morphology (shape, size, texture) of a feature describes its physical attributes (e.g., lobate mound). While valuable on their own, these morphological attributes also provide an indication of the environmental origin of a feature, that is, its geomorphology (e.g., lava flow). Geomorphology describes features with discrete bathymetric expression. Accurately characterizing the geomorphology of features provides an indication of substrate composition (e.g., hard vs. soft, mineral content) and competence (resistance to erosion). Some organisms may preferentially inhabit carbonate platforms vs. fossil lava flows, for example, whereas areas supporting significant mobile sedimentation are less likely to support abundant/biodiverse habitats.

Characterizing the distribution and thickness of mobile sediment will further improve the accuracy of the substrate classifications. The substrate maps are snapshots of the substrate condition when the data or imagery were acquired. To fully understand the substrate characteristics, it is important to examine variation over oceanographic cycles, such as tidal, storm, and seasonal.

Expand expertise

These environmental covariates (i.e., seabed metrics) could be incorporated in species distribution models (e.g., GAMS, MAXENT, boosted regression trees). For coral reef and fisheries science, this is where the theory is put to practical use (e.g., sensitivity tests to determine relative influence of individual metrics, such as substrates vs. rugosity) in predicting distribution/abundance of benthos, sessile biomass, etc. Initial results from Atlantis suggest significant influence.

Further method development

New clustering methods may improve accuracy (e.g., DBSCAN). Additionally, if/when OBIA is democratized (it may be available on new free platforms such as the statistical program, R), it would be useful to compare results from the clustering approach presented herein with results from an OBIA approach.

References

Asner, G.P., Martin, R.E. and Mascaro, J., 2017. Coral reef atoll assessment in the South China Sea using Planet Dove satellites. Remote Sensing in Ecology and Conservation, 3(2), pp.57–65.

Battista, T.A., Costa, B.M., and S.M. Anderson, S.M. 2007. Shallow-Water Benthic Habitats of the Main Eight Hawaiian Islands. NOAA Technical Memorandum NOS NCCOS 61, Biogeography Branch. Silver Spring, MD.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS journal of photogrammetry and remote sensing, 65(1), pp.2–16.

Brown, C.J., Smith, S.J., Lawton, P. and Anderson, J.T., 2011. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuarine, Coastal and Shelf Science, 92(3), pp.502–520.

Calvert, J., Strong, J.A., Service, M., McGonigle, C. and Quinn, R., 2015. An evaluation of supervised and unsupervised classification techniques for marine benthic habitat mapping using multibeam echosounder data. ICES Journal of Marine Science, 72(5), pp.1498–1513. doi:10.1093/icesjms/fsu223

Costa, B., Kendall, M.S., Parrish, F.A., Rooney, J., Boland, R.C., Chow, M., Lecky, J., Montgomery, A. and Spalding, H., 2015. Identifying suitable locations for mesophotic hard corals offshore of Maui, Hawai‘i. PLoS ONE, 10(7), p.e0130285.

Davies, A.J. and Guinotte, J.M., 2011. Global habitat suitability for framework-forming cold-water corals. PLoS ONE, 6(4), p.e18483.

Diesing, M., Green, S.L., Stephens, D., Lark, R.M., Stewart, H.A. and Dove, D., 2014. Mapping seafloor sediments: comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Continental Shelf Research, 84, pp.107–119.

Dolan, M.F.J., 2012. Calculation of slope angle from bathymetry data using GIS-effects of computation algorithms, data resolution and analysis scale. NGU Report.

Drăguţ, L., Tiede, D., and Levick, S.R. 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data, International Journal of Geographical Information Science, 24:6, pp. 859–871. DOI: 10.1080/13658810903174803.

Dunn, D.C. and Halpin, P.N., 2009. Rugosity-based regional modeling of hard-bottom habitat. Marine Ecology Progress Series, 377, pp.1–11.

Ehses, J.S., Rooney, J.J., 2015. Depth derivation using multispectral WorldView-2 satellite imagery. U.S. Dept. of Commerce, NOAA Technical Memorandum, NOAA-TM-NMFS-PIFSC-46, 24 p. doi:10.7289/V5668B40.

Friedman, S.P. and Robinson, D.A., 2002. Particle shape characterization using angle of repose measurements for predicting the effective permittivity and electrical conductivity of saturated granular media. Water Resources Research, 38(11).

Goff, J.A., Kraft, B.J., Mayer, L.A., Schock, S.G., Sommerfield, C.K., Olson, H.C., Gulick, S.P. and Nordfjord, S., 2004. Seafloor characterization on the New Jersey middle and outer shelf: correlatability and spatial variability of seafloor sediment properties. Marine Geology, 209(1–4), pp.147–172.

Harris, Peter T., and Elaine K. Baker. 2012. "GeoHab Atlas of Seafloor Geomorphic Features and Benthic Habitats: Synthesis and Lessons Learned." In *Seafloor Geomorphology as Benthic Habitat*, edited by Peter T. Harris and Elaine K. Baker, 871–90. Elsevier. https://doi.org/10.1016/B978-0-12-385140-6.00064-5.

Hasan, R.C., Ierodiaconou, D. and Monk, J., 2012. Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. Remote Sensing, 4(11), pp.3427–3443.

Hochberg, E.J., Atkinson, M.J. and Andréfouët, S., 2003. Spectral reflectance of coral reef bottom-types worldwide and implications for coral reef remote sensing. Remote Sensing of Environment, 85(2), pp.159–173.

Jain, A.K., 2010. Data clustering: 50 years beyond K-means. Pattern recognition letters, 31(8), pp.651–666.

Knowlton, N., 2001. The future of coral reefs. Proceedings of the National Academy of Sciences, 98(10), pp.5419–5425.

Le Bas, T.P. 2016. "RSOBIA - A New OBIA Toolbar and Toolbox in ArcMap 10.x for Segmentation and Classification." In *GEOBIA 2016 : Solutions and Synergies*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). https://doi.org/10.3990/2.448.

Lecours, V., Brown, C.J., Devillers, R., Lucieer, V.L. and Edinger, E.N., 2016a. Comparing selections of environmental variables for ecological studies: a focus on terrain attributes. PLoS ONE, 11(12), p.e0167128.

Lecours, V., Dolan, M.F., Micallef, A., and Lucieer, V.L., 2016b. A review of marine geomorphometry, the quantitative study of the seafloor. Hydrology and Earth System Sciences, 20(8), p.3207.

Lundblad, E.R., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Naar, D.F., Donahue, B.T., Anderson, S.M. and Battista, T., 2006. A benthic terrain classification scheme for American Samoa. Marine Geodesy, 29(2), pp.89–111.

McArthur, M.A., Brooke, B.P., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D. and Radke, L.C., 2010. On the use of abiotic surrogates to describe marine benthic biodiversity. Estuarine, Coastal and Shelf Science, 88(1), pp.21–32.

Miller, J., Maragos, J., Brainard, R., Asher, J., Vargas-Ángel, B., Kenyon, J., Schroeder, R., Richards, B., Nadon, M., Vroom, P. and Hall, A., 2008. The state of coral reef ecosystems of the Pacific Remote Island Areas. The state of coral reef ecosystems of the United States and Pacific Freely Associated States, 353–386.

Mumby, P.J. and Edwards, A., 2000. Water column correction techniques. Remote sensing handbook for tropical coastal management, pp.121–128.

Mumby, P.J., Skirving, W., Strong, A.E., Hardy, J.T., LeDrew, E.F., Hochberg, E.J., Stumpf, R.P. and David, L.T., 2004. Remote sensing of coral reefs and their physical environment. Marine pollution bulletin, 48(3–4), pp.219–228.

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. International Journal of Remote Sensing, 33(12), pp.3768–3797.

Pittman, Simon J., and Kerry A. Brown. 2011. "Multi-Scale Approach for Predicting Fish Species Distributions across Coral Reef Seascapes." Edited by Brian Gratwicke. *PLoS ONE* 6 (5). Public Library of Science: e20583. https://doi.org/10.1371/journal.pone.0020583.

Porskamp, P., Rattray, A., Young, M. and Ierodiaconou, D., 2018. Multiscale and Hierarchical Classification for Benthic Habitat Mapping. Geosciences, 8(4), p.119.

Pyle, R.L., Boland, R., Bolick, H., Bowen, B.W., Bradley, C.J., Kane, C., Kosaki, R.K., Langston, R., Longenecker, K., Montgomery, A. and Parrish, F.A., 2016. A comprehensive investigation of mesophotic coral ecosystems in the Hawaiian Archipelago. PeerJ, 4, p.e2475.

Richards, Benjamin L., S. G. Smith, J. S. Ault, G. T. DiNardo, D. Kobayashi, R. Domokos, J. Anderson, J. Taylor, W. Misa, L. Giuseffi, A. Rollo, D. Merritt, J. C. Drazen, M. E. Clarke, and C. Tam. 2016. Design and Implementation of a Bottomfish Fishery-independent Survey in the Main Hawaiian Islands. U.S. Dept. of Commerce, NOAA Technical Memorandum, NOAA-TM-NMFS-PIFSC-53, 54p.

Richards, B.L., John R. Smith, Steven G. Smith, Jerald S. Ault, Christopher D. Kelley, Virginia N. Moriwake. In Review. Development and Use of a Novel Main Hawaiian Islands; Bathymetric and Backscatter Synthesis in a Stratified Fishery-Independent Bottomfish Survey. NOAA Technical Memorandum.

Rooney, J., Donham, E., Montgomery, A., Spalding, H., Parrish, F., Boland, R., Fenner, D., Gove, J. and Vetter, O., 2010. Mesophotic coral ecosystems in the Hawaiian Archipelago. Coral Reefs, 29(2), pp.361–367.

Samsudin, S.A. and Hasan, R.C., 2017. Assessment of Multibeam Backscatter Texture Analysis for Seafloor Sediment Classification. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp.177–183.

Sekund, Linn, and Simon Pittman. 2017. "Explaining Island-Wide Geographical Patterns of Caribbean Fish Diversity: A Multi-Scale Seascape Ecology Approach." *Marine Ecology* 38 (3). Wiley/Blackwell (10.1111): e12434. https://doi.org/10.1111/maec.12434.

Suka, R. and Rooney, J., 2017. Acoustic Characterization of Mesophotic Coral Reef Ecosystems of West Hawai'i. U.S. Dept. of Commerce, NOAA Technical Memorandum NOAA-TM-NMFS-PIFSC-61, 19 p.

Taylor, P.D. and Wilson, M.A., 2003. Palaeoecology and evolution of marine hard substrate communities. Earth-Science Reviews, 62(1–2), pp.1–103.

Weijerman, M., Fulton, E.A., Kaplan, I.C., Gorton, R., Leemans, R., Mooij, W.M. and Brainard, R.E., 2015. An integrated coral reef ecosystem model to support resource management under a changing climate. PLoS one, 10(12), p.e0144165.

Weiss, J., Miller, J. and Rooney, J., 2008. Seafloor characterization using multibeam and optical data at French Frigate Shoals, Northwestern Hawaiian Islands. In Proceedings of the 11th International Coral Reef Symposium (Vol. 17).

Wilson, M.F., O'Connell, B., Brown, C., Guinan, J.C. and Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Marine Geodesy, 30(1–2), pp.3–35.