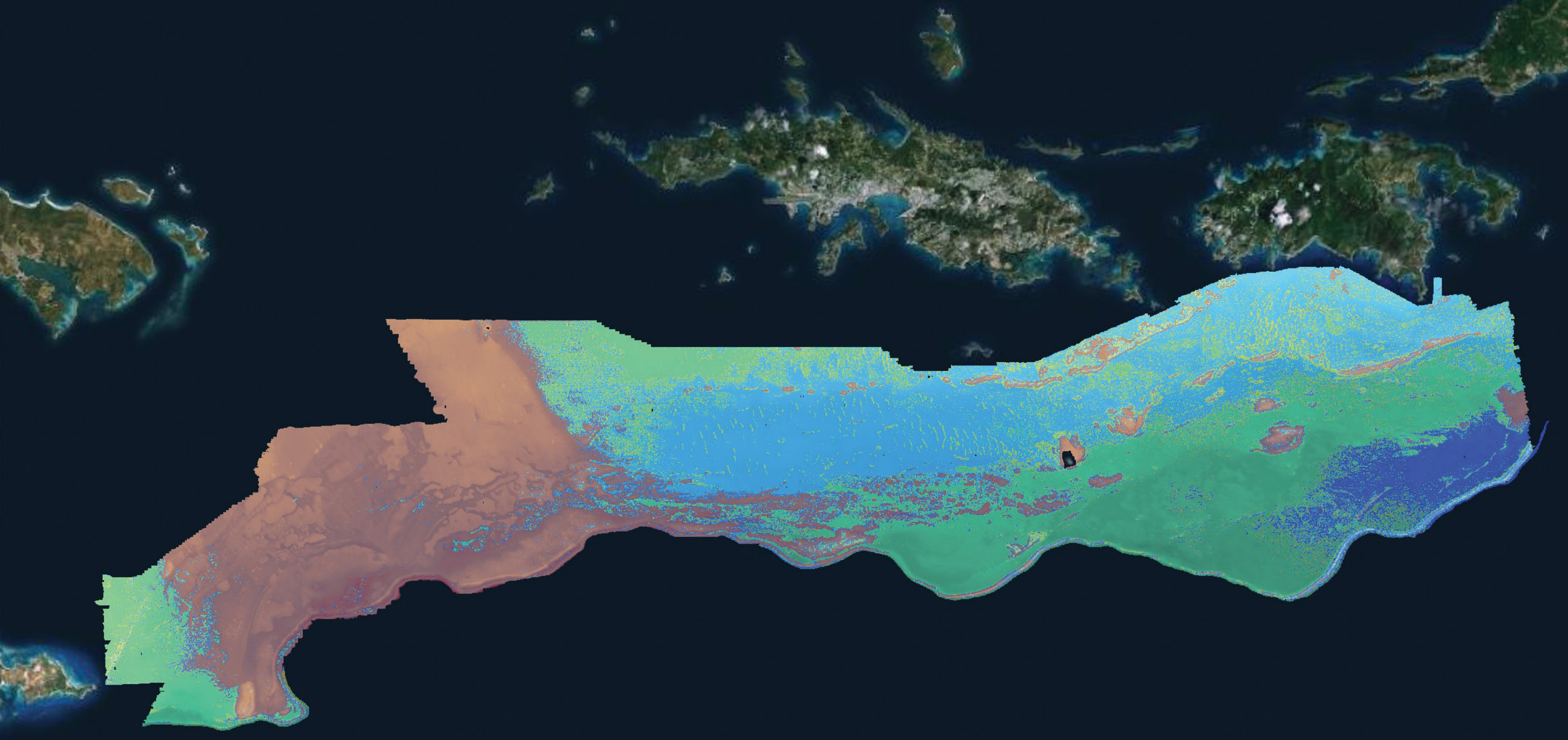
https://doi.org/10.7289/V5/TM-NOS-NCCOS-241

Benthic Habitat Maps for the Insular Shelf South of St. Thomas and St. John



Authors

Bryan Costa

Laura Kracker

Tim Battista

Will Sautter

Ayman Mabrouk

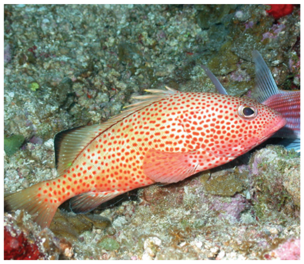
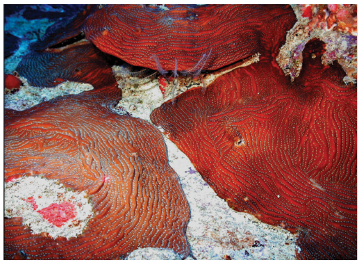
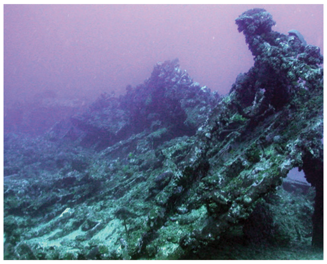
Kimberly Edwards

Chris Taylor

Erik Ebert

December 2017



NOAA TECHNICAL MEMORANDUM NOS NCCOS 241

NOAA National Centers for Coastal Ocean Science

Suggested Citation

Costa, B., L.M. Kracker, T. Battista, W. Sautter, A. Mabrouk, K. Edwards, C. Taylor, and E. Ebert. 2017. Benthic Habitat Maps for the Insular Shelf South of St. Thomas and St. John. NOAA Technical Memorandum NOS NCCOS 241. Silver Spring, MD. 59 pp. https://doi.org/10.7289/V5/TM-NOS-NCCOS-241

Acknowledgments

This habitat characterization work was funded by NOAA's Coral Reef Conservation Program (CRCP) and National Centers for Coastal Ocean Science (NCCOS). The map products provided in this report represent the culmination of an extensive mapping campaign in the USVI conducted by NOAA's National Centers for Coastal Ocean Science (NCCOS), led by Tim Battista. These surveys were conducted in collaboration with several agencies, including U.S. Geological Survey (USGS), NOAA's Office of Marine and Aircraft Operations (OMAO) NOAA ship *Nancy Foster*, University of North Carolina at Wilmington (UNCW) Undersea Vehicles Program, NOAA's Office of Coast Survey (OCS), National Park Service (NPS), Caribbean Fishery Management Council (CFMC), U.S. Virgin Islands (USVI) Department of Planning and Natural Resources (DPNR), Puerto Rico's Department of Natural and Environmental Resources (DNER), University of the Virgin Islands (UVI), and University of Puerto Rico (UPR). The lead hydrographers for these acoustic surveys were Sean Rooney (NOAA) in 2004, Jay Lazar (NOAA) in 2005, and Mike Stecher (Solmar Hydro) in 2006 to 2011. A special thanks to Mike Stecher for his efforts to always collect the highest-quality MBES data possible; and to Lance Horn, Glen Taylor and Jason White (UNCW Undersea Vehicles Program) for expertly and safely piloting the ROV. Chris Taylor led the fisheries acoustic data collection effort, with support from Laura Kracker and Erik Ebert.

A research permit was obtained from the National Park Service for field work completed within the USVI National Monument (Permit #: VICR-2016-SCI-0003). Review of the draft map products was conducted by Leslie Henderson (USVI DPNR), Caroline Rogers (USGS), Lee Richter (NPS), Marilyn Brandt (UVI), Graciela Garcia-Moliner (CFMC) and Tyler Smith (UVI). Tyler Smith also provided valuable insights regarding the habitat classes. Matthew Kendall, Arliss Winship and Simon Pittman (NOAA NCCOS) reviewed this report, and provided many helpful comments and suggestions. Captain Kevin McCartney (CrewZen VI) safely navigated us to field sites for the 2016 ground validation and 2017 accuracy assessment missions. Many thanks to Ken Buja (NOAA NCCOS) for creating the online data viewer, Jessica Morgan for helping to archive the GIS products, and Jamie Higgins for compiling, editing and formatting this final report. All cover photos were taken by T. Battista and B. Costa. Editorial content was improved by Kevin McMahon (NOAA). CSS employees were supported under NOAA Contract No. EA133C17BA0062 with CSS, Inc..

For more information on NOAA's National Centers of Coastal Ocean Science, please visit:
https://coastalscience.noaa.gov/

For more information on this project, please visit:
https://coastalscience.noaa.gov/project/habitat-map-insular-shelf-south-of-st-thomas-st-john/

Or direct questions and comments to:

NOAA/NOS/National Centers for Coastal Ocean Science

1305 East West Highway, SSMC4, N/SCI-1

Silver Spring, MD 20910

Bryan Costa
bryan.costa@noaa.gov

or

Laura Kracker
240-533-0362
laura.kracker@noaa.gov

https://doi.org/10.7289/V5/TM-NOS-NCCOS-241

Benthic Habitat Maps for the Insular Shelf South of St. Thomas and St. John

Prepared by:
Biogeography Branch
NOAA National Centers for Coastal Ocean Science (NCCOS)
Silver Spring, MD
USA

December 2017

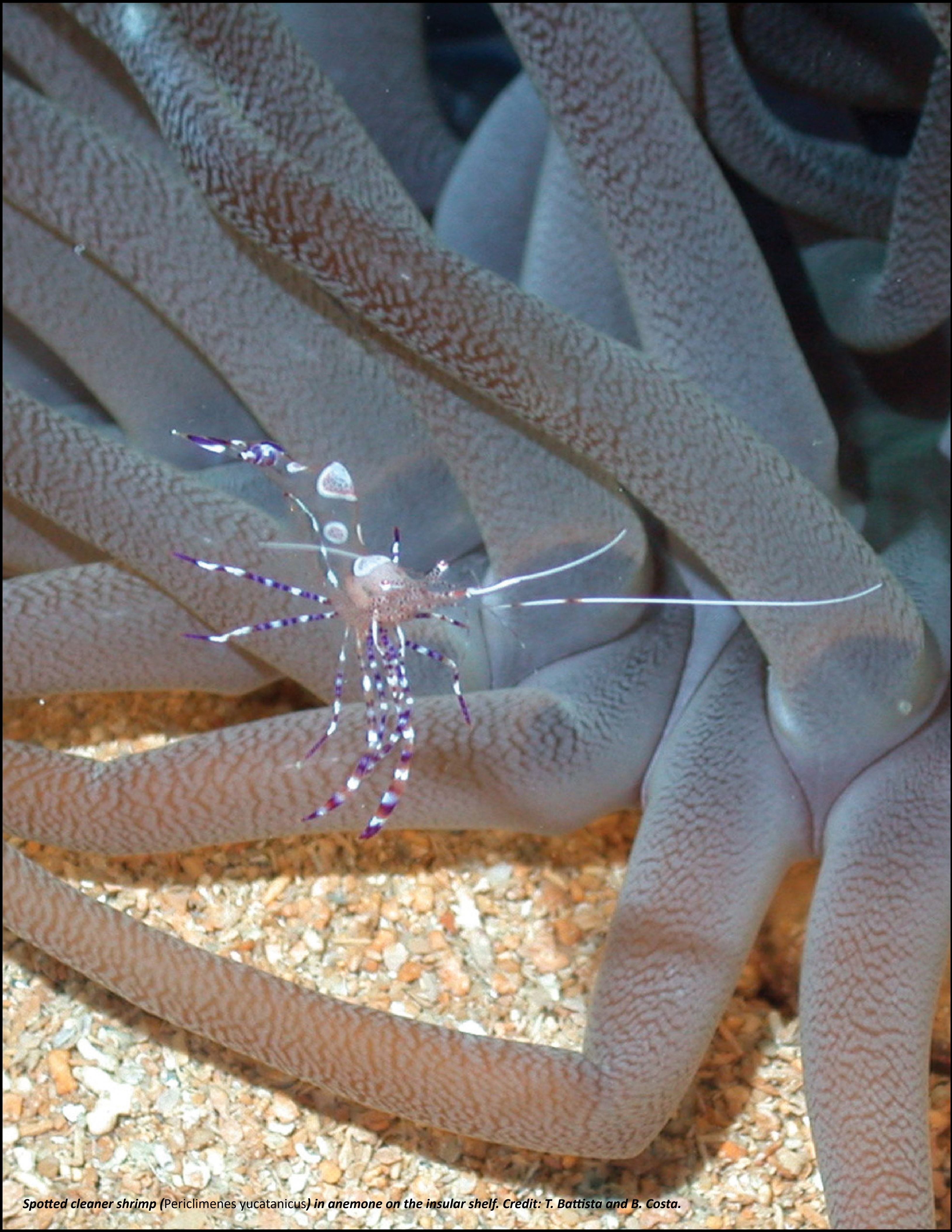
Authors

Bryan Costa[1], Laura M. Kracker[1], Tim Battista[1], Will Sautter[1,2], Ayman Mabrouk[1,2], Kimberly Edwards[3,4], Chris Taylor[3], Erik Ebert[3,4]

[1] NOAA National Ocean Service/National Centers for Coastal Ocean Science, Marine Spatial Ecology Division, Silver Spring, MD, U.S.A.
[2] CSS. NOAA National Centers for Coastal Ocean Science, Marine Spatial Ecology Division, Silver Spring, MD, U.S.A
[3] NOAA National Centers for Coastal Ocean Science, Marine Spatial Ecology Division, Beaufort, NC, U.S.A.
[4] JHT. NOAA National Centers for Coastal Ocean Science, Marine Spatial Ecology Division, Beaufort, NC, U.S.A.

NOAA Technical Memorandum NOS NCCOS 241

United States Department of Commerce	National Oceanic and Atmospheric Administration	National Ocean Service
Wilbur L. Ross, Jr. Secretary	Timothy Gallaudet Administrator, Acting	Russell Callender Assistant Administrator



Spotted cleaner shrimp (Periclimenes yucatanicus ***) in anemone on the insular shelf. Credit: T. Battista and B. Costa.***

Executive Summary

The insular shelf south of St. Thomas and St. John, in the U.S. Virgin Islands (USVI) is an expansive geomorphologic feature with extensive mesophotic coral reefs occurring at 30 to 150 meter depths. These mesophotic reefs are part of a broader coral reef community in the U.S. Caribbean that provide a wide range of ecosystem goods and services estimated to provide $187 million per year (in 2007 US dollars) to local communities. While these reefs provide great economic value to local communities, they are increasingly under threat from multiple human-caused stressors, making it critical for the USVI and Puerto Rico jurisdictions to find ways to preserve and sustainably manage them. Ideally, the first step in any marine management process is to comprehensively map and inventory the location of coral reef resources. The habitat map products provided in this report represent the first complete habitat map for the insular shelf south of St. Thomas and St. John. These products also represent the culmination of an extensive seafloor mapping campaign conducted by NOAA's National Centers for Coastal Ocean Science (NCCOS) in collaboration with regional partners.

Model-based mapping techniques have advanced considerably since this seafloor mapping campaign began, and the map products created here take advantage of these improvements. These habitat map products are 'pixel-based,' developed from 11x11 meter resolution raster images and mathematical modeling techniques called boosted regression and classification trees. These modeling techniques generate continuous spatial predictions by finding relationships between the geographic distribution of habitats (extracted from 1,005 underwater videos) and environmental conditions that may be influencing or correlated with these distributions (extracted from 20 raster images describing the oceanography, geography, and seafloor topography of the area). The classification scheme used here is based on broad benthic assemblages and functional groups, including substrate and biological cover types. The final habitat map products describe the probability of occurrence and associated precision for four individual substrata (i.e., 'Coral Reef', 'Pavement', 'Rhodoliths', and 'Sand') and two biological cover types (i.e., 'Live Hard Coral' and 'Live Soft Coral'). These six layers were also combined to create a single, composite benthic map with five habitat classes including: 'Coral reef colonized with live coral', 'Rhodoliths with macroalgae', 'Bare sand', 'Rhodoliths with macroalgae and bare sand', and 'Pavement colonized with live coral.' The performance and thematic accuracy of these products were evaluated using an independent set of underwater videos (n=348).

In total, 652 km^2 of seafloor were characterized on the insular shelf south of St. Thomas and St. John. The performance of the habitat models was considered to be good to outstanding based on four quantitative metrics, including percent deviance explained (PDE), area under the receiver operating characteristic curve (AUC), mean error (bias) and root mean square error (RMSE). For the habitat models, PDE ranged from -3.9% to 83.3% ($\bar{x}$ =37.4% ±16.2 SE); AUC values ranged from 0.70 (good) to 0.99 (outstanding) ($\bar{x}$ =0.86 ±0.05 SE); bias was small to moderate, ranging between -0.17 to +0.04 ($\bar{x}$ =-0.04 ±0.03 SE); and RMSE values ranged from 0.2 to 0.46 ($\bar{x}$ =0.33 ±0.04 SE). The 'Coral Reef' model and prediction performed the best overall. The relative importance of each environmental predictor differed among the substrate and cover models, although the geographic and seafloor predictors were important for all the models. The composite map performed equally as well as the individual habitat models. The overall accuracy and tau value for the composite habitat map was high at 85.6% and 0.82 ±0.05, respectively.

'Coral Reef colonized with live coral' was the most abundant habitat type on the insular shelf, comprising 34.3% (223.6 km^2) of the area. It was dominant throughout the Virgin Passage, Hind Bank, along the upper shelf of French Cap Bank, and along ridge features south of St. Thomas. 'Bare sand' was often found around the edges of these coral reef structures, particularly on the eastern portion of the insular shelf. Linear patches of bare sand were also common in the middle of the insular shelf between Hind Bank and Frenchcap Bank. Interestingly, these patches were fairly regularly spaced, and commonly oriented in the north-south direction, suggesting they were formed by consistent bottom currents in the area. 'Rhodoliths with macroalgae' was the second most abundant habitat, comprising 30.2% (196.8 km^2) of the insular shelf. This habitat was more common in deeper depths (greater than approximately 45 m), especially near Frenchcap Bank, Tampo Bank, and west of El Seco. 'Rhodoliths with macrolagae' often transitioned into 'Rhodoliths with macroalgae and bare sand' at this depth threshold. This transition was particular sharp south of Frenchcap Cay and the Mid-shelf Reef. Lastly, 'Pavement colonized with Live Coral' was the least abundant habitat type, comprising 4.8% (31.3 km^2) of the area. It occurred primarily in the southeast region on the insular shelf.

As these complex spatial patterns suggest, seafloor habitats vary widely across the insular shelf south of St. Thomas and St. John. The products described here capture these complex patterns and translate them into maps that can be used to plan for potential activities both onshore nearby islands and offshore on the insular shelf. They can also be integrated with other spatially coincident datasets collected by NCCOS and regional partners, including underwater videos collected using remotely operated vehicles and acoustic information describing the size, abundance and distribution of fish. The best way to access and use these products is through geographic information systems (GIS) software that allows users to zoom in and create custom maps. These GIS-ready layers are archived at NOAA's National Centers for Environmental Information (NCEI) and are freely available for viewing and download from the following project page: https://coastalscience.noaa.gov/project/habitat-map-insular-shelf-south-of-st-thomas-st-john/. For users that do not have GIS software or expertise, these map products are accessible as print-ready maps at the end of this report, and through an online map viewer linked to the above project page.

NOAA Ship Nancy Foster *collecting data on insular shelf. Credit: A. Clarke.*

Horse-eye Jacks (Caranx latus) on insular shelf. Credit: T. Battista and B. Costa.

1.0 INTRODUCTION

The insular shelf south of St. Thomas and St. John, USVI (subsequently referred to as the insular shelf) (Figure 1.1) is an expansive geomorphologic feature with extensive mesophotic coral reefs, known spawning aggregations, and several promontories. This portion of the insular shelf is approximately 71 km in length, averages approximately 14 km in width, and extends from the islands of Culebra and Vieques in Puerto Rico east to the island of St. John in the USVI. The insular shelf is characterized by a broad, low slope platform with localized areas of high relief areas. The outer perimeter is defined by a shelf edge and steep slope margins at 75-100 m depths (Smith et al., 2016). Our focus in this project was the mesophotic depths on the shelf from approximately 30 m to 100 m deep, although some shallower areas were characterized around Sail Rock and Frenchcap Cay in the process.

The coral reefs on the insular shelf (and in the broader U.S. Caribbean) provide a wide range of ecosystem goods and services. These include coral reef associated tourism, fisheries, coastal protection, real estate value, recreation and cultural value, and research and education value (Brander and van Beukering, 2013). Combined, the total economic value (TEV) of coral reef services in the USVI is estimated at $187 million per year (in 2007 US dollars) (van Beukering et al., 2011). Local communities in the USVI are some of the primary beneficiaries of reef-related tourism and ecosystem services (van Beukering et al., 2011), both directly and indirectly. The

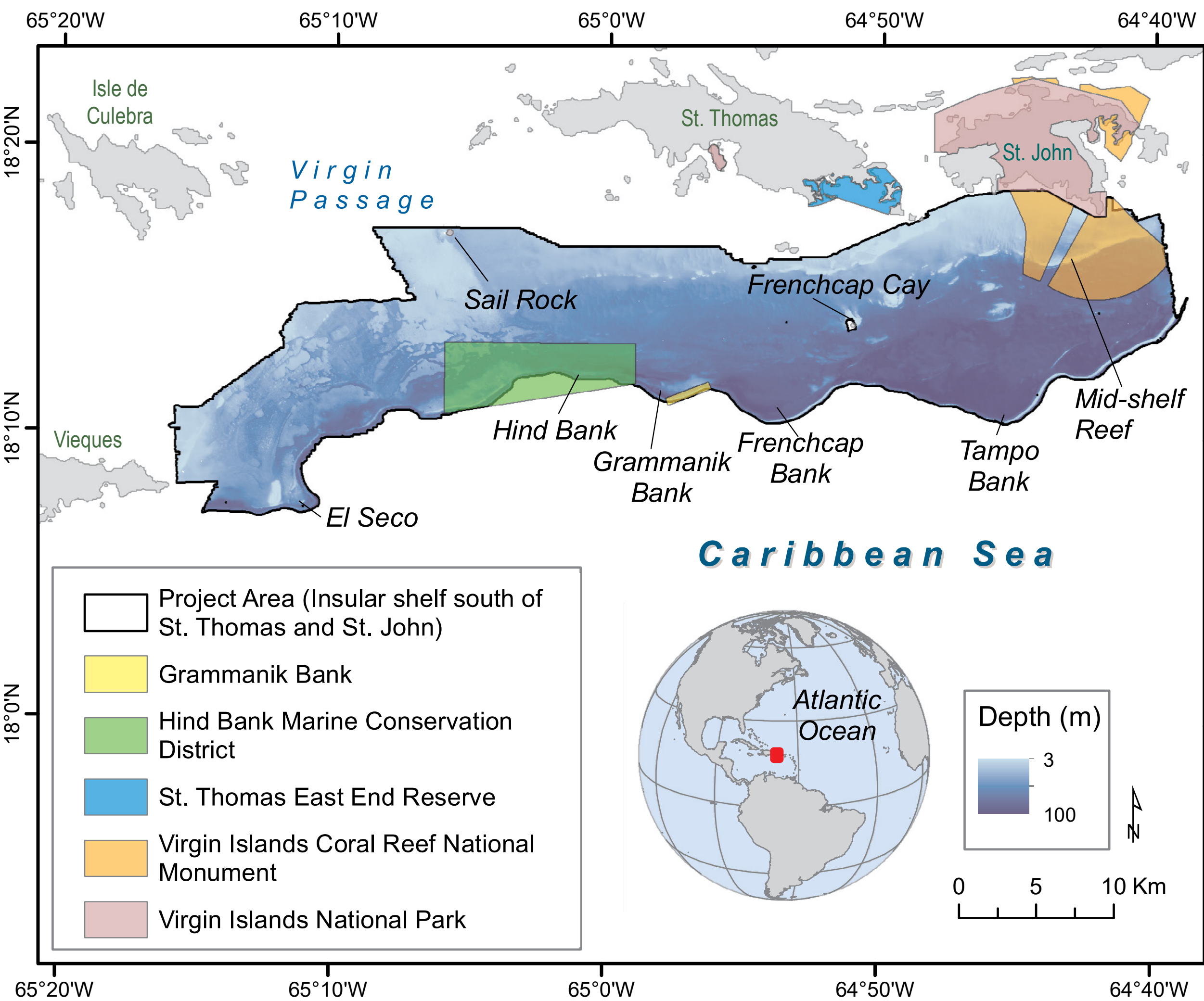


Figure 1.1. Extent of project area and geographic points of interest on the insular shelf south of St. Thomas and St. John, USVI.

Introduction

coral reefs that provide these services are increasingly under threat from multiple human-caused stressors (Rothenberger et al., 2008), making it critical for the USVI and Puerto Rico jurisdictions to work with local communities to preserve and sustainably manage them.

Ideally, the first step in any marine management process is to comprehensively map and inventory the location of resources in a timely, accurate, and consistent way. In 2000, the United States Coral Reef Task Force charged the National Oceanic and Atmospheric Administration (NOAA) with leading federal efforts to produce benthic habitat maps of all U.S. coral reef ecosystems (NOAA, 2002). Initial efforts focused on characterizing shallow-water reef environments (< 30 m deep) using photographs acquired from airplanes and satellites (Monaco et al., 2012). In 2004, a concerted effort also began to map and characterize reefs deeper than 30 m using ship-based sound navigation and ranging (SoNAR) systems and underwater cameras on remotely operated vehicles (ROVs) and other towed platforms (NOAA NCCOS, 2017). These surveys were led by NOAA's National Centers for Coastal Ocean Science (NCCOS) and the U.S. Geological Survey (USGS), in collaboration with several other agencies, including NOAA's Office of Marine and Aircraft Operations (OMAO), NOAA Office of Coast Survey (OCS), the National Park Service (NPS), Caribbean Fishery Management Council (CFMC), USVI Department of Planning and Natural Resources (DPNR), Puerto Rico's Department of Natural and Environmental Resources (DNER), the University of the Virgin Islands (UVI), the University of Puerto Rico (UPR), and the University of North Carolina Wilmington (UNCW). The products provided in this report represent the culmination of this extensive mapping campaign, conducted over eight years and seven missions on the insular shelf.

The NOAA ship *Nancy Foster*, operated by NOAA's OMAO, was the primary platform used by NCCOS to map reefs > 30 m deep on the insular shelf and throughout U.S. Caribbean. Seafloor mapping was conducted using high-resolution SoNAR systems called multibeam echosounders (MBES) and phase differencing bathymetric sonars (PDBS), which collect both bathymetry (depth) and backscatter (acoustic intensity) data (Figure 1.2). The spatial distribution, size and abundance of fish were mapped at the same time using splitbeam echosounders (SBES) (Section 3.6). These acoustic systems have been used elsewhere to support a wide variety of ecosystem-

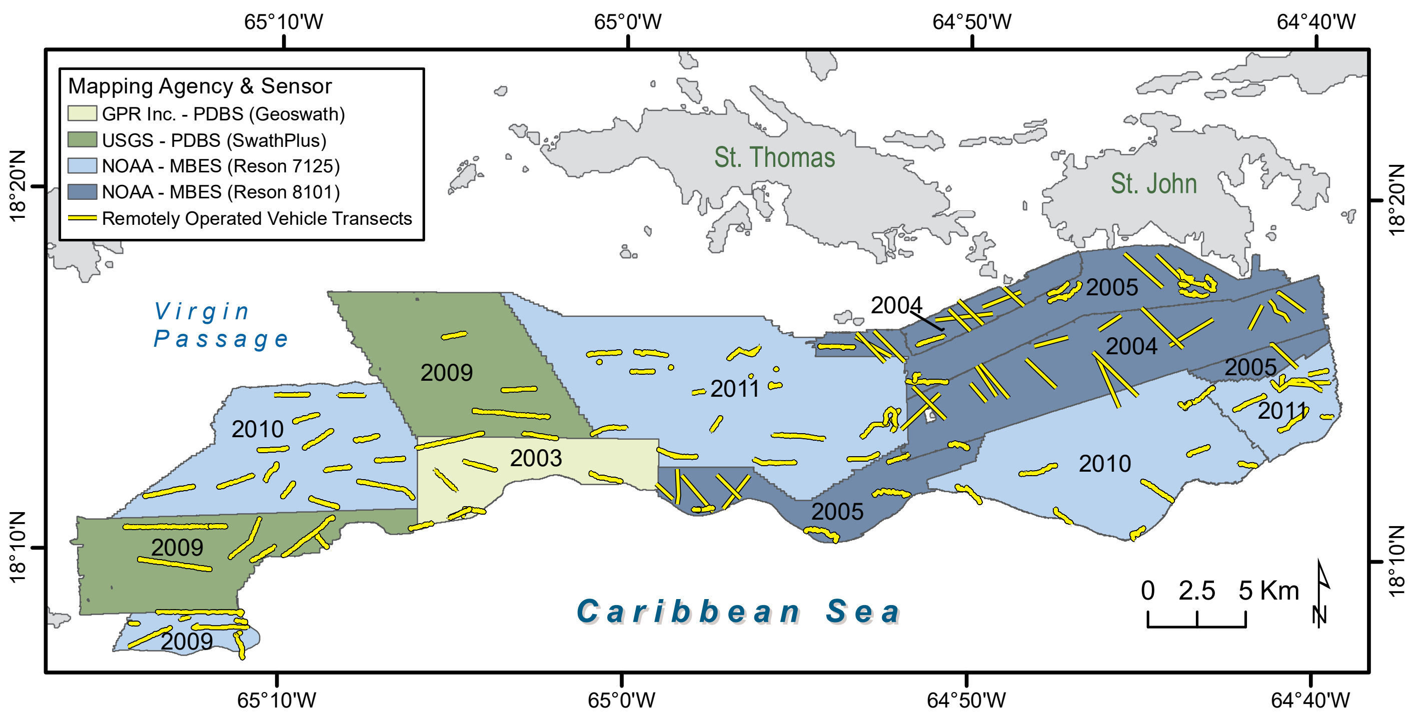


Figure 1.2. Geographic area mapped on the insular shelf by agency and sensor type. Mapping data were collected between 2003-2011. Remotely Operated Vehicle (ROV) surveys were conducted during NOAA mapping missions between 2005-2011. Spatially coincident information describing the spatial distribution, size and abundance of fish was also collected during several of these surveys (Appendix C.).

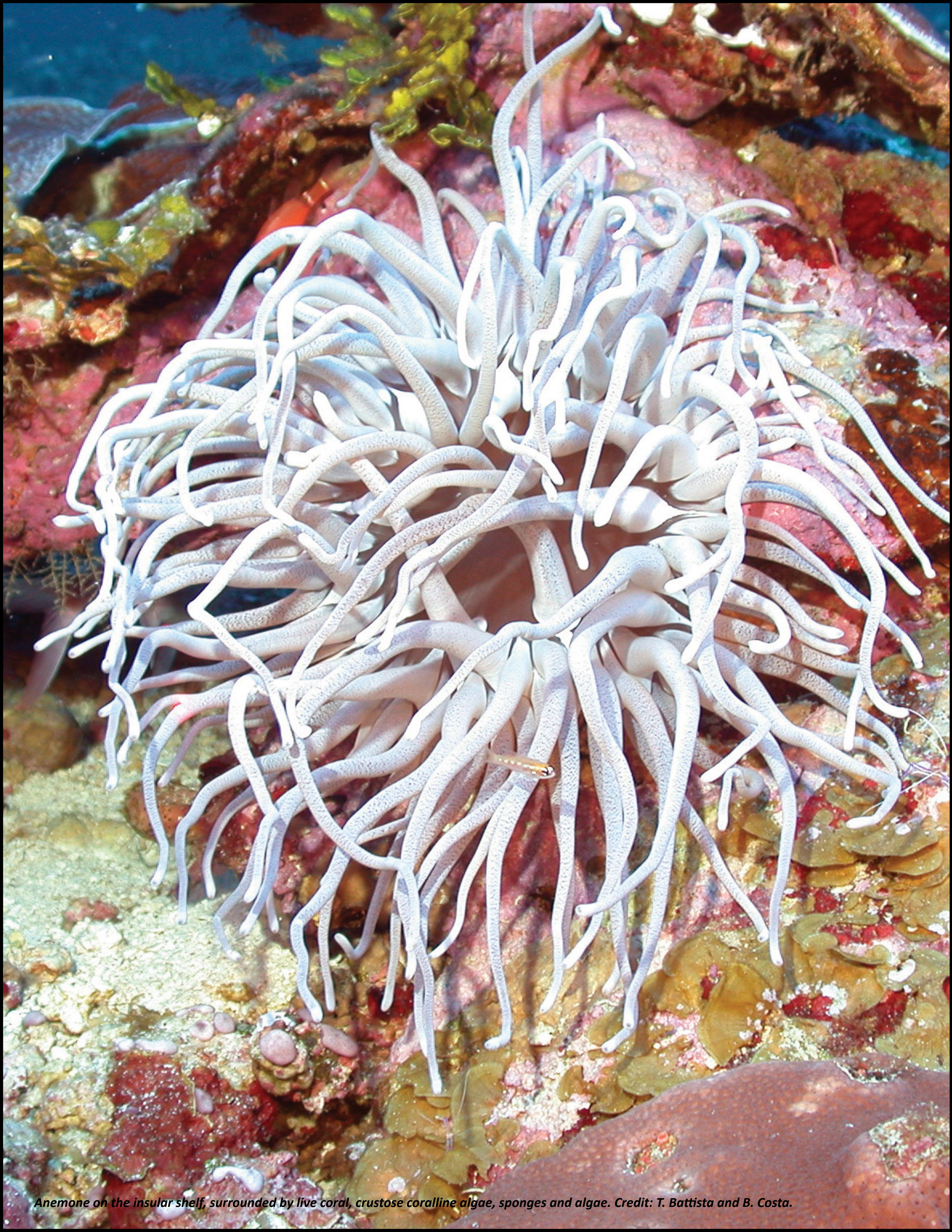
based management applications (Trenkel et al., 2011), and are capable of mapping deeper depths and in more turbid locations compared to previous mapping efforts using photographs. These datasets can be analyzed and combined with other spatial information to describe the topography, substrate, and biological cover on the seafloor (Kenny et al., 2003). Mapping surveys conducted by NOAA NCCOS and the USGS have provided data at greater spatial detail than previously available, revealing unknown topographic and geomorphologic features on the insular shelf. These high-resolution datasets were the basis for this project, with the goal of developing detailed benthic habitat maps for the entire insular shelf.

For this project, the availability of high-resolution bathymetry, advances in mathematical modeling, increases in computing power, the collection of high definition video, and the development of new mapping methods (Kendall et al., 2017) were crucial for producing a new generation of benthic habitat maps for the insular shelf. These advances resulted in benthic habitat maps that are more spatially resolved, preserve habitat gradients, and are more scalable and flexible compared to previous maps in the region (Costa et al., 2009, 2013). This new approach also quantifies and describes the uncertainty associated with the habitat map products, which when included in the decision-making process, can have profound effects in marine management actions and conservation outcomes (Guisan et al., 2013; Tulloch et al., 2013).

Six separate habitat predictions were developed for the insular shelf. These predictions describe the probability of occurrence of four substrate and two biological cover types. These layers were also combined to create a single, composite benthic habitat map depicting the most likely combination of substrate and cover types that create distinct habitats on the seafloor. The performance and thematic accuracy of these products were evaluated using independent field data. The following products are freely available for viewing and are downloadable in GIS compatible formats from: https://coastalscience.noaa.gov/project/habitat-map-insular-shelf-south-of-st-thomas-st-john/:

1) One composite benthic habitat map;
2) Probability of occurrence maps for four substrate and two biological cover types;
3) High resolution depth and topographic layers;
4) Oceanographic layers describing euphotic depth, turbidity, and ocean temperature anomalies;
5) Georeferenced underwater videos used to develop habitat predictions and evaluate their performance;
6) A technical report (this document) describing the methods, results, and limitations for scientific and management applications.

In the USVI, NCCOS's benthic habitat maps have provided important baseline spatial information that has been used by resource managers to help design and implement a variety of monitoring and conservation measures. These measures include: (1) constructing sampling designs for coral reef ecosystem monitoring and assessment programs (Menza et al., 2006; NOAA NCRMP, 2017), (2) evaluating the efficacy of existing marine protected areas (MPAs) (Pittman et al., 2013, 2014), (3) planning for and designing new MPAs (Pittman et al., 2013), (4) identifying and prioritizing the most important and resilient coral reefs in the region (Pittman et al., 2007; Knudby et al., 2014), (5) targeting research to better understand the socioeconomic, oceanographic and ecological processes affecting coral reef ecosystem function and condition (Pittman et al., 2017), (6) providing a baseline to evaluate the changes in ecosystems over time (Menza et al., 2012) and (7) helping to prioritize areas for further study and protection (Pittman et al., 2008; Friedlander et al., 2013). The benthic habitat maps and products contained in this report will provide similarly valuable information to support the management of coral reefs on the insular shelf from the USVI to Puerto Rico.



Anemone on the insular shelf, surrounded by live coral, crustose coralline algae, sponges and algae. Credit: T. Battista and B. Costa.

2.0 METHODS

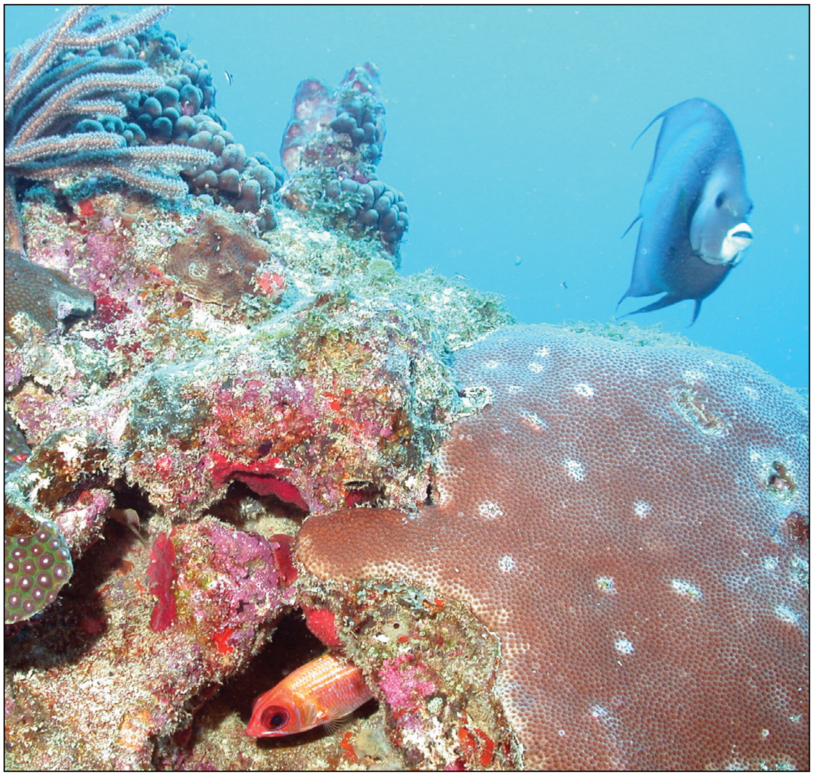
Several tasks were integrated to develop habitat map products for the insular shelf. These steps closely follow those used in other recent mapping activities (Kendall et al., 2017) including: (1) creating a benthic habitat classification scheme (Section 2.1); (2) preparing oceanographic, topographic, and geographic datasets to be used as predictors in the habitat models (Section 2.2); (3) collecting underwater videos to train the habitat models and evaluate their performance (Section 2.3); and (4) using the habitat models to create spatial predictions and a composite habitat map (Section 2.4).

2.1 BENTHIC HABITAT CLASSIFICATION SCHEME

The classification scheme used here is intended to characterize broad benthic assemblages and functional groups, including substrate type (e.g., sand) and biological cover type (e.g., macroalgae). It was developed based on our mapping team's expertise gained over fifteen years of underwater surveys in the region, and was limited to those habitats that were likely to be detected and differentiated in SoNAR imagery. This classification scheme was refined through consultation with local researchers and resource managers, and was used to annotate underwater videos collected on the insular shelf. This habitat classification scheme was the basis for all the habitat map products developed here.

The initial classification scheme included five substrate types and four biological cover types (Table 2.1). However, models and predictions were not created for all substrate and cover types. In cases where habitat types were very rare (e.g., artificial had a prevalence = 0.3%), the model failed because there were not enough presences to find a clear mathematical relationship between the habitat type and predictors. Where habitat types were very common (e.g., macroalgae and sponge had prevalences of 97% and 93%, respectively), the model also failed because there was no distinct set of predictors that explained the spatial distribution of the respective habitat types.

The spatial distributions of the remaining substrate and cover types were predicted across the insular shelf using mathematical models. These substrate and cover types included: 'Coral Reef', 'Pavement', 'Rhodoliths', 'Sand', 'Live Hard Coral', and 'Live Soft Coral' (Figure 2.1). These six habitat predictions were then used to create a composite (i.e., single layer) habitat map depicting five habitat classes (Table 2.2). These five habitat classes depict commonly occurring combinations of substrate and cover types. The individual habitat predictions, as well as the composite habitat map, were translated into the Coastal and Marine Ecological Classification (CMECS) standard to meet Federal Geographic Data Committee (FGDC) requirements (Tables 2.1 and 2.2; CMECS, 2017).



*Live hard corals, soft corals, sponges and algae on the insular shelf. Squirrelfish (*Holocentrus *spp.) and Gray angel fish (*Pomacanthus arcuatus*) look on. Credit. T. Battista and B. Costa.*

Table 2.1. Substrate and biological cover types used to describe benthic habitats across the insular shelf. Equivalent CMECS classifications are suggested.

Habitats			Definition	CMECS IDs	CMECS Class
SSubstrate	**1**	Coral Reef	Highly rugose, hard bottom comprised of live scleractinian (hard) corals and/or dead corals. Typically colonized by algae, sponges and soft coral.	310, 421, 501	Coral Reef Substrate (Substrate) & Mixed Shallow/ Mesophotic Coral Reef Biota (Biotic) & Colonized Shallow/ Mesophotic Reef (Biotic Group)
SSubstrate	**2**	Pavement	Flat, low-relief or sloping hard bottom comprised of live hard corals and/or dead corals with little or no fine-scale rugosity. Typically colonized by algae, sponges and soft coral.	137, 310	Pavement Area (Geoform) & Coral Reef Substrate (Substrate)
SSubstrate	**3**	Rhodoliths	Rhodoliths are cylindrical or irregularly shaped calcareous nodules typically ranging from 3-15 cm in diameter (Foster 2001). They are formed by coralline red algae, and are often found on top of sand and colonized by macroalgae.	309	Rhodolith Substrate (Substrate Subclass)
SSubstrate	**4**	Sand	Sand with particle sizes range from 0.0625 millimeters to < 2 millimeters (Wentworth, 1922). Bare or colonized by sparse algae. Sand was considered to be present when it occurred in ≥20% of the camera's field of view.	279	Sand (Substrate)
SSubstrate	**5**	Artificial*	Man-made structures including derelict fishing gear and shipwrecks.	362	Anthropogenic Substrate (Substrate)
Cover	**1**	Live Hard Coral	Live scleractinian (hard) coral. Hard corals often had plating morphologies.	421	Mixed Shallow/Mesophotic Coral Reef Biota (Biotic)
Cover	**2**	Live Soft Coral	Live soft coral. Mainly gorgonians including sea whips, sea rods, sea fans and sea plumes.	849, 873	Soft Coral Colonized Shallow/ Mesophotic Reef (Biotic), Attached Soft Corals (Biotic)
Cover	**3**	Macroalgae**	Macroalgae that is fleshy and may be green, brown or red. Macroalgae was considered to be present when it occurred in ≥20% of the camera's field of view.	433	Benthic Macroalgae (Biotic Subclass)
Cover	**4**	Sponges**	Sponges. May be rope, encrusting, tubular, or barrel shaped.	517	Attached Sponges (Biotic Group)
Cover	**5**	Unknown	Habitat unknown because of poor video and/or multibeam data quality.	No Equivalent	NULL

* Habitat class too rare for successful model runs, and not used to create the composite habitat map.
** Habitat class too prevalent for successful model runs, and not used to create the composite habitat map.

Table 2.2. The five habitat classes that make up the composite habitat map for the insular shelf. These five classes were chosen by clustering the substrate and cover types from the annotated underwater videos. Equivalent CMECS classifications are suggested.

Code	Habitat	Definition	CMECS ID (Geoform or Substrate)	CMECS Class (Geoform or Substrate)	CMECS ID (Biotic)	CMECS Class (Biotic)
1	Coral Reef Colonized with Live Coral	Coral reef colonized by live hard corals. Soft corals, algae and sponges also likely to be present.	310	Coral Reef Substrate	421, 873, 433, 565, 560, 561	Mixed Shallow/Mesophotic Coral Reef Biota, Attached Soft Corals, Benthic Macroalgae, Turf Algal Bed, Coralline/Crustose Algal Bed, Filamentous Algal Bed
2	Pavement Colonized with Live Coral	Pavement colonized by live hard corals. Soft corals, algae and sponges also likely to be present.	137, 310	Pavement Area and Coral Reef Substrate	421, 873, 433, 565, 560, 561	Mixed Shallow/Mesophotic Coral Reef Biota, Attached Soft Corals, Benthic Macroalgae, Turf Algal Bed, Coralline/Crustose Algal Bed, Filamentous Algal Bed
3	Rhodoliths with Macroalgae	Rhodoliths often colonized by fleshy macroalgae.	309	Rhodolith Substrate	433	Benthic Macroalgae
4	Bare Sand	Sand with no biological cover.	279	Sand	NULL	NULL
5	Rhodoliths with Macroalgae & Bare Sand	Rhodoliths often colonized by fleshy macroalgae with patches of bare sand.	309, 279	Rhodolith Substrate, Sand	433, NULL	Benthic Macroalgae, NULL

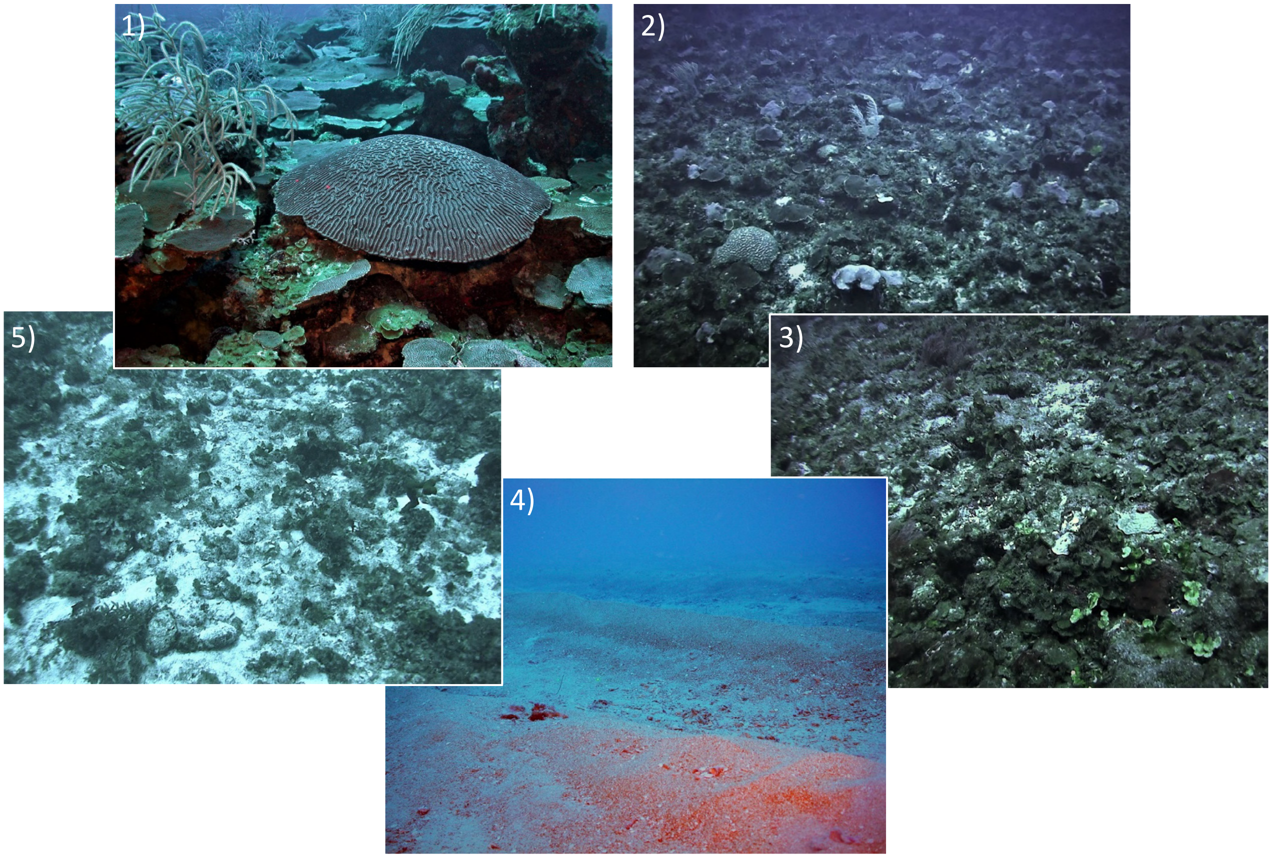


Figure 2.1. Five benthic habitats mapped throughout the insular shelf. 1) Coral reef colonized with live coral, 2) Pavement colonized with live coral, 3) Rhodoliths with macroalgae, 4) Bare sand, and 5) Rhodoliths with macroalgae and bare sand.

2.2 ENVIRONMENTAL PREDICTORS

Twenty environmental predictors were used to create models and spatial predictions for the individual substrate and cover types in the region. These environmental predictors described spatial patterns related to the oceanography, geography, and topography of the seafloor on the insular shelf. The genesis and synthesis of these environmental predictors is described in detail below.

2.2.1 Seafloor Predictors

Seafloor depth, roughness, hardness and topography are known to be useful predictors for identifying specific habitat types such, as sand, pavement, and coral reefs (Costa et al., 2009; Costa and Battista, 2013). SoNARs can collect depth data (i.e., bathymetry), which can be analyzed to describe the topography of the seafloor. Some SoNARs can also collect backscatter data, which describe the roughness and hardness of the substrate on the seafloor. Bathymetry and backscatter were collected on the insular shelf between 2003 and 2011 by several organizations using a variety of SoNAR systems (Figure 2.2 and 2.3). NOAA NCCOS collected depth and backscatter data using the Reson 8101 Extended Range, Reson 7125, and Konsberg Simrad 1002 multibeam echosounders (NOAA NCCOS, 2017). USGS and Geophysics GPR International collected depth information using phase differencing bathymetric systems, including the Geoswath and SwathPlus systems (GPR, 2003; ten Brink and Andrews, 2009).

Since the insular shelf was mapped by various organizations using different sensors, the spatial resolution and horizontal reference systems for bathymetry and backscatter varied across the project area. These datasets were standardized by converting them to the North American Datum 1983 (NAD83) Universal Transverse Mercator (UTM) Zone 20 North horizontal coordinate system (performed using ArcGIS 10.4 Reproject tool; ESRI, 2016). The bathymetry was also referenced to Mean Lower Low Water (MLLW) vertical tidal datum. The spatial resolutions of the bathymetry and backscatter datasets varied from 0.5x0.5 m to 8x8 m, and were standardized to 11x11 m (performed in CARIS HIPS and SIPS and Geocoder software). This spatial resolution was chosen because ±11 m was the estimated, maximum positional uncertainty associated with our underwater video sites (see Section 2.3).

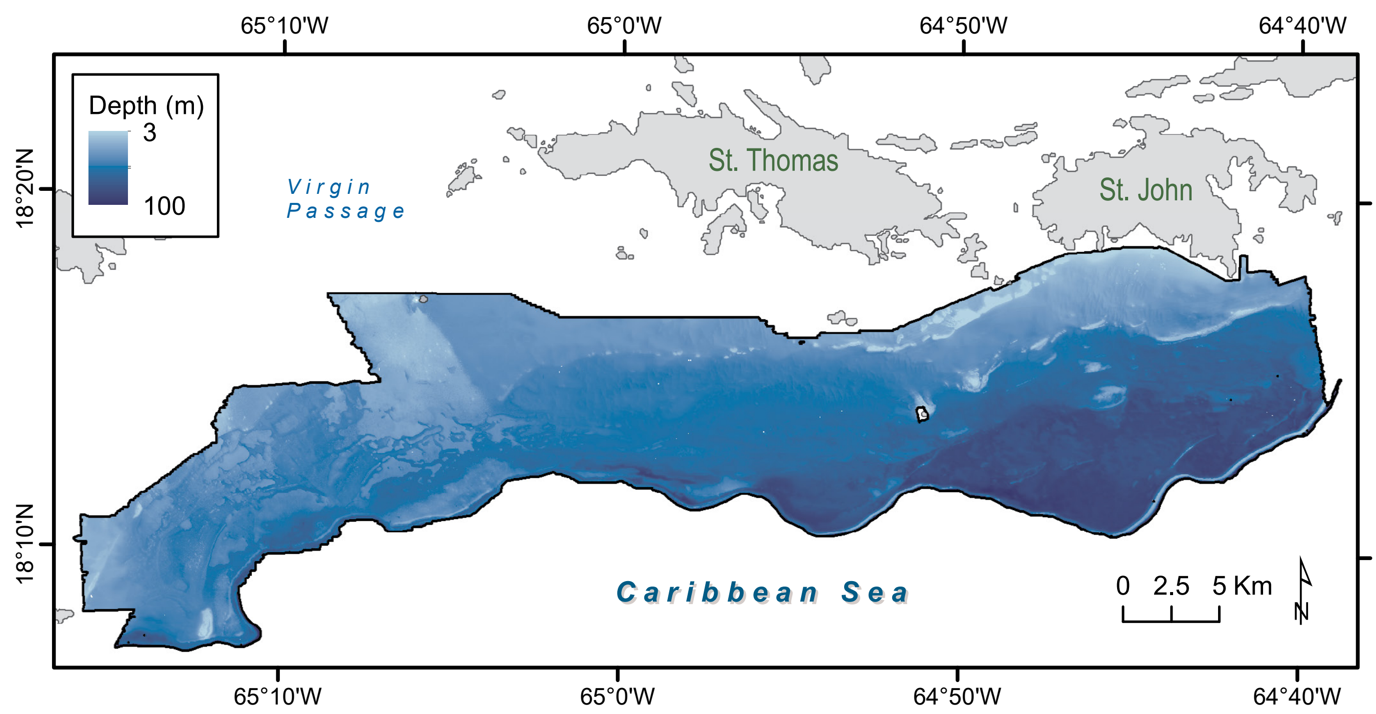


Figure 2.2. Depth data (bathymetry) for the insular shelf.

While these datasets would have allowed for finer resolution mapping, resampling these layers to 11x11 m allowed us to be confident that the underwater videos were located in the correct pixel, and that the habitats seen in the underwater videos were being associated with the correct seafloor information. Associating a habitat type with the wrong information about the seafloor would have degraded the accuracy of the final habitat products. Once standardized, the depth layers were merged and clipped to the 100 m depth contour to create a single, unified depth layer for the project area (Figure 2.2). The backscatter layers were not included because they were collected by systems (i.e., MBES versus PDBS) with different frequencies, acquisition parameters and geometries. These differences meant the quality of the backscatter data varied widely across the insular shelf, and was especially poor where PDBS were used (Figure 2.3). For these reasons, backscatter was excluded from our modeling workflow.

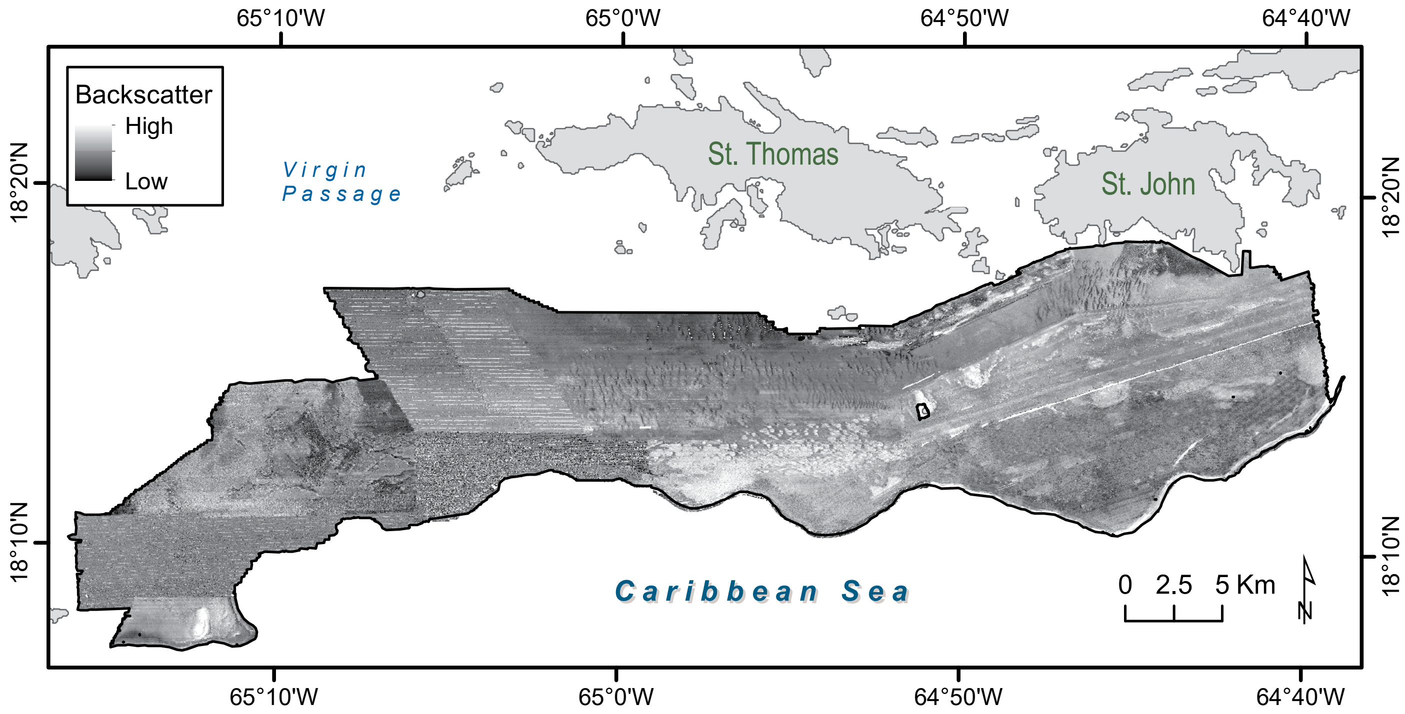


Figure 2.3. Backscatter data (intensity) for the insular shelf. Higher values generally indicate harder and/or smoother bottom types (and vice versa).

The unified depth layer was critical to the benthic habitat modeling process. This layer was used to derive the predictors describing the topographic complexity of the seafloor. These topographic predictors included: (1) Depth (Standard Deviation), (2) Total Curvature, (3) Plan Curvature, (4) Profile Curvature, (5) Rugosity, (6) Slope, and (7) Slope Rate of Change (Figure 2.4). Each topographic surface was calculated using a square 3x3 cell neighborhood, where the pixel in the center of the neighborhood was assigned the calculated value (performed using ArcGIS 10.4: DEM surface tools, see Jenness, 2016). These topographic predictors were chosen because previous research has shown their utility for characterizing benthic and essential fish habitats (Pittman et al., 2009; Pittman and Brown, 2011; Costa et al., 2009; Costa et al., 2013; Costa et al., 2014; Diesing et al., 2014; Hasan et al., 2014).

2.2.2 Oceanographic Predictors

Four oceanographic predictors were developed and included in the modeling process. These predictors were included to account for environmental conditions in the water column that may affect or be correlated with the distribution of benthic habitats, particularly live hard and soft corals. They included: (1) euphotic depth, (2) turbidity (at 547 nm), (3) sea surface temperature anomalies frequency (SSTA), and (4) thermal stress anomalies frequency (TSA). Euphotic depth and turbidity were derived from Aqua MODIS 4x4 km annual climatologies

Methods

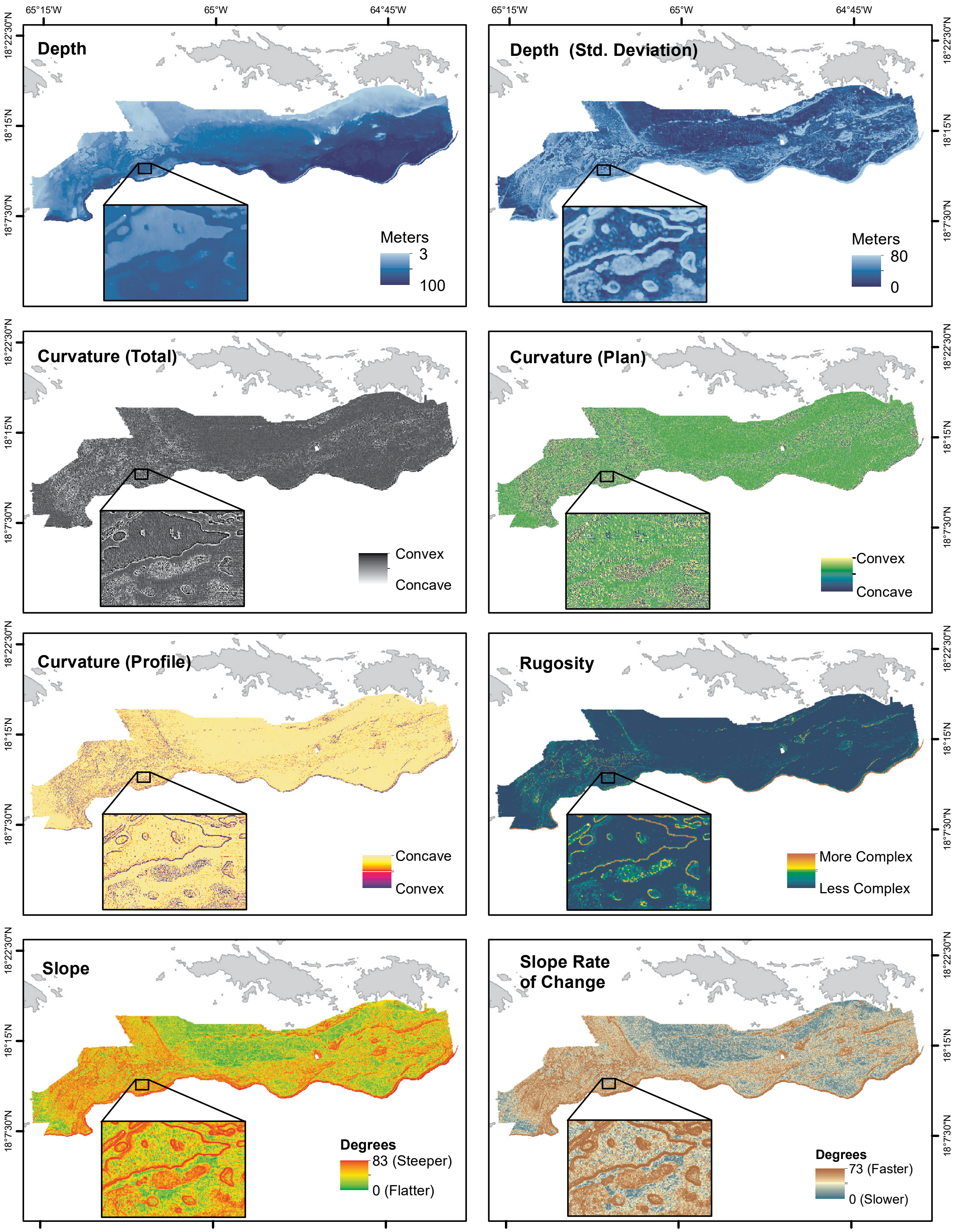


Figure 2.4. Maps depicting the depth and topography of the seafloor. These maps were used as predictors to create the substrate and cover models and spatial predictions.

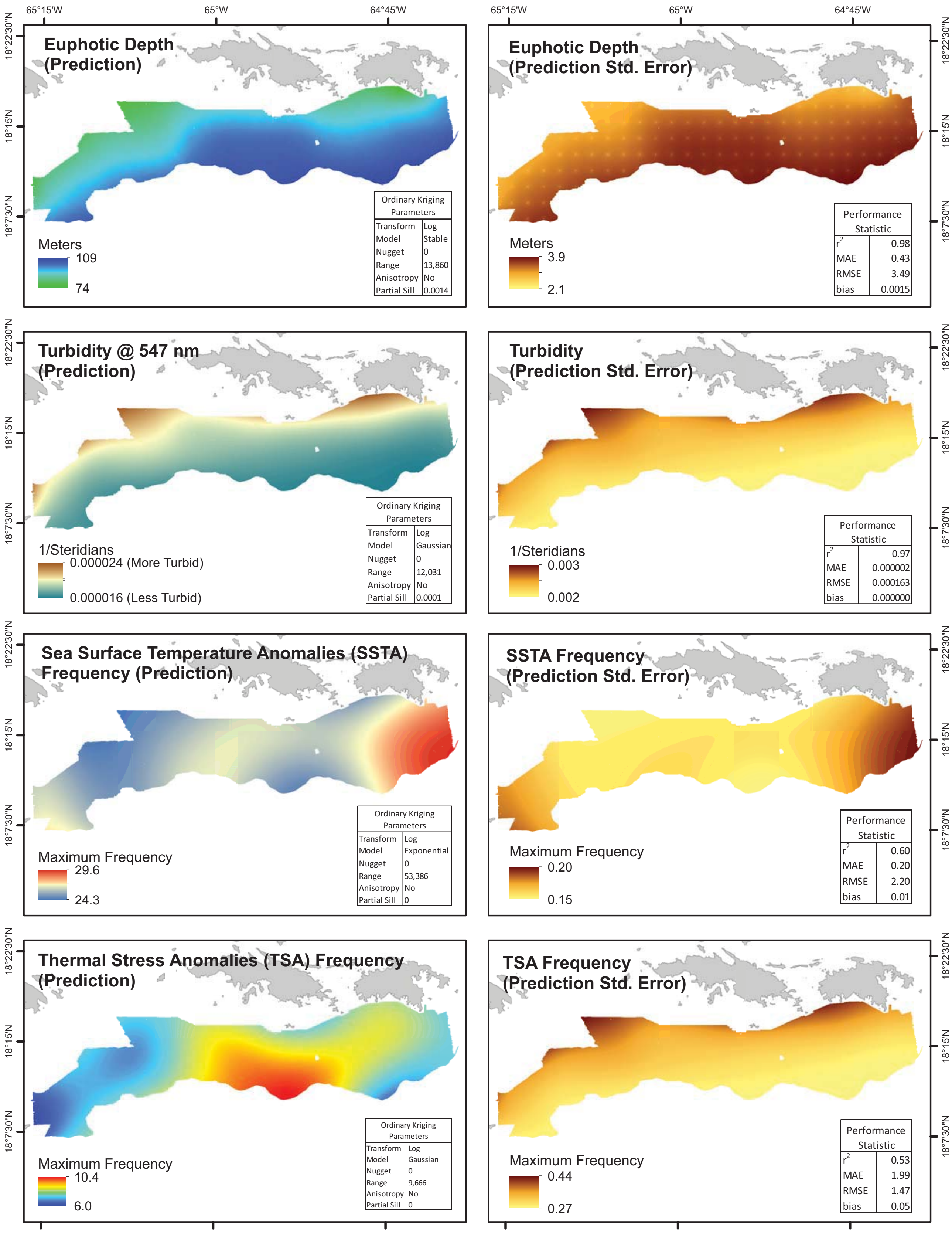


Figure 2.5. Maps depicting broad oceanographic conditions on the insular shelf. These maps were used as predictors to create the substrate and cover models and spatial predictions. The maps on the left contain the kriging parameters and prediction for each climatology. The maps on the right depict the associated prediction error layers and performance statistics of these krigged layers.

downloaded from NASA's Ocean Color website (NASA, 2016). SSTA and TSA surfaces were derived from AVHRR Pathfinder ~4x4 km annual climatologies downloaded from NOAA's Coral Reef Temperature Anomaly Database (CoRTAD) version 5 (Casey et al., 2015). The time period of these annual climatologies ranged from 2004-2014, roughly matching the dates when SoNAR data and underwater videos were collected in the project area. The 10 annual climatologies were averaged (for each environmental predictor) to produce single climatologies for euphotic depth, turbidity, SSTA and TSA. Ordinary kriging was used to resample these single climatologies from ~4x4 km to 11x11 m (performed using ArcGIS 10.4 Geostatistical Extension; ESRI, 2016; Figure 2.5). The performance of these predicted layers was evaluated using cross validation during the kriging process. This validation suggested that these krigged layers were reasonable representations of broad oceanographic conditions on the insular shelf.

2.2.3 Geographic Predictors

Four geographic predictors were used to account for spatial variation in benthic habitats that was not explained by the other environmental predictors. These included distance to shelf edge, distance to shore, latitude (y), and longitude (x) (Figure 2.6). They were created at a spatial resolution of 11x11 m to match the spatial resolution of the other predictors (performed using ArcGIS 10.4 Euclidean Distance tool and Marine Geospatial Ecology Tools 0.8a64; MGET, 2016). The shoreline was extracted from previous NCCOS habitat maps (Kendall et al., 2001). The shelf edge was defined as the 100 m depth contour, and was extracted from the unified depth layer (Figure 2.2). This depth contour was chosen because depths and the associated benthic habitats change rapidly around 100 m on the insular shelf (Smith et al., 2016).

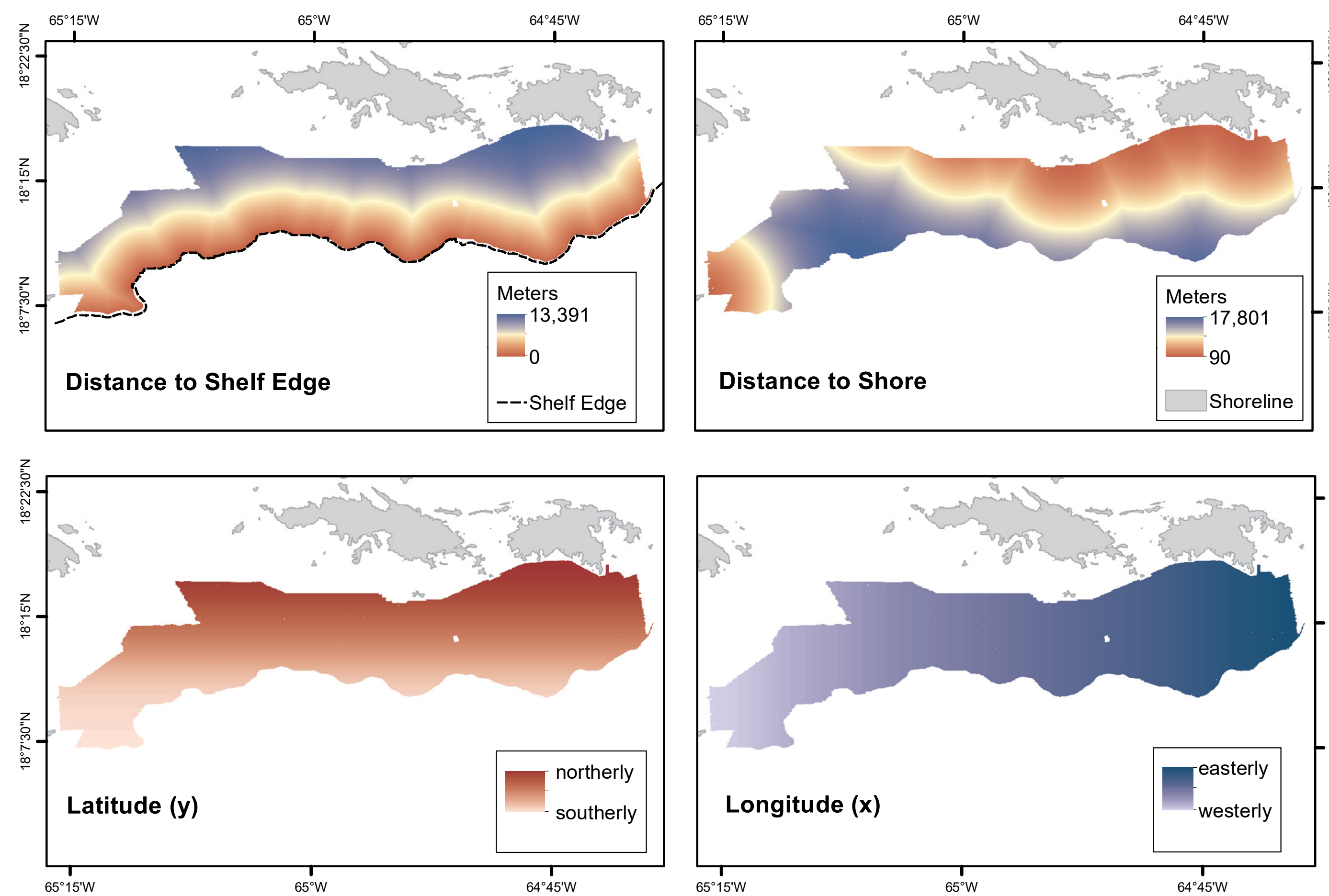


Figure 2.6. Maps depicting the geographic predictors used to create the substrate and cover models and spatial predictions. The shelf edge is defined as the 100 m depth contour. The shoreline for the region was extracted from NOAA's 2001 benthic habitat maps (Kendall et al., 2001).

2.3 FIELD DATA

Underwater video was used to identify and document substrate and biological cover types (Table 2.1) at 1,353 locations on the insular shelf. At each site, a Seaviewer high definition (HD) underwater drop camera was lowered from a small boat to approximately one meter above the seafloor. High definition video was collected at 30 frames per second and 1080x1920 resolution. A portion of these underwater videos (n=1,005) were used to develop the habitat predictions and the composite habitat map. A separate subset of these underwater videos (n=348) were used to assess the performance of the habitat predictions and accuracy of the composite habitat map. More information about the collection of these datasets is provided below.

2.3.1 Ground Validation (GV) Training Dataset

GV data are the basis for finding relationships between the observed substrate and cover types and the values in the environmental predictor datasets. Ultimately, the GV data were used to train and optimize the mathematical models that were used to predict habitats throughout the region. Locations of the GV sites were selected manually by visually reviewing the predictors, and selecting locations that included the full range of habitats, depths, and environmental settings found in the region. Underwater HD videos were collected at 688 GV sites between August 20 and September 2, 2016. This dataset was combined with standard definition drop camera video data at 104 sites collected during previous NCCOS field efforts in 2010, and 213 sites in 2011 from the NOAA ship *Nancy Foster*. In all, GV data from 1,005 GV sites were used to develop and optimize the habitat models for the insular shelf (Figure 2.7).

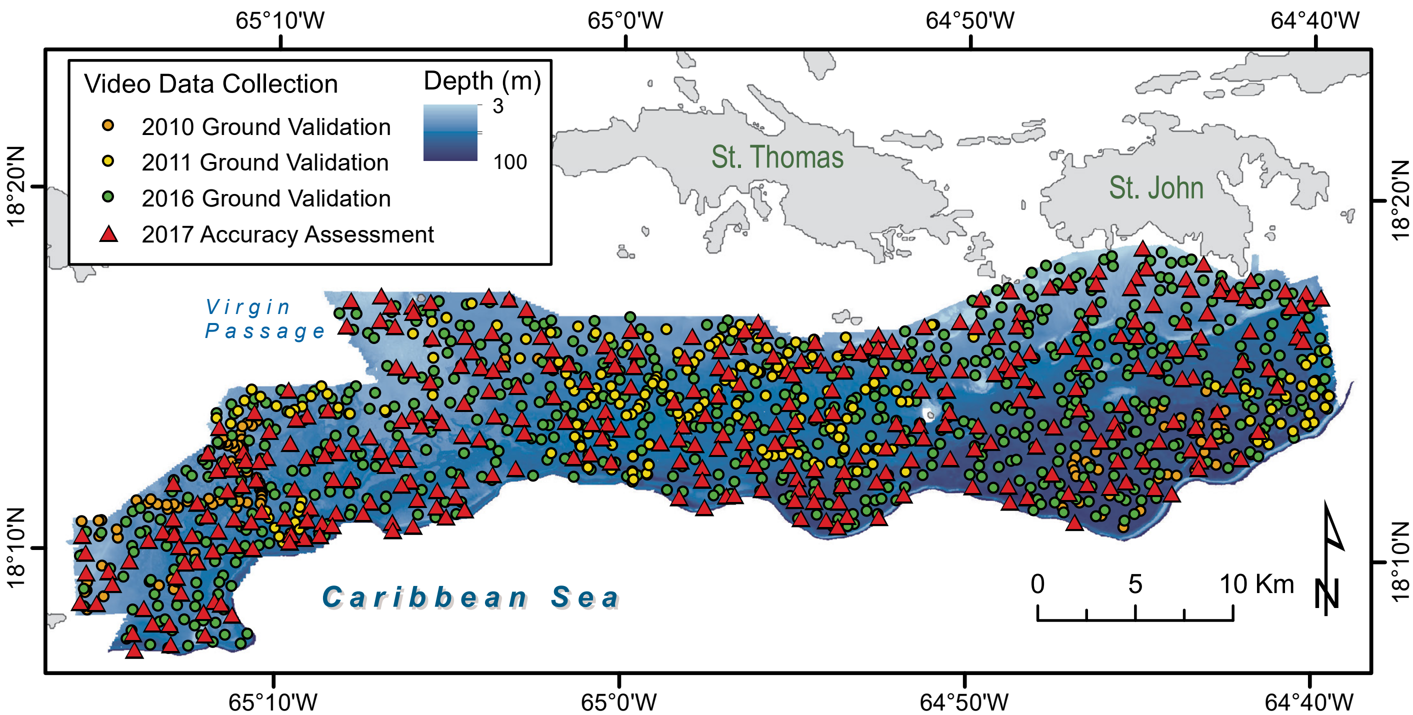


Figure 2.7. Locations of GV and AA underwater videos used to develop and evaluate the habitat map products.

2.3.2 Accuracy Assessment (AA) Test Dataset

AA data were used to independently evaluate the performance of the habitat predictions and the accuracy of the composite habitat map. Underwater HD videos were collected at 348 sites from January 30 to February 13, 2017 (Figure 2.7). The locations of these sites were randomly assigned within six habitat strata based on a preliminary draft habitat map (performed using ArcGIS 10.4 Create Accuracy Assessment Points Tool; ESRI, 2016). The purpose of this stratification was to ensure that at least 30 sites were randomly distributed in each habitat class. Additional sites were randomly assigned in habitats that were more common (i.e., covered larger amounts of area) on the insular shelf. All sites had a minimum distance of 300 m between them to ensure statistical independence. The distance threshold was identified using variograms. Sites were also limited to depths shallower than 80 m (i.e., the depth limit for the drop camera system). This limitation had minimal impact on the accuracy assessment, since 0.2% (1.3 km^2) of the project area was deeper than 80 m.

2.3.3 Field Data Collection

The process for collecting both the GV and AA data was identical at each site in the field. Sites were accessed via small boat using a hand-held Garmin 76 global positioning system (GPS) unit. Once the drop camera was deployed, the boat's location was recorded every five seconds using a Trimble GeoXH GPS receiver for the duration of the underwater video (Figure 2.8). The presence and absence of the five substrate and four cover types were noted at each site. Once back in the office, GPS data were post-processed and differentially corrected using the closest Continually Operating Reference System (CORS) station, which were on St. Thomas (station STVI), Culebra Island, Puerto Rico (station CUPR) and St. Croix (station CR01). All underwater videos were reviewed a second time by benthic experts and the presence (1) and absence (0) of each substrate and cover type were confirmed in the office. Multiple substrate and cover types were often present at each site.

Figure 2.8. Equipment used to collect GV and AA data in the field: (left) small boat used to access sites, (middle) Seaviewer HD video drop camera setup, and (right) topside system, including a GPS, monitor and video recording.

2.4 PREDICTING AND CLASSIFYING BENTHIC HABITATS

Boosted regression trees (BRTs) and boosted classification trees (BCTs) model complex ecological relationships by developing many (hundreds to thousands) simple classification or regression (tree) models. Classification and regression trees (Breiman et al., 1984) relate a response (i.e., habitat type) to environmental predictors by iteratively splitting the data into two homogenous groups. These models are built in a stage-wise fashion, where existing trees are left unchanged and the variance remaining from the last tree is used to fit the next one. These simple models are then combined linearly to produce one final combined model (Friedman, 2002; Elith et al., 2006; Elith et al., 2008).

Separate spatial predictions were developed using BRTs for each substrate and cover type. BCTs were used to combine these individual predictions into a single composite habitat map. Three main steps were used to create habitat predictions and a composite habitat map for the insular shelf: (1) preparing the model input data, (2) creating and evaluating substrate and cover models and spatial predictions, and (3) creating and evaluating a composite habitat map (Figure 2.9). This work was conducted primarily in ArcGIS 10.4 (ESRI, 2016) and R software (R Core Team, 2016) using the dismo (Hijmans et al., 2014), caret (Kuhn, 2016), and raster (Hijmans, 2014) packages.

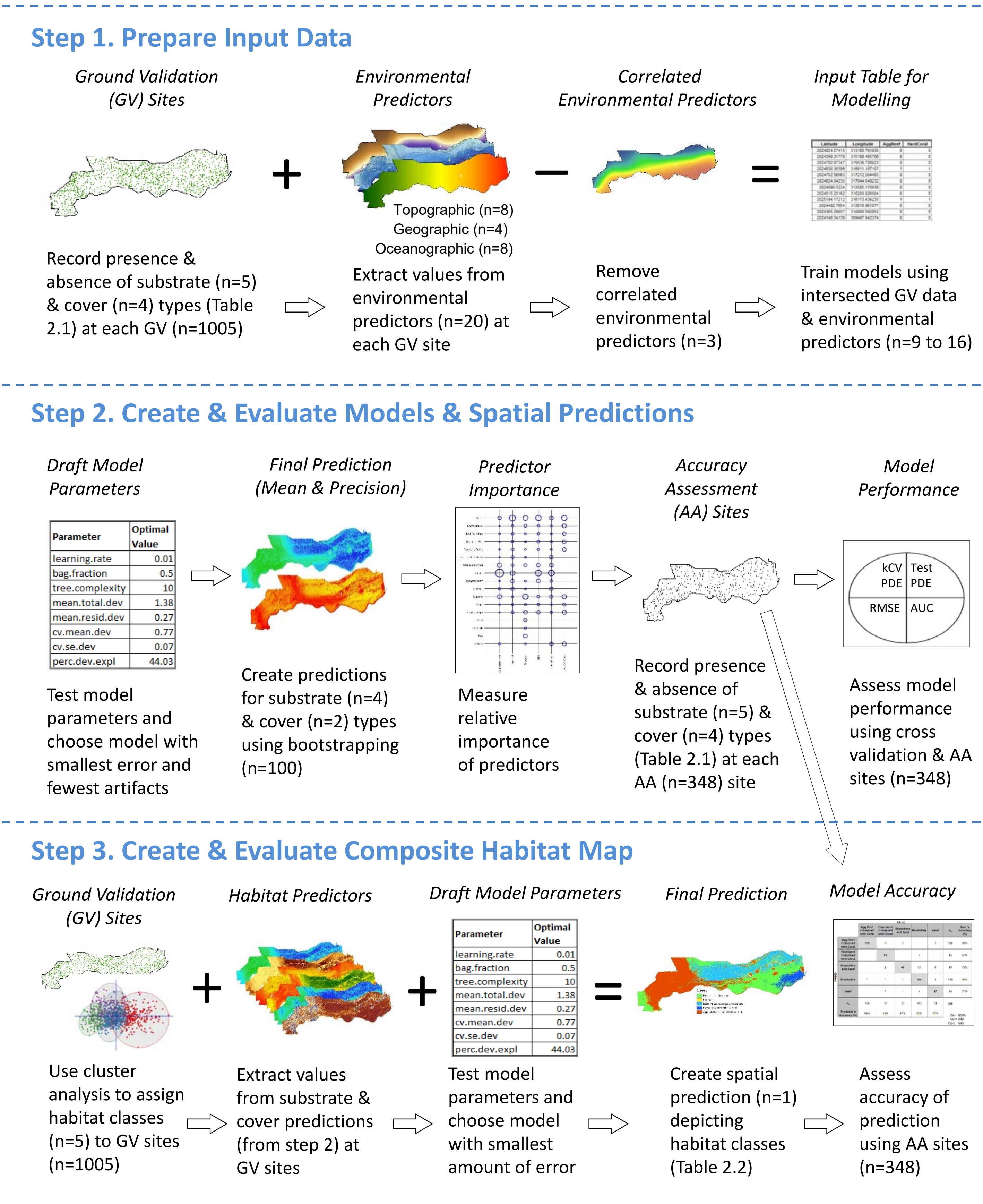


Figure 2.9. Diagram depicting steps in modeling process to predict substrate and cover distributions and develop a composite habitat map.

Methods

Step 1 – Prepare Input Data. The presence and absence of the five substrate and four cover types based on the ground validation data was used as the response variable in the BRT modeling process. This binary response variable was modeled using a binomial (two groups) distribution. No transformations were applied. All of the original 20 environmental (i.e., oceanographic, geographic, and topographic) predictors were numeric. We conducted pairwise testing to identify predictors that were highly correlated using Spearman's Rank correlation coefficient (i.e., $\rho \geq 0.95$ or $\rho \leq -0.95$). Two predictors (TSA Frequency Std. Error and Turbidity Std. Error) met this criterion and were removed. An additional predictor (Euphotic Depth Std. Error) was also removed from the modeling process because it contained regularly spaced artifacts (see Figure 2.5) that may have negatively affected the model outputs. The GV sites were intersected with the remaining 17 predictors to extract their value at each location. This spatial intersection combined the GV and predictor datasets into a single table.

Step 2 – Create Models and Spatial Predictions. The BRT models were fit and optimized in R (R Core Team, 2016) using the dismo package (Hijmans et al., 2014). Several model parameters were tested during this process, including the learning rate (*lr*), tree complexity (*tc*), and bag fraction (*bf*). Learning rate (*lr*) controls how much each tree contributes to the model. The faster the learning rate, the more each tree contributes to the model. Tree complexity (*tc*) dictates how many nodes (splits) there are in a tree. The greater the number of splits, the more complex the model. Bag fraction (*bf*) specifies the proportion of data randomly chosen at each step. The larger the bag fraction, the more data available to train the model at each step.

For each of the five substrate and four cover types, we tested 36 combinations of *lr, tc,* and *bf* (Table 2.3). We used *k*-fold cross validation (kCV) to identify the combinations of *lr, tc,* and *bf* that created the model with the smallest amount of error. Here, the kCV process divided the input table into 10 folds (i.e., 10 data subsets). Nine of these folds were used to create models, while the remaining one was used to evaluate the model's performance. This process was repeated 10 times (i.e., one time for each fold) x 36 model parameter combinations x 9 substrate and cover types (n=3,240 models total). We measured model performance using the percent deviance explained (PDE) averaged across the 10 folds. PDE is the amount (%) of variation explained in the response data. PDE values normally range between 0 and 100% with higher values indicating better model performance and lower error.

Table 2.3. Suite of boosted regression tree (BRT) model parameters and values tested.

Regularization Parameters	Parameters Tested	Definition	Impact	Example
Learning Rate (*lr*)	0.01, 0.001, 0.005	Determines contribution of each tree to the growing model	Decreasing (slowing) *lr* increases the number of trees required for optimal prediction	*lr* = 0.005 will grow more trees than *lr* = 0.01
Tree Complexity (*tc*)	2, 3, 4, 5, 10, 15	Controls the number of decision nodes in a tree	Decreasing *tc* will reduce the size (number of nodes) in a tree	*tc* = 15 will grow larger trees (with more nodes) than *tc* = 2
Bag Fraction (*bf*)	0.5, 0.75	Controls proportion of data randomly selected to build each tree	Decreasing *bf* will reduce the number of points randomly used to build a tree	*bf* = 0.5 will randomly sample 25% fewer data points than *bf* = 0.75

The model with the highest kCV PDE and the fewest visible artifacts in the predicted surface was selected for each substrate and cover type. Environmental predictors were dropped from certain substrate or cover models because they introduced artifacts that reduced the models' kCV PDE and degraded the overall prediction. As a result, the number of predictors included in each substrate and cover model varied, ranging from nine for the 'Soft Coral' model to sixteen for the 'Rhodoliths' model. The relative importance of these remaining predictors (Elith et al., 2008) was quantified and presented using a bubble plot (Section 3.2). In other cases, the prevalence of a substrate or cover type was too low (e.g., <0.2% for 'Artificial') or too high (e.g., >93% for 'Macroalgae' and 'Sponges') to be successfully modeled using BRTs. Consequently, spatial predictions were note created for these three substrate and cover types, and they were excluded from the composite habitat map.

The best models for the remaining four substrate and two cover types were used to predict their spatial distribution across the insular shelf (performed using the raster package in R; Hijmans, 2014; R Core Team, 2016). These predictions describe the probability-of-occurrence for each habitat, i.e., the likelihood (%) that a particular substrate or cover type is present. Larger probabilities indicate that a habitat is more likely to be present. The precision associated with each probability of occurrence prediction was also quantified and reported as the coefficient of variation (CV). CV is a measure of model precision and represents the standard error as a proportion of the mean (Leathwick et al., 2006). Larger CVs indicate greater uncertainty associated with the spatial prediction. For each substrate and cover type, these probabilities and precisions represent the average of, and variation in, 100 model iterations created using bootstrapping (see Glossary).

Step 2 – Evaluate Performance of Models and Spatial Predictions. The performance of the four substrate and two cover predictions were evaluated using four different metrics: (1) kCV PDE, (2) test PDE, (3) RMSE, and (4) Receiver Operating Characteristic (ROC) Area Under the Curve (AUC). These four metrics were calculated because they describe model performance in different ways, and when viewed together, provide a more thorough understanding of the models limitations. For example, models with higher kCV PDE, test PDE, and AUC values, but lower RMSE, can be used with greater confidence because they correctly explain more variation in the response data with lower amounts of error. Conversely, models with higher kCV PDE and low test PDE may be fit too closely to the response data, and may not generalize well and accurately predict species distributions across the entire study area.

kCV PDE was calculated during *k*-fold cross validation by comparing the observed values (in one randomly chosen validation fold) to the predicted values (from the models developed using the remaining nine training folds). Test PDE, AUC and RMSE were independently calculated using the independent AA dataset (Section 2.3). Test PDE, like kCV PDE, is the amount (%) of variation explained in the response data. PDE values normally range between 0 and 100% with higher values indicating better model performance. Conversely, RMSE measures the error associated with a model by calculating the square root of the average squared difference between the predicted values (extracted from the model) and the observed values (extracted from the underwater videos). Lower RMSE denotes less error.

ROC curves measure a model's predictive performance differently compared to PDE and RMSE. Specifically, ROC curves compare a model's sensitivity (i.e., true positive prediction rate) to its specificity (i.e., true negative prediction rate). This rate depends on the choice of a particular probability of occurrence threshold above which substrate or cover types are classified as "present" and below which they are classified as "absent." Area Under the ROC Curve (AUC) does not require selecting a threshold, and can be used to measure the overall predictive performance of a model (compared to a random guess). AUC values ranging from 0.7 to 0.8 denote "good" model performance; values from 0.8 to 0.9 denote "excellent" model performance, and values greater than 0.9 denote "outstanding" model performance (Hosmer and Lemeshow, 2000). AUC values at or below 0.5 indicate that the model's prediction was no better than one created by chance alone.

Step 3 – Create Composite Habitat Map. BCTs were used to develop a single composite habitat map. The raw GV data (from Step 1) were grouped into five commonly co-occurring substrate and biological cover types using hierarchical cluster analysis. This type of clustering grouped the 1,005 GV sites with similar substrate and cover types into progressively larger and larger clusters (until all sites were in one group). We chose the resulting substrate and cover clusters that were most ecologically relevant for modeling (Table 2.2). We intersected the 1,005 GV sites (each of which were assigned one of the five habitat types) with the 'Coral Reef', 'Pavement', 'Rhodoliths', 'Sand', 'Live Hard Coral', and 'Live Soft Coral' probability of occurrence predictions to extract their value at each location. This spatial intersection combined the GV and probability of occurrence values into a single table.

Next, model parameters for BCTs were fitted and optimized in R (R Core Team, 2016) using the caret package (Kuhn, 2016). We tested 36 combinations of *lr, tc,* number of trees (*n.trees*) and minimum terminal node size (*n.minobsinnode*) (Table 2.4). The *lr* and *tc* parameters are the same as those used to develop BRTs above. Number of trees denotes the number of classification trees that are fitted to the response data. The minimum terminal node size tells the modelling process when to stop splitting the response data and denotes the number of observations (e.g., 5 or 10) for each end point in a classification. *k*-fold cross validation (kCV) was used to identify the combinations of *lr*, *tc, n.trees* and *n.minobsinnode* that created the model with the smallest amount of error. We calculated kCV PDE to identify the parameter combination that created the highest performing model. This highest performing model was then applied spatially to create the composite habitat map throughout the region. Only minor manual editing was needed to improve this composite map. These edits included adjusting a site along the shelf edge that had been misclassified, and editing locations where the MBES data were noisy, creating artifacts in the final habitat map. Approximately 1.5% of the total mapped area was edited manually.

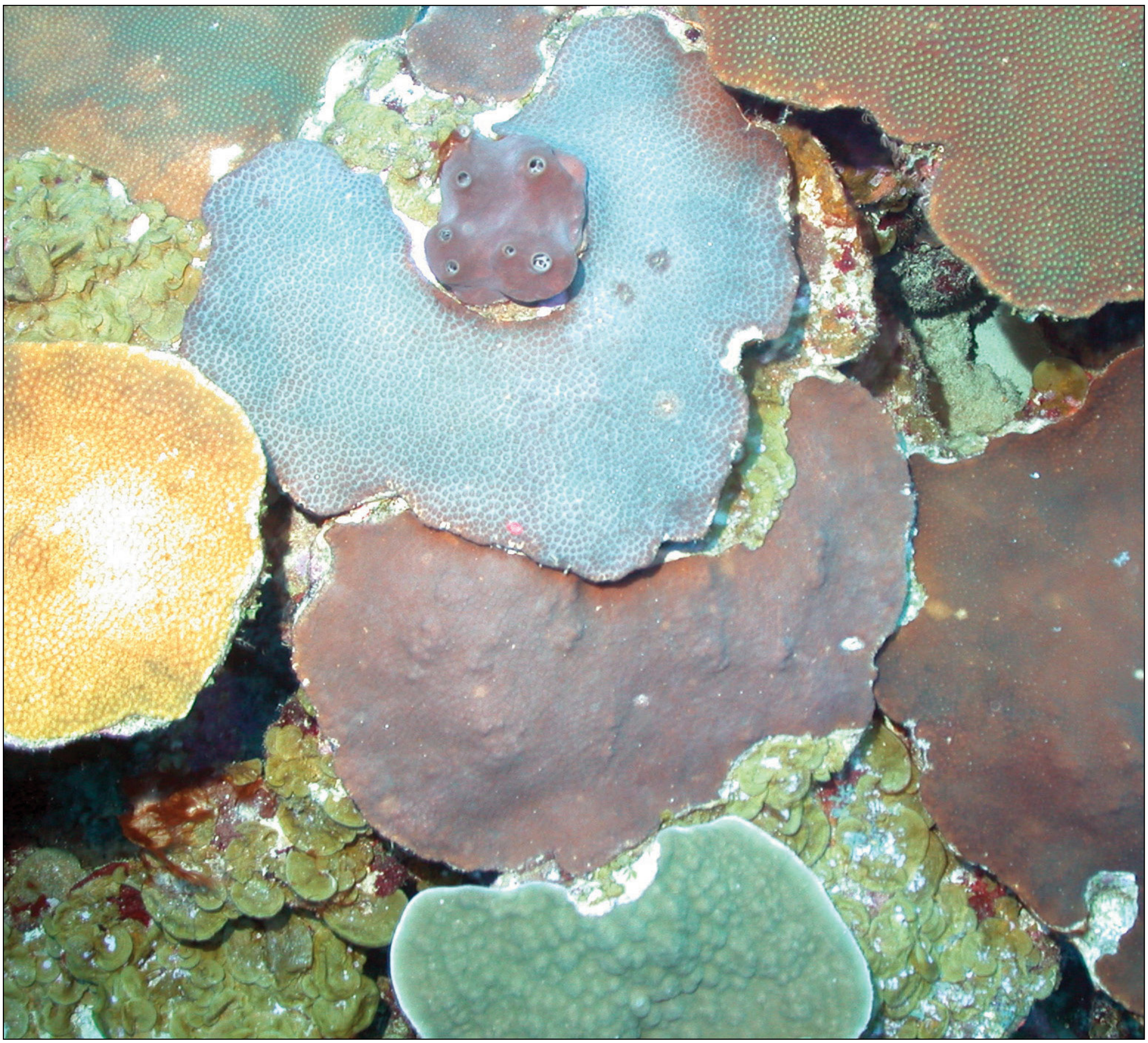
Table 2.4. Suite of boosted classification tree (BCT) model parameters and values tested.

Regularization Parameters	Parameters Tested	Definition	Impact	Example
Learning Rate (*lr*)	0.01, 0.001, 0.005	Determines contribution of each tree to the growing model	Decreasing (slowing) *lr* increases the number of trees required for optimal prediction	*lr* = 0.005 will grow more trees than *lr* = 0.01
Tree Complexity (*tc*)	2, 3, 4, 5, 10, 15	Controls the number of decision nodes in a tree	Decreasing *tc* will shrink the size (number of nodes) in a tree	*tc* = 15 will grow larger trees (with more nodes) than *tc* = 2
Number of Trees (*n.trees*)	500	Describes the number of classification trees that are fitted to the response data	More classification trees will create more complex models (at the risk of overfitting the data)	*n.trees* = 500 will grow 500 classification trees
Minimum Terminal Node Size (*n.minobsinnode*)	5, 10	Describes the number of observations at each endpoint in a classification tree	A lower number of observations will increase the risk of overfitting the model	*n.minobsinnode* = 3 will stop fitting when a classification tree has 3 observations

Step 3 – Evaluate Accuracy of Composite Habitat Map. The thematic accuracy of the composite habitat map was assessed qualitatively by a panel of local experts from NOAA, the NPS, US Virgin Islands DPNR, and the University of the Virgin Islands. It was also evaluated quantitatively by NOAA NCCOS using an independent set of underwater videos (i.e., n=348 AA sites). The AA sites were grouped into the same five habitats identified by the cluster analysis. Sites were considered correct if the same habitat was present within 11 meters (one pixel or raster cell) of the AA site. A matrix was developed from the 348 AA points describing the composite maps' overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA) (Story and Congalton, 1986). This matrix was constructed as a square array of numbers arranged in rows (map classification) and columns (accuracy assessment classification).

The OA was calculated as the sum of the major diagonal (i.e. correct classifications), divided by the total number of AA samples. The PA and UA were calculated to describe the thematic accuracy of individual map categories. PA describes errors due to omission and is a measure of how often habitats were incorrectly excluded from their correct habitat class. UA describes commission errors, and is a measure of how often certain habitats were incorrectly included in another (often similar looking) habitat class. Each diagonal element was divided by the column total (n_i) to yield a producer's accuracy and by the row total (n_j) to yield a user's accuracy. We also calculated the Tau coefficient to account for the random, chance agreement between the AA data and composite habitat map (Ma and Redmond, 1995). The probability of random agreement decreases as the number of habitat classes increases.

While stratification helps ensure all habitat classes are adequately evaluated, it has the undesired effect of introducing bias into the confusion matrix. This bias is due to the different amounts of area (km^2) occupied by each habitat class (Card, 1982), causing rarer habitats (e.g., 'Live Coral') to be sampled at a greater density than common habitats (e.g., 'Sand'). This sampling bias was removed using the method of Card (1982), which uses the proportion (%) of the map occupied by each habitat to correct the thematic accuracies. These proportions were also used to compute confidence intervals for the overall accuracy (Card, 1982; Congalton and Green, 1999). For more information about these calculations and the equations, please see Costa et al., 2009. As a last step, any errors identified during the accuracy assessment were also corrected in the composite habitat map.



Live hard coral, sponges and macroalgae on the insular shelf. Credit: T. Battista and B. Costa.

Invasive red lionfish (Pterois volitans) on the insular shelf. Credit: T. Battista and B. Costa.

3.0 RESULTS AND DISCUSSION

We characterized 652 km^2 of seafloor on the insular shelf. BRTs were used to predict the presence of individual substrate and cover types, and BCTs were used to generate a composite habitat map. In this section, we present the results from these models and map, highlight some of the main features of the habitat predictions, convey performance and accuracy of the maps, and discuss potential applications and ways to customize the products to meet particular research and management needs.

3.1 MODEL PERFORMANCE

Six BRT models and resulting spatial predictions describe the probability-of-occurrence for four substrate and two cover types. Model performance was generally considered 'good to outstanding' based on four evaluation metrics. The predictive model with the highest kCV PDE, test PDEs, and AUCs also had low bias and error ('Coral Reef'). For all the models, kCV PDE ranged from 20.9% to 67.4% ($\bar{x}$ =40.2% ±8.3 SE), and test PDE ranged from -3.9% to 83.3% ($\bar{x}$ =37.4% ±16.2 SE). The 'Coral Reef' model had the highest kCV and test PDEs (67.4% and 83.3%). The 'Live Soft Coral' model had the lowest kCV PDE (18.0%), and the 'Sand' model had the lowest test PDE (-3.9%). AUC values ranged from 0.70 (good) to 0.99 (outstanding) ($\bar{x}$ =0.86 ±0.05 SE) for all the models. The 'Coral Reef' model had the highest AUC (0.99), and the 'Live Soft Coral' model had the lowest (0.70). Bias was small to moderate for all models, ranging between -0.17 to +0.04 ($\bar{x}$ =-0.04 ±0.03 SE). Bias indicates whether the model under predicted (-) or over predicted (+) the probability-of-occurrence. The presence of sand was consistently under-predicted by 17%, and Rhodolith presence was over-predicted by 4%. The presence of Coral Reef, Pavement, and Live Hard Coral showed little systematic bias (≤5%), while Soft Live Coral was under predicted by 6%. Lastly, RMSE values ranged from 0.2 to 0.46 ($\bar{x}$ =0.33 ±0.04 SE). The 'Sand' model had the largest amount of error (0.46), while the 'Coral Reef' model had the least (0.2). Some of the models were affected by the noisy depth data collected in 2009 and in 2004. Linear artifacts oriented east-west can be seen where a SwathPlus PDBS was used for data collection. This noise is the result of the poor performance of the SwathPlus system near nadir (Lurton, 2000). Linear artifacts oriented northeast-southwest can also be seen just south of St. John where a pole-mounted MBES was used for data collection. This noise is the result of the pole vibrating while the ship was collecting data. The linear artifacts in these two areas (Figure 3.1) propagated through the modeling process, and are visible in the spatial predictions and precision surfaces. They are most noticeable in the 'Coral Reef', 'Rhodoliths', 'Live Hard Coral' and 'Live Soft Coral' models.

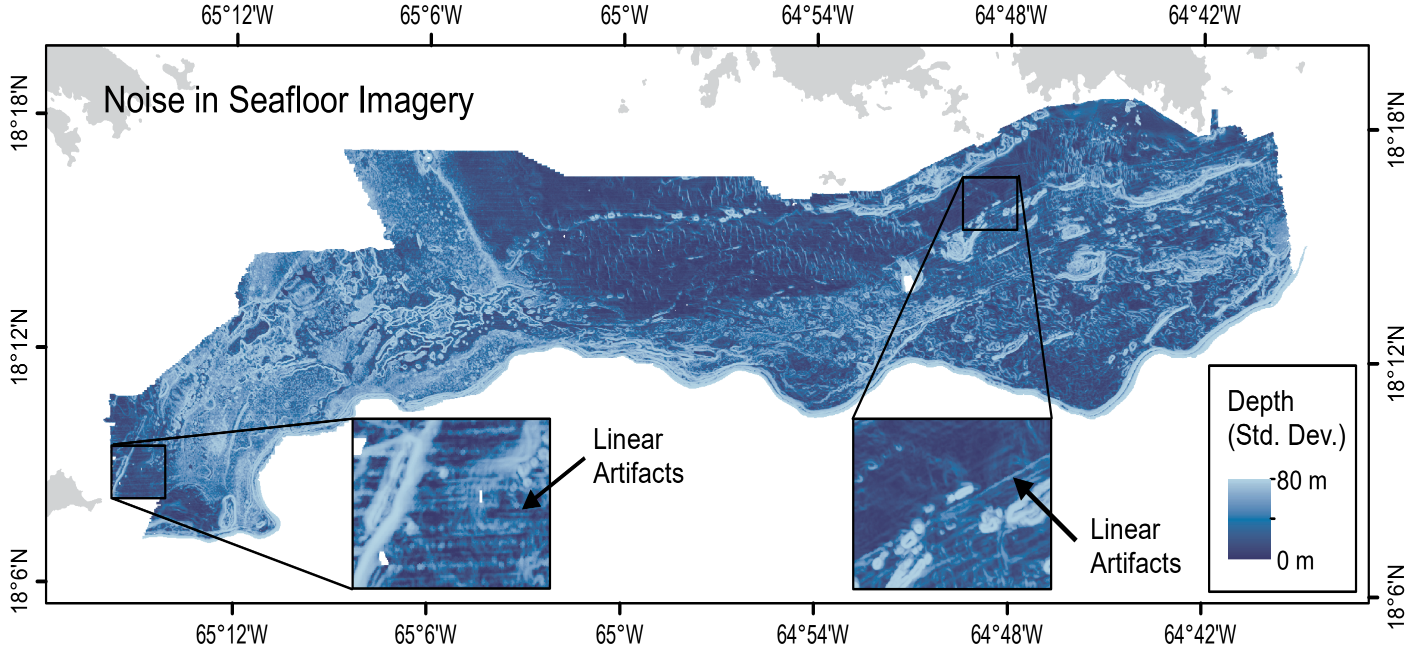


Figure 3.1. Noise (artifacts) visible in the acoustic data. These artifacts affected the accuracy and precision of the spatial predictions.

Results and Discussion

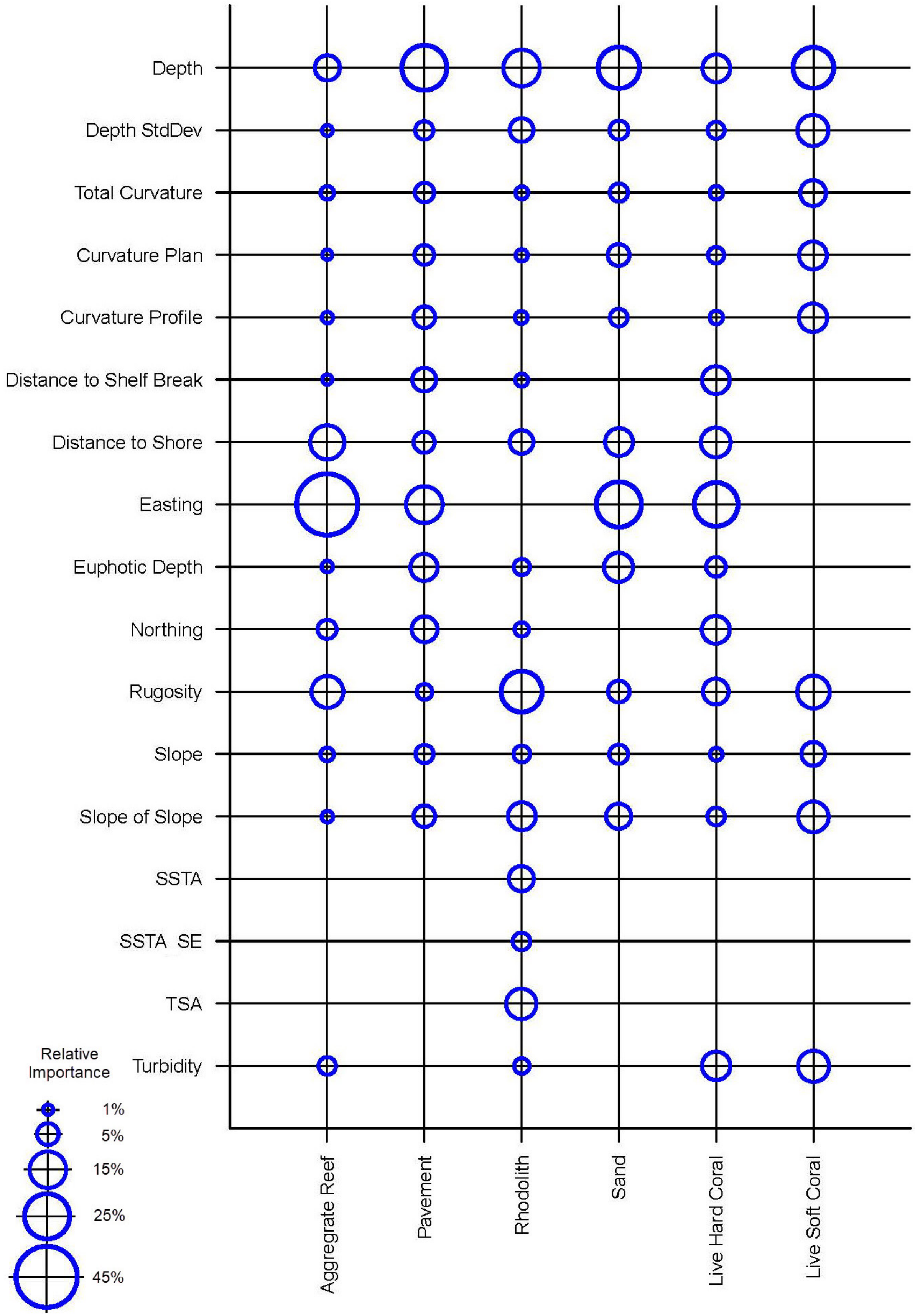


Figure 3.2. Relative importance of the environmental predictors used to develop the six habitat models and predictions. Circle size is proportional to a predictor's relative importance averaged across 100 model iterations. The larger the circle, the more important the predictor. No circle indicates predictor was not used in the model. In these cases, predictors were dropped because they introduced artifacts that degraded the substrate or cover prediction.

3.2 PREDICTOR IMPORTANCE

The relative contribution of each environmental predictor differed among the substrate and cover models (Figure 3.2). Relative contribution (also known as relative importance) describes how often a predictor is used for tree splitting (Elith et al., 2008), and can provide insight into environmental physical drivers that influence the distribution of habitats. Geographic predictors (especially easting) were important for the Substrate models ‘Coral Reef’ (46%), ‘Pavement’ (16%), and ‘Sand’ (25%). Distance to shore contributed to the ‘Coral Reef’ model (14%) and depth was a strong driver in the ‘Pavement’ (25%), ‘Rhodolith’ (16%) and ‘Sand’ (21%) models. Of the Topographic metrics, rugosity was an important contributor to the “Rhodoliths’ (21%) and ‘Coral Reef’ (12%) models. The ‘Live Hard Coral’ model was primarily driven by geographic location, specifically easting (23%) and distance to shore (10%). The ‘Live Soft Coral’ model was largely driven by depth (20%) with predictors related to complexity (rugosity, depth std. dev., and slope rate of change) each contributing >11% to the model. Turbidity was also important for explaining variance in both the ‘Live Hard Coral’ model (10%) and ‘Live Soft Coral’ model (11%). It is important to note that geographic predictors (i.e., easting) are only proxies for other ecological mechanisms that drive habitat distributions, and do not directly explain the ecological forces driving the distribution of the above substrate and cover types.

3.3 GEOGRAPHIC PATTERNS OF SUBSTRATE AND COVER TYPES

Substrate: Coral Reef. 'Coral Reef' (Figure 3.3a) was particularly prevalent in the area southwest of St. Thomas, across Sail Rock Bank, through Virgin Passage, and south to El Seco. Coral Reef was observed at 34% of the GV sites (339/1005) (Figure 3.3b). Other observations were at Hind Bank, Grammanik Bank and French Cap Bank. There were very few observations of coral reef at GV sites in the eastern third of the insular shelf south of St. John. The predicted surface produced from the ‘Coral Reef’ model reflects this same spatial distribution (Figure 3.3e). The highest probability of coral reef occurrence is concentrated in the Virgin Passage and to the south and west of Hind Bank. CV values were lowest (<0.25) in these same locations (Figure 3.3f), indicating higher precision and lower uncertainty for places where coral reef is more likely to be present. An exception to this is at Sail Rock Bank where there is >50% probability of occurrence, but also low precision (CV>1). Precision is high (CV<0.25) across much of the central and eastern regions of the shelf (Figure 3.3f), where there is a very low likelihood of occurrence (Figure 3.3e). The performance statistics for the ‘Coral Reef’ model (Figure 3.3d) show it to be one of the best performing habitat models.

Although this model was one of the top performers, its quality was degraded in certain locations because of noisy depth data. This noise propagated through to the other seafloor predictors (e.g., depth (standard deviation), slope rate of change). This noise is clearly visible in Figures 3.3e and 3.3f on the western and eastern sides of the insular shelf. Map insets in these figures provide a closer look at the noise in the PDBS data, and at the seam between the clean and noisy depth datasets collected in 2005 and 2004, respectively. The noise appears as linear stripes (artifacts) oriented in the east-west, and northeast-southwest. BRTs falsely classified this noise as locations where coral reefs were likely to be present. This result is in keeping with other research showing that the presence of noise in acoustic data will decrease the thematic accuracy of habitat maps, and increase the need for manual editing (Costa and Battista, 2013).

Substrate: Pavement. ‘Pavement’ (Figure 3.4a) was not very common across the insular shelf and was present at 16% (161/1005) of the GV sites (Figure 3.4b). ‘Pavement’ was unevenly distributed throughout the insular shelf, occurring primarily on Tampo Bank and northeast of Tampo. Other instances occur on French Cap Bank, Grammanik Bank, and El Seco. These locations are known spawning aggregation sites for commercially important fish species (Kadison et al., 2017). There are very few occurrences of pavement in Virgin Passage. The ‘Pavement’ model (Figure 3.4c) showed similar spatial patterns, with the highest likelihood of the presence of pavement on Tampo Bank and to the northeast of Tampo (Figure 3.4e). Probability-of-occurrence values were also high on Tampo Bank and to the northeast of Tampo. Precision for the pavement prediction are lowest (CV>2) in the areas where the model predicted high probability of occurrence, and highest in the areas where there is a low probability of occurrence. This pattern suggests that the prediction is more precise where pavement is more likely to be absent (Figure 3.4f).

Substrate: Rhodolith. 'Rhodolith' (Figure 3.5a) was the most commonly observed habitat in the GV data, present at 56% of sites (565/1005) (Figure 3.5b). Rhodoliths were observed primarily in the eastern two thirds of the insular shelf, directly south of the islands. There were scattered occurrences of rhodoliths west of Virgin Passage. This habitat was sparse through the Virgin Passage. The 'Rhodolith' model (Figure 3.5e) showed similar spatial patterns, with the highest likelihood of presence to the south of St. Thomas and St. John including on French Cap and Tampo Banks, and between Sail Rock and Frenchcap Cay. Like for the 'Coral Reef' model, CV values were lowest (<0.25) in these same locations (Figure 3.5f), indicating higher precision for places where rhodoliths were more likely to be present. CV values were also low (<0.25) where probability of occurrence was low (<25%) indicating precision was also high for locations where rhodoliths were likely to be absent. The 'Rhodolith' model was the second best performing model for the insular shelf. Like for the 'Coral Reef' model, noisy depth data degraded the quality of the 'Rhodolith' spatial prediction in the SwathPlus and 2004 MBES acquisition areas (Figure 3.5e and f inset maps).

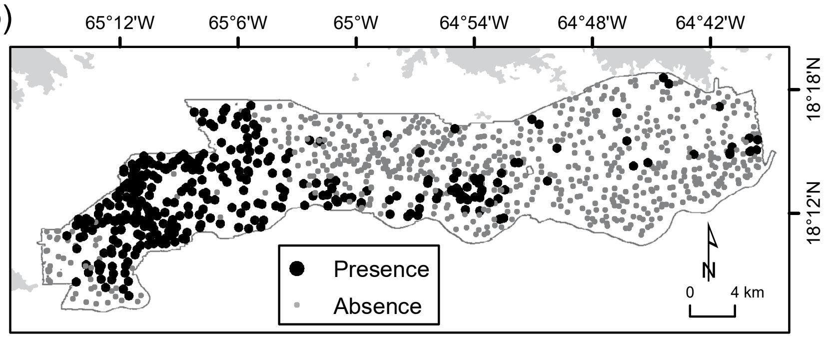
Substrate: Sand. 'Sand' (Figure 3.6a) was prevalent throughout the insular shelf. It was present at 42% (424/1005) of GV sites (Figure 3.6b). Sand was fairly evenly distributed throughout the insular shelf, especially just south of St. Thomas and St. John. However, sand was notably less abundant closer to the shelf edge and just west of Sail Rock Bank. The 'Sand' model (Figure 3.6e) showed similar spatial patterns, with the highest likelihood of sand just south of the islands and between coral reefs on the west side of the insular shelf. The highest precision occurs on French Cap and Tampo Banks and west of Virgin Passage (Figure 3.6f), where the probability of occurrence was lowest (<10%). These patterns suggest that the 'Sand' model is more precise where sand is more likely to be absent.

*Vibrant coral reef community on the insular shelf, including pillar corals (*Dendrogyra cylindrus*). Pillar corals were listed as threatened under the Endangered Species Act in 2014. Credit: T. Battista and B. Costa.*

Substrate: Coral Reef

a)

b)



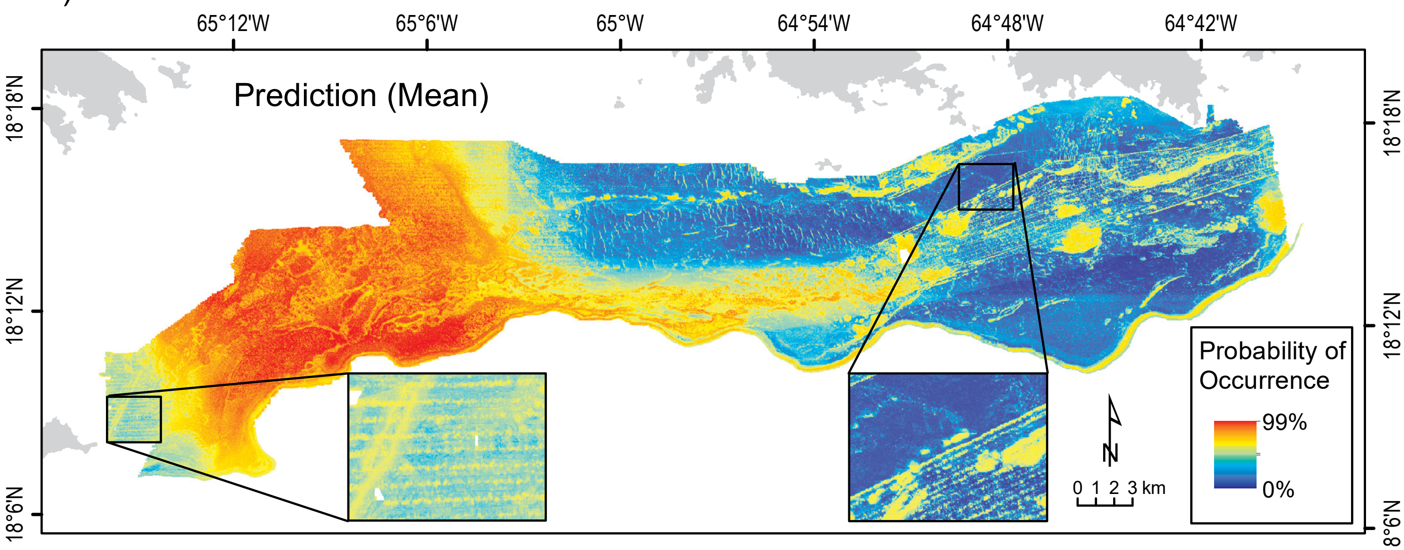
c) ***Model Inputs***

Tree Complexity = 4
Learning Rate = 0.01
Bag Fraction = 0.5
\# Trees = 800

d) ***Model Performance***

CV PDE = 67.4 ±3.0
Test PDE = 83.3
Test AUC = 0.99
Test Bias = -0.03
Test RMSE = 0.20

e)



f)

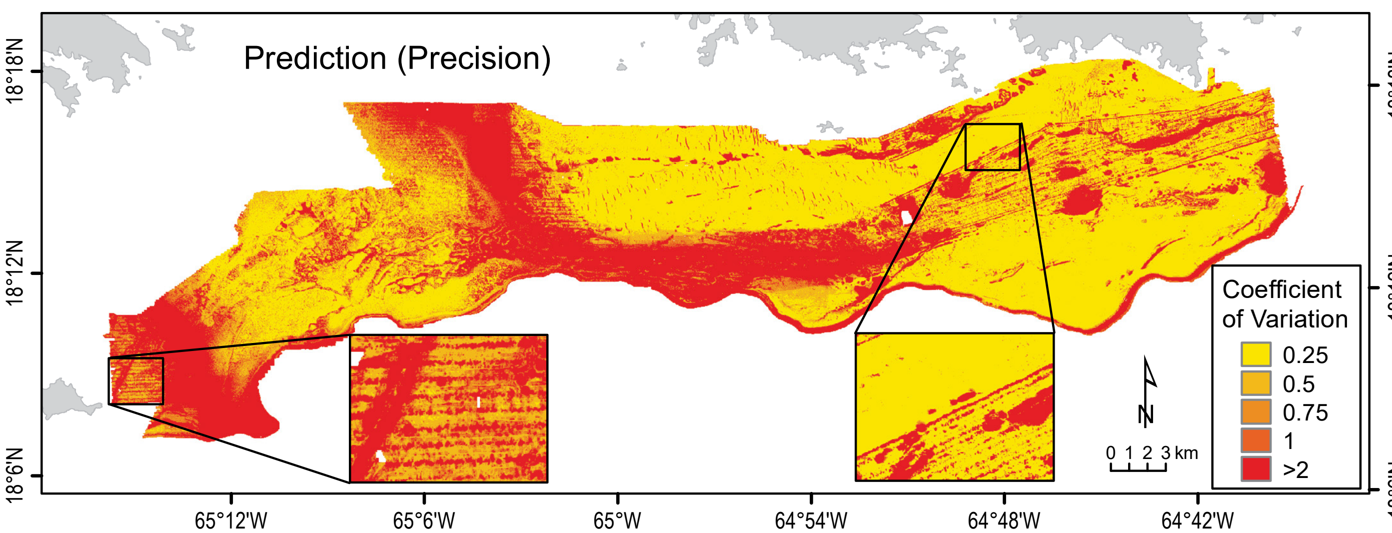


Figure 3.3. Predicted probability of occurrence for 'Coral Reef'. Figure panels depict: a) a photo of coral reef; b) a map denoting the presences and absences of coral reef in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Coral Reef' model.

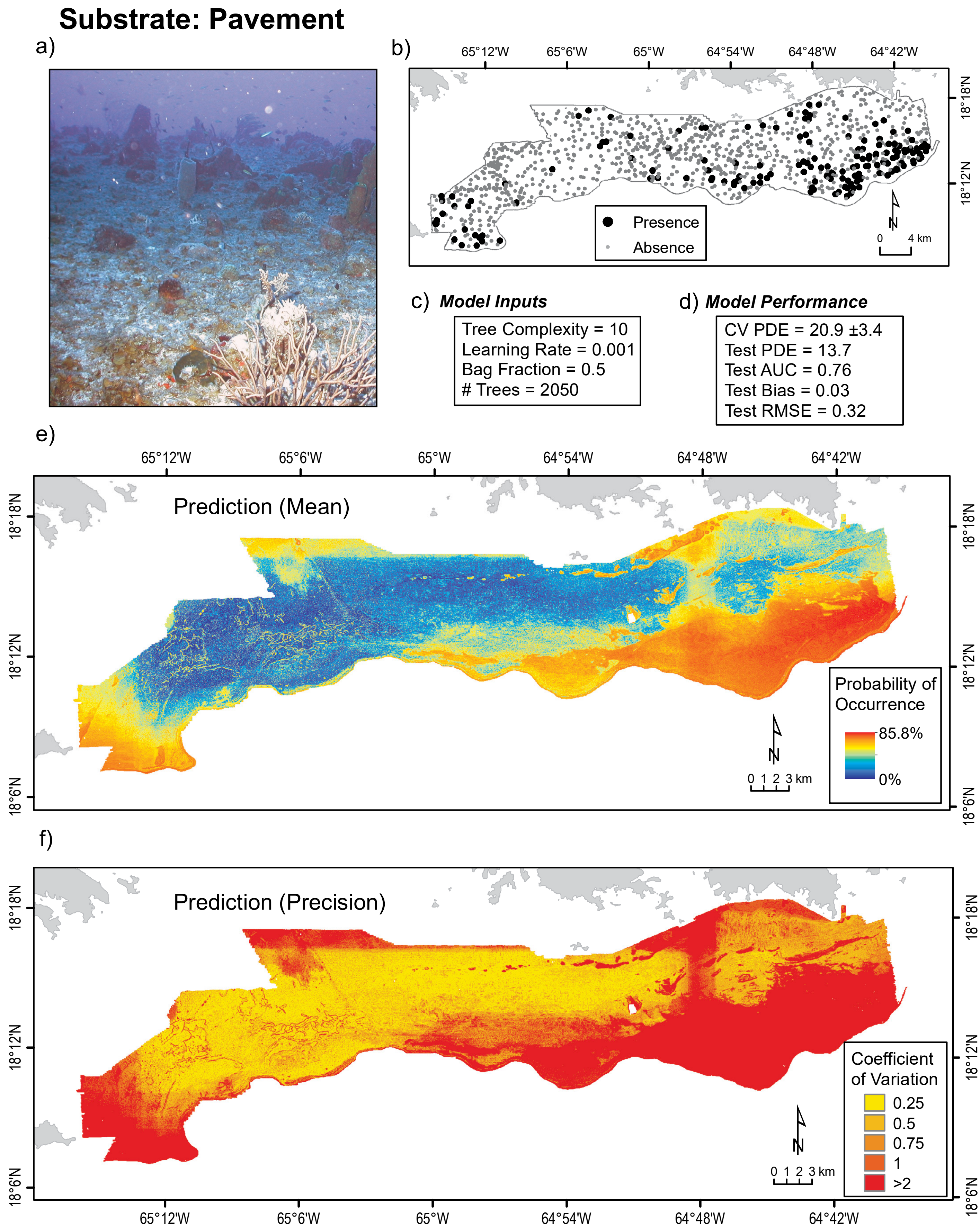


Figure 3.4. Predicted probability of occurrence for 'Pavement'. Figure panels depict: a) a photo of pavement; b) a map denoting the presences and absences of pavement in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Pavement' model.

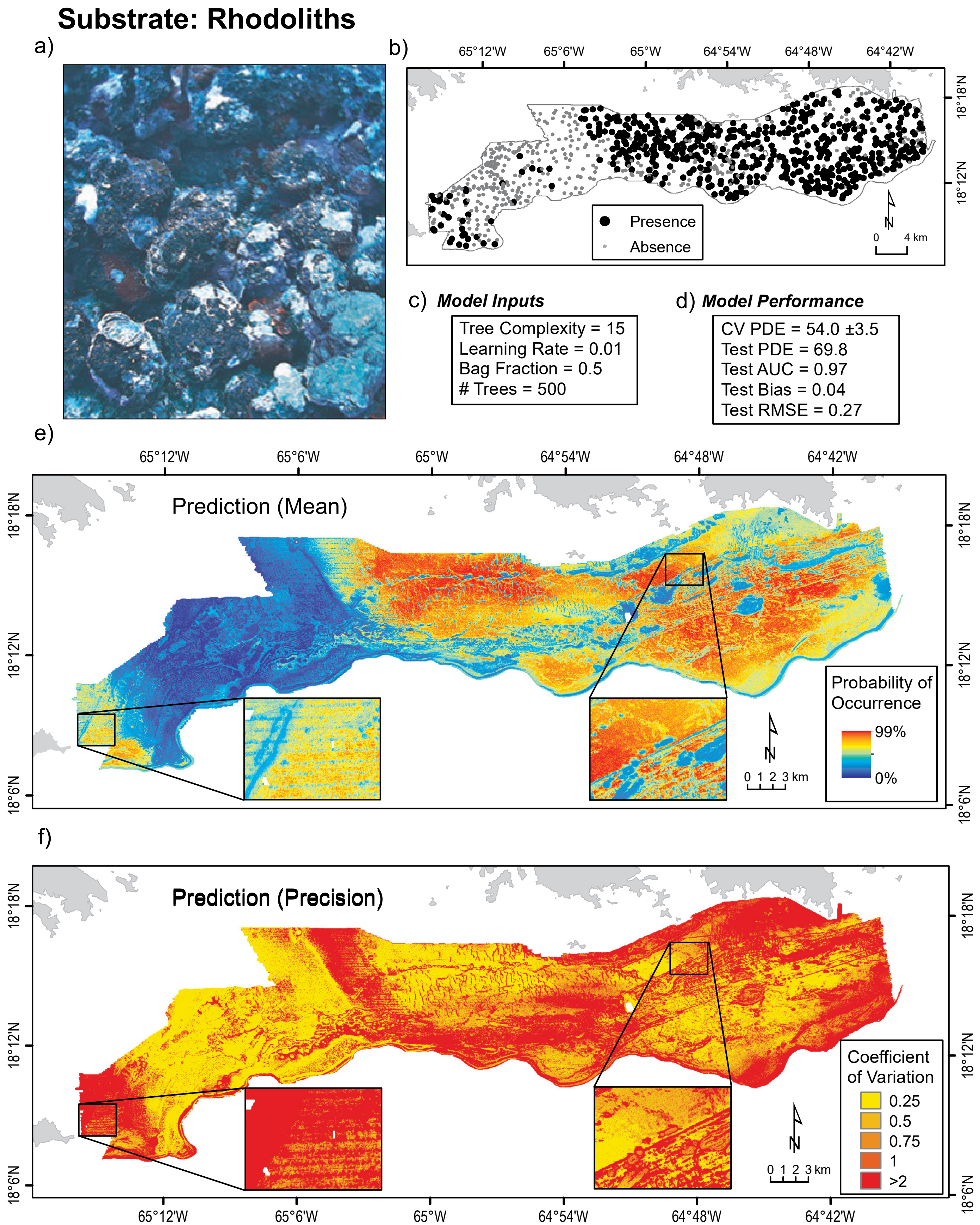


Figure 3.5. Predicted probability of occurrence for 'Rhodoliths'. Figure panels depict: a) a photo of rhodoliths; b) a map denoting the presences and absences of Rhodolith in the GV data; c) the input parameters used to create the final model; d) the performance metrics of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Rhodoliths' model. The insets in e) and f) show the impact of the noisy depth data on the model prediction and associated precision.

Results and Discussion

Substrate: Sand

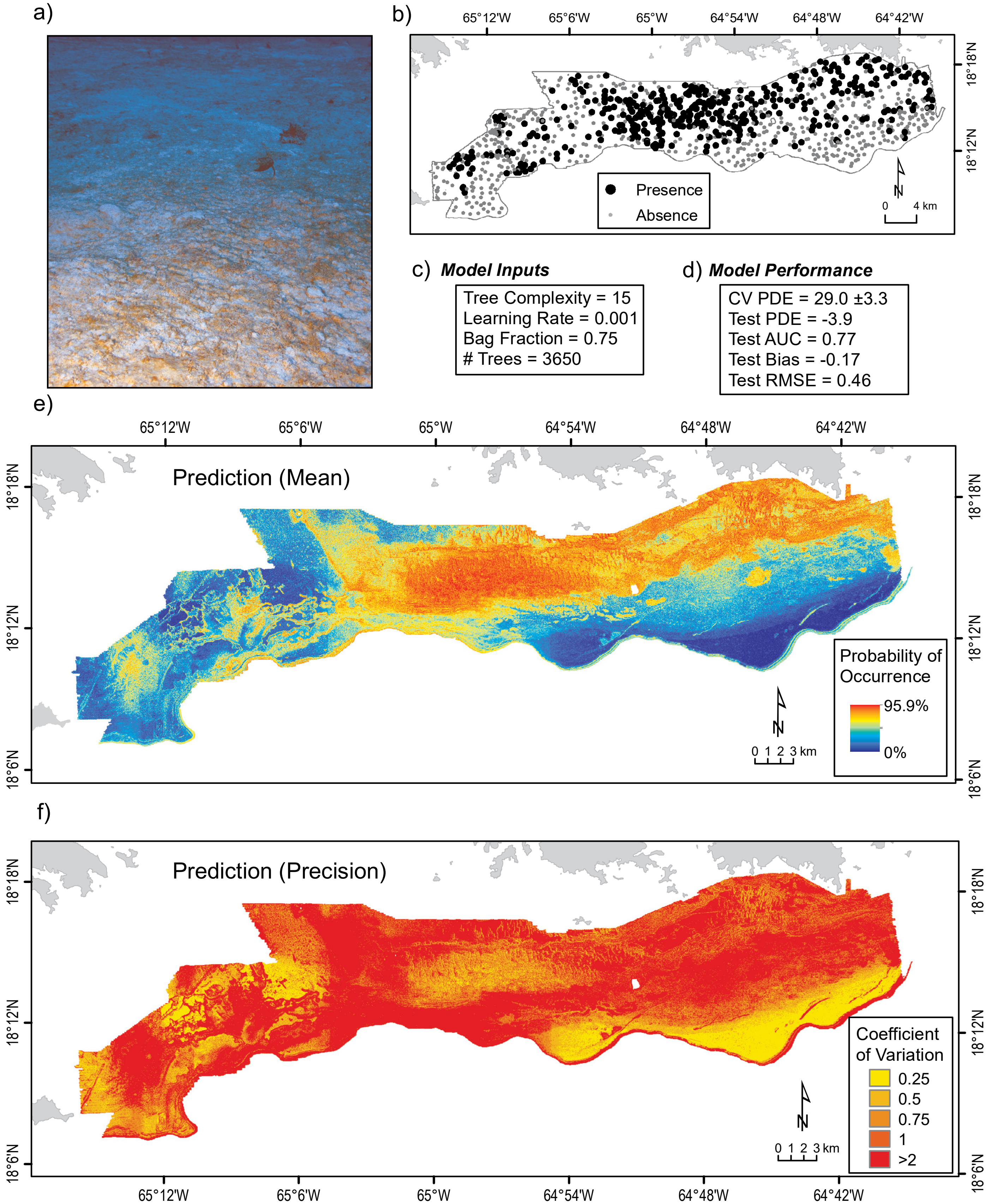
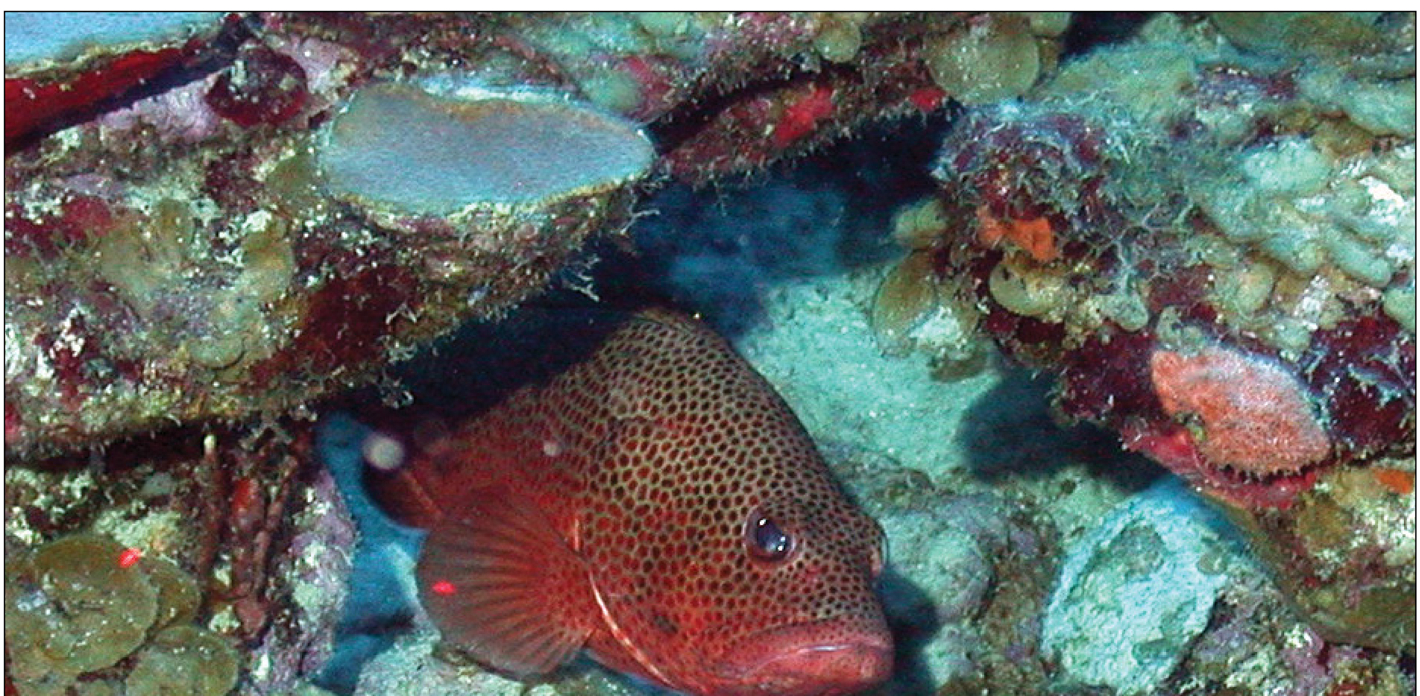


Figure 3.6. Predicted probability of occurrence for 'Sand'. Figure panels depict: a) a photo of sand; b) a map denoting the presences and absences of sand in the GV data; c) the input parameters used to create the final model; d) the performance metrics of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Sand' model.

Cover: Live Hard Coral. 'Live Hard Coral' (Figure 3.7a) was prevalent across the insular shelf, occurring at 52% (524/1005) of the GV sites (Figure 3.7b). Live Hard Coral occurred more frequently in the western half of the project area and near the shelf edge. The 'Live Hard Coral' prediction (Figure 3.7e) showed similar spatial patterns, with the highest likelihood of live hard coral occurring west of Virgin Passage and near the shelf edge across Hind Bank, Frenchcap Bank and Tampo Bank. Similar to the 'Coral Reef' and 'Rhodolith' models, the CV values were lowest (<0.25) where the model predicted high probabilities-of-occurrence (Figure 3.7f). This pattern suggests the 'Live Hard Coral' model is more precise where live hard coral is more likely to be present. A notable exception to this pattern is at Frenchcap Bank, Tampo Bank and the Mid-shelf Reef, where areas of high probability of occurrence also have low precision (CV>1). One explanation for the low precision in these areas is noise in the underlying depth and derived topographic predictors (as described above). Sensor noise is visible as linear stripes (artifacts) oriented in the northeast-southwest direction (Figure 3.7e and f inset maps). These noisy locations are the same areas where the 'Coral Reef' model was degraded. The BRT models falsely classified the noise in these locations as habitats where live corals were likely to be present.

Cover: Live Soft Coral. Live soft coral (Figure 3.8a) was relatively less common on the insular shelf, occurring at 17% (175/1005) of the GV sites (Figure 3.8b) primarily between El Seco and Sail Rock Bank and at Hind Bank. The 'Live Soft Coral' prediction (Figure 3.8e) reflected the spatial patterns observed in the GV data, with the highest likelihood of occurrence west of Sail Rock Bank. Coral reef areas south of St. Thomas and St. John also had high probability of occurrences for live soft coral. Precision was the lowest (CV>1) for the 'Live Soft Coral' model where its predicted probabilities-of-occurrence were highest, and highest (<0.25) where it was less likely to be present (Figure 3.8f). This pattern indicates that there is higher model precision for locations where these organisms are likely to be absent, and lower model precision for locations where they are likely to be present. This inverted pattern is due to the rarity of live soft coral, similar to the spatial patterns seen with the precision of the 'Pavement' prediction. Lastly, noisy depth data degraded the quality of the "Live Soft Coral' spatial prediction in the SwathPlus and 2004 MBES acquisition on areas (Figure 3.8e and f inset maps).



*Red Hind (*Epinephelus guttatus*) on the insular shelf. Credit: T. Battista and B. Costa.*

Results and Discussion

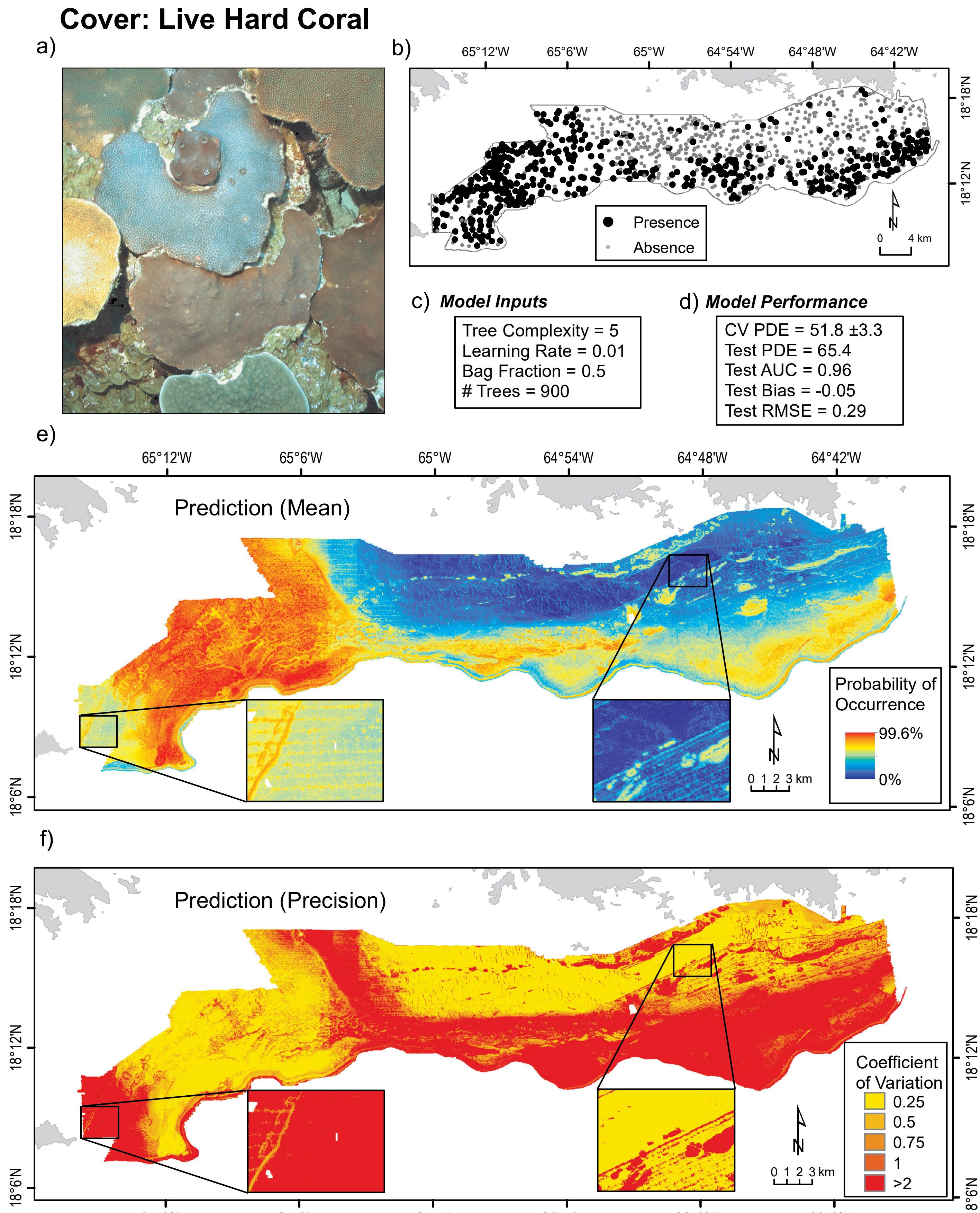


Figure 3.7. Predicted probability of occurrence for 'Live Hard Coral'. Figure panels depict: a) a photo of live hard coral; b) a map denoting the presences and absences of live hard coral in the GV data; c) the input parameters used to create the final model; d) the performance metrics of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Live Hard Coral' model.

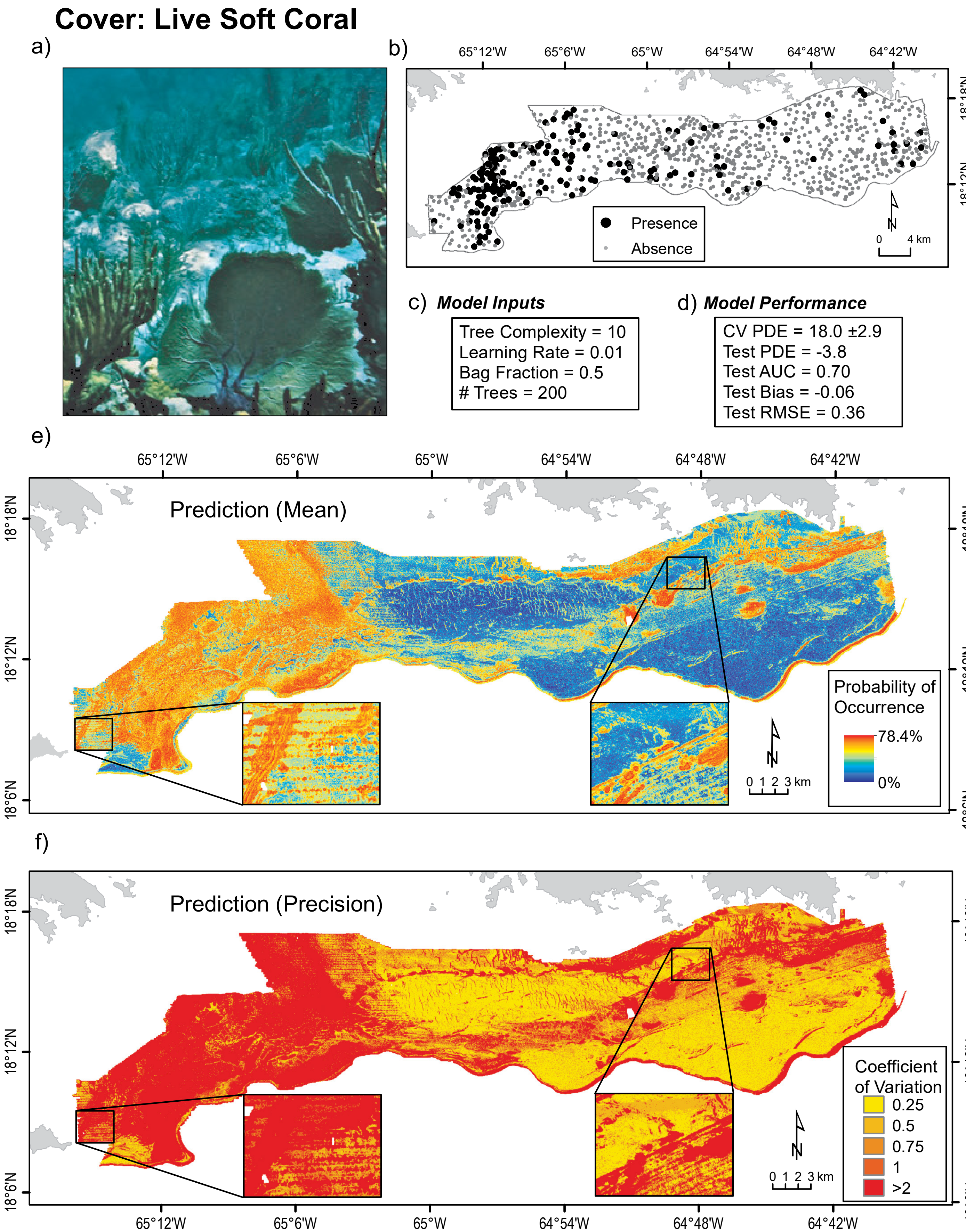


Figure 3.8. Predicted probability of occurrence for 'Live Soft Coral'. Figure panels depict: a) a photo of live soft coral; b) a map denoting the presences and absences of live soft coral in the GV data; c) the input parameters used to create the final model; d) the performance metrics of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation for the 'Live Soft Coral' model.

3.4 COMPOSITE HABITAT MAP

We characterized approximately 652 km^2 of the insular shelf seafloor using BCTs. The composite habitat map displays the predicted spatial distribution of five commonly co-occurring combinations of substrate and cover types (Figure 3.9). 'Coral Reef colonized with Live Coral' was the most abundant habitat class, comprising 34.3% (223.6 km^2) of the area and was dominant throughout the Virgin Passage. It also occurs at Hind Bank, along the upper shelf of French Cap Bank, and along ridge features south of St. John. 'Rhodoliths with Macroalgae' was the next most abundant habitat class mapped, comprising 30.2% (196.8 km^2) of the area. This class was dominant on Tampo Bank and portions of French Cap Bank. 'Rhodoliths with Macroalgae' also occurs in the western-most portion of the project area northwest of El Seco. 'Rhodoliths with Macroalgae and Bare Sand' was the third most abundant habitat, comprising 25.2% (164.2 km^2) of the area. This habitat was most abundant in a swath across the northern section of the project area west of Sail Rock Bank. 'Bare Sand' comprised 5.5% (36.0 km^2) of the insular shelf, and was found primarily around the edges of these coral reef structures. Linear patches of bare sand were also common in the middle of the insular shelf between Hind Bank and Frenchcap Bank. Interestingly, these patches were fairly regularly spaced, and commonly oriented in the north-south direction, suggesting they were formed by consistent bottom currents in the area. Lastly, 'Pavement colonized with Live Coral' is the least abundant habitat class, comprising 4.8% (31.3 km^2) of the area and occurs almost exclusively in the most southeasterly region on the insular shelf project area.

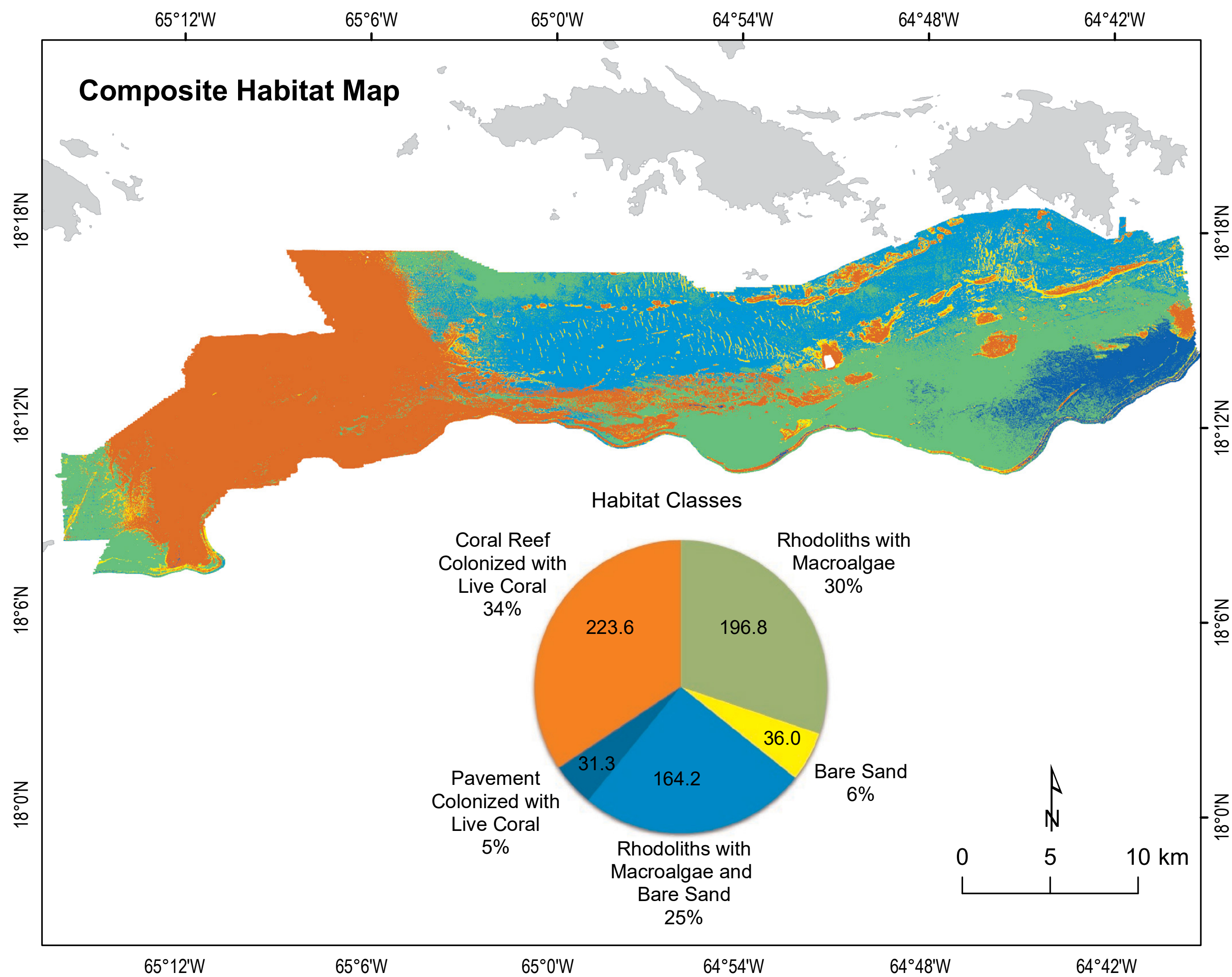


Figure 3.9. This map depicts the five benthic habitat classes mapped throughout the insular shelf. The numbers inside and outside the pie chart denote the amount of area (km^2) and overall proportion of area occupied by each habitat class, respectively.

3.5 MAP ACCURACY

Agreement between the field data and habitat map suggests the BCTs were able to describe the relationships among the habitats and environmental predictors reasonably well. The overall accuracy and tau value for the composite habitat map (quantified using the AA points) was high at 85.6% and 0.82 ±0.05, respectively (Table 3.1). The user's accuracies were also moderate to high, ranging from 70% to 95% for the individual habitat classes. The overall accuracy was very similar after correcting for proportional biases (84.8% ± 4.0% at the 95% confidence level). Two habitat classes were more commonly confused and misclassified than other habitat classes. Specifically, the 'Coral Reef with Live Corals' class was more commonly confused with the 'Pavement with Live Corals' class than with any of the other habitat classes. Similarly, the 'Rhodoliths with Macroalgae and Bare Sand' class was more frequently confused with the 'Rhodoliths with Macroalgae' and 'Bare Sand' classes than with any other habitat types. These results are not surprising, given that these pairs of habitat classes look very similar in the seafloor predictors. It is likely that the presence and inclusion of high quality MBES backscatter would have helped BRTs more reliably and accurately differentiate among these confused pairs of habitat classes. Flat, hard habitats (like pavement) would in theory be brighter and have a different texture in the MBES backscatter than high relief, hard habitats like coral reefs. Besides these two classes, habitat misclassifications were evenly distributed in the confusion table. This suggests that the habitat characterization process did not consistently confuse the remaining habitat types. Overall, the user accuracies of these classes and overall accuracy of the composite map are similar to the other benthic habitat maps created by NOAA NCCOS in the U.S. Caribbean Region (Kendall et al., 2001; Zitello et al., 2009; Costa et al., 2009). As a result, this habitat map can be used with high levels of confidence for most research and management applications (Roelfsema and Phinn, 2013; Tulloch et al., 2017). Their quantitative measures of accuracy and precision can also be incorporated in the management decision-making process to help ensure that resource targets and conservation outcomes are met (Tulloch et al., 2013; Tulloch et al., 2017).

Table 3.1. The confusion matrix for the composite habitat map. Accuracy Assessment sites are listed as columns and corresponding mapped habitats, as rows. Diagonal cell values denote the number of correctly classified sites. Off-diagonal cell values denote the number of incorrectly classified sites.

Map (i) \ AA (i)	Coral Reef Colonized with Live Coral	Pavement Colonized with Live Coral	Rhodoliths with Macroalgae and Bare Sand	Rhodoliths with Macroalgae	Bare Sand	n_j	User's Accuracy (%)
Coral Reef Colonized with Live Coral	**119**	8	2		5	**134**	89%
Pavement Colonized with Live Coral		**10**		1		**11**	91%
Rhodoliths with Macroalgae and Bare Sand		2	**48**	13	6	**69**	70%
Rhodoliths with Macroalgae	1	1	1	**104**	3	**110**	95%
Bare Sand		2	1	4	**17**	**24**	71%
n_i	120	23	52	122	31	**348**	
Producer's Accuracy (%)	99%	43%	92%	85%	55%		

OA = 85.6%
Tau = 0.82
CI (±) = 0.05

3.5 APPLICATIONS

Spatial data describing the distribution of species and habitats is a necessary component of marine spatial planning (Foley et al., 2010; Pittman et al., 2011; Caldow et al., 2015). These datasets and maps can be used for a variety of applications, including identifying and quantifying essential fish habitat, calculating damage and costs of ship groundings, minimizing the impacts of development on important habitats, designing future monitoring or scientific studies, planning for areas of particular concern, and for educational purposes. In the past, NCCOS's benthic habitat maps have been used for many similar applications in USVI and Puerto Rico jurisdictions, enhancing local efforts to preserve and manage coral reef ecosystems. In particular, for the last 15 years, these maps have been critical for the planning of coral reef monitoring missions funded by NOAA's Coral Reef Conservation Program (CRCP). The objective of these missions is to characterize and monitor the distribution of benthic communities, and the abundance and size of reef fish and macroinvertebrates (i.e., conch, lobster, sea urchin, etc.) that rely on those communities (Pittman and Knudby, 2014).

Data from these monitoring missions, led by NCCOS and NOAA's Southwest Fisheries Science Center (SWFSC), have helped managers in the U.S. Caribbean estimate damages from vessel groundings, evaluate zoning strategies (Pittman et al., 2017); evaluate the efficacy of existing marine protected areas, including no-take areas where fishing is prohibited (e.g., Grammanik Bank and Hind Bank Marine Conservation District) (Pittman et al., 2013; Pittman et al., 2014); develop management plans for MPAs (STEER, 2011); and understand the impacts of land based sources of pollution on marine animals and plants (Whitall et al., 2015; Whitall et al., 2016). Also, in St. Thomas, scientists and planners have used the benthic maps to help evaluate the efficacy of sediment reduction efforts where high-density commercial, residential and recreational development or use threatens the surrounding marine ecosystem (Costa et al., 2013; Pait et al., 2013), and identify high-value coral reef ecosystems that spatially coincide with proposed infrastructure development sites (Pait et al., 2013).

The new benthic habitat maps (described here) can be similarly used to plan for potential activities both onshore and across the insular shelf. One of the map's first anticipated applications is to compare the distribution of mesophotic reefs (30-100 m deep) with shallow-water (<30 m deep) reefs around St. Thomas and St. John (Smith, Personal Communication, 18 August 2017). This comparison will help quantify the amount of mesophotic reefs south of these islands, as local and global stressors continue to impact and degrade shallow-water reefs. This map may also be used to evaluate the impacts and changes to marine ecosystems due to Hurricane Irma and Maria. These hurricanes ravaged the U.S. Caribbean in September 2017, and their long-term economic and ecological effects are not known. Another anticipated application is to include this new habitat map in NOAA NCCOS's decision support framework tool for the USVI, developed in partnership with USVI's DPNR, UVI and the Nature Conservancy (Knudby et al., 2014; NOAA NCCOS, 2016). This decision support framework is designed to help resource managers to identify and rank coral reef ecosystems in the USVI, so they can better prioritize and more efficiently focus their resources and conservation efforts (Pittman et al., 2017; Pittman et al., In Press). The tool works by integrating spatial information on benthic habitats, associated biodiversity, and human use values related to coral reefs to objectively rank sites into several categories based on their resilience. Please see the following website for more information: https://coastalscience.noaa.gov/project/prioritizing-sites-coral-reef-conservation-us-virgin-islands/.

Previous Benthic Habitat Maps in the Region

In the last several years, computing power and model-based mapping techniques have advanced considerably. The map products (described here) take maximum advantage of those improvements and advancements to create highly-resolved raster maps. Existing users of NCCOS's habitat maps may be initially challenged by the transition to these new, raster maps (Figure 3.10). To help users with this transition, a composite habitat map was created during this mapping process, describing commonly co-occurring substrate and cover types. The look and feel of this composite map is similar to NCCOS's previous polygon-based maps (Kendall et al., 2001; Costa et al., 2009; Zitello et al., 2009; Costa et al., 2013). The underlying habitat predictions, however, are fundamentally different from NCCOS's previous maps. Previously, heterogeneous pixels within a polygon were aggregated together, and all given the same habitat

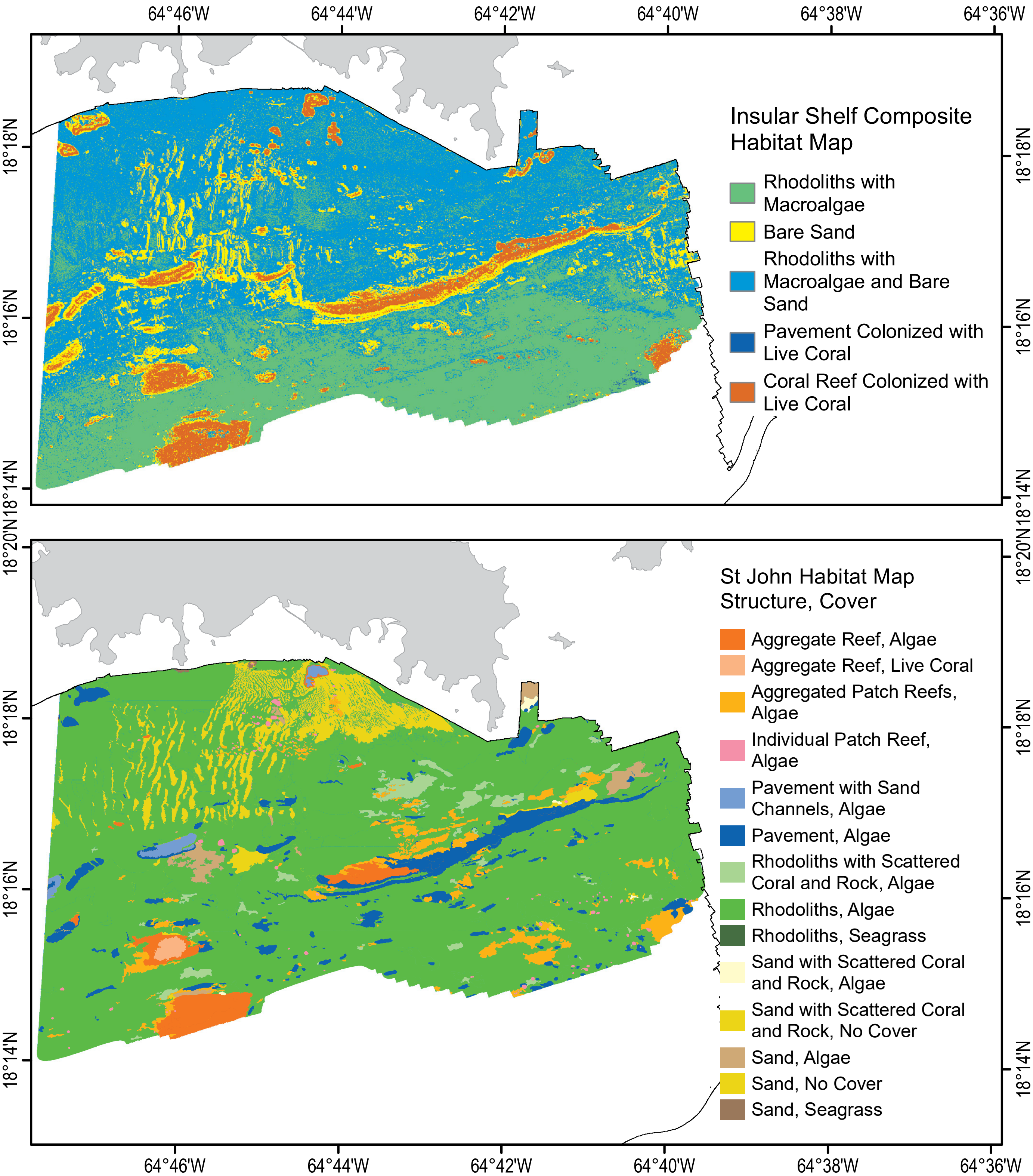


Figure 3.10. Comparison of pixel-based and polygon-based maps south of St. John, USVI. (Top) Pixel-based map developed in 2017 with a spatial resolution of 121 m^2. (Bottom) Polygon-based habitat map developed in 2009 with a spatial resolution of 1,000 m^2 (Costa et al., 2009).

class name. These new, next generation habitat maps preserve fine-scale heterogeneity, associated habitat gradients, and the smaller benthic features present across the seascape. They also move away from absolute classifications, and instead predict the likelihood that a substrate or cover type (e.g., live coral, sand, etc.) is present at a specific location. These predictions also quantify the uncertainty associated with each pixel in the map, helping users understand the precision of these predictions. Combined, these fundamental changes are designed to generate habitat maps that are more repeatable and flexible, moving habitat characterization closer to becoming a monitoring tool.

These different map products can be customized for specific applications using the basic tools in GIS or other freely available software. Some options for customization may include using spatial filters to identify dominant habitats, enhance priority bottom types, smooth out variability in heterogeneous areas or resample the map to different spatial scales. Habitat classes can also be aggregated (similar to the polygon based approach) to simplify the number of habitat classes that are visualized (e.g., hard bottom and soft bottom). They can also be translated into other classification systems that are commonly used (e.g., CMECS, 2017). Predictions for individual substrate and cover types can be converted from the continuous probability-of-occurrence surfaces, which are useful for examining gradients, into classified maps. The threshold used to convert these maps can be tailored to the specific management applications. For example, if the goal is to protect 95% of a species' distribution, the ROC curve for this species can be used to identify the probability of occurrence threshold where sensitivity equals 0.95. The precision associated with each habitat prediction can also be classified and applied to understand the sensitivity of a management decision to varying levels of uncertainty (Costa et al., In Prep). It can also be used to identify data gaps and areas with low precision, and prioritize them for future data collection efforts. Combined, these options make the new generation habitat maps more readily customizable for a wider range of marine applications.

Download and Access Products

Since there are a wide range of potential applications for these next generation maps, the best way to access and use them is through GIS or other software that allows users to zoom in and create custom products. These GIS-ready map products are available from the project page: https://coastalscience.noaa.gov/project/habitat-map-insular-shelf-south-of-st-thomas-st-john/. We also recognize some users do not have GIS software, and can better utilize a hard copy or online map. Hard copy maps are provided in an atlas format at the end of this report (see Appendix A). The atlas consists of six color maps printed at a scale of 1:70,000. In addition to the atlas, an online map viewer was created for this project, allowing users to view and query the habitat map products online, and to view and download the underwater videos without any specialized software. Access to this online data viewer is also available through the project page (listed above).

3.6 SUMMARY

The map products provided in this report represent the culmination of this extensive mapping campaign in the USVI. This mapping and characterization effort included not only the collection of high resolution MBES imagery, but also the acquisition of underwater videos (using ROVs, see Appendix B), and data describing the spatial distribution, size and abundance of fish (using Splitbeam Echosounders (SBES), see Appendix C). The geographic and temporal extents and potential uses of these ancillary datasets are described in more detail in the appendices. Together, MBES, ROV and SBES data provide rich spatio-temporal information across the insular shelf, describing the presence, abundance and condition of coral reef ecosystems and their associated fish assemblages. The integration of these datasets can lead to a better understanding of the potential environmental drivers of habitat use by fish and fish aggregations, and may suggest particularly sensitive habitats on the insular shelf (Costa et al., 2014). Furthermore, these datasets provide a rare snapshot in time of these resources, and can be analyzed (or certain sites revisited) to compare and quantify changes over time or after catastrophic events like Hurricanes Irma and Maria. This baseline information is increasingly important and critical for researchers and managers, as a combination of local and global stressors impact and fundamentally change coral reef ecosystems around the world.

*Live coral (*Porites porites*) with polyps extended on the insular shelf. Credit: T. Battista and B. Costa.*

Red hind (Epinephelus guttatus) on the insular shelf. Credit: T. Battista and B. Costa.

Brander, L. and P. van Beukering. 2013. The total economic value of US coral reefs: a review of the literature. [US Department of Commerce], National Oceanic and Atmospheric Administration, Coral Reef Conservation Program. 32 pp.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.I. Stone. 1984. Classification and regression trees. Taylor & Francis, Belmont, CA. 368 pp.

Caldow, C., M.E. Monaco, S.J. Pittman, M.S. Kendall, T.L. Goedeke, C. Menza, B.P. Kinlan, and B.M. Costa. 2015. Biogeographic assessments: a framework for information synthesis in marine spatial planning. Marine Policy, 51, pp. 423-432.

Card, D.H. 1982. Using known map categorical marginal frequencies to improve estimates of thematic map accuracy. Photogrammetric Engineering and Remote Sensing 48: 431-439.

Casey, K.S., E.R. Selig, D. Zhang, K. Saha, A. Krishnan, and E. McMichael. 2015. The Coral Reef Temperature Anomaly Database (CoRTAD) Version 5 - Global, 4 km Sea Surface Temperature and Related Thermal Stress Metrics for 1982-2012 (NCEI Accession 0126774). Data downloaded 4 March, 2016. Data website: https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0126774 (Site Accessed 1 September 2017).

CMECS. 2017. Coastal and Marine Ecological Classification Standard, Catalog of Units. Online: https://www.cmecscatalog.org/ (Site Accessed 1 September 2017).

Congalton, R.G. and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC/Lewis Press, Boca Raton, FL. 137 pp.

Costa, B.M., L.J. Bauer, T.A. Battista, P.W. Mueller, and M.E. Monaco. 2009. Moderate-Depth Benthic Habitats of St. John, U.S. Virgin Islands. NOAA Technical Memorandum NOS NCCOS 105. Silver Spring, MD. 57 pp.

Costa, B.M. and T.A. Battista. 2013. The semi-automated classification of acoustic imagery for characterizing coral reef ecosystems. International journal of remote sensing 34:6389-6422.

Costa, B.M., M.S. Kendall, K. Edwards, G. Kågesten, and T.A. Battista. 2013. Benthic Habitats of Fish Bay, Coral Bay and the St. Thomas East End Reserve. NOAA Technical Memorandum NOS NCCOS 175. Prepared by the NCCOS Center for Coastal Monitoring and Assessment Biogeography Branch. Silver Spring, MD. 68 pp.

Costa, B.M, J.C. Taylor, L. Kracker, T.A. Battista, and S.J. Pittman. 2014. Mapping Reef Fish and the Seascape: Using Acoustics and Spatial Modeling to Guide Coastal Management. PLoS ONE 9(1): e85555. doi:10.1371/journal.pone.0085555.

Costa, B., M. Kendall, and S. McKagan. In Prep. Managers, modelers and measuring the impact of prediction uncertainty on marine zoning in Saipan Lagoon, CNMI. PLoS ONE (Target Journal).

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

References

Elith, J., C.H. Graham, R.P. Anderson, M. Dudı´k, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. Overton, A.T. Peterson, S.J. Phillips, K.S. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N.E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

Elith J., J.R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802–81.

ESRI. 2016. ArcGIS (Version 10.4). Environmental Systems Research Institute, Redlands, California. Software Online: http://www.esri.com/ (Site Accessed 18 July 2017).

Foley, M.M., B.S. Halpern, F. Micheli, M.H. Armsby, M.R. Caldwell, C.M. Crain, E. Prahler, N. Rohr, D. Sivas, M.W. Beck, and M.H. Carr. 2010. Guiding ecological principles for marine spatial planning. Marine Policy 34(5): 955-966.

Foster, M.S., 2001. Rhodoliths: between rocks and soft places. Journal of phycology, 37(5): 659-667.

Friedman, J.H. 2002. Stochastic gradient boosting. Computational Statistics and Data Analysis 38: 367-378.

Friedlander, A.M., C.F.G. Jeffrey, S.D. Hile, and S.J. Pittman. 2013. Coral Reef Ecosystems of St. John, U.S. Virgin Islands: Spatial and Temporal Patterns in Fish and Benthic Communities (2001-2009). NOAA Technical Memorandum 152. Silver Spring. MD. 150 pp.

GPR. 2003. Marine habitat mapping offshore St. Thomas & St. Croix, U.S. Virgin Islands EEZ – Final Report Volume 2 (M-03704) to the Caribbean Fishery Management Council. Geophysics GPR International Inc. for the Caribbean Fisheries Management Council. 34 pp.

Guisan, A., R. Tingley, J.B. Baumgartner, I. Naujokaitis-Lewis, P.R. Sutcliffe, A.I. Tulloch, and T.G. Martin. 2013. Predicting species distributions for conservation decisions. Ecology letters, 16(12): 1424-1435.

Hasan, R.C., D. Ierodiaconou, L. Laurenson, and A. Schimel. 2014. Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. PLoS ONE 9(5): e97339.

Hijmans, R.J. 2014. R package, Raster: Geographic data analysis and modeling. Software Downloaded March 2016. Software Website: http://CRAN.R-project.org/package=raster (Site Accessed 25 August 2017).

Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2014. R package, dismo: Species distribution modeling. Software Downloaded March 2016. Software website: http://CRAN.R-project.org/package=dismo (Site Accessed 25 August 2017).

Hosmer, D.W. and S. Lemeshow. 2000. Applied logistic regression (2nd ed.). John Wiley & Sons, pp. 375.

Jenness, J. 2016. DEM Surface Tools (Version 2.1.399). Jenness Enterprises, Flagstaff, AZ. Software Online: http://www.jennessent.com/arcgis/surface_area.htm (Site Accessed 31 August 2017).

Kadison, E., M. Brandt, R. Nemeth, J. Martens, J. Blondeau, and T. Smith. 2017. Abundance of commercially important reef fish indicates different levels of over-exploitation across shelves of the US Virgin Islands. PloS one, 12(7), p.e0180063.

Kendall, M.S., C.R. Kruer, K.R. Buja, J.D. Christensen, M. Finkbeiner, R.A. Warner, and M.E. Monaco. 2001. Methods Used to Map the Benthic Habitats of Puerto Rico and the U.S. Virgin Islands. NOAA NCCOS CCMA. Silver Spring, MD.

Kendall, M.S., B. Costa, S. McKagan, L. Johnston, and D. Okano. 2017. Benthic Habitat Maps of Saipan Lagoon. NOAA Technical Memorandum NOS NCCOS 229. Silver Spring, MD. 77 pp.

Kenny, A.J., I. Cato, M. Desprez, G. Fader, R.T.E. Schüttenhelm, and J. Side. 2003. An overview of seabed-mapping technologies in the context of marine habitat classification. ICES Journal of Marine Science 60: 411–418

Knudby A., S.J. Pittman, J. Maina, and G. Rowlands. 2014. Chapter 5: Remote Sensing and Modeling of Coral Reef Resilience. pp. 103-134. In (C.W. Finkl and C. Makowski, eds.). Remote Sensing and Modeling: Advances in Coastal and Marine Resources, Springer: Switzerland. 501 pp.

Kracker, L.M., J.C. Taylor, E.F. Ebert, T.A. Battista, and C. Menza. 2011. Integration of fisheries acoustics surveys and bathymetric mapping to characterize midwater-seafloor habitats of US Virgin Islands and Puerto Rico (2008–2010). NOAA Technical Memorandum NOS NCCOS 130. Silver Spring, MD. 44 pp.

Kuhn, M. 2016. R package, Caret: Classification and Regression Training. Software. Downloaded October 2015. Software Website: https://cran.r-project.org/web/packages/caret/index.html (Site Accessed 25 August 2017).

Leathwick, J.R., J. Elith, M.P. Francis, T. Hastie, and P. Taylor. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology Progress Series 321: 267-281.

Lurton, X. 2000. Swath bathymetry using phase difference: Theoretical analysis of acoustical measurement precision. IEEE Journal of Oceanic Engineering 25: 351-363.

Ma, Z., and R.L. Redmond. 1995. Tau coefficients for accuracy assessment of classification of remote sensing data. Photogrammetric Engineering and Remote Sensing 61: 435-439.

Menza, C., J. Ault, J. Beets, J. Bohnsack, C. Caldow, J. Christensen, A. Friedlander, C. Jeffrey, M. Kendall, J. Luo, M. Monaco, S. Smith, and K. Woody. 2006. A Guide to Monitoring Reef Fish in the National Park Service's South Florida / Caribbean Network. NOAA Technical Memorandum NOS NCCOS 39. Silver Spring, MD. 169 pp.

Menza, C., L.M. Wedding, R. Clark, C. Caldow, and M.E. Monaco. 2012. A Baseline Assessment of Fish and Coral Bays, St. John, U.S. Virgin Islands in Support of Watershed Restoration Activities. Part I: Fish, Coral and Benthic Habitats. NOAA Technical Memorandum NOS NCCOS 154. Silver Spring, MD. 29 pp.

MGET (Marine Geospatial Modeling Tools). 2016. Marine Geospatial Ecology Tools: Open source geoprocessing for marine research and conservation (version 0.8a68). Duke Nicholas School of the Environment, Marine

References

Geospatial Ecology Lab. Software Online: http://mgel2011-kvm.env.duke.edu/mget/download/ (Site Accessed 25 August, 2017).

Miller, J., R. Waara, E. Muller, and C. Rogers. 2006. Coral bleaching and disease combine to cause extensive mortality on reefs in U.S. Virgin Islands. Coral Reefs, 25(3): 418.

Miller, J., E. Muller, C. Rogers, R. Waara, A. Atkinson, K.R.T. Whelan, M. Patterson, and B. Witcher. 2009. Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. Coral Reefs, 28(4): 925.

Monaco, M.E., S.M. Anderson, T.A. Battista, M.S. Kendall, S.O. Rohmann, L.M. Wedding, and A.M. Clarke. 2012. National Summary of NOAA's Shallow-water Benthic Habitat Mapping of U.S. Coral Reef Ecosystems. NOAA Technical Memorandum NOS NCCOS 122. Prepared by the NCCOS Center for Coastal Monitoring and Assessment Biogeography Branch. Silver Spring, MD. 83 pp.

NASA (National Aeronautics and Space Administration). 2016. Ocean Color Web. Standard, Aqua MODIS Remote sensing reflectance at 547 nm, Annual composite, 4 km. Data downloaded 4 March, 2016. Data website: https://oceancolor.gsfc.nasa.gov/cgi/l3 (Site Accessed 30 August 2017).

NOAA (National Oceanic and Atmospheric Administration). 2002. A national coral reef action strategy. Report to congress on implementation of the Coral Reef Conservation Act of 2000 and the National Action Plan to Conserve Coral Reefs in 2002-2003. Silver Spring, Maryland. 156 pp. Report website: https://www.coris.noaa.gov/activities/actionstrategy/action_reef_final.pdf (Accessed 1 September, 2017).

NOAA NCCOS (National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science). 2016. Prioritizing Sites for Coral Reef Conservation in the U.S. Virgin Islands. Project website: https://coastalscience.noaa.gov/project/prioritizing-sites-coral-reef-conservation-us-virgin-islands/ (Site Accessed 31 August, 2017).

NOAA NCCOS (National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science). 2017. Seafloor Characterization of the U.S. Caribbean. Project Products, Data Collections: 2004 USVI (NCEI Acc. 0131850); 2005 USVI (NCEI Acc. 0131860); 2009 PR (NCEI Acc. 0131857); 2010 USVI (NCEI Acc. 0131854); 2011 USVI (NCEI Acc. 0131858). Data website: https://coastalscience.noaa.gov/project/seafloor-characterization-caribbean/ (Site Accessed 30 August 2017).

NOAA NCRMP (National Oceanic and Atmospheric Administration, National Coral Reef Monitoring Program). 2017. NOAA's National Coral Reef Monitoring Program. Data website: https://www.coris.noaa.gov/monitoring/ (Site Accessed 30 August, 2017).

Pait, A.S., S.I. Hartwell, A.L. Mason, R.A. Warner, C.F. Jeffrey, A.M. Hoffman, D.A. Apeti, F.R. Galdo Jr, and S.J. Pittman. 2013. An assessment of chemical contaminants, toxicity and benthic infauna in sediments from the St. Thomas East End Reserves (STEER). NOAA Technical Memorandum NOS NCCOS 202. Silver Spring, MD. 219 pp.

Pittman, S.J., J.D. Christensen, C. Caldow, C. Menza, and M.E. Monaco. 2007. Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. Ecological Modelling, 204: 9-21.

Pittman, S.J., S.D. Hile, C.F.G. Jeffrey, C. Caldow, M.S. Kendall, M.E. Monaco, and Z. Hillis-Starr. 2008. Fish assemblages and benthic habitats of Buck Island Reef National Monument (St. Croix, U.S. Virgin Islands) and the surrounding seascape: a characterization of spatial and temporal patterns. NOAA Technical Memorandum NOS NCCOS 71, Silver Spring, MD. 96 pp.

Pittman SJ, Costa B, Battista T. 2009. Using Lidar bathymetry and boosted regression trees to predict the diversity and abundance of fish and corals. J Coast Res Spec Issue 53: 27–38

Pittman, S.J. and K.A. Brown. 2011. Multi-scale approach for predicting fish species distributions across coral reef seascapes. PloS one, 6(5): e20583.

Pittman, S.J., D.W. Connor, L. Radke, and D.J. Wright. 2011. Application of estuarine and coastal classifications in marine spatial management. 163–205 In Treatise on estuarine and coastal science, (E. Wolanski and D. McLusky, eds.). Elsevier.

Pittman, S.J., D.S. Dorfman, S.D. Hile, C.F.G. Jeffrey, M.A. Edwards, and C. Caldow. 2013. Land-Sea Characterization of the St. Croix East End Marine Park, U.S. Virgin Islands. NOAA Technical Memorandum NOS NCCOS 170. Silver Spring, MD. 119 pp.

Pittman, S.J., L. Bauer, S.D. Hile, C.F.G Jeffrey, E. Davenport and C. Caldow. 2014. Marine protected Areas of the U.S. Virgin Islands: Ecological Performance Report. NOAA Technical Memorandum NOS NCCOS 187. Silver Spring, MD. 100 pp.

Pittman S.J. and A. Knudby. 2014. Predictive mapping of coral reef fish species and communities. In: Bortone, editor. Interrelationships between coral reefs and fisheries. Florida: CRC Press. pp. 219-234.

Pittman, S.J., C.F.G. Jeffrey, C. Menza, G. Kågesten, A. Orthmeyer, D.S. Dorfman, D. Mateos-Molina, A. Mabrouk, S.D. Hile, V. Ransibrahmanakul, and A. Ramos Álvarez. 2017. Mapping Ecological Priorities and Human Impacts to Support Land-Sea Management of Puerto Rico's Northeast Marine Corridor. NOAA Technical Memorandum NOS NCCOS 218. Silver Spring, MD. 71 pp.

Pittman S.J., M. Poti, C.F.G. Jeffrey, L.M. Kracker, and M. Ayman. In Press. Decision support framework for the prioritization of coral reefs in the U.S. Virgin Islands. Ecological Informatics).

R Core Team. 2016. R: A language and environment for statistical computing (Version 3.3.0). R Foundation for Statistical Computing, Vienna, Austria. Software Online: https://www.r-project.org/ (Site Accessed 1 September, 2017).

Roelfsema, C.M. and S.R. Phinn. 2013. Validation. Pages 375–401 in J. A. Goodman, S. J. Purkis and S. R. Phinn, editors. Coral reef remote sensing: A guide for multi-level sensing mapping and assessment. Springer, Dordrecht, Netherlands.

Rothenberger, P., J. Blondeau, C. Cox, S. Curtis, W.S. Fisher, V. Garrison,Z. Hillis-Starr, C.F.G. Jeffrey, E. Kadison, I. Lundgren, and W.J. Miller. 2008. The state of coral reef ecosystems of the US Virgin Islands. In The state of coral reef ecosystems of the United States and Pacific Freely Associated States (Vol. 73, pp. 29-73). US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Centers for Coastal Ocean Science.

References

Smith, T.B., V.W. Brandtneris, M. Canals, M.E. Brandt, J. Martens, R.S. Brewer, E. Kadison, M. Kammann, J. Keller, and D.M. Holstein. 2016. Potential Structuring Forces on a Shelf Edge Upper Mesophotic Coral Ecosystem in the US Virgin Islands. Frontiers in Marine Science 3: 115.

Smith, T.B. University of the Virgin Islands. Personal Communication. 18 August, 2017.

STEER. 2011. St. Thomas East End Reserves Management Plan. Edited by J. Brown, A. Holecek, J. Murray, L. Noori, J.P. Oriol, J. Pierce, R. Platenberg, S. Romano, D. Russo, and C. Settar. 99 pp. Report website: http://www.horsleywitten.com/STEERwatersheds/pdf/WatershedManagementPlan/1305031_STEERWatershedPlan.pdf (Site Accessed 31 August 2017).

Story, M. and R. Congalton. 1986. Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing 52: 397-399.

ten Brink, U. and B. Andrews. 2009. Field Activity Details for 2009-001-FA (Cruise ID T-1-09-PR). https://walrus.wr.usgs.gov/infobank/t/t109pr/html/t-1-09-pr.meta.html (Site Accessed 1 September 2017).

Trenkel, V.M., P.H. Ressler, M. Jech, M. Giannoulaki, and C. Taylor. 2011. Underwater acoustics for ecosystem-based management: state of the science and proposals for ecosystem indicators. Marine Ecology Progress Series 442: 285-301.

Tulloch, V.J., H.P. Possingham, S.D. Jupiter, C. Roelfsema, A.I. Tulloch, and C.J. Klein. 2013. Incorporating uncertainty associated with habitat data in marine reserve design. Biological Conservation 162: 41-51.

Tulloch, V.J., C.J. Klein, S.D. Jupiter, A.I. Tulloch, C. Roelfsema, and H.P. Possingham. 2017. Trade-offs between data resolution, accuracy, and cost when choosing information to plan reserves for coral reef ecosystems. Journal of Environmental Management 188: 108-119.

van Beukering, P., L. Brander, B. Van Zanten, E. Verbrugge, and K. Lems. 2011. The economic value of the coral reef ecosystems of the United States Virgin Islands. IVM Report number: R-11/06. 160 pp. Online: http://www.ivm.vu.nl/en/Images/The_Economic_Value_of_the_Coral_Reefs_of_the_USVI_tcm234-232341.pdf (Accessed 27 September 2017).

Wentworth, C.K. 1922. A scale of grade and class terms for clastic sediments. The Journal of Geology 30:377-392.

Whitall, D., A. Pait, and S.I. Hartwell. 2015. Chemical contaminants in surficial sediment in Coral and Fish Bays, St. John, U.S. Virgin Islands. Marine Environment Research 112: 1-8.

Whitall, D., A. Ramos, D. Wehner, M. Fulton, A. Mason, E. Wirth, B. West, A. Pait, E. Pisarski, B. Shaddrix, and L.A. Reed. 2016. Contaminants in Queen Conch (Strombus gigas) in Vieques, Puerto Rico. Regional Studies in Marine Science 5: 80-86.

Zitello, A.G., L.J. Bauer, T.A. Battista, P.W. Mueller, M.S. Kendall, and M.E. Monaco. 2009. Shallow-Water Benthic Habitats of St. John, U.S. Virgin Islands. NOAA Technical Memorandum NOS NCCOS 96. Silver Spring, MD. 53 pp.

Modelling Glossary

Note: many of these definitions are specific to the context of this project.

Bag fraction (bf) – In a boosting context, a parameter that defines the fraction of the data drawn at random, without replacement, from the full training dataset at each iteration.

Boosted classification tree model – A modeling approach that combines a machine learning technique, boosting, with traditional tree-based statistical modeling. In this approach, a large number of classification trees are fit stagewise (i.e., after each tree is fit, the remaining variation in the data is used to fit the next tree) and then combined to generate a final, combined (i.e., “ensemble”) model.

Boosted regression tree model – A modeling approach that combines a machine learning technique, boosting, with traditional tree-based statistical modeling. In this approach, a large number of regression trees are fit stagewise (i.e., after each tree is fit, the remaining variation in the data is used to fit the next tree) and then combined to generate a final, combined (i.e., “ensemble”) model.

Boosting – A technique for fitting models that employs an iterative approach. Models are built in a stage-wise fashion, where existing trees are left unchanged and the variance remaining from the last tree is used to fit the next one.

Bootstrapping – A data re-sampling technique used to estimate the statistical precision associated with model predictions. Bootstrapping is a technique in which the data are randomly re-sampled and the model is refit multiple times. The precision of the model predictions can then be assessed from the variability in predictions across the bootstrap replicates.

Coefficient of variation (CV) – Measure of dispersion for a distribution, representing the standard deviation as a proportion of the mean. In the context of a model prediction, a larger CV indicates more variation or uncertainty and a lower precision for a prediction. A smaller CV indicates less variation or uncertainty and higher precision for a prediction.

Environmental predictor – An independent dataset in a model that is used to explain spatial variation in the response data. Here, the environmental predictors were geographic, oceanographic and topographic, and the response data were the benthic habitat types extracted from the underwater videos and photos.

K-fold cross-validation (kCV) – A technique for evaluating the predictive ability of a fitted model. The data are divided into 10 data subsets (i.e., folds). Nine of these folds were used to create models, while the remaining one fold was used to evaluate the model’s performance.

K-fold cross-validation percent deviance explained (kCV PDE) – Percent deviance explained calculated using one cross validation fold.

Learning rate (lr) – In a boosting context, the degree to which each tree contributes to the final model. The optimal learning rate is one that minimizes prediction error in the fewest number of boosting iterations.

Percent deviance explained (PDE) – A measure of the variation in the data explained by a model (beyond that explained by a model without predictor variables). Values normally range between 0 and 100%, although negative values are possible. Higher PDE values indicate better model performance.

Modelling Glossary

Receiver operating characteristic (ROC) Area under the curve (AUC) – An ROC curve is a graphical representation of how well a model can discriminate between (or predict) the presence and absence of benthic habitat types. ROC curves can be used to identify the "optimal" classification thresholds (e.g., presence/absence) for specific management applications. The AUC is the integral of a ROC curve. AUC values range between 0 and 1 where a value >0.5 indicates performance better than a random guess. Higher AUC values indicate better model performance.

Resampling – A method of using randomly drawn subsets of data to estimate statistical precision (e.g., variation in model predictions), to perform a significance test (e.g., permutation test of predictor importance), or to perform model validation (e.g., cross-validation). ArcGIS uses the term re-sampling to describe the interpolation methods used to change the resolution of a raster dataset.

Root Mean Square Error (RMSE) – RMSE measures the error associated with a model by calculating the difference between the predicted data (extracted from the model) and the response data (extracted from the underwater videos).

Spatial predictive modeling – Modeling technique whereby relationships between environmental predictors and a response (e.g., benthic habitat type) are estimated for areas with survey data. These relationships are then used to predict the response in areas without response data.

Sensitivity – A measure of model performance for binary classification models (e.g., presence/absence) that measures the proportion of presences that are correctly predicted as presences. This measure can be used to identify optimal thresholds for specific management applications. For example, if the goal is to design a new marine protected area that includes 95% of a species' distribution, than managers could identify the probability of occurrence threshold where sensitivity equals 0.95.

Specificity – A measure of model performance for binary classification models (e.g., presence/absence) that measures the proportion of absences that are correctly predicted as absences. Like for sensitivity, this measure can be used to identify optimal thresholds for specific management applications. For example, if the goal is to identify anchoring areas that exclude 95% of a species' distribution, than managers could identify the probability of occurrence threshold where specificity equals 0.95.

Test data – Data that are excluded during model fitting and used to test the predictive performance of the fitted model. Here, the test data were collected independently during the Accuracy Assessment (AA) mission.

Test percent deviance explained (PDE) – PDE calculated for a fitted model using test data from the AA mission.

Training data – Data to which a model is fitted collected during the Ground Validation (GV) mission.

Tree complexity (tc) – In boosted regression and classification tree models, a parameter that controls the number of allowable nodes in a tree. This limits the number of possible interactions between predictor variables. In general, greater tree complexity results in fewer iterations needed for model convergence.

*Macroalgae (*Udotea cyanthiformis*) and cyanobacteria on the insular shelf. Credit: T. Battista and B. Costa.*

*Fire coral (*Millepora spp.*) on the insular shelf. Credit: T. Battista and B. Costa.*

APPENDIX A. COMPOSITE HABITAT MAPS

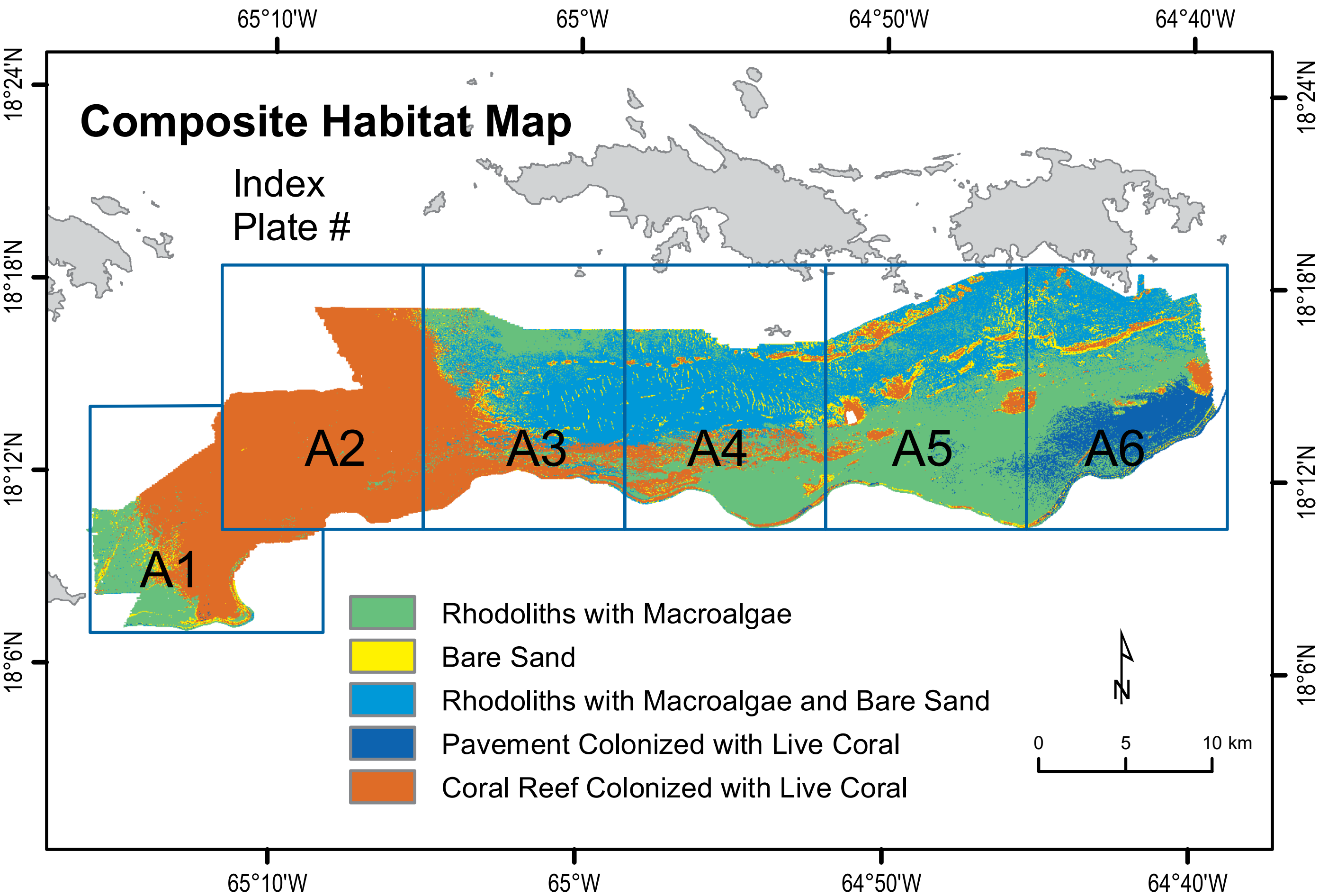


Figure. A.0. Overview of benthic habitat map sections for the insular shelf.

Appendix

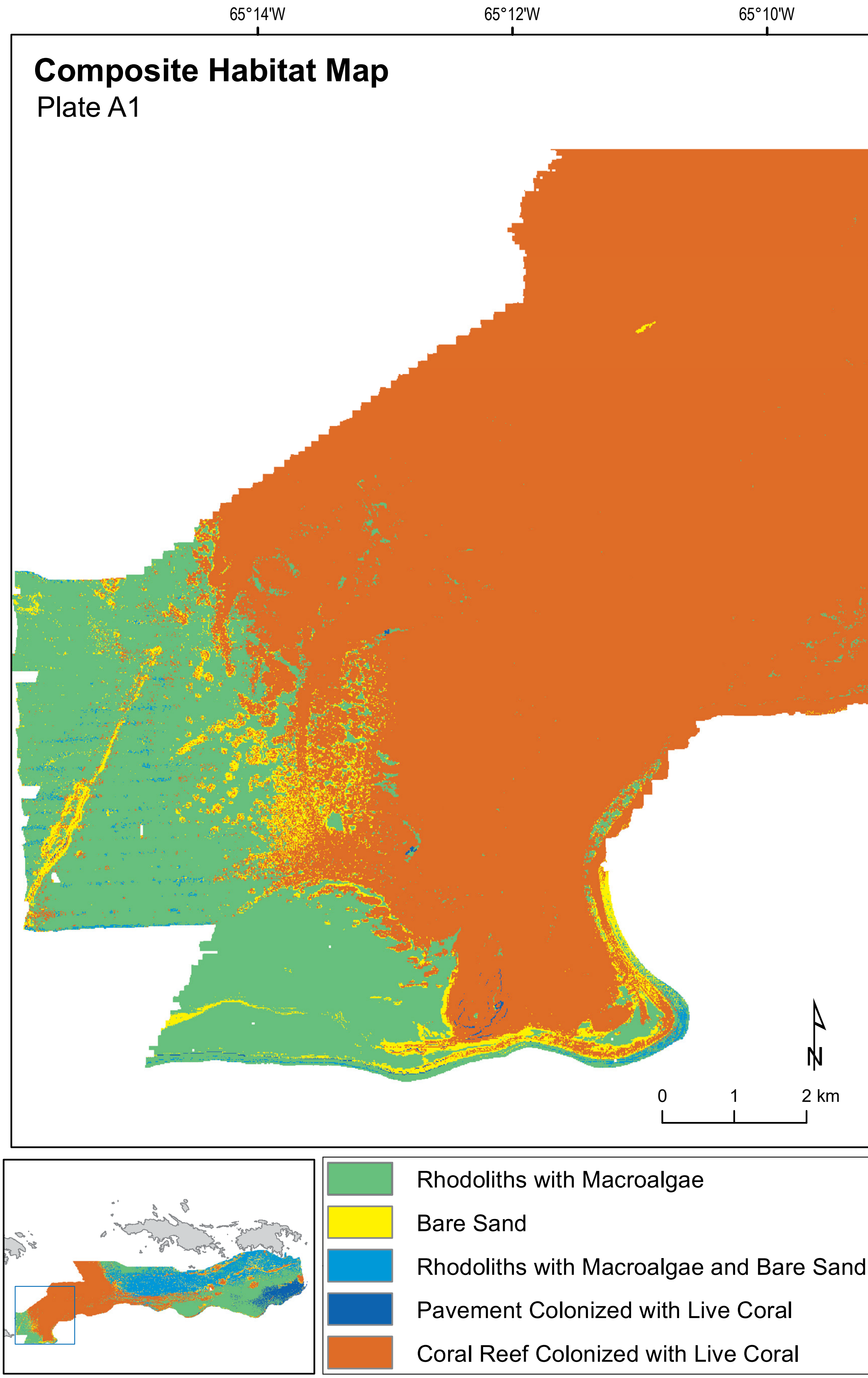


Figure A.1. Plate A1 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.

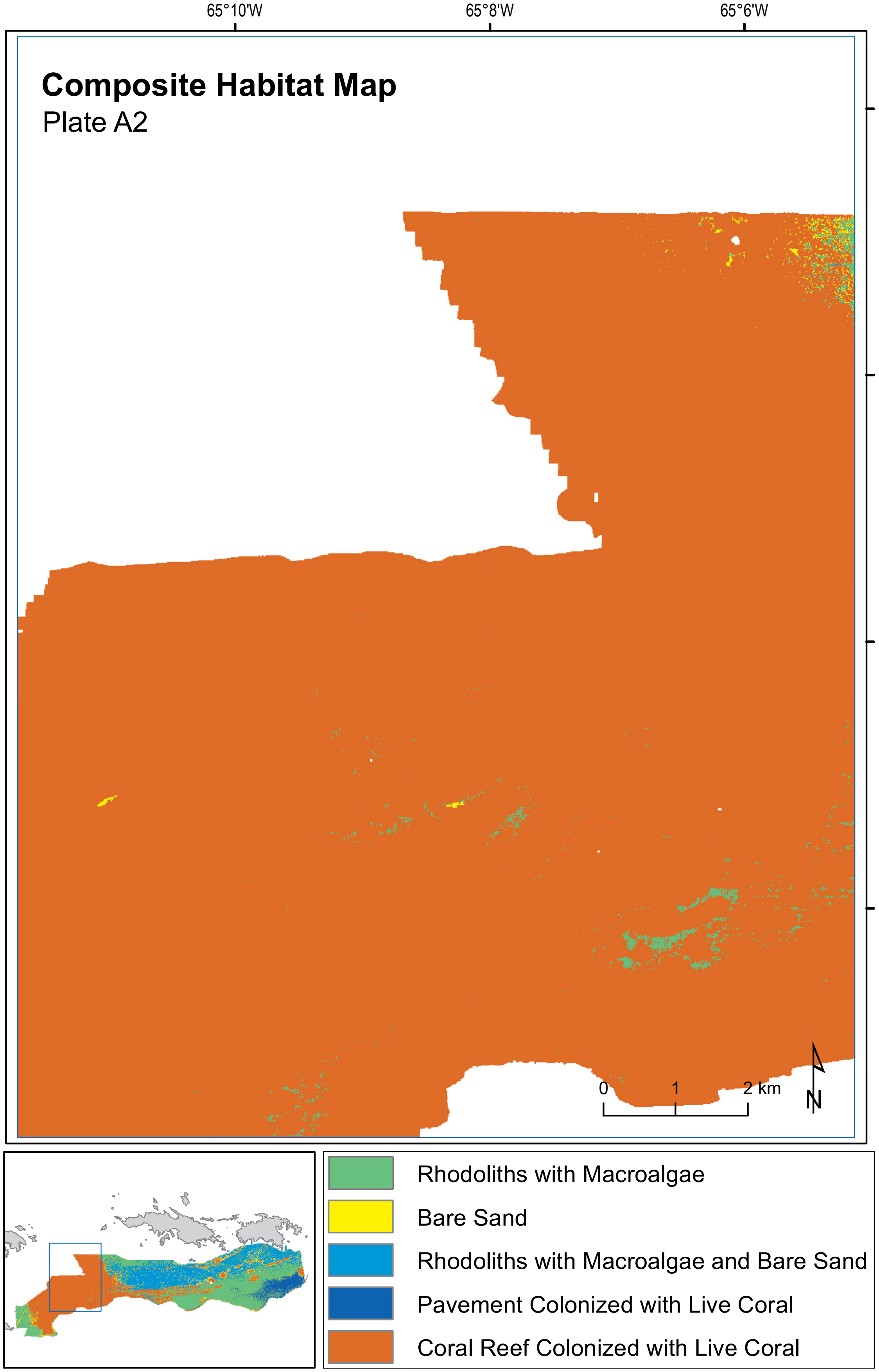
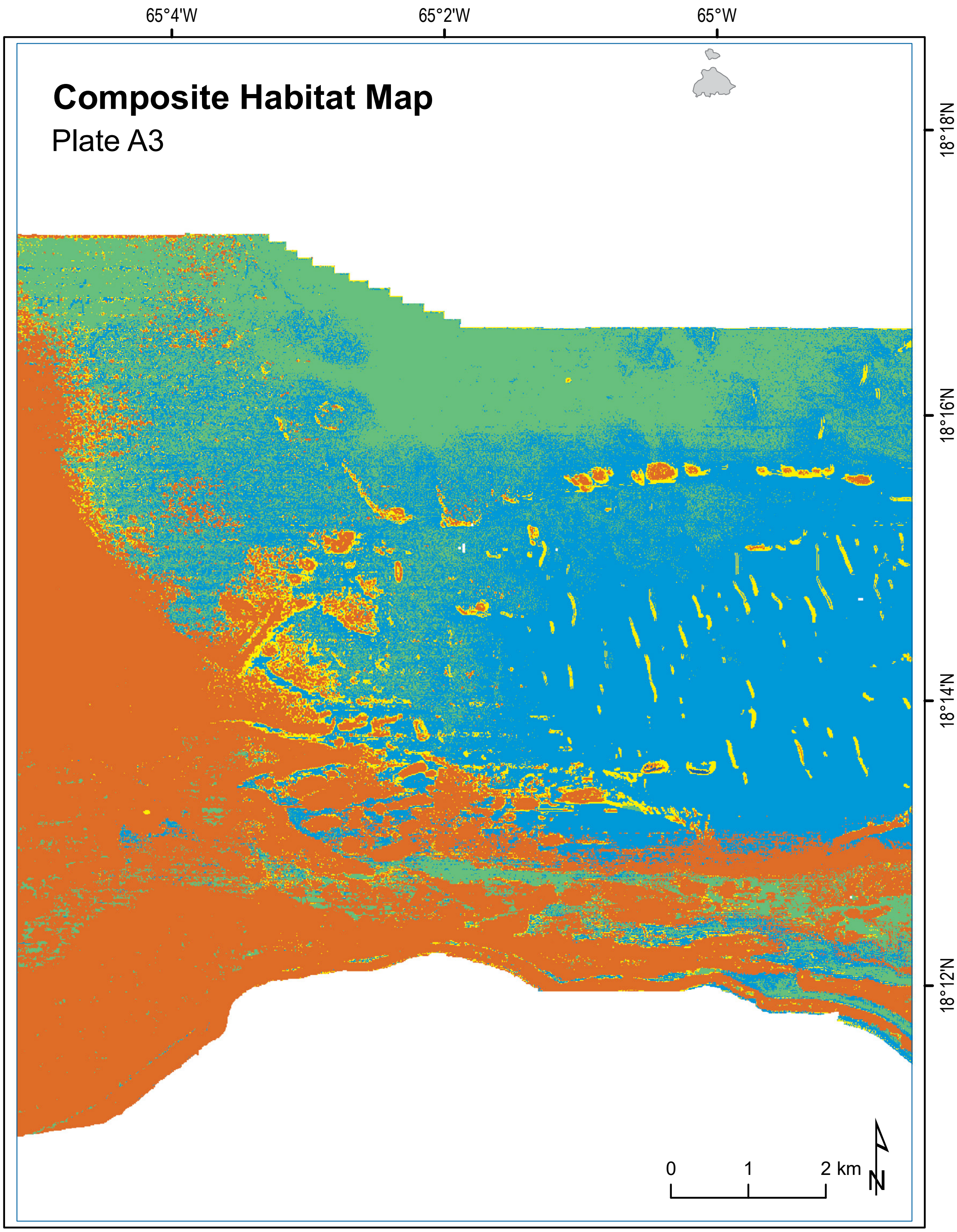
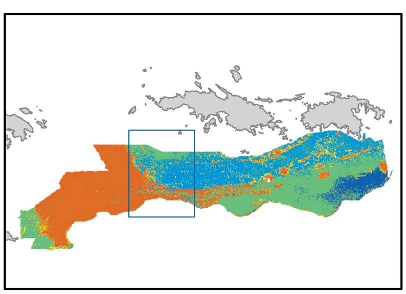


Figure A.2. Plate A2 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.





	Rhodoliths with Macroalgae
	Bare Sand
	Rhodoliths with Macroalgae and Bare Sand
	Pavement Colonized with Live Coral
	Coral Reef Colonized with Live Coral

Figure A.3. Plate A3 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.

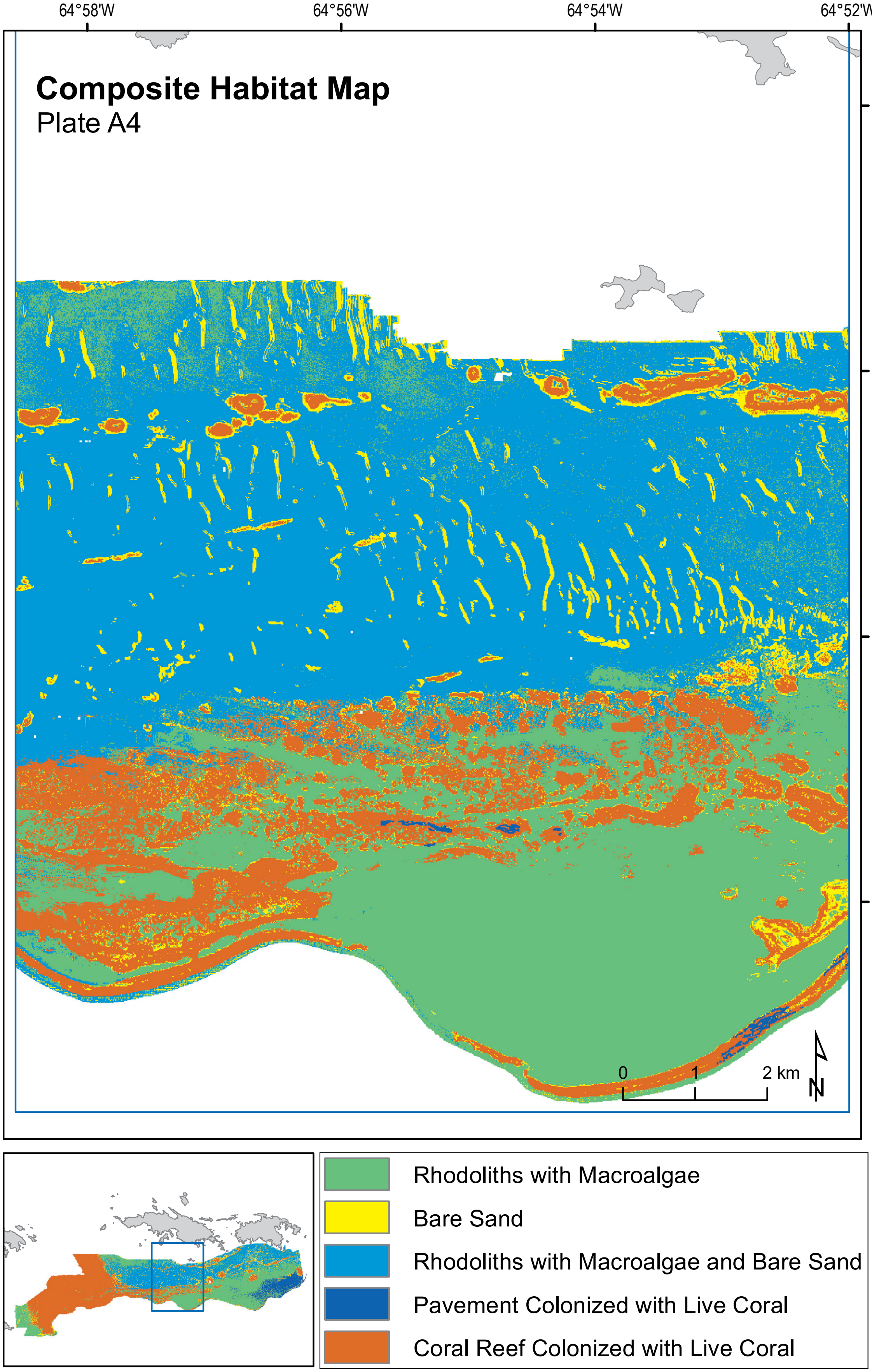


Figure A.4. Plate A4 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.

Appendix

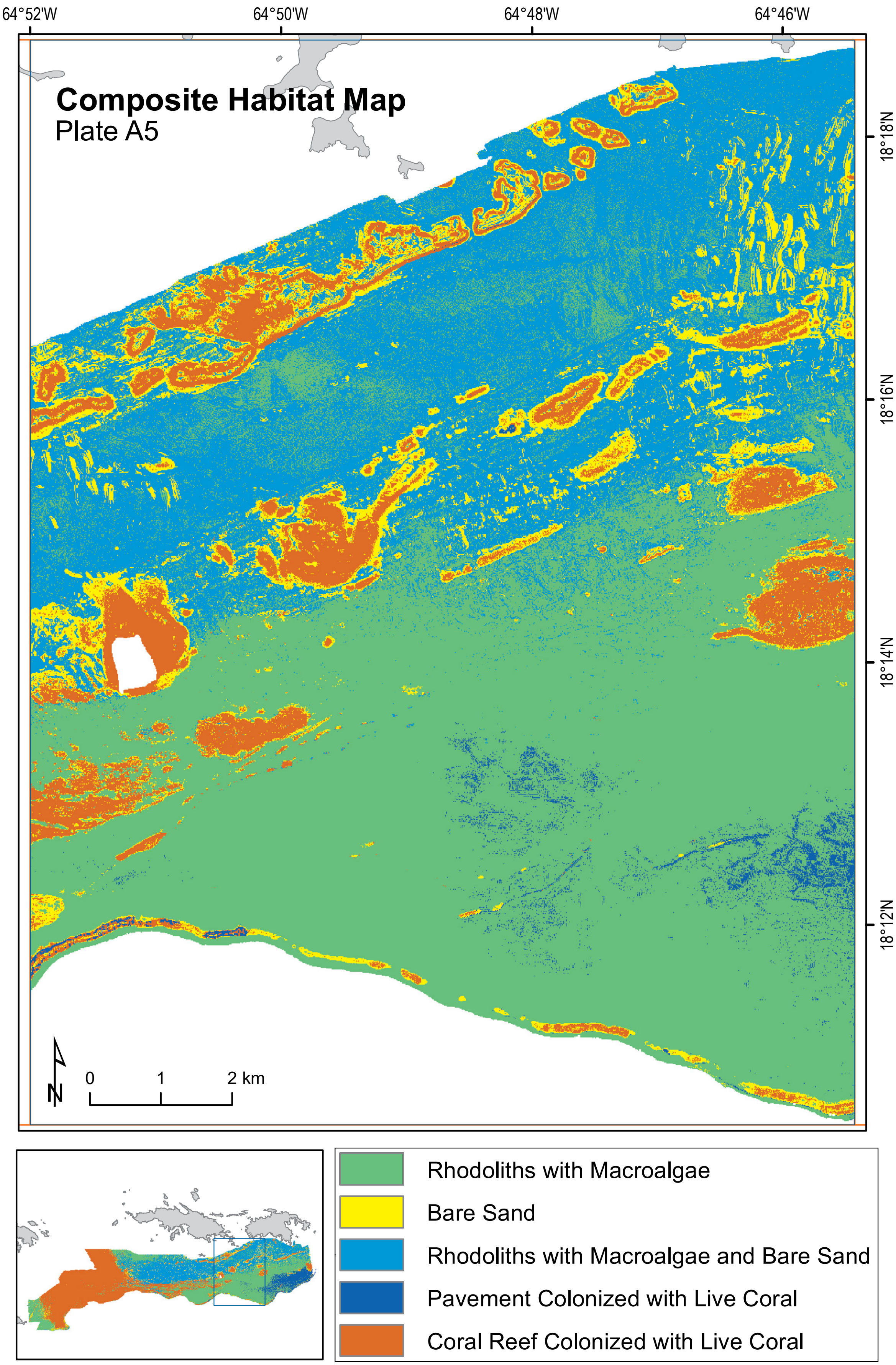


Figure A.5. Plate A5 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.

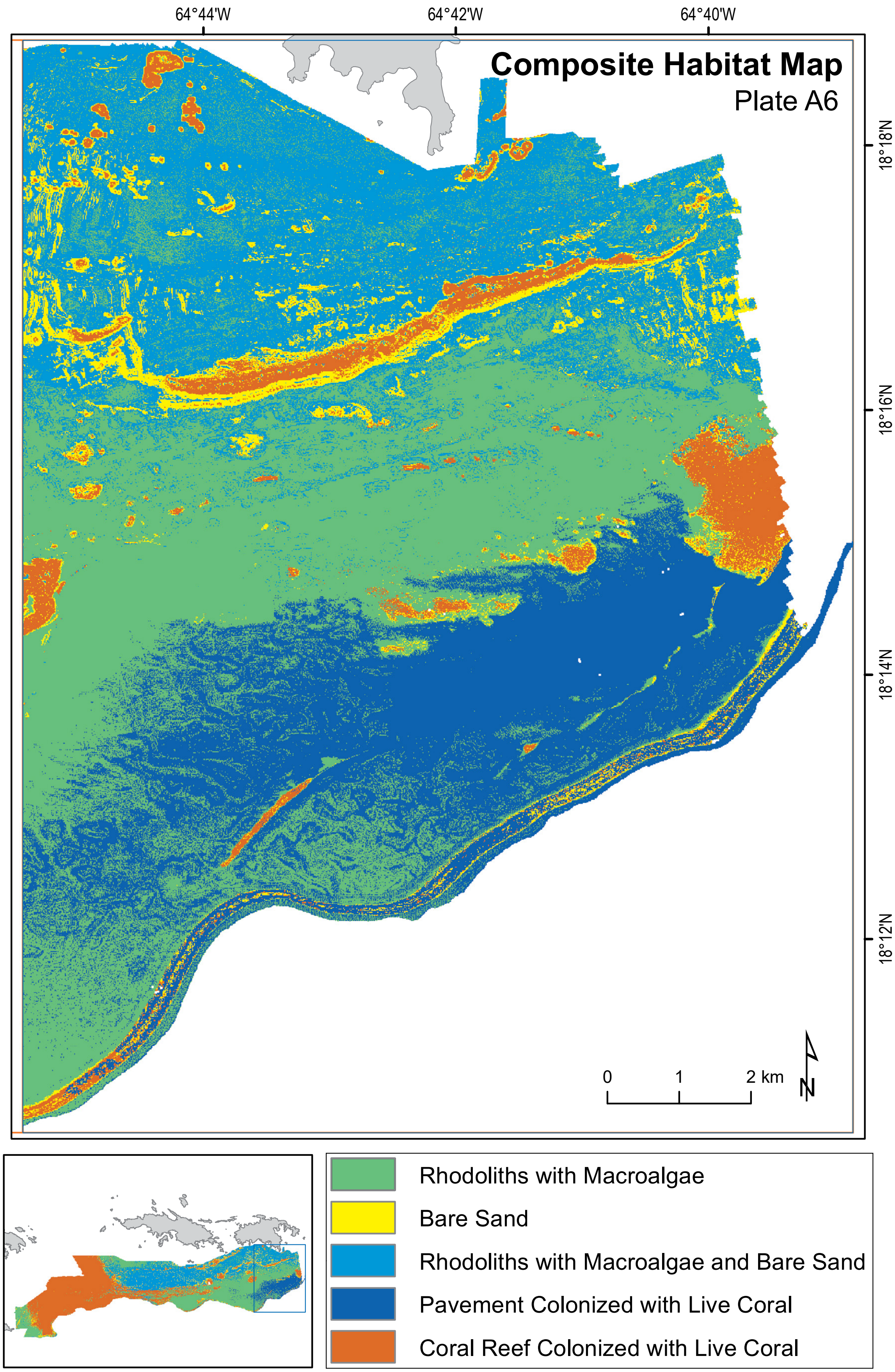


Figure A.6. Plate A6 showing benthic habitats on the insular shelf. The map inset shows the location of this plate in the project area.

Appendix

APPENDIX B. UNDERWATER VIDEOS (ROV)

Between 2004 and 2011, several different ROVs and underwater camera platforms were used to collect underwater videos and photos across a diverse set of habitats on the insular shelf. These platforms included: a MiniBAT towed camera system, a Phantom S2 ROV and a SeaViewer Drop Camera. Over this seven year period, the insular shelf was explored and documented at seven MiniBAT dives sites, 122 Phatom S2 ROV dive sites (31 in 2005, 8 in 2009, 33 in 2010, 50 in 2011) and 310 sites visited with a drop camera (104 in 2010, 206 in 2011). These dive sites span the entire insular shelf from El Seco near Isla de Vieques to the mid-shelf reef south of St. John (Figure B.1).

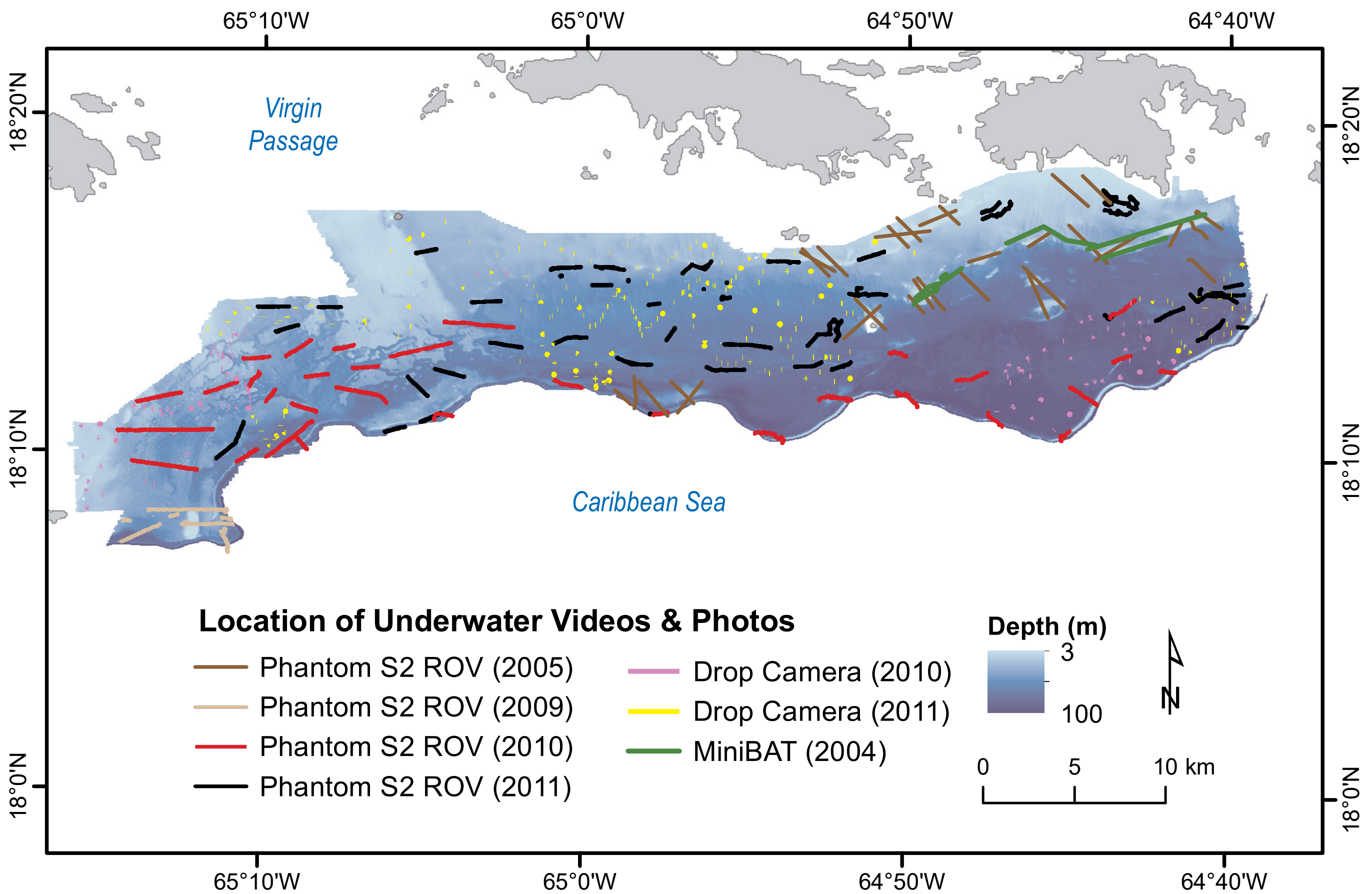


Figure B.1. Location of underwater videos and photos collected between 2004-2011 on the insular shelf using a towed MiniBAT camera system, Phantom S2 ROV and Seaviewer Drop camera.

A variety of natural patterns, maritime heritage resources, and anthropogentic impacts were documented during these ROV, towed camera, and drop camera dives on the insular shelf. Notably, the resulting videos and photos recorded the presence of strange looking species, like the short-nosed batfish (*Ogcocephalus nasutus*) (Figure B.2). They also recorded and documented species protected under the Endangered Species Act, including Nassau Grouper (*Epinephelus striatus*) (Figure B.3) and *Orbicella* coral species. These species are both considered "Threatened," requiring additional protections to prevent them from becoming endangered.

The ROV dives also confirmed the presence of fish spawning aggregation sites of commercial species, including snappers and groupers. Protecting these spawning aggregation sites is important to ensure fish populations are replenished and remain sustainable. These underwater videos also documented mesophotic coral habitats before and after the 2005 bleaching event in the U.S. Caribbean. This bleaching event caused the loss of up to 60% of the live hard coral cover in the USVI (Miller et al., 2006; Miller et al., 2009). ROV dives over the years also documented the rapid geographic expansion of the invasive Red Lionfish (*Pterois volitans*)—a native to the Pacific Ocean (Figure B.4). This invasive species has the potential to harm reef ecosystems because it does not have any natural predators in the Atlantic, and competes for food and space with overfished native stocks (such as snapper and grouper). Lastly, the ROV dives documented anthropogenic impacts to the insular shelf as well including the presence of marine debris, such derelict fish traps (Figure B.5). In some special cases, the ROV was used to explore shipwrecks that were previously uncharted and in one case, of significant historical value to the jurisdiction (Figures B.6, B.7). Combined, this extensive catalogue of georeferenced underwater videos and photos are an important permanent record and baseline for the distribution and health of coral reef ecosystems and associated fish assemblages on the insular shelf.

*Figure B.2. Two short-nosed batfish (*Ogcocephalus nasutus*). Batfish are bottom dwelling fish mostly found at depths between 200 and 1,000 meters. Credit: T. Battista and B. Costa.*

*Figure B.3. A Nassau Grouper (*Epinephelus striatus*), which is protected under the Endangered Species Act. Credit: T. Battista and B. Costa.*

*Figure B.4. An invasive Red Lionfish (*Pterois volitans*). Red lionfish are originally from the Pacific Ocean and threaten coral reef ecosystems in the Caribbean. Credit: T. Battista and B. Costa.*

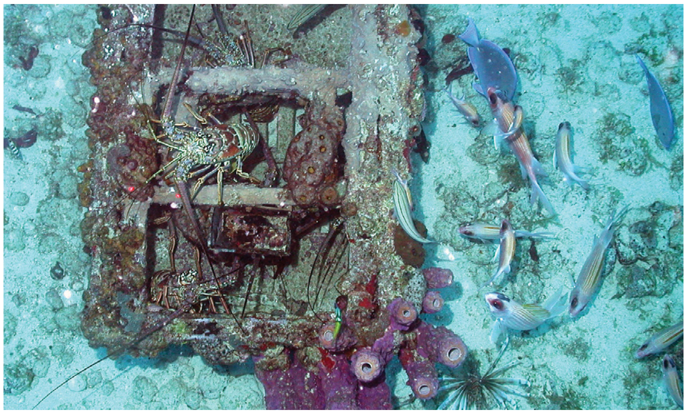


*Figure B.5. A derelict trap which has become inhabited by a Red Lionfish and several Caribbean spiny lobsters (*Panulirus argus*). Credit: T. Battista and B. Costa.*

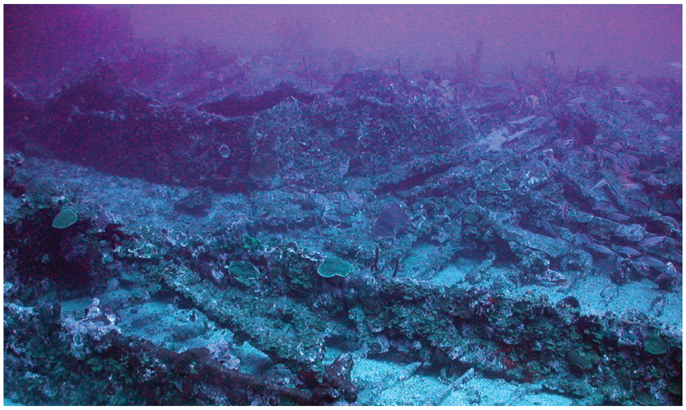


Figure B.6. A historic ship wreck found on the insular shelf. This ship is thought to have sank sometime in the late 19th century or early 20th century. Credit: T. Battista and B. Costa.

*Figure B.7. Caribbean Reef Shark (*Carcharhinus perezii*) on shipwreck on the insular shelf. Credit: T. Battista and B. Costa.*

APPENDIX C. SPLITBEAM ECHOSOUNDERS (SBES) FISHERIES ACOUSTICS

During 2009-2011, fisheries acoustics surveys were conducted in tandem with seafloor mapping and ROV operations (Kracker et al., 2011). These concurrent surveys optimize ship time and resources, and lend themselves to an integrated approach to characterizing bottom and mid-water habitats and the fishes associated with them. These fisheries acoustic surveys were conducted using SBES. SBES use a relatively narrow acoustic beam (5% of the swath of the MBES) to sample the entire watercolumn and detect backscatter from plankton to large fish and fish schools (Figure C.1). These animals have air pockets in them that reflect sound. Larger animals have more air and therefore reflect more sound than smaller ones. Data are interpreted using signal processing algorithms as described in Costa et al. (2014) to count individual fish and estimate density of fish schools. It is not possible to assign species to the acoustic targets. Instead, fish densites are estimated from the approximate target strength or fish size based on a known relationship between the amount of sound returned and the size of the fish.

Densities of fish mapped over the seafloor show possible hotspots of fish biomass. For example, the outer reef near the insular shelf break south of St. John was a consistent hotspot of fish density (Figure C.2). These outer reef ridges, at the edge of the insular shelf and the deep Caribbean Sea appear to be areas of high productivity for benthic corals and fish communities, as well as documented fish aggregations (Figure C.3; Costa et al., 2014). Where habitats appear similar on the outer reef, the densities of fish are not uniform and suggest links to oceanography. While this provides a snapshot of fish density distributions, SBES surveys also present a way to repeatedly survey these habitats to assess changes in distribution resulting from behaviors of fishes over day, night and seasons, as well as overall changes in abundance caused by potential fishing pressure, environmental impacts, or conservation measures.

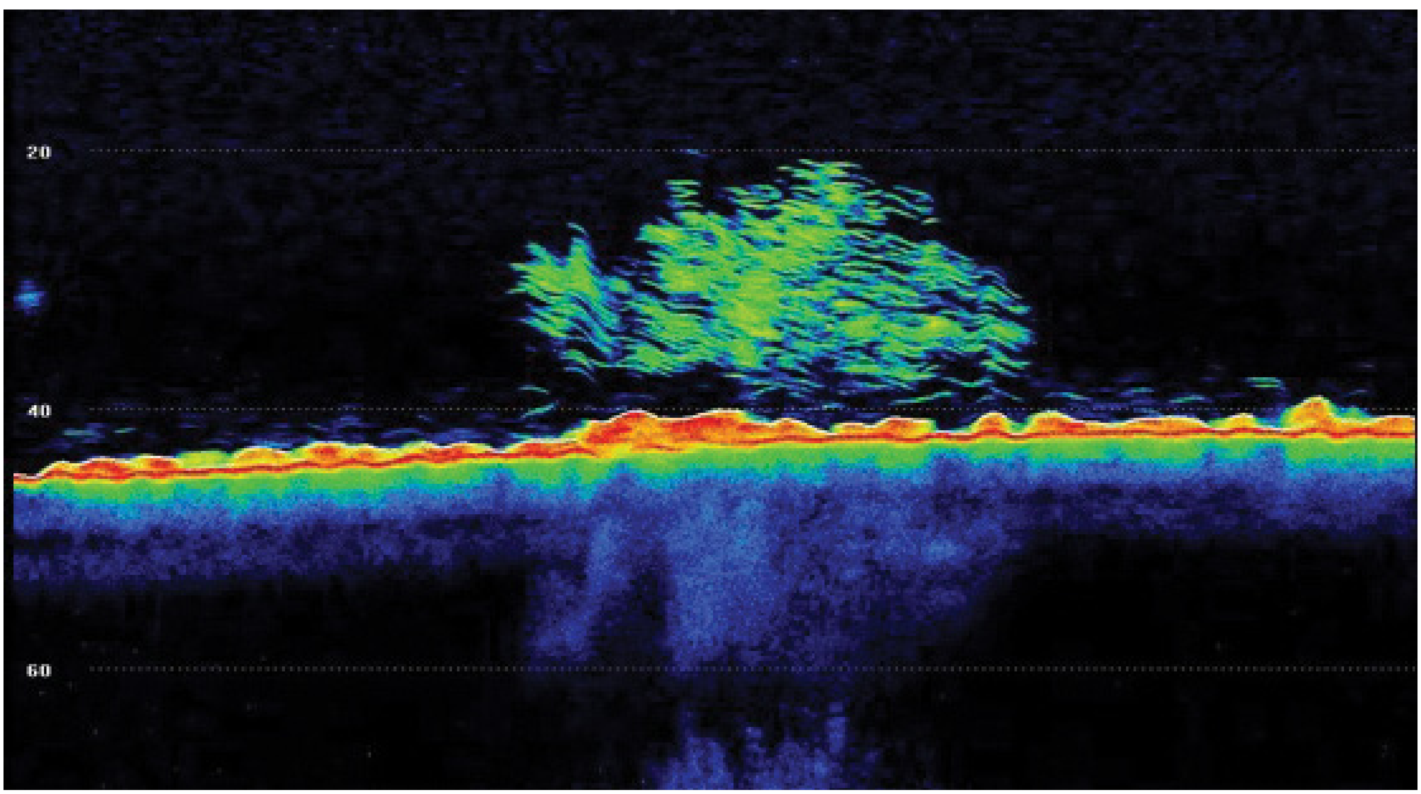


Figure C.1. Example SBES echogram showing a rough seafloor (orange-red) with individual fish (blue traces) and a large fish school (green traces).

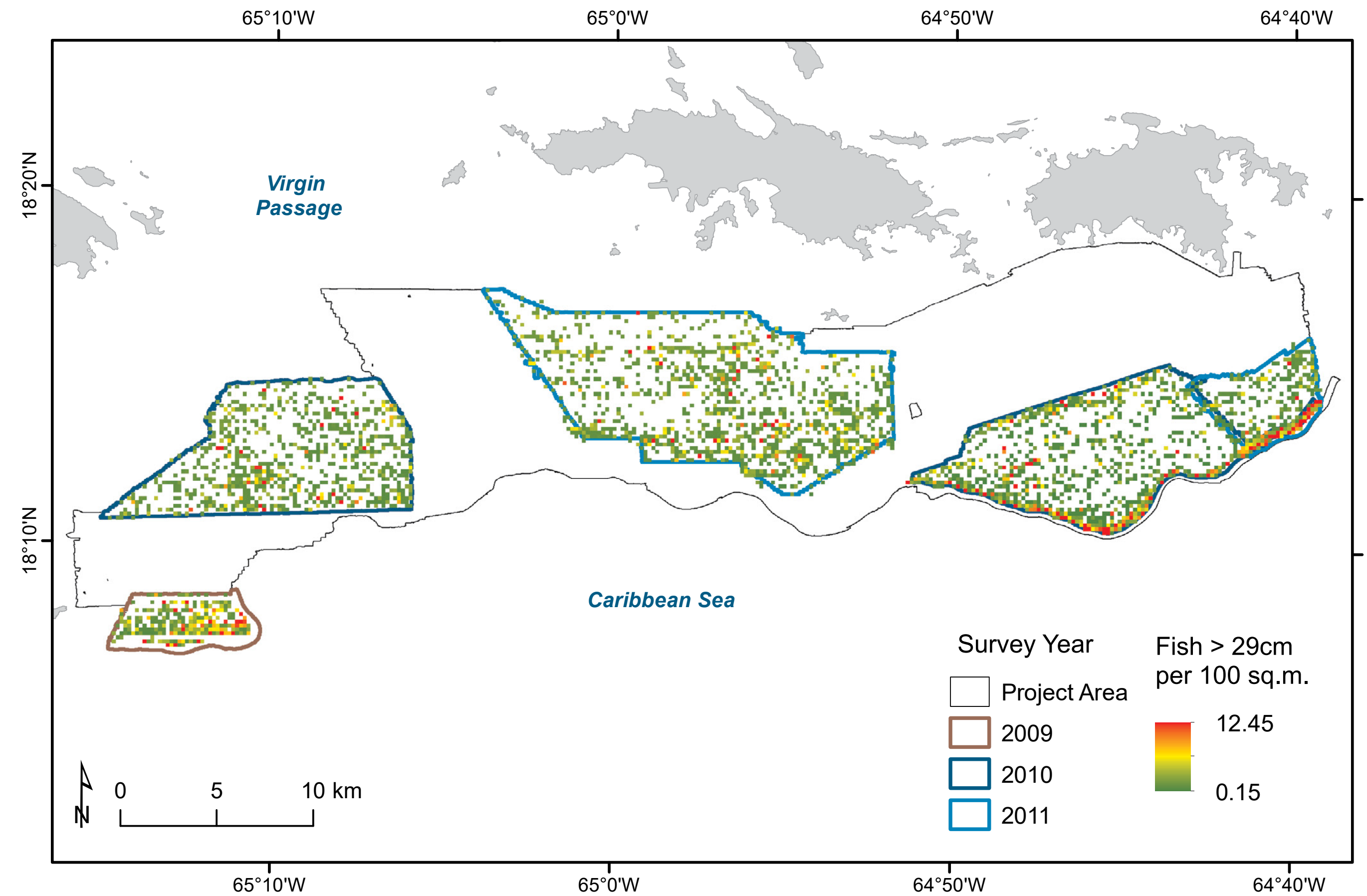


Figure C.2. Densities of large fish (estimated size >29 cm) in three areas of the insular shelf surveyed during years indicated by colored polygon. Fish densities presented as number of fish per 100 m², with zero fish not displayed.

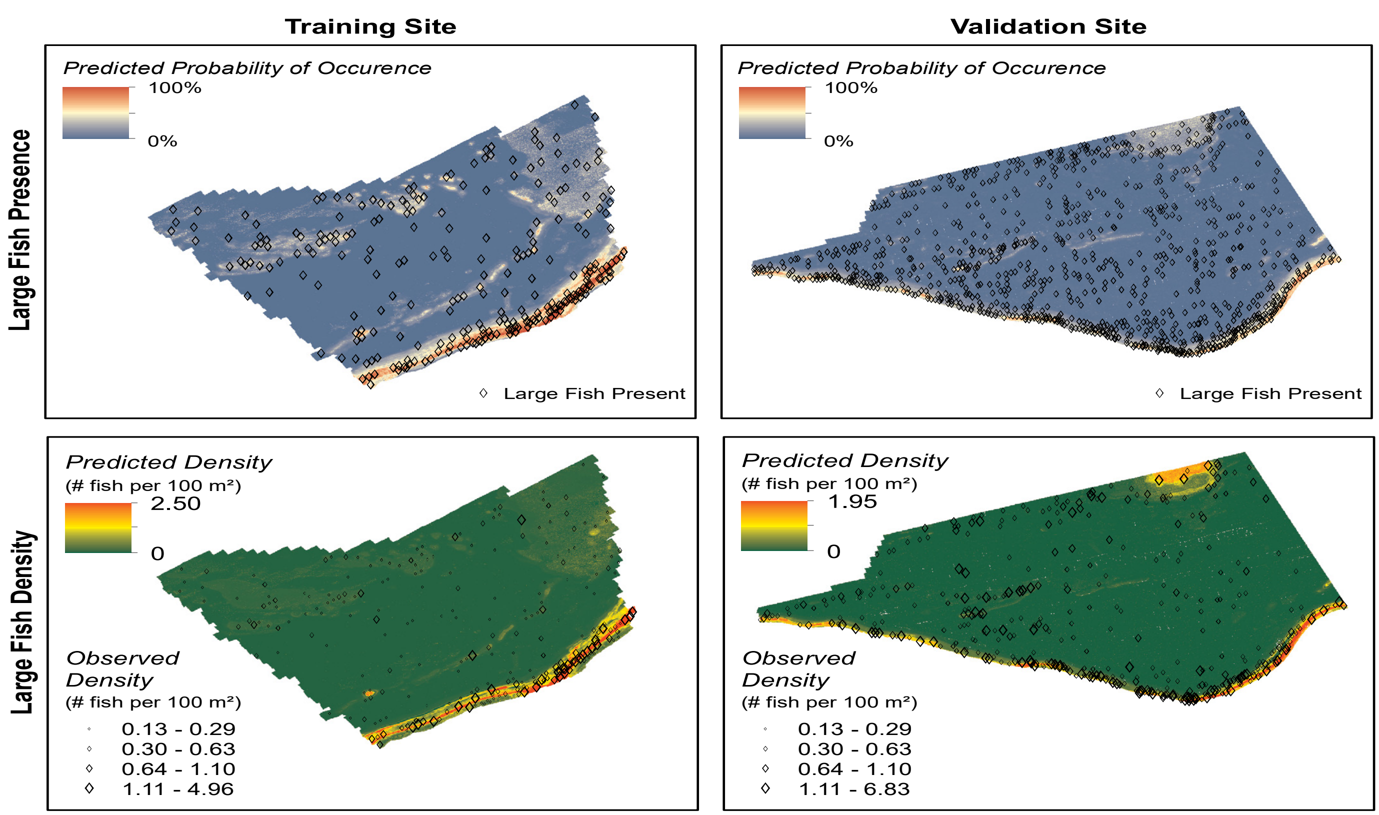


Figure C.3. Predicted probabilities of occurrence and densities of large fish on the eastern portion of the insular shelf. Figure adapted from Costa et al., 2014.

U.S. Department of Commerce
Wilbur L. Ross, Jr., *Secretary*

National Oceanic and Atmospheric Administration
Timothy Gallaudet, *Acting Under Secretary for Oceans and Atmosphere*

National Ocean Service
Russell Callender, *Assistant Administrator for National Ocean Service*

The mission of the National Centers for Coastal Ocean Science is to provide managers with scientific information and tools needed to balance society's environmental, social, and economic goals. For more information, visit: https://coastalscience.noaa.gov/