Development and Use of a Novel Main Hawaiian Islands Bathymetric and Backscatter Synthesis in a Stratified Fishery-independent Bottomfish Survey

Benjamin L. Richards, John R. Smith, Steven G. Smith, Jerald S. Ault, Christopher D. Kelley, Virginia N. Moriwake

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Pacific Islands Fisheries Science Center

NOAA Technical Memorandum NMFS-PIFSC-87
https://doi.org/10.25923/bh8v-0184

August 2019

Development and Use of a Novel Main Hawaiian Islands Bathymetric and Backscatter Synthesis in a Stratified Fishery-independent Bottomfish Survey

Benjamin L. Richards[1], John R. Smith[2], Steven G. Smith[3], Jerald S. Ault[3], Christopher D. Kelley[2], Virginia N. Moriwake[2]

[1]Pacific Islands Fisheries Science Center
National Marine Fisheries Service
1845 Wasp Boulevard
Honolulu, Hawaii 96818

[2]Department of Oceanography
University of Hawaii
1000 Pope Road
Honolulu, Hawaii 96822

[3] Rosenstiel School for Marine and Atmospheric Science
University of Miami
4600 Rickenbacker Causeway
Miami, Florida 33149

NOAA Technical Memorandum NMFS-PIFSC-87
August 2019

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

National Oceanic and Atmospheric Administration
Neil A. Jacobs, Ph.D., Acting NOAA Administrator

National Marine Fisheries Service
Chris Oliver, Assistant Administrator for Fisheries

Recommended citation:

Richards BL, Smith JR, Smith SG, Ault JS, Kelley CD, Moriwake VN. 2019. Development and use of a novel main Hawaiian Islands bathymetric and backscatter synthesis in a stratified fishery-independent bottomfish survey. NOAA Tech Memo. NMFS-PIFSC-87, 48 p.
doi: 10.25923/bh8v-0184

Copies of this report are available from:

Science Operations Division
Pacific Islands Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
1845 Wasp Boulevard, Building #176
Honolulu, Hawaii 96818

Or online at:

https://repository.library.noaa.gov/

Table of Contents

List of Tables

List of Figures

List of Figures (cont'd)

Executive Summary

The Hawaii deep-slope bottomfish fishery preferentially targets seven high value species (i.e. six snappers and one grouper, hereafter referred to as Deep7) and represents the major insular commercial fishery in the state. The NOAA Pacific Islands Fisheries Science Center (PIFSC) Stock Assessment Program (SAP) is responsible for regularly assessing this stock complex.

To augment and improve the data used in the main Hawaiian Islands Deep7 Stock Assessment, PIFSC began developing a multi-gear Bottomfish Fishery-Independent Survey for the Deep7 in 2011. Benthic habitat has been shown to play a key role in determining the spatial distribution of demersal living marine resources. Accordingly, the BFISH survey employs a stratified random sampling design based on depth and habitat.

Prior to the work presented here, full development of the BFISH survey was impeded by a lack of usable backscatter data across the entire MHI domain. Backscatter data had been collected throughout the MHI using various platforms and sensors, thus producing a data set that included numeric backscatter values representing different levels of substrate hardness in different areas.

To remedy this situation, and to improve stratification of the BFISH survey, the NOAA Pacific Islands Fisheries Science Center collaborated with the University of Hawaii at Mānoa, to create a comprehensive, high-resolution, bathymetric and backscatter synthesis for the MHI. We describe this effort in two parts:

1. Part I: Developing the 5-m resolution bathymetric and backscatter synthesis covering the MHI at depths of 75 to 400 m;

2. Part II: Using these new products to refine and improve stratification for the NOAA PIFSC Bottomfish Fishery-independent Survey in Hawaii.

Introduction

Commercial and recreational fishing are extremely important to the economy and culture of Hawaii (Haight et al. 1993) and the "Deep7" bottomfish stock, comprising six snapper and one grouper species, represents one of the more important insular fisheries in the region (Western Pacific Regional Fishery Management Council 2010). The NOAA Pacific Islands Fisheries Science Center (PIFSC) Stock Assessment Program (SAP) is responsible for regularly assessing this stock, requiring reliable time-series of catch, abundance, and life history demographics to estimate trends and to evaluate sustainability benchmarks.

In its efforts to continually improve the data used in the Deep7 assessment, PIFSC has developed a main Hawaiian Islands (MHI) multi-gear Bottomfish Fishery-Independent Survey (BFISH) (Richards et al., 2016). Essential fish habitat (EFH) for the Deep7 stock is defined as all MHI waters within the 75–400 m depth range (United States Department of Commerce 2016). Within this domain, the BFISH survey employs a stratified-random experimental design to improve survey precision and efficiency. As benthic habitat has been shown to play an important role in structuring the spatial distribution of demersal living marine resources (Richards et al. 2012), including the Deep7 (Moore 2008), sampling gears are randomly allocated to primary sampling units (PSUs) stratified by depth and habitat type (e.g. high slope or low slope, hardbottom or softbottom).

Each BFISH PSU is defined as a 500-m × 500-m grid cell. Defining appropriate strata designations for any given PSU entails two aspects: 1) designating a threshold value that distinguishes "hardbottom" from "softbottom" or high slope from low slope at the individual pixel level; and (2) determining criteria for the number of pixels (amount of area) within a PSU necessary to assign the most appropriate strata.

Pilot BFISH experiments conducted from 2011 to 2015 (Richards et al., 2016) used a stratification scheme developed from existing 20-m resolution multi-beam bathymetry and backscatter data available for the Maui-Nui region (Figure 1). Each 20-m pixel was classified as "hardbottom" if it exceeded a backscatter threshold value of 187 n scaling from pixels to PSU, PSUs were classified as "hardbottom" if one or more hardbottom pixels were present within the PSU. Similarly, PSUs with one or more high slope pixels were classified as high slope (Table 1).

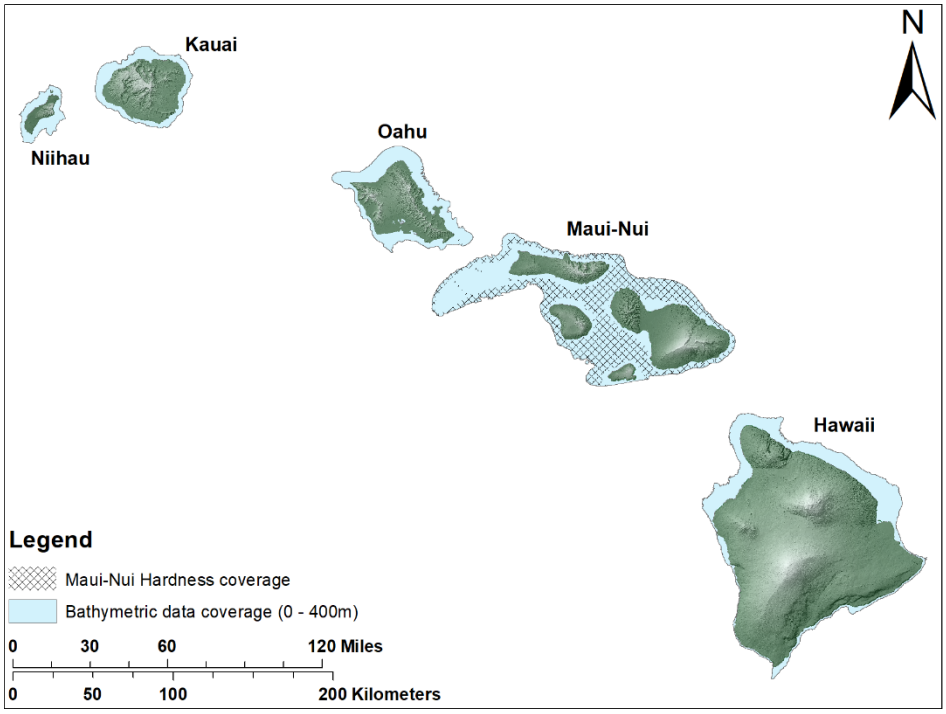


Figure 1. Map of the main Hawaiian Islands (MHI) divided into five island groups. Bathymetric data (blue shading) exists for all zones. Prior to the current project, benthic habitat data existed only for the Maui-Nui island region (hatched area).

Table 1. Definitions for pilot hardbottom/softbottom and high slope/low slope for 20-m mapping data for the Maui-Nui region.

Variable Category	Threshold Value	Threshold PSU Area
Hardbottom	Backscatter ≥ 187	≥ 0.16% (1/625 pixels)
Softbottom	Backscatter <1 87	100.0% (625/625 pixels)
High Slope	Slope ≥ 20 degrees	≥ 0.16%
Low Slope	Slope < 20 degrees	100.0%

The 2011–2015 pilot BFISH experiments conducted in the Maui-Nui region showed that habitat-based stratification using available 20-m resolution bathymetric and backscatter information effectively partitioned spatial variance. However, expansion of these pilot experiments into an operational survey was hampered by a lack of comprehensive, comparable, high-resolution bathymetric and backscatter data outside of the Maui-Nui region. Bathymetric mapping surveys outside of Maui-Nui have been conducted, but using a variety of platforms and sensors operating at a range of different frequencies. Hence, similar numeric values from disparate regions do not represent similar levels of substrate hardness, precluding their use in the development of consistent strata across the full survey domain.

To remedy this situation, researchers at PIFSC and the University of Hawaii at Mānoa conducted a 1-year project to synthesize comprehensive and consistent, high-resolution depth, slope, and seafloor hardness data for the MHI. This paper describes that effort in two parts:

1. Part I: Methods and results pertaining to the development of a comprehensive, 5-m resolution MHI bathymetric and backscatter synthesis covering depths of 75 to 400 m;
2. Part II: Novel methods and results by which this synthesis was used to quantitatively define an effective habitat stratification for the BFISH survey.

Part I: Development of a Bathymetric and Backscatter Synthesis for the Main Hawaiian Islands

Backscatter data for the majority of the 100 to 400 m depth in the MHI (and to a lesser extent from the shoreline down to 100 m) have been collected and archived at the University of Hawaii Pacific Benthic Habitat Mapping Center (University of Hawaii 2019a) and Hawaii Mapping Research Group (University of Hawaii 2019b). However, these data sets have been collected from a number of ships, at different times and spatial scales, using a variety of multibeam systems (Table 2) and under the direction of multiple organizations and research groups (Table 3).

Table 2. Data collection sensors and vessels.

Sensor	Frequency	Vessel
Kongsberg EM 120	12 kHz	R/V *Kilo Moana*
Kongsberg EM 122	12 kHz	R/V *Kilo Moana*
Kongsberg EM 300	30 kHz	M/V *Ocean Alert* & NOAA Ship *Hi'ialakai*
Kongsberg EM 302	30 kHz	NOAA Ship *Okeanos Explorer &* R/V *Falkor*
Kongsberg EM 710	40 – 100 kHz	R/V *Kilo Moana* & R/V *Falkor*
Kongsberg EM 1002	95 kHz	R/V *Kilo Moana*
Kongsberg EM 3002d	300 kHz band	NOAA Ship *Hi'ialakai*
Reson 8101-ER	240 kHz	R/V *AHI*

Table 3. Data collection organization or research group.

Entity Name
University of Hawai'i Undersea Research Laboratory (HURL)
University of Hawai'i Mapping Research Group (HMRG)
U.S. Geological Survey (USGS), Western Coastal and Marine Geology Program
NOAA/PIFSC Coral Reef Ecosystem Division (CRED)
NOAA/PIFSC Pacific Islands Benthic Habitat Mapping Center (PIBHMC)
NOAA Undersea Research Program (NURP)
NOAA Office of Ocean Exploration and Research (OER)
NOAA Pacific Islands Regional Office (PIRO)
Monterey Bay Aquarium Research Institute (MBARI)
Schmidt Ocean Institute (SOI)

We developed a GIS protocol (Appendix A) to reprocess and standardize backscatter values from these data sets to a uniform scale of intensity (0–256). This protocol was first used to create a 60-m resolution backscatter synthesis of the entire MHI from the coastline to abyssal depths (Kelley & Smith, 2014). We applied this method to create a more detailed 5-m resolution backscatter synthesis for the 75–400 m depth range necessary for the BFISH survey. A corresponding 5-m multi-beam bathymetry synthesis was also developed and used to correct the backscatter data for topographic slope. The depth range for both products extends from as shallow as data were available to at least 500 m.

The quality and resolution of multi-beam bathymetry data vary substantially between different platforms and systems. However, various bathymetric data sets can be meaningfully combined with the application of appropriate weighting factors. Newer multi-beam systems collect coincident bathymetry and backscatter data in the same data file or file system. By changing options in the post-processing software, either component can be further processed, edited, gridded, plotted, and/or merged with other like data. However, the backscatter component cannot be meaningfully merged with backscatter data from other systems, frequencies, and/or sometimes platforms, in the same way as the bathymetric component.

Raw bathymetric and backscatter point data can be "gridded" at user-defined resolutions by merging information from all points that exist within grid cells at the desired resolution. To generate grids of backscatter and bathymetry at 5-m cell size, each island group (Figure 1) was divided into smaller sub-domains that together comprised the 0–400 m depth range. This was done to increase gridding speed, remain within computer memory limits, and simplify data management for the numerous multibeam input files. This subdivision resulted in 19 sub-domains: 3 for Niihau and Kauai (combined), 3 for Oahu, 9 for Maui Nui, and 4 for Hawai‘i Island (Figure 2). The resulting sub-domain-level grids were subsequently merged to one layer for each island group, and finally into a seamless layer for the entire MHI.

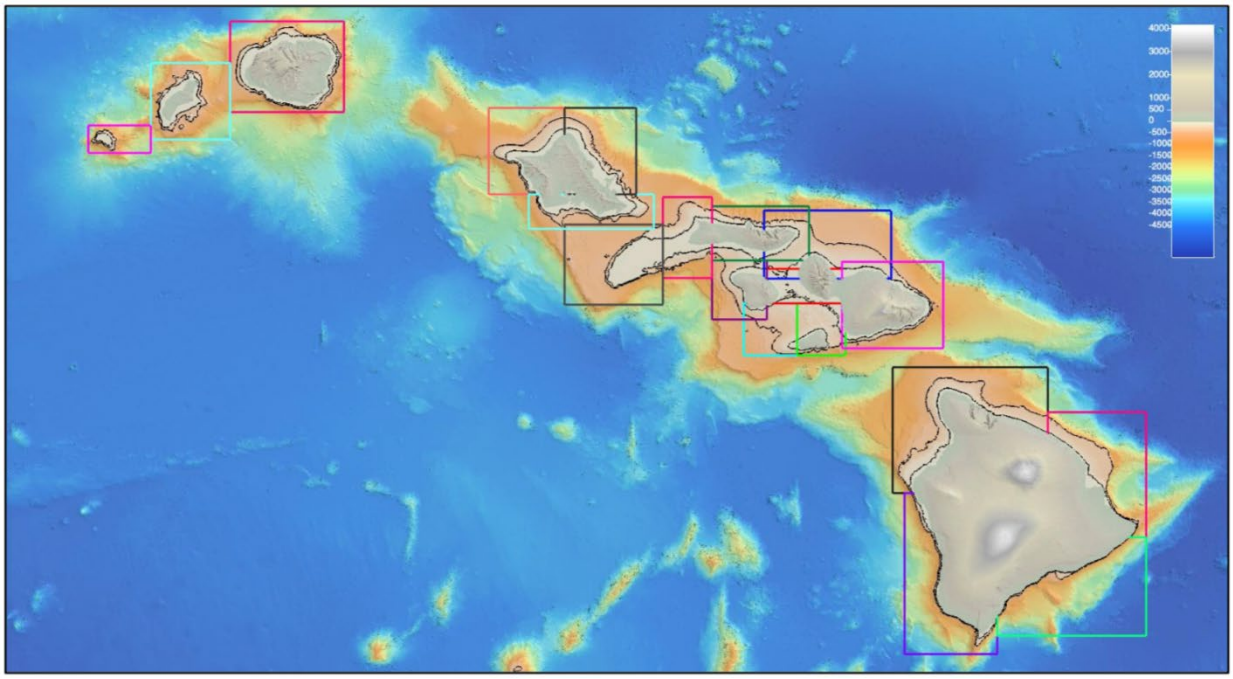


Figure 2. Bathymetry and topography of the main Hawaiian Islands with variously colored lines delineating processing sub-domains.

To create the backscatter synthesis, 34,618 individual multibeam bathymetry files from hundreds of dedicated bathymetric surveys were selected by examining track line pattern, extent of coverage, and data quality. While selection criteria were mostly qualitative (e.g. what looked the best overall) some quantitative information (e.g. which multibeam systems should produce the best data in the given depth range) was available. Opportunistic surveys, where bathymetric mapping was not a priority, were only used when no other data was available. We assigned

weighting values to each raw data file based on the above information prior to incorporating the data into the grid for the sub-domain. Methods Used to Produce the Syntheses

Bathymetry (Depth)

Bathymetric data were acquired with the systems listed in Table 2, along with additional multibeam systems from other ships, including SeaBeam classic, SeaBeam 2000, SeaBeam 2112, Hydrosweep. Where no multibeam data existed, LIDAR data were also incorporated where available.

The backscatter synthesis is a subset of the bathymetric synthesis. We used all available files of good quality in the bathymetric synthesis with appropriate weighting. Files, or surveys, used in the backscatter synthesis were manually chosen, reflecting the highest quality, geometries, and coverage possible.

Not all bathymetric data files included appropriate backscatter data for several reasons:

1) Older multibeam systems only collected bathymetric data. The aforementioned multibeam systems are those without backscatter data, or of poor quality. This includes LiDAR which, not being a SONAR-based system, does not include backscatter data.
2) For the most part, only dedicated surveys were used in the backscatter synthesis, rather than adding in all available files that may include transit lines of lower quality.
3) Backscatter data is more susceptible to quality issues than bathymetry, such as survey geometry, look direction, environmental, etc.

Raw multibeam "swath" files commonly contain outliers and extraneous data that are edited and removed ("cleaned") by the survey teams while aboard ship and/or after returning to the lab using a variety of open source and commercial software including, but not limited to, MB-System (Paduan, 2016, Caress and Chayes, 1996), SABER ("SABER," 2018), CARIS ("HIPS and SIPS 11.0 | Teledyne CARIS," n.d.), and Fledermaus ("Fledermaus," n.d.). Outliers can result from a number of sources including bad sonar ping returns from rapidly changing seafloor terrain, ship or sea noise interference from high transit speeds, and turns in the survey track. When available, cleaned versions of the original swath files were used as input for gridding. When cleaned data were not available, raw files were used but were typically assigned a low weighting so that, when data overlap occurred, cleaned data would take precedence. Similar weighting was used to ensure that data from dedicated surveys took precedence over data from opportunistic ones.

While detailed editing of bathymetry data was beyond the scope of this project, gridded results were visually examined for obvious nadir, stair-step, or other artifacts and, if found, further analysis was carried out to remove the corresponding input data and rerun the gridding process. This method, while somewhat subjective, permitted the use of nearly all available data and creation of complete grids with limited data gaps. A smoothing filter was run over the final raw grid to further reduce the number and severity of outliers.

Bathymetric gridding was exclusively carried out using the industry standard MB-System (Caress & Chayes, 1996) and Generic Mapping Tools (GMT) (Wessel & Smith, 1991) open source software packages. Examples of the gridding and filtering commands are provided in Appendix A.

As mentioned earlier, processing and gridding was carried out at the sub-domain level before merging to the island and domain level (Figure 3).

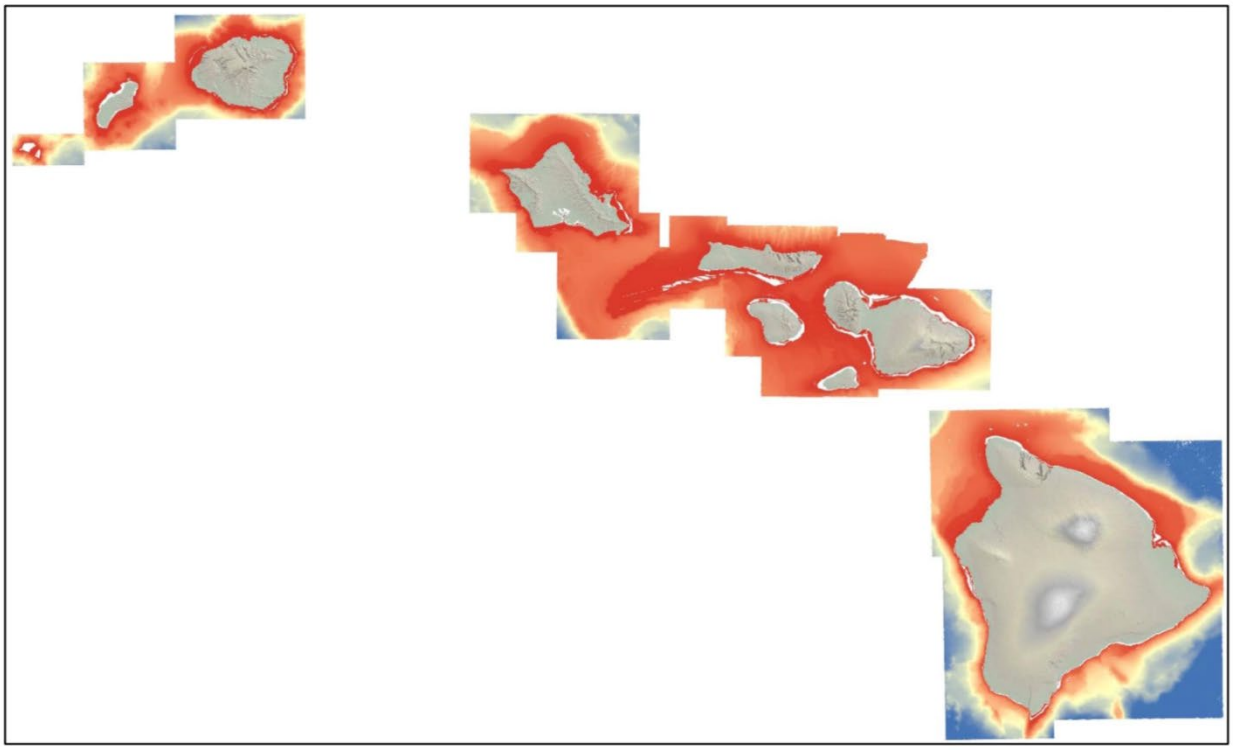


Figure 3. The 5-m resolution multibeam bathymetric synthesis for the MHI with island topography in gray shades. Cooler colors (i.e., blue) indicate deeper depths, while hotter colors (i.e., red) indicate shallower depths.

Backscatter (Hardness)

We used raw swath files as input data for backscatter processing as this resulted in a more consistent product across surveys, systems, and ships. Use of raw data also allowed us to use the Fledermaus Geocoder Toolbox (FMGT) 7.8.4 software module (Fonseca and Calder 2009), which will only accept *raw.all* or certain Generic Sensor Format (GSF) multibeam data. This is a relatively new technique and, to the best of our knowledge, had not yet been applied to any of the existing data. To compensate for bathymetric artifacts, filtered bathymetric "reference" grids were read into the FMGT project to correct the backscatter data. We should note that FMGT is far more sensitive than MB-System with respect to corrupted or missing data or metadata in the files. MGT is also more memory-limited program that will refuse to run until memory issue are addressed and which will crash repeatedly when it encounters a significant error.

In all but one case, we used the default settings for FMGT to process the backscatter data. The software reads the swath files' metadata and automatically adjusts the processing scheme to produce the best result. The metadata contains information such as geographic extents, datagram packet numbers, sonar modalities, and sonar type. After the metadata are extracted, FMGT computes the coverage and extracts the navigation of each line. FMGT first adjusts the backscatter data by extracting the backscatter then executing radiometric corrections based on sonar type and bottom topography. Next, it performs filter processing that includes angle varying

gain (AVG) adjustments along with anti-aliasing of the backscatter data. As this stage advances, it transmits the results of the backscatter adjustment to the project hierarchy. Finally, FMGT builds the desired mosaic using the pre-calculated or manually assigned resolution. This value is pre-computed by FMGT during coverage processing and is estimated based on the sonar beam configuration and along track coverage. When processing EM 1002 data from R/V *Kilo Moana*, the Backscatter Source was changed from the default Beam Time Series to Beam Average (calibrated) to produce a more consistent result for the entire survey. We also manually assigned a uniform gridding resolution of 5 m, overriding automatic resolution estimates.

Once initial backscatter layers had been created, iterative and somewhat subjective decisions were made to edit and/or toggle off files because of turns, data artifacts, or overlapping swaths. Backscatter mosaics were then re-rendered and output to ArcGIS v10.5.1 (ESRI Inc. 2017) (Figure 4).

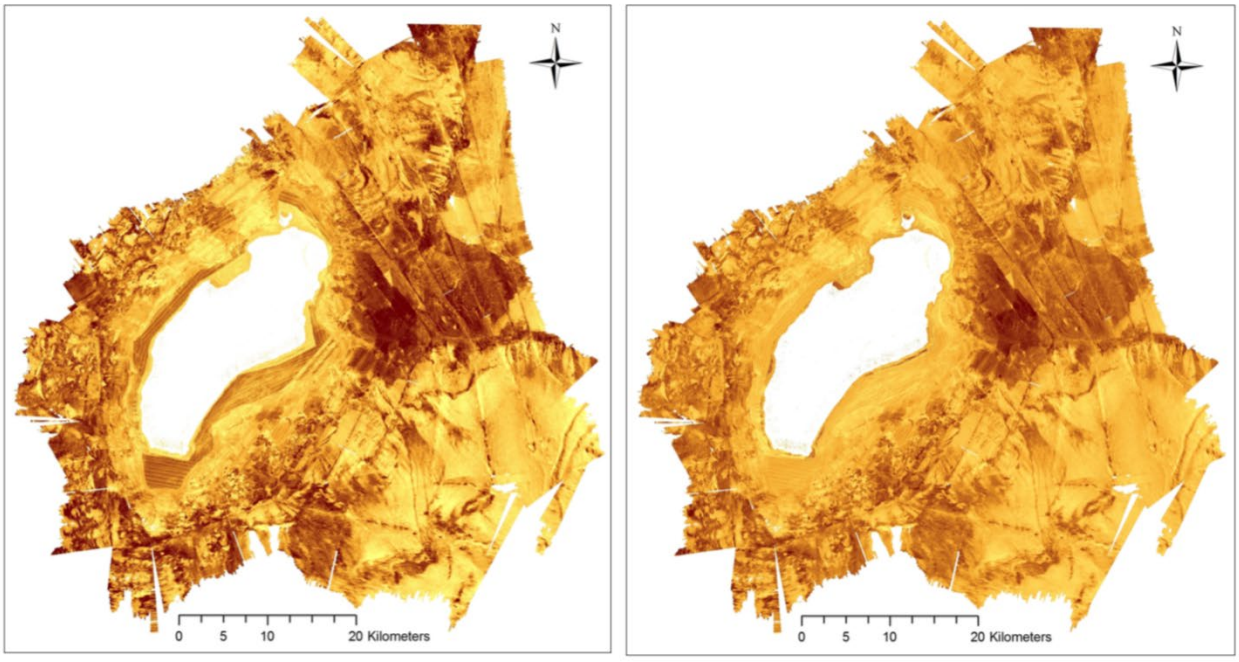


Figure 4. Backscatter data around Ni'ihau. Brown color indicates harder seafloor, while yellow/gold denotes softer seafloor. Left side image is a stack of 5 'raw' data sets. Right side are same data sets after going through reclassification and synthesizing process described in text and depicted in following figures.

Data files that generated unresolvable errors in FMGT were processed using the mbgrid module of MB- System (Caress & Chayes, 1996), which is extremely robust and able to read, process, and grid nearly every available sonar system format. Mbgrid allows manual setting of various options and does not rely on file metadata to automatically optimize processing options and flow. However, the result is a "rougher" product that includes more noticeable nadir artifacts and false "hard bottom" classifications. Jobs can be batched and run autonomously once the options are chosen, leading to a high degree of computational efficiency. Example command lines using MB-System for backscatter gridding and conversion to ArcGIS format are given in Appendix A.

Within ArcGIS, we converted the FMGT and mbgrid output ASCII grids to projected (WGS 1984 UTM 4N) floating point raster layers and generated a histogram of values (Figure 5). We

then visually inspected the data, visually compared results among weighted overlapping grid layers, and manually trimmed the upper and lower tails of the associated histograms, to achieve reasonable visual consistency across data sets.

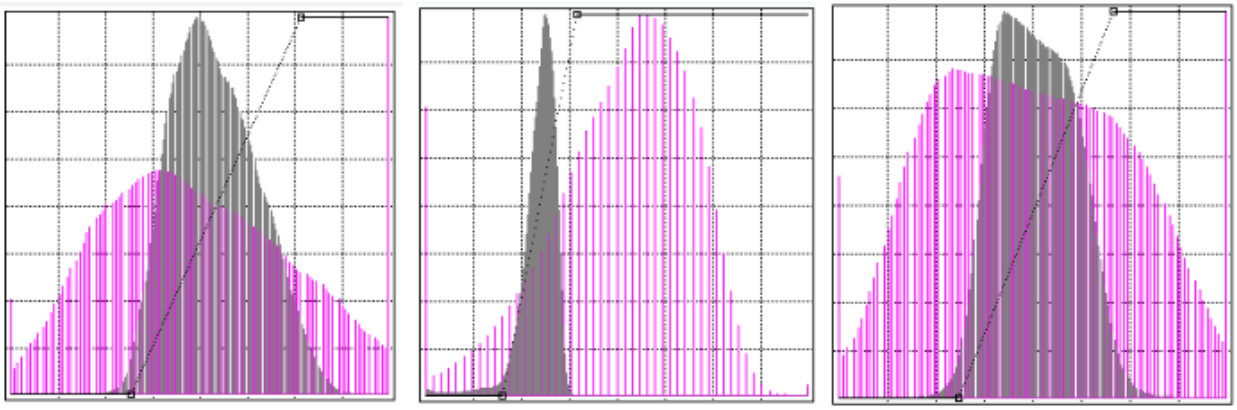


Figure 5. Three examples from the Ni'ihau data sets showing how the histograms from the raw grids (gray) are adjusted (magenta) in ArcGIS by using the sliders on top and bottom of graph. *x*-axis is backscatter value, *y*-axis is pixel count. See Figure 6 for more examples and explanation.

When we were satisfied with the visual consistency of grid layers in the display, we applied the selected reclassification values to the raw data using the reclassification tool within the RasterTools toolbox in ArcGIS 10.0 (Figure 6). It should be noted that ESRI has removed the reclassification too in ArcGIS v10.5 and newer. Users of ArcGIS v10.5 and newer should use the Image Renderer tool to create a Format TIFF file (32 to 8 bit), and then generate a grid from the TIFF (8 to 32 bit).

The reclassification step, also referred to as "outlier adjustment," is critical but arguably the most subjective stage in this process. Reclassification allows users to visually inspect and adjust overlapping backscatter data from different multibeam systems so that they have a similar, if not identical, appearance. This is regardless of the actual numerical range of the original grid values, which are then rescaled from 1 to 256.

After properly ordering the stack of grid layers based on weighting factors and criteria described, we used the Mosaic to New Raster tool within the RasterTools toolbox in ArcGIS 10.0 (Figure 7) to generate a combined mosaic for each scale from sub-domain to the full MHI survey domain (Figure 8).

The final MHI multibeam bathymetric and backscatter synthesis layers are available for download from the University of Hawaii, School of Ocean and Earth Science and Technology, Hawaii Mapping Research Group website at http://www.soest.hawaii.edu/HMRG/multibeam/index.php

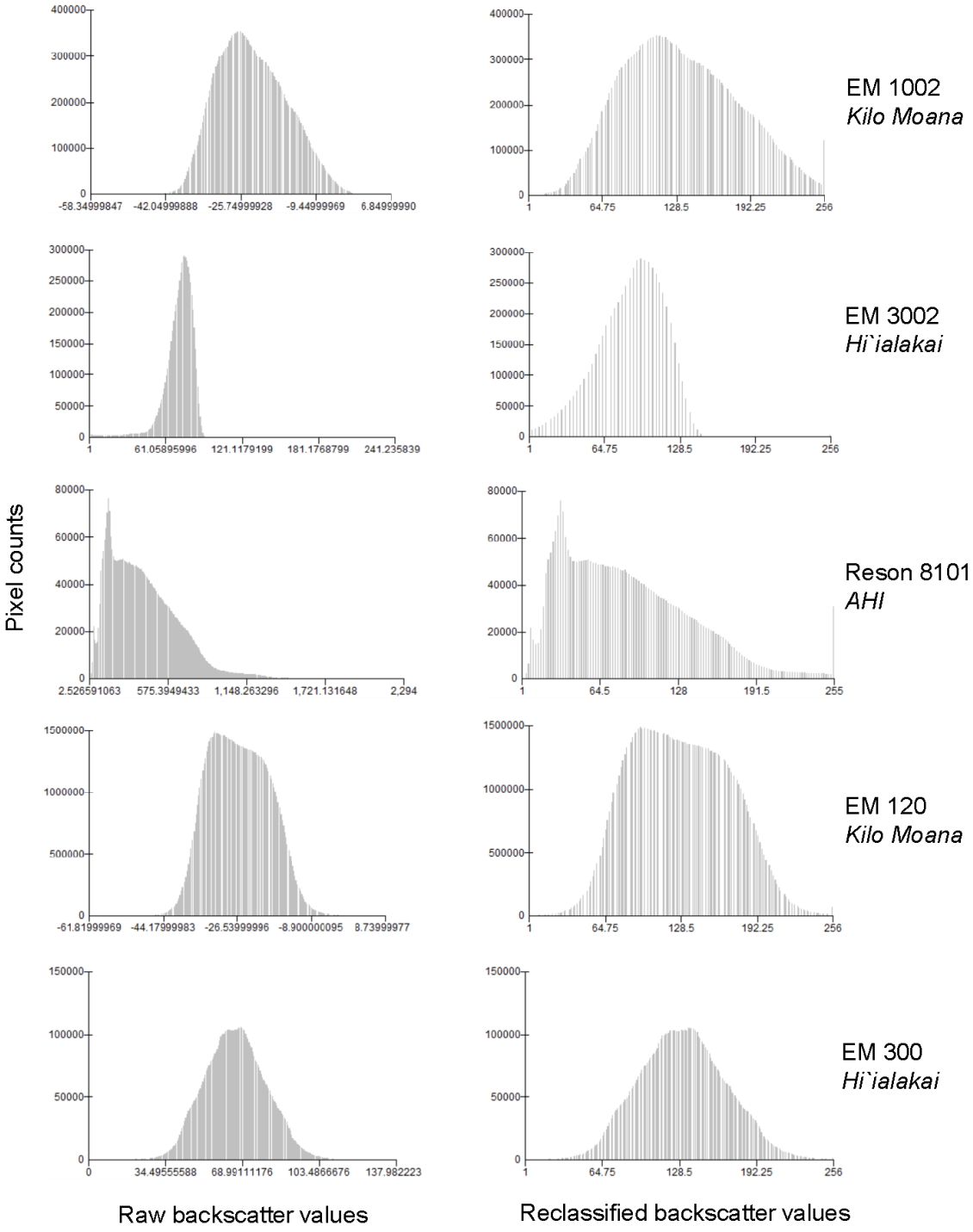


Figure 6. Histograms extracted from all 5 data sets making up the Ni‘ihau mosaic, as shown in Fig. 4. Left column shows raw grids prior to reclassification, and afterwards (right column). The multibeam systems and vessels are indicated on right side, and were mosaicked in this order, from top to bottom. Note the wide range of backscatter values on the *x*-axis within the left column and how they become standardized in right column (*y*-axis shows number of pixels at each backscatter value).

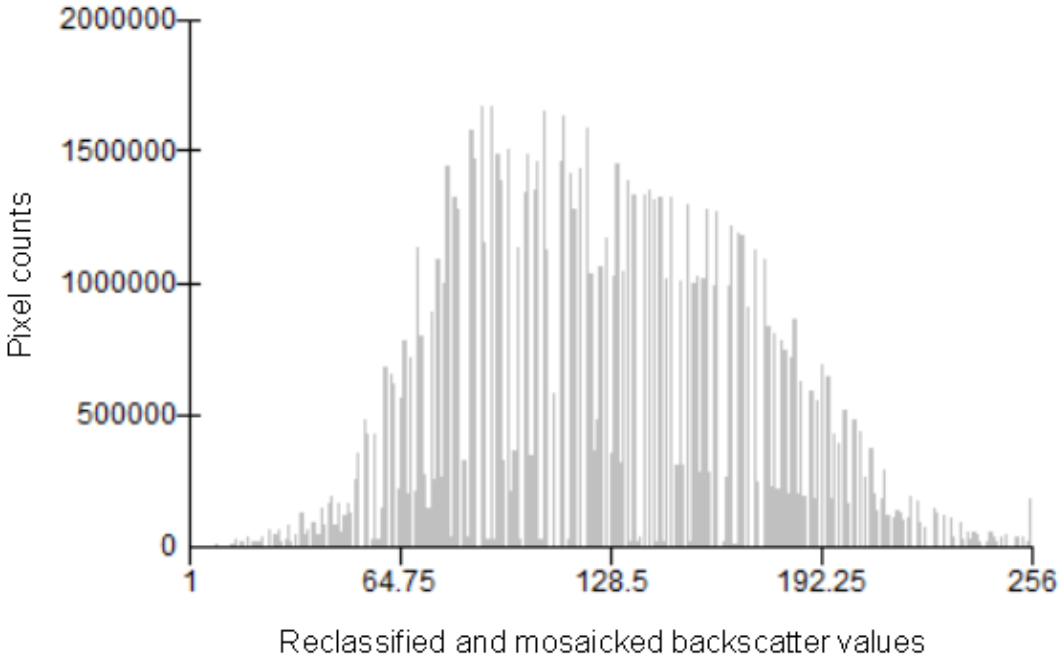


Figure 7. Final histogram of the 5 reclassified and synthesized Ni'ihau data sets. Note the appearance of overlapping histograms within the same range of values and a more Gaussian shape overall.

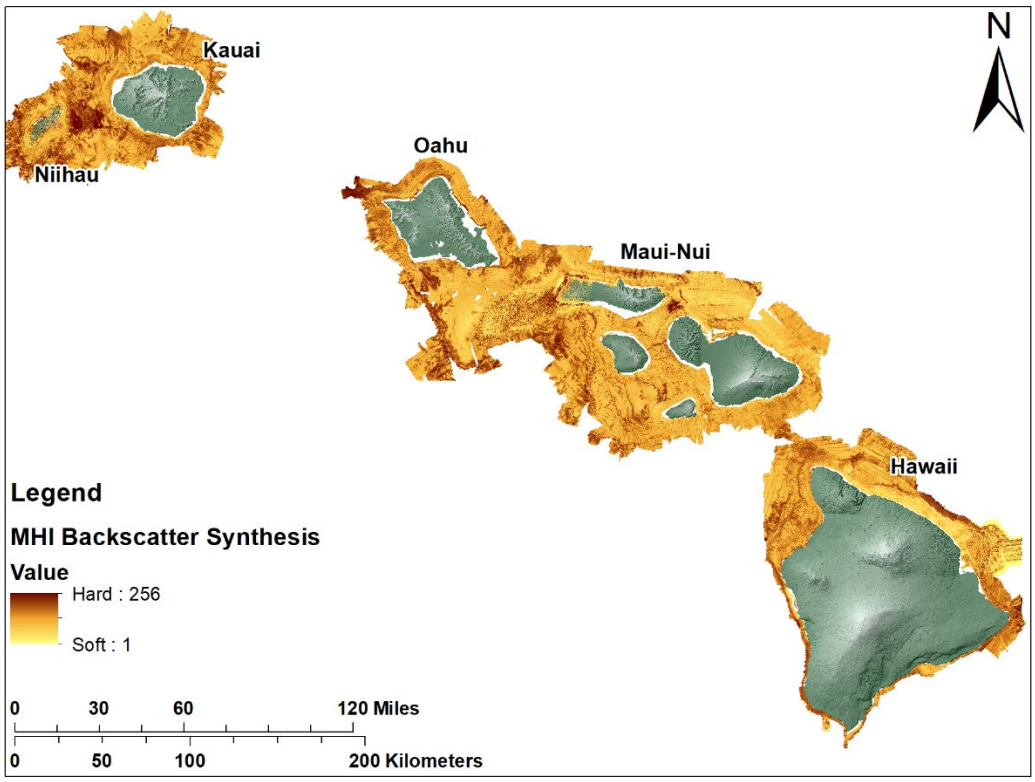


Figure 8. Resulting 5-m resolution multibeam backscatter synthesis for the MHI. Brown color indicates harder seafloor, while yellow/gold denotes softer seafloor.

While it currently represents the only high-resolution bathymetric and backscatter synthesis for the MHI, this product is not without error. Backscatter data collection relies on slant range sonar returns. Nadir lines (within the black boxes in Figure 9 and Figure 10) represent the ship where no useable backscatter are acquired. Nadir artifacts appear darker (harder) than the swath to either side of the ship as, at nadir, only a very small portion of the signal is lost due to angled reflectance. When possible, FMGT/Geocoder was used to process backscatter data. FMGT/Geocoder did a good job of blending the nadirs, but failed on several of the data types, including data from the NOAA Ship *Hi'ialakai* (EM 300, EM 3002), USGS 1998 data from the M/V *Ocean Alert* (EM 300), and data from the R/V*AHI* (Reson 8101). For these data types, MB-System (mbgrid) was utilized. MB-System (mbgrid) is more robust, but has less complexity in its processing and options.

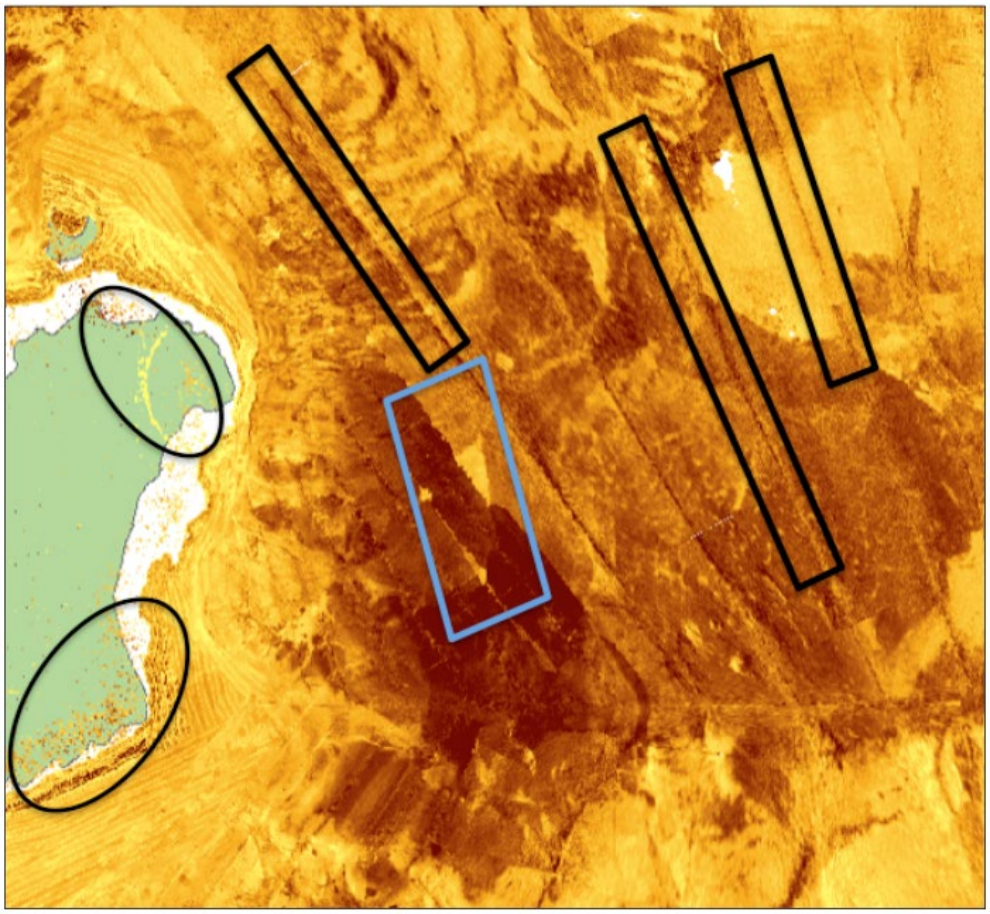


Figure 9 . Close-up view of backscatter data off the northeastern end of Ni'ihau showing three different artifacts: nadir lines (indicated by black rectangles), variable power/gain (blue polygon), on-land chatter (ellipses).

The blue box in Figure 9 shows a probable automatic power and/or gain increase in one of the shallow/intermediate depth multibeam systems operating near its depth limit (e.g., R/V *Kilo Moana's* EM 1002). Some of these variable power/gain artifacts show up in waters deeper than 500 m, well below the bottomfish range of interest and thus beyond the boundaries of the synoptic survey stratification grid cells. In order to reduce or eliminate the presence of such artifacts, the survey layer would have to be broken up into separate pieces and stacked in a

different order with another survey overlapping the offending piece(s). This would add additional complexity and variability into the consistency of the final mosaic and was beyond the scope of this project.

The circled artifacts in Figure 9 that overlap the terrestrial domain resulted from processing the *AHI* backscatter data in MB-System's mbgrid. In some cases with R/V *AHI* Reson 8101 data, Geocoder could be used and such artifacts did not appear. The depth range in which these artifacts occur is generally shallower than the depth range of interest for which this project was developed. We have elected to retain these data, allowing users to create custom masks using the bathymetry and/or topography to exclude and lessen the effects of these artifacts.

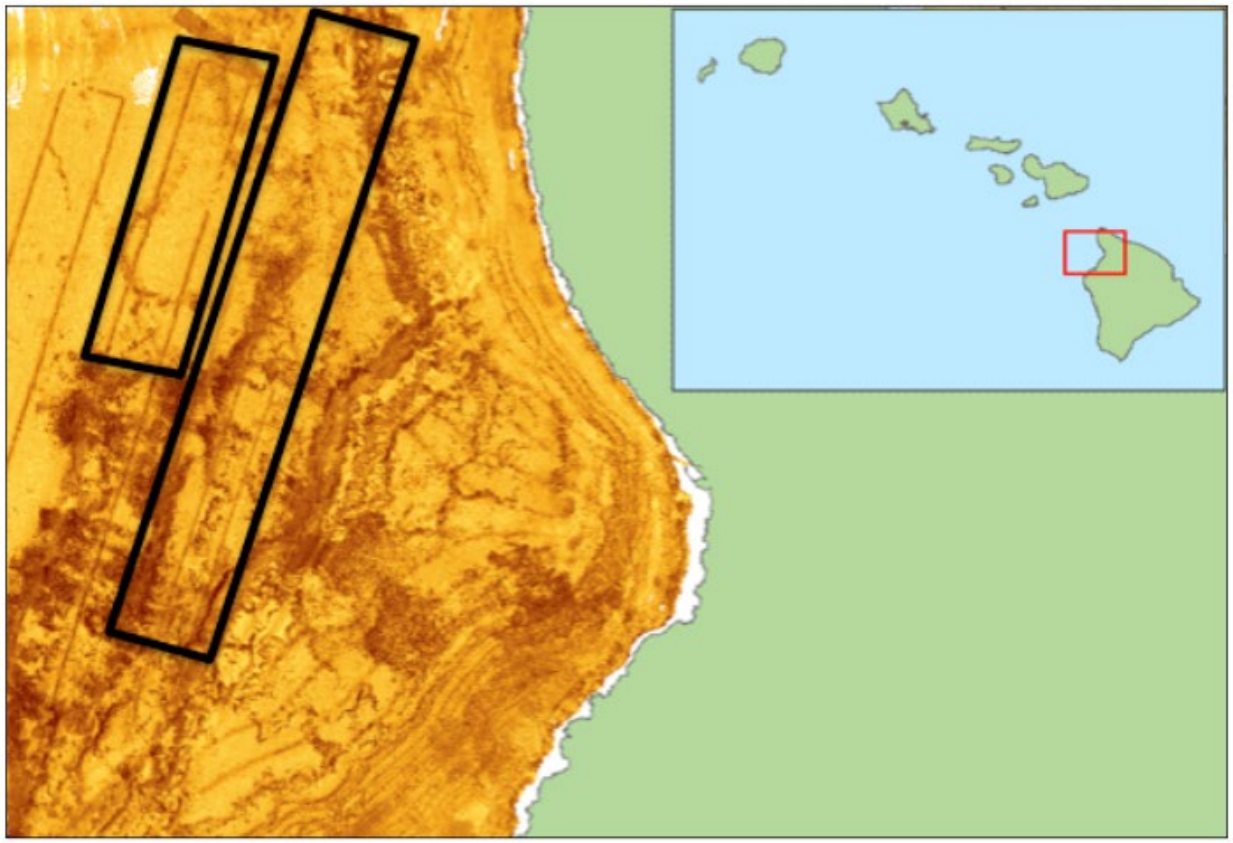


Figure 10. Close-up view of backscatter data off the northwestern side of the Big Island of Hawai'i showing nadir artifacts (indicated by black rectangles).

Finally, bathymetric slope can affect the strength of backscatter signal return and resulting hardness estimates (Lurton and Lamarche 2015). In some cases, the ship must cover the same steep area from numerous angles to get sufficient coverage of the high slope. If the slope is facing away from the swath, little to no acoustic return is received, erroneously suggesting softer substrates. Likewise, even if the ship is navigating along an ideal heading to map the morphology, extremely steep slopes can cause shadowing of the fan-shaped swath downslope, leaving gaps in both the bathymetric and backscatter data. Repeated passes from adjacent lines are the only way to achieve more complete coverage, at the expense of more ship time.

Part II: Using Synthesis Data to Define Habitat Strata for the Bottomfish Fishery-independent Survey in Hawaii

Methods

We used generalized linear regression analysis to guide development of stratification variables for substrate hardness and slope for the bottomfish survey. Combined CPUE data for all Deep7 species and both BFISH survey gears (research fishing and camera gears) was used as the response variable the regression analyses. We used CPUE data from pilot sampling conducted from 2011 to 2015 in Maui-Nui compiled and calibrated for gear type (Richards et al. 2016). This data set is made up of 1610 observations from 715 unique PSUs as well as the aforementioned synthesized bathymetric and backscatter data matched with each individual PSU.

We evaluated several aspects of design performance (e.g. CPUE variance partitioning, survey precision and cost (i.e., sample size), sampling efficiency) for principal Deep7 species (i.e. opakapaka, ehu, onaga) as compared to a simple random design. Hapu'upu'u, gindai, and lehi were excluded due to low data density. Kalekale was excluded, as it is not a primary fishery target.

For the regression analysis, two components of survey CPUE, presence-absence (i.e., occurrence) and catch when present were evaluated with respect to relationships with benthic habitat explanatory variables. In this approach, separate regression models were developed for occurrence (p) and catch when present (u) as functions of habitat variable X, and then multiplied together to yield CPUE as a function of X. This approach alleviated the problem of zero-inflated CPUE values often encountered with fishery-independent survey data, but also provided ecological insight on the nature of influence of a given habitat variable on relative abundance (e.g., does the habitat variable affect the probability of occurrence (presence-absence), the magnitude of abundance when present, or both?)

Regression models were developed in two steps. In the first step we developed exploratory models for p = f(X) and u = f(X) to provide insight on: (i) the model form of the relationship between a given response and explanatory variable (e.g. linear, quadratic, asymptotic); and (ii) an appropriate probability density function (pdf) for describing model error. We subsequently used the model form and error pdf identified in step 1 to fit final models for p = f(X) and u = f(X).

Results

Regression Analysis

We initially developed exploratory models for occurrence p and catch when present u using only depth as the explanatory variable. This allowed us to identify the principal depth range for subsequent analyses of substrate hardness and slope. A key aspect of exploratory model-building was converting continuous explanatory variables (e.g. depth) into categorical interval variables (e.g. 200–210 m, 210–220 m) along the range of X. We used logistic regression to evaluate relationships between Deep7 occurrence p and depth intervals I = 1,2,...,k using the model

$$\text{logit}(p) = \alpha + b_1 X_1 + \cdots + b_{k-1} X_{k-1} + \varepsilon \qquad (1)$$

where $(p) = \left(\frac{p}{1-p}\right)$, X_i are discrete categorical variables for depth intervals, α is the intercept, b_i are the depth interval coefficients, and ε is the residual error. Estimates of p_i, the mean occurrence for depth interval i, are given by

$$p_i = \frac{e^{\alpha+b_i}}{1+e^{\alpha+b_i}} \quad (2)$$

The overall depth range of the observations was 75–400 m, however sampling intensity varied along this range. We specified discrete depth intervals to facilitate our understanding of the general relationship between mean occurrence and depth, while also taking into account observation density along the depth range. Initially, depth intervals of 10 m were specified. These were subsequently adjusted based on sample size beginning at the shallowest interval and grouping successive 10 m intervals until each interval contained 40 or more observations.

Point estimates of Deep7 logit(p) were generally highest and fairly uniform for depth range 100–340 m, equating to mean occurrence (p) levels above 20% (Figure 11). Estimates of logit(p) were markedly lower at depths shallower and deeper than this range, equating to occurrence levels below 10%. Subsequent analyses focused on the 100–340 m depth range to minimize the effect of depth on the evaluation of substrate and slope variables.

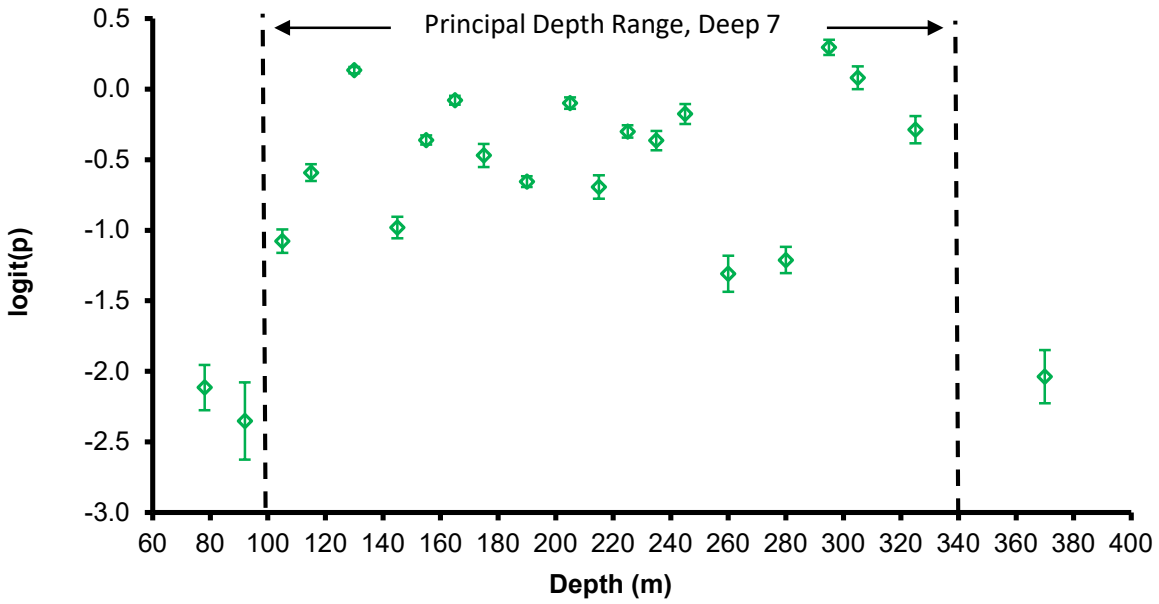


Figure 11. Logistic regression point estimates of logit(p) (±SE) by depth intervals for combined Deep7 species. The minimum depth interval for a given point estimate was 10 m, increased as necessary to achieve a sample size of 40 or more observations.

We also used logistic regression (Equation 1, Equation 2) to explore relationships between Deep7 occurrence and percent hardbottom, based on various backscatter threshold values (Figure 12) and prevalence of hardbottom habitat within a PSU (% of PSU area). Interval specification for percent hardbottom followed the aforementioned procedure for depth. A minimum interval was specified (usually 1%, but 0.5% in some cases), which was then adjusted by grouping intervals to obtain at least 40 observations per interval.

We initially used a backscatter threshold value of 140 to delineate softbottom (< 140) from hardbottom (≥ 140). Visual comparison of the reclassified 5-m backscatter synthesis with the original 20-m Maui-Nui backscatter products suggested a backscatter value of 140 in the 5-m synthesis would be analogous to the threshold value of 187 given by Kelley et al. (2006) (Figure 6). However, Deep7 occurrence was 10% or higher at very low levels of hardbottom area based on this threshold (Figure 12A). This suggested that the initial 140 threshold was likely too high (i.e., pixels that are actually hardbottom may be miss-classified as softbottom), as previous studies (Richards et al., 2016) found very low occurrence in softbottom substrates. Backscatter threshold values of 120, 110, and 100 were subsequently evaluated (Figure 12). A threshold value of 110 showed an abrupt large change in occurrence below and above 1% hardbottom area, suggesting that this threshold may be appropriate for delineating substrate hardness preferred by Deep7 species. A similar distinct change in occurrence was observed for the 100 threshold; however, in contrast to the results for the 110 threshold, this change was detected above and below 2% hardbottom area, and the magnitude of change was lower. This suggested that the 100 threshold may be too low (i.e., pixels that are actually softbottom may be miss-classified as hardbottom). We therefore selected the threshold value of 110 as optimal for delineating hardbottom from softbottom.

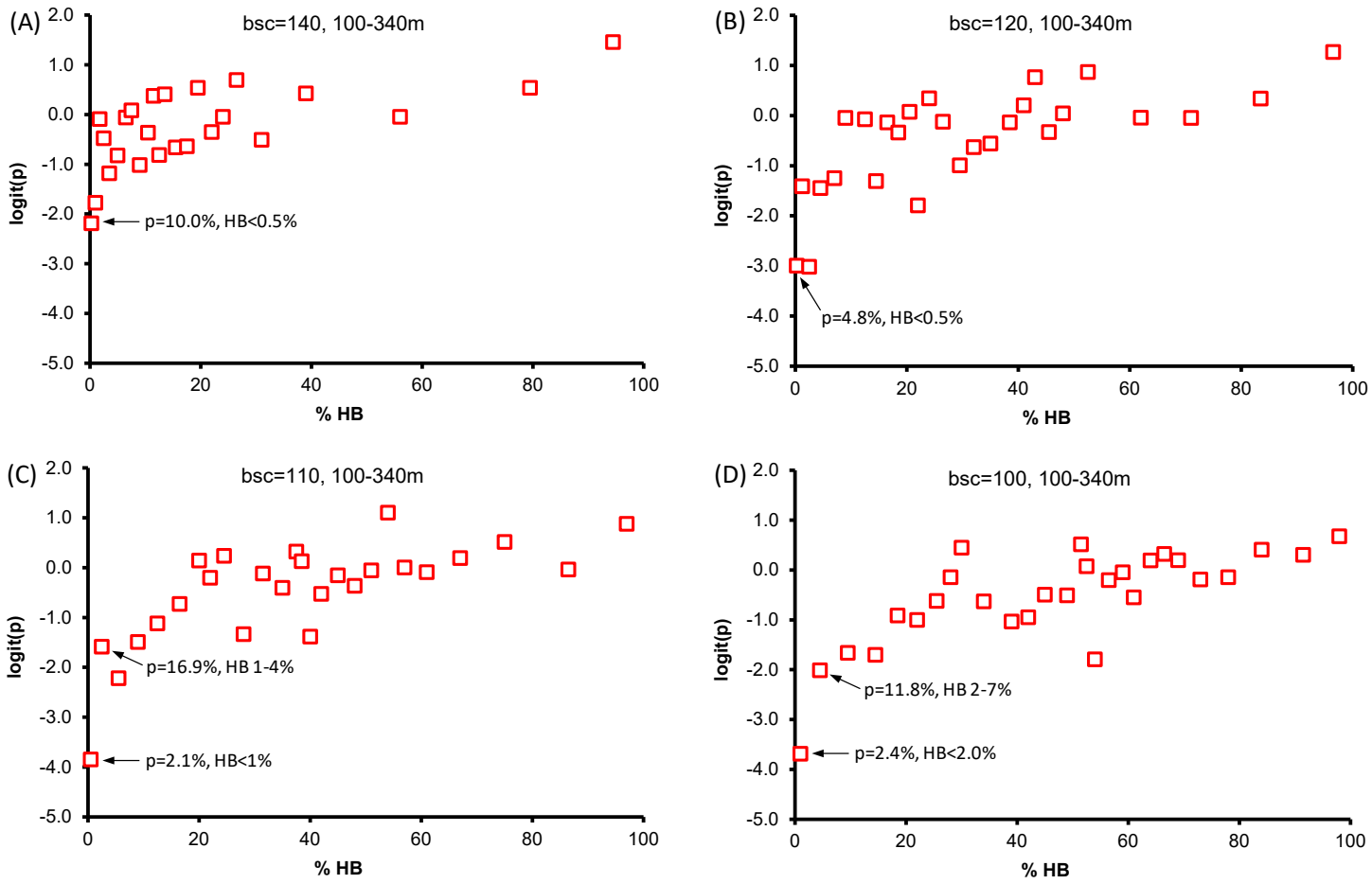


Figure 12. Logistic regression point estimates of Deep7 logit(*p*) by intervals of percent hardbottom (HB), the percentage of sample unit area exceeding a backscatter (bsc) threshold value. Four threshold values were evaluated: (A) 140, (B) 120, (C) 110, and (D) 100. Data were restricted to the principal Deep7 depth range (100–340 m), and each point estimate was based on 40 or more observations. Note that the spacing between point estimates, corresponding to percent hardbottom intervals, changes with changing threshold values, reflecting the richness (closely spaced, panel A) or sparseness (widely spaced, panel D) of observations along the *x*-axis.

Using the backscatter threshold value of 110, we used standard least-squares multiple linear regression to evaluate relationships between Deep7 catch when present (u) (i.e. positive survey CPUE) and depth intervals $i = 1,2,...,k$ using the model

$$\log(u) = \alpha + b_1 X_1 + \cdots + b_{k-1} X_{k-1} + \varepsilon \quad (3)$$

where $\log(u)$ is log-transformed positive survey CPUE, and the coefficients and variables are defined as in Equation 1. Log-transformation yielded approximately normally-distributed error residuals (ε). Intervals of percent hardbottom for $\log(u)$ were specified to comprise at least 20 observations. The full set of final regression models for occurrence, catch—when present, and survey CPUE as a function of percent hardbottom area are shown in Figure 13. Point estimates of $\log(u)$ varied somewhat across the gradient of percent hardbottom, but there appeared to be no distinct relationship (Figure 13B). The abrupt change in occurrence at 1% hardbottom area was modeled as a discontinuity in the regression function (Figure 13C). This carried over to the CPUE=f (percent hardbottom) regression function (Figure 13E), showing an abrupt increase in CPUE at 1% hardbottom area, and then a more gradual increase with increasing hardbottom area.

The same logistic regression procedures (Equation 1, Equation 2) were applied to evaluate relationships between Deep7 occurrence and slope, using various thresholds to delineate high vs. low slope (Figure 14) as well as prevalence of high slope habitat with in a PSU (% of PSU area). Intervals of percent high slope area for logit(p) were specified to comprise at least 40 observations. During pilot studies, a threshold slope value of 20 degrees was used to delineate high from low slope (Merritt et al., 2011). However, using the 5-m resolution synthesis data, applying a threshold slope value of 20 degrees resulted in point estimates of Deep7 occurrence as high as 45.9% below 1% high slope area (Figure 14A). This suggested that a threshold value of 20 degrees was set too high. A threshold of 10 degrees resulted in occurrence levels of about 10% below 1% high slope area, and an abrupt increase in occurrence above 1% high slope area (Figure 14B). For a slope threshold of 5 degrees, occurrence was as high as 13.7% and the change in occurrence was less distinct at low percent areas of high slope (Figure 14C). Hence, we chose a threshold value of 10 degrees to develop regression models for occurrence, catch when present, and CPUE as a function of percent high slope area (Figure 15).

As before, intervals of percent high slope area for $\log(u)$ were specified to comprise at least 20 observations. Point estimates of $\log(u)$ and the subsequent fitted continuous function showed no distinct relationship along the gradient of percent high slope area (Figure 15B). Similar to the relationships for hardbottom area, the functions for occurrence (Figure 15D) and CPUE (Figure 15E) exhibited an abrupt increase at 1% high slope area, and then a more gradual increase with increasing high slope area.

These regression analysis results, using the newly developed 5-m bathymetric and backscatter synthesis data, suggest the definitions provided in Table 4 are appropriate for delineating hardbottom/softbottom and high slope/low slope attributes for PSUs used in the BFISH survey.

Table 4. Definitions for hardbottom/softbottom and high slope/low slope based on 5-m resolution mapping data in the Maui-Nui region.

Variable Category	Threshold Value	Threshold PSU Area
Hardbottom	Backscatter ≥ 110	≥ 1% (100/10,000 pixels)
Softbottom	Backscatter < 110	➢ 99.0% (9,901/10,000)
High Slope	Slope ≥ 10 degrees	≥ 1% (100/10,000 pixels)
Low Slope	Slope < 20 degrees	➢ 99.0% (9,901/10,000)

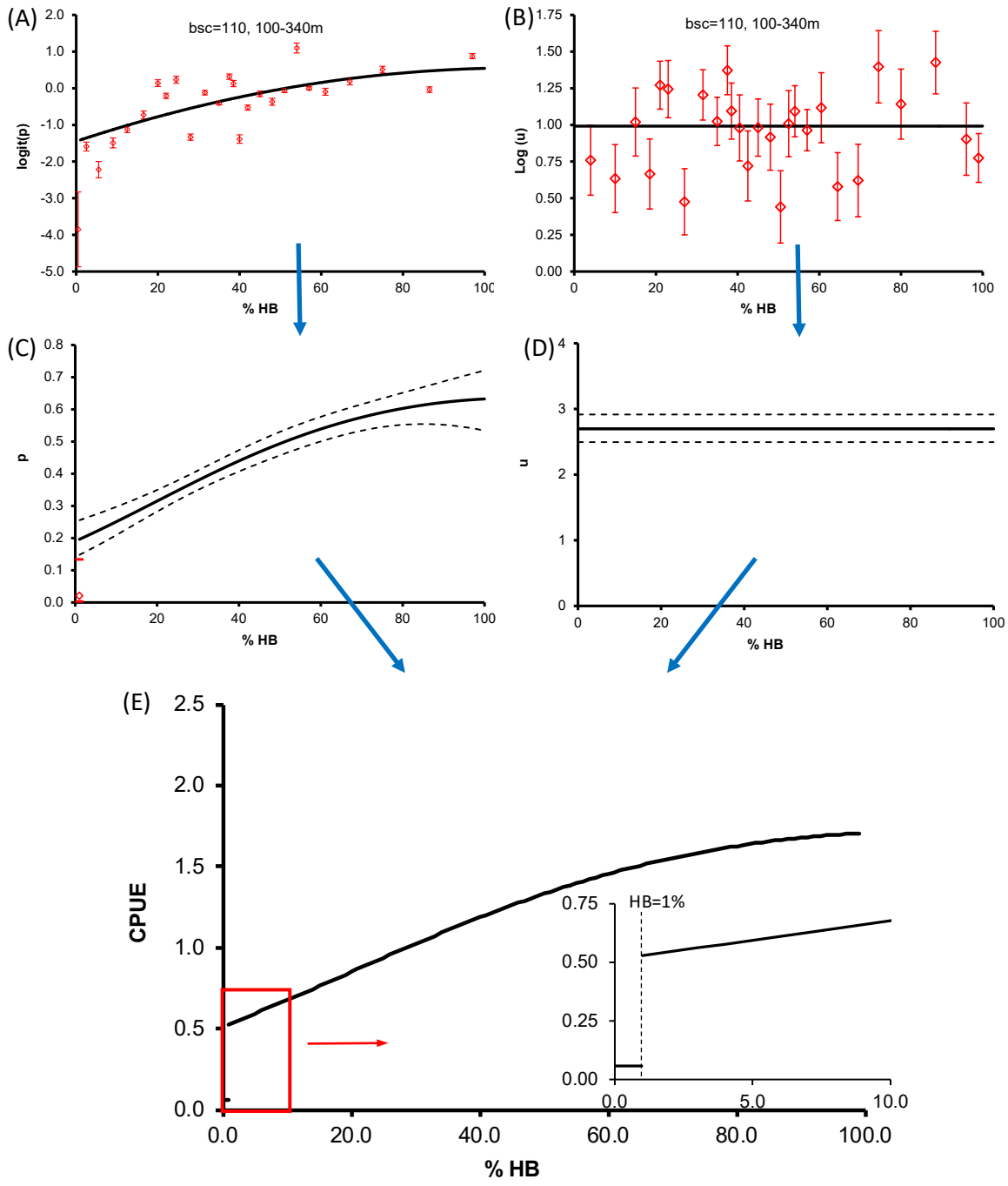


Figure 13. Regression modeling for Deep7 CPUE as a function of percent hardbottom (HB) (backscatter threshold value 110, depth range 100–340 m). (A–B) Exploratory model point estimates and final model regression functions for (A) logit(p) and (B) log(u) dependent upon percent hardbottom. (C–D) Back-transformed regression functions for (C) occurrence *p* and (D) catch when present *u* dependent on percent hardbottom. (E) Regression function for CPUE dependent upon percent HB, created by multiplying functions for *p* (panel C) and *u* (panel D); the detail for the red boxed area shows the discontinuity in the relationship below and above 1% HB.

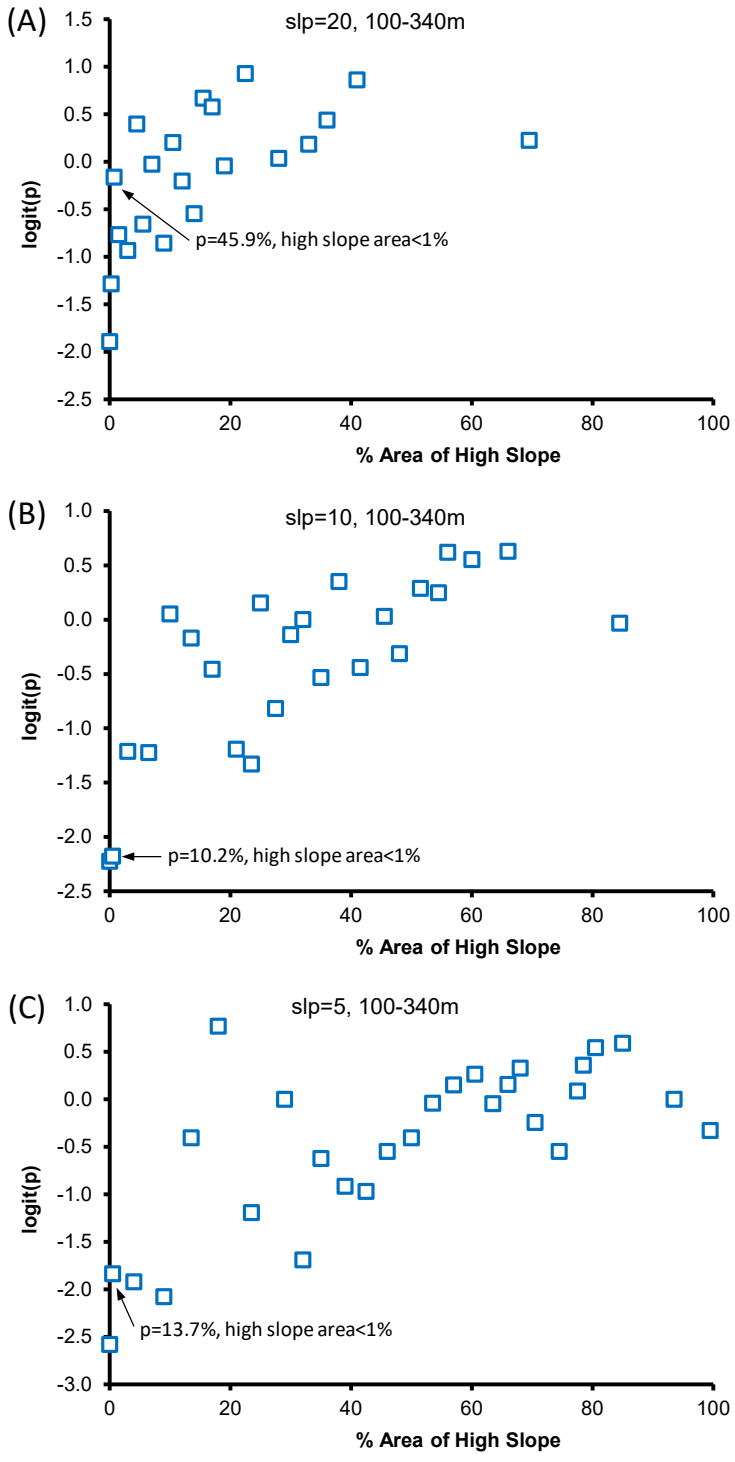


Figure 14. Logistic regression point estimates of Deep7 logit(*p*) by intervals of percent high slope area, i.e. the percentage of sample unit area exceeding a slope threshold value. Three slope (slp) threshold values were evaluated: (A) 20 degrees, (B) 10 degrees, and (C) 5 degrees. Data were restricted to the principal Deep7 depth range (10–-340 m), and each point estimate was based on 40 or more observations.

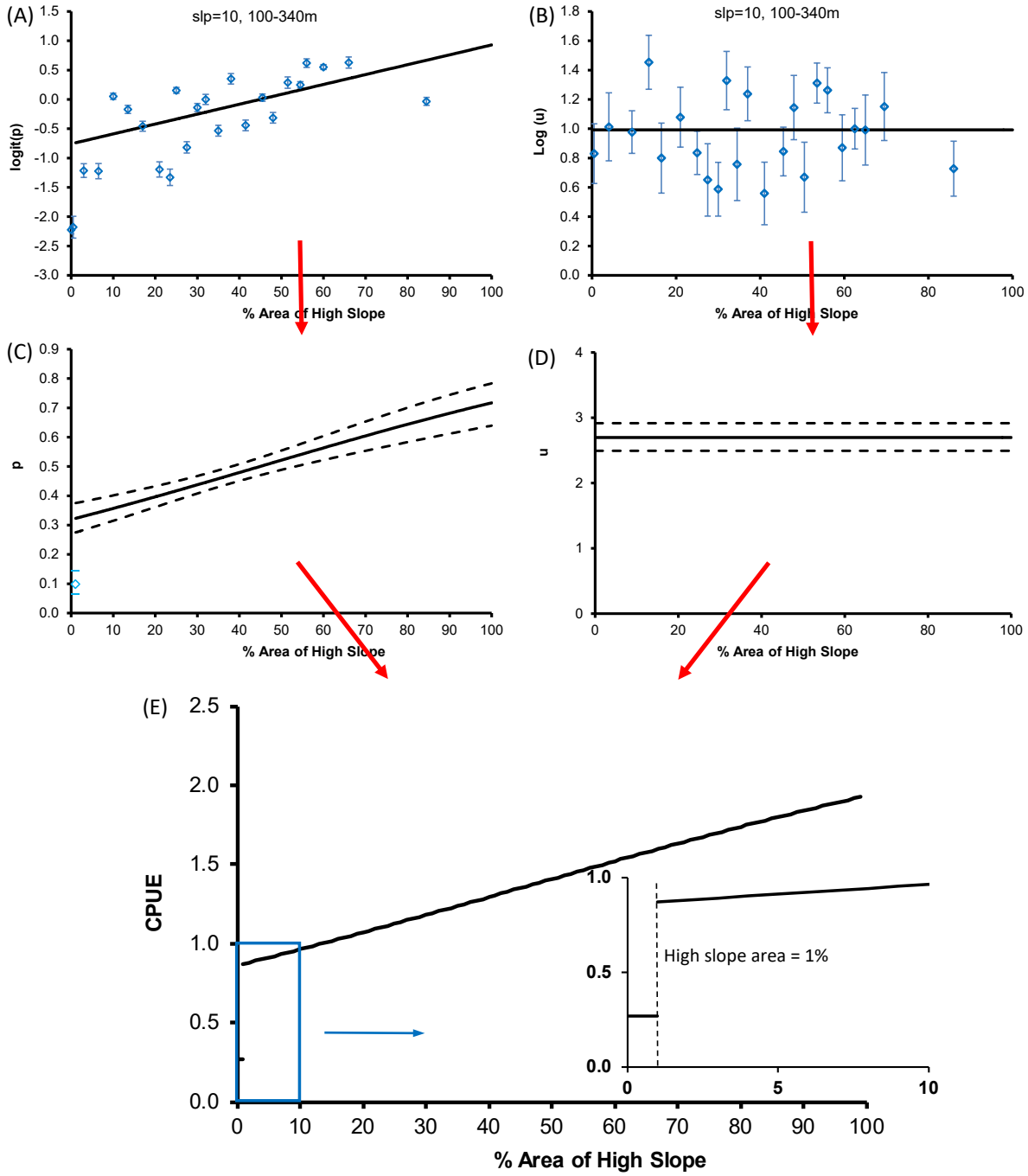


Figure 15. Regression modeling for Deep7 CPUE as a function of percent high slope area (slope threshold value 10 degrees, depth range 100–340 m). (A–B) Exploratory model point estimates and final model regression functions for (A) logit(*p*) and (B) log(*u*) dependent upon percent high slope area. (C–D) Back-transformed regression functions for (C) occurrence *p* and (D) catch when present *u* dependent upon percent high slope area. (E) Regression function for CPUE dependent upon percent high slope area, created by multiplying functions for *p* (panel C) and *u* (panel D); the detail for the blue boxed area shows the discontinuity in the relationship below and above 1% high slope area.

We conducted final exploratory regression analyses to evaluate dependencies among the potential stratification variables (depth, percent hardbottom area, percent high slope area). Substrate characteristics (hardness, complexity) appeared to be independent from depth. The relationships between percent hardbottom area and depth and between percent slope and depth were examined by estimating mean percent hardbottom area and mean percent high slope area for depth intervals across the range 75 to 400 m. Depth intervals were again specified to comprise at least 40 observations. There appeared to be no distinct relationship between percent hardbottom and depth or between percent high slope and depth (Figure 16A, B).

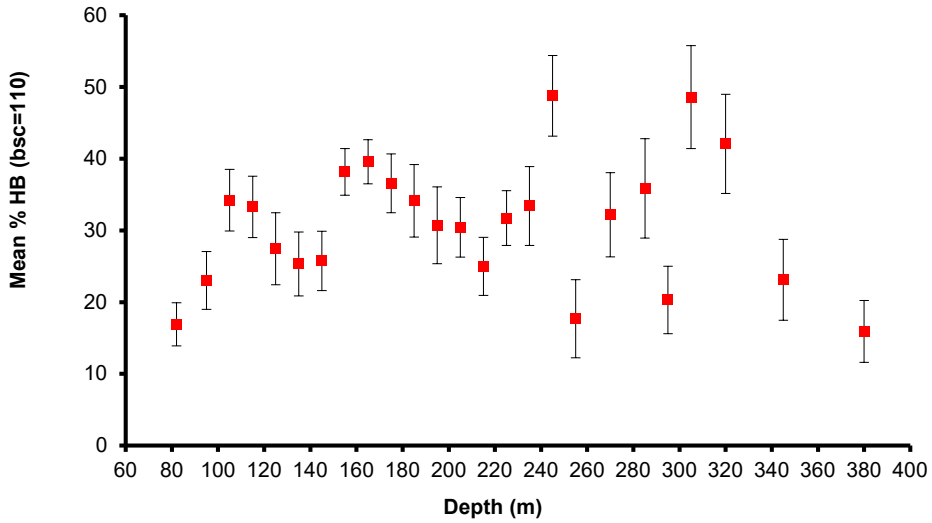
Relationships between percent hardbottom area and percent high slope area were examined from two perspectives: we first evaluated mean percent hardbottom area with respect to intervals of percent high slope area (Figure 17A), and then by evaluating mean percent high slope area with respect to intervals of percent hardbottom area (Figure 17B). The results suggest some correspondence between percent hardbottom area and percent high slope area, where low values of mean percent hardbottom occurred at low values of percent high slope area (Figure 17A), and low values of mean percent high slope area occurred at low values of percent hardbottom area (Figure 17B). This relationship was, perhaps, to be expected as areas of high slope appear to require a minimum amount of hardbottom area and softbottom habitats are not often able to support high slopes.

Stratification Variables for the Bottomfish Survey

We used the newly developed bathymetric and backscatter synthesis and above analyses to redefine stratification variables for hardness and slope for the BFISH survey (Table 5). Depth stratification remains defined according to the previous analysis of Richards et al. (2016). Our reclassification resulted in a substantial decrease in area classified as softbottom habitat as well as a substantial increase in area classified as hardbottom, high slope habitat (Table 6).

The new stratification resulted in effective spatial partitioning of mean survey CPUE and variance for ehu, onaga, and opakapaka, the principal survey design species (Figure 18). Mean relative abundance and variance were higher for both ehu (Figure 18A) and onaga (Figure 18B) in mid-depth and deep hardbottom, high slope strata compared to other habitats. Opakapaka mean relative abundance and variance were also higher in shallow hardbottom strata, and low or zero in other habitats (Figure 18C). Our results also indicate that softbottom-high slope habitat is relatively rare (Table 4) and seldom occupied by any of the three design species (Figure 18). Hence, we merged the softbottom-high and -low slope strata, resulting in a nine-class stratification scheme for substrate-slope-depth: 3 substrate classes (hardbottom high slope, hardbottom low slope, softbottom all slopes) crossed with 3 depth classes (Table 8).
We compared the expected design performance of the new nine-class stratification against that of a simple random design using the metric n*(15%), the projected sample size required to achieve a 15% CV of mean CPUE. For this analysis, we combined data for ehu and onaga as a proxy for mid-depth target species compared to opakapaka as the shallow target proxy. Our new stratification is likely considerably more effective, requiring sample sizes less than half that required under simple random sampling to achieve the same CV (Table 7). This new 9-class stratification (Table 8), based on the newly developed 5-m bathymetric and backscatter synthesis has now been applied to the full sampling domain for the MHI to serve as the basis for the Bottomfish Fishery-independent Survey in Hawaii (BFISH).

(A)



(B)

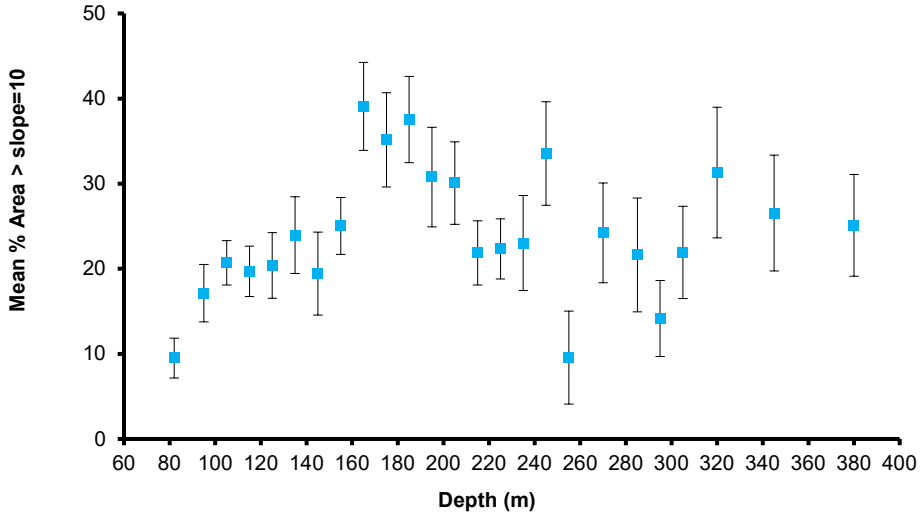
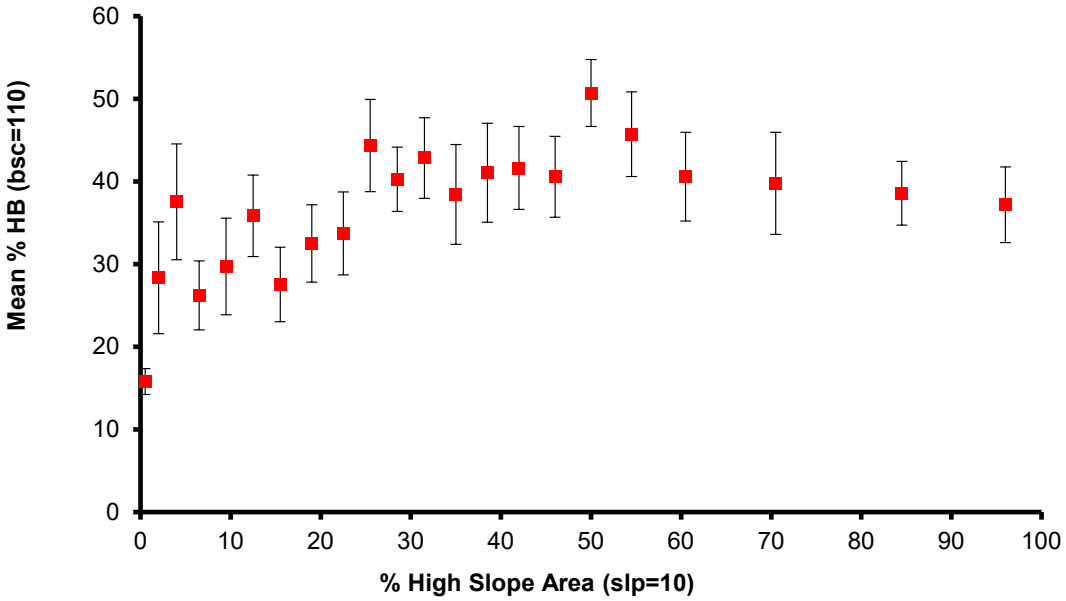


Figure 16. (A) **Mean percent hardbottom (HB, backscatter threshold =110) by depth intervals, and** (B) **mean percent high slope area (slope threshold = 10 degrees) by depth intervals, for full depth range 75–400 m. Each point estimate was based on 40 or more observations.**

(A)



(B)

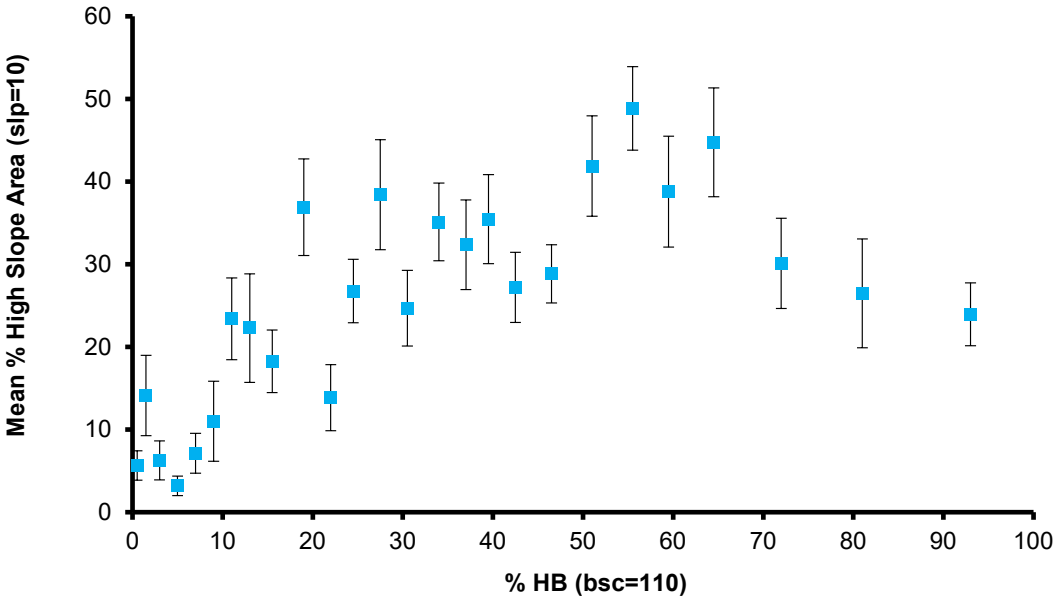


Figure 17. (A) **Mean percent hardbottom (HB, backscatter threshold = 110) by intervals of percent high slope area (slope threshold = 10 degrees).** (B) **Mean percent high slope area (slope threshold = 10 degrees) by intervals of percent hardbottom (backscatter threshold = 110). Each point estimate was based on 40 or more observations.**

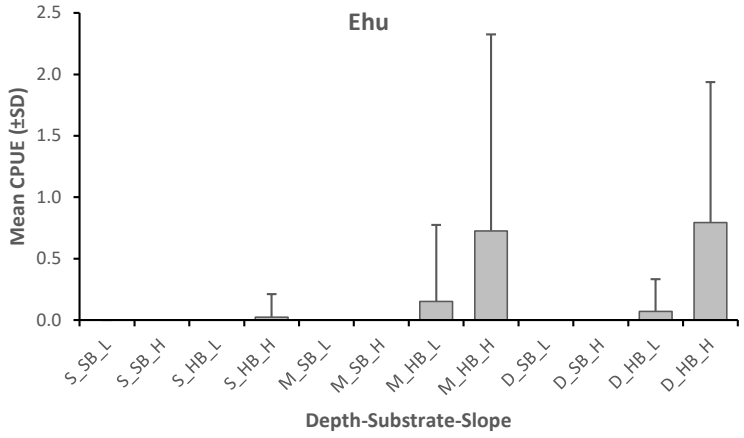
Table 5. Definitions of stratification variables for substrate hardness, slope, and depth for the bottomfish survey based on the 5-m resolution MHI bathymetric and backscatter synthesis.

Hardness (Backscatter)	Slope	Depth (m)
Hardbottom: ≥ 1 % >110	High: ≥ 1 % > 10 deg	Shallow: 75–200
Softbottom: < 1 % >110	Low: < 1 % > 10 deg	Medium: 200–300
		Deep: 300–400

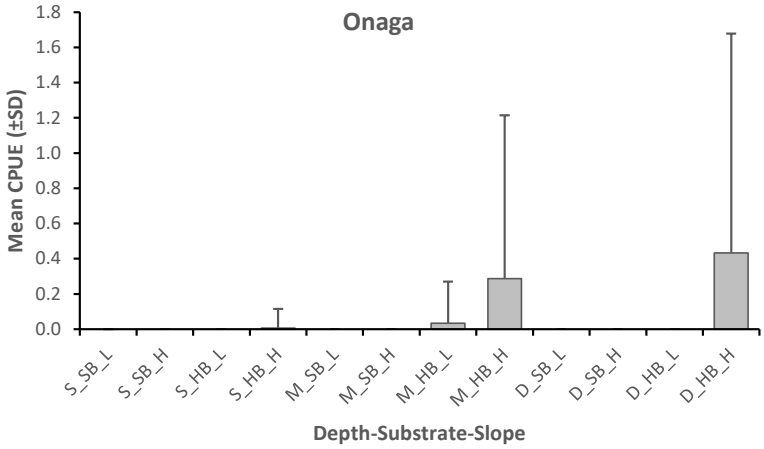
Table 6. Comparison of number of 500 x 500 m grid cells (PSUs) by depth-substrate-slope categories for the 20 m and 5-m pixel maps in the Maui-Nui region. Difference (Diff) is 5m PSUs minus 20 m PSUs. The reclassification with the 5-m mapping data resulted in a substantial decrease in softbottom habitat and a substantial increase in hardbottom, high slope habitat.

Depth	Substrate	Slope	Strata Code	PSUs, 20 m	PSUs, 5 m	Diff
Shallow (75–200 m)	SB (Softbottom)	L (Low)	SB_L_ S	1487	800	-687
Shallow	SB	H (High)	SB_H_ S	416	111	-305
Shallow	HB (Hardbottom)	L	HB_L_ S	2407	2118	-289
Shallow	HB	H	HB_H_ S	1105	2386	1281
Mid-depth (200–300 m)	SB	L	SB_L_ M	820	642	-178
Mid-depth	SB	H	SB_H_ M	228	93	-135
Mid-depth	HB	L	HB_L_ M	2076	1954	-122
Mid-depth	HB	H	HB_H_ M	690	1125	435
Deep (300–400 m)	SB	L	SB_L_ D	1077	766	-311
Deep	SB	H	SB_H_ D	136	131	-5
Deep	HB	L	HB_L_ D	1326	1339	13
Deep	HB	H	HB_H_ D	429	732	303

(A)



(B)



(C)

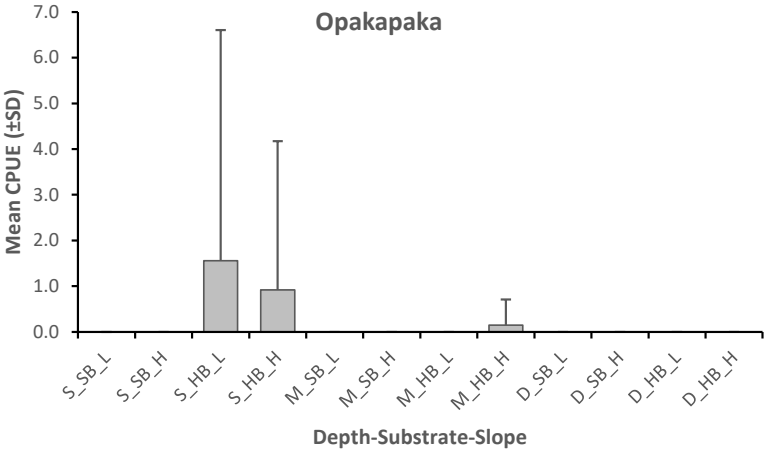


Figure 18. Mean survey CPUE and standard deviation (both gears combined) by depth-substrate-slope categories (see Table 6 for codes) for Ehu, Onaga, and Opakapaka.

Table 7. Estimates of n*(15%), the projected sample size to achieve a 15% CV for mean CPUE, for Ehu-Onaga and Opakapaka comparing simple random sampling and the 9-strata depth-hardness-slope design based on the new 5-m mapping data.

		n*(15%)	
Design	No. of Strata	Ehu-Onaga	Opakapaka
Simple random	1	477.7	1047.2
Depth-Substrate-Slope	9	216.0	457.0

Table 8. Final stratification of substrate (hard-soft), slope (high-low), and depth (shallow-medium-deep) for the bottomfish survey in the MHI.

Substrate	Slope	Depth	Strata Code	PSUs	Area (km^2)
SB (softbottom)	A (all)	Shallow (75-200 m)	SB_A_S	1863	465.75
HB (hardbottom)	L (low)	Shallow	HB_L_S	4562	1140.5
HB	H (high)	Shallow	HB_H_S	4777	1194.25
SB	A	Mid-depth (200-300 m)	SB_A_M	1449	362.25
HB	L	Mid-depth	HB_L_M	2688	672
HB	H	Mid-depth	HB_H_M	2412	603
SB	A	Deep (300-400 m)	SB_A_D	1591	397.75
HB	L	Deep	HB_L_D	3801	950.25
HB	H	Deep	HB_H_D	2749	687.25
		Total		25892	6473

Discussion

Stock assessments, including that for the MHI Deep7 bottomfish complex (Langseth et al. 2018), have largely relied on long-term fisheries-dependent time series to create estimates of relative abundance and trends through time. In many cases, development of robust fishery-independent surveys (Rotherham et al. 2007; Yoklavich et al. 2007; Clarke et al. 2009; Stallings 2009; Smith et al. 2011; Ault, Smith, Richards, Yau, B.J. Langseth, et al. 2018) can help to improve data used in stock assessments. Such surveys can help mitigate some of the biases (e.g. imposed length and catch limits, variable gear types, market forces, and fisher behavior) associated with fishery-dependent data (Hilborn and Walters 1992; Maunder and Punt 2004; Ault et al. 2014; Ault, Smith, Richards, Yau, B. Langseth, et al. 2018). Establishing robust fishery-independent methods to both understand and accurately predict fine-scale stock distributions is critical to such survey design (Moore, Harvey, & Van Niel, 2009). As the field of species distribution modeling continues to grow, incorporation of detailed habitat information, such as that provided here, can contribute to the continual improvement of stock assessments. A clear understanding of species-habitat relationships allows for more accurate and precise prediction of the spatial distribution of a species across its geographic domain, which in turn allows for estimation of absolute abundance that includes unsampled locations.

With the comprehensive bathymetric and backscatter synthesis for the MHI described herein, the Pacific Islands Fisheries Science Center has been able to improve the design of the BFISH survey. Of key importance is a change in the method for designating a primary sampling unit (PSU) as “hardbottom” or “softbottom”. Prior to this study, the majority of available data was gridded at 20-m resolution and was available only for the Maui-Nui island region. PSU were classified based on a hierarchical classification scheme where a PSU was classified as “hardbottom” or “high slope” if it contained even a single “hardbottom” or “high slope” pixel (Richards et al., 2016). This method over-classified these strata and was particularly sensitive to artifacts common to multibeam sonar data. We believe that the methods for PSU classification outlined in this study result in more accurate classification of PSU and improve efficiency of variance-based sampling of each strata. This improved stratification should also improve the precision of abundance estimates.

Using this newly developed bathymetric and backscatter synthesis, researchers have the ability to conduct more detailed studies of species-habitat relationships across the MHI. Such studies may be useful in continually improving the data used in stock assessments, definition of essential fish habitat, and in furthering ecosystem-based management approaches. Moore et al. (2009) have demonstrated the importance of detailed habitat maps in understanding the spatial ecology and predicting distribution of Australian demersal fish species. In their study, strong associations were detected between species distributions and habitat variables including depth, aspect, curvature, slope, and rugosity. For the majority of species, over half of the variance was explained by habitat variables alone. Among the most important were continuous measures of depth, substratum, and broad-scale measures of topographic complexity, the general importance of which have been well documented in the literature (Luckhurst and Luckhurst 1978; Choat and Ayling 1987; McCormick and Choat 1987; Friedlander and Parrish 1998; Priede and Merrett 1998; Rueda 2001; Harman et al. 2003; Jones et al. 2003; Anderson and Millar 2004)).

Kelley et al (2006) used multibeam sonar data complete an initial classification and description of essential fish habitat for the Deep7 bottomfish in the Northwestern Hawaiian Islands while Moore et al (2013) and Misa et al. (2013) demonstrated the importance of high-profile reef, low-profile reef, high-profile sediment, and low-profile sediment in the distribution of Deep7 bottomfish in the MHI. Moore et al. (2009; 2010) and Monk et al. (2010) have demonstrated the use of multibeam sonar derivatives including slope, backscatter, rugosity, curvature, and hypsometric or bathymetric position index. Lundblad et al. (2006) laid out methods for computing important derived variables including bathymetric position index and habitat zones such as crests, shelves, depressions, and slopes, many of which may be important in determining bottomfish distribution. Development of these types of derived features from the new MHI bathymetric and backscatter synthesis and evaluating their influence on the distribution of Deep7 bottomfish and other important fishery resources represent an important area of research.

Stratified sampling designs, such as that described here, are not novel (Smith et al. 2011). However, strata are typically developed based on human observation and intuition (Kelley et al. 2006). There are several issue with such methods. At the pixel level, the threshold between "hardbottom" and "softbottom" areas is typically a subjective and based on the observer (human or fish). Habitat strata are also rarely homogenous. When scaling mapping data up to the resolution of sampling strata, the threshold for prevalence is often as important as that delineating the categorical variable. This determination is also usually subjective.

The two-step regression analysis we describe offers a more objective method to define habitat strata from the perspective of the organism under study, which may be different from that of a human observer and even among species. These regression techniques are commonly used in survey data analysis, but we have not seen them used to objectively define survey stratification a priori. The data-derived threshold value of 110 to delineate "soft" from "hard" bottom was lower than the 187 threshold previously defined by human observers (Kelley et al., 2006). The data-derived threshold value of 10 degree to delineate "high" from "low" slope was also lower than the prior human-defined threshold of 20 degrees. The data-derived finding that a PSU should be classified as "hardbottom" or "high slope" if 1 percent of the enclosed pixels were defined as such was also surprising. These findings suggest that Deep7 bottomfish may be more sensitive to their environment, and to small-scale features within that environment, than previously thought. This conclusion supported recent work by (Oyafuso et al. 2017) and (Moore et al. 2016) and provides a methodology through which their findings may be used to quantitatively evaluate and define future BFISH strata.

While creating this bathymetric and backscatter synthesis, we attempted to identify and rectify noticeable artifacts. For the majority of the MHI, multiple overlapping multibeam and backscatter data sets allowed poor quality or inferior data files to be omitted. Ideally, we would want compare our newly reclassified synthesis products directly with ground-truth data (e.g. submersibles, ROVs, towed/drop cameras, substrate samplers) taken from a wide variety of locations, geomorphologies, and water depths throughout the MHI. However, such a task was beyond the scope of the project. In general, the hard (high) values correspond with expected geologic and morphologic features such as reef terraces, steep scarps, debris channels with coarse sediment, exposed lava flows and cones, current swept surfaces, headlands, submarine canyon walls, and carbonate platforms. Soft values (low) values corresponded with features

including sediment basins and other catchments, low-slope abyssal seafloor, and nearshore areas in close proximity to high terrestrial runoff.

Conclusion

This synthesis provides the first comprehensive high-resolution backscatter data with a consistent scale range across the entire MHI. This information is vital for generating properly stratified sampling designs based on habitat, investigating detailed species-habitat relationships, and as the basis for ecosystem-based fishery management. To our knowledge, the analyses we present represent the first use of data from a fishery-independent survey to delineate thresholds for hardbottom and high slope habitat preference from the perspective of a Deep7 bottomfish. As more data become available from future BFISH surveys and new methods are developed, these same methods can be used to refine the stratification to better define additional habitat variables affecting the distribution of Deep7 bottomfish in the MHI. Finally, the synthesis described herein is not without error and many significant issues have been described. At present, a comprehensive and focused multibeam backscatter survey of the MHI appears the best means to address these issues in the future.

Acknowledgements

This project was funded by NOAA's Pacific Islands Fisheries Science Center and was carried out in collaboration with the University of Hawaii, Mānoa, School of Ocean and Earth Science and Technology (SOEST) and the University of Miami Rosenstiel School of Marine and Atmospheric Science. The authors greatly benefited from the efforts of many known and unknown reviewers, researchers, technicians, officers, and crew of the institutions and research vessels used to collect the raw multibeam data.

The findings and conclusions in the paper are those of the authors and do not necessarily represent the views of the NOAA Fisheries, NOAA, or the Government of the United States. The use of trade, firm, or corporation names in this publication is for the convenience of the reader and does not constitute an official endorsement or approval of any product or service to the exclusion of others that may be suitable. The mapping data presented in this paper are for research purposes only and are not for use as navigational aids.

Literature Cited

Anderson MJ, Millar RB. 2004. Spatial variation and effects of habitat on temperate reef fish assemblages in northeastern New Zealand. Journal of Experimental Marine Biology and Ecology. 305:191–221.

Ault JS, Smith SG, Browder JA, Nuttle W, Franklin EC, Luo J, DiNardo GT, Bohnsack JA. 2014. Indicators for assessing the ecological dynamics and sustainability of southern Florida's coral reef and coastal fisheries. Ecological Indicators. 44:164–172.

Ault JS, Smith SG, Richards BL, Yau AJ, Langseth B, Humphreys R, Boggs CH, DiNardo GT. 2018. Towards Fishery-Independent Biomass Estimation for Hawaiian Deep7 Bottomfish. Honolulu, HI: NOAA Pacific Islands Fisheries Science Center Report No.: NMFS-PIFSC-67.

Ault JS, Smith SG, Richards BL, Yau AJ, Langseth BJ, O'Malley JM, Boggs CH, Seki MP, DiNardo GT. 2018. Towards fishery-independent biomass estimation for Hawaiian Islands deepwater snappers. Fisheries Research. 208:321–328. doi:10.1016/j.fishres.2018.08.012.

Choat JH, Ayling AM. 1987. The relationship between habitat structure and fish faunas on New Zealand reefs. Journal of Experimental Marine Biology and Ecology. 110:257–284. doi:DOI: 10.1016/0022-0981(87)90005-0.

Clarke ME, Tolimieri N, Singh H. 2009. Using the Seabed AUV to Assess Populations of Groundfish in Untrawlable Areas The Future of Fisheries Science in North America. In: Beamish RJ, Rothschild BJ, editors. 31. Northwest Fisheries Science Center NOAA Fisheries Service 2725 Montlake Blvd E. WA Seattle 98136 USA: Springer Netherlands. (Fish & Fisheries Series). p. 357–372.

ESRI Inc. 2017. ArcGIS Desktop. ESRI Inc. http://www.esri.com.

Fonseca LE, Calder BR. 2009. Geocoder: An Efficient Backscatter Map Constructor. In: Efficient Backscatter Map Constructor. In Proceedings of the U.S. Hydrographic Conference. San Diego, CA, USA. p. 11.

Friedlander AM, Parrish JD. 1998. Habitat characteristics affecting fish assemblages on a Hawaiian coral reef. Journal of Experimental Marine Biology and Ecology. 224:1–30.

Haight WR, Kobayashi DR, Kawamoto KE. 1993. Biology and management of deepwater snappers of the Hawaiian Archipelago. Marine Fisheries Review. 55:20–27.

Harman N, Harvey ES, Kendrick GA. 2003. Differences in fish assemblages from different reef habitats at Hamelin Bay, south-western Australia. Marine and Freshwater Research. 54:177–184.

Hilborn R, Walters CJ. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Boston/Dordrecht/London: Kluwer Academic Publishers (books.google.com).

Jones EG, Tselepides A, Bagley PM, Collins MA, Priede IG. 2003. Bathymetric distribution of some benthic and benthopelagic species attracted to baited cameras and traps in the deep eastern

Mediterranean. Marine Ecology Progress Series. 251:75–86.

Kelley C, Moffitt R, Smith JR. 2006. Mega- to micro-scale classification and description of bottomfish essential fish habitat on four banks in the Northwestern Hawaiian Islands. Atoll Research Bulletin. 543:319–332.

Langseth B, Syslo J, Yau A, Kapur M, Brodziak, Jon K. T. (Jon Kenton Tarsus). 2018. Stock assessment for the main Hawaiian Islands Deep 7 bottomfish complex in 2018, with catch projections through 2022. Honolulu, HI Report No.: NMFS-PIFSC-69. [accessed 2019 Mar 8]. ftp://ftp.library.noaa.gov/noaa_documents.lib/NMFS/PIFSC/TM_NMFS_PIFSC/NOAA_Tech_Memo_PIFSC_69.pdf.

Luckhurst BE, Luckhurst K. 1978. Analysis of the influence of substrate variables on coral reef fish communities. Marine Biology. 49:317–323.

Maunder MN, Punt AE. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research. 70:141–159.

McCormick MI, Choat JH. 1987. Estimating total abundance of a large temperate-reef fish using visual strip-transects. Marine Biology. 96:469–478.

Moore C, Drazen JC, Radford BT, Kelley C, Newman SJ. 2016. Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. Marine Policy. 69:32–41. doi:10.1016/j.marpol.2016.03.021.

Moore CH. 2008. Defining and predicting species-environment relationships: understanding the spatial ecology of demersal fish communities. The University of Western Australia.

Oyafuso ZS, Drazen JC, Moore CH, Franklin EC. 2017. Habitat-based species distribution modelling of the Hawaiian deepwater snapper-grouper complex. Fisheries Research. 195:19–27. doi:10.1016/j.fishres.2017.06.011.

Priede IG, Merrett NR. 1998. The relationship between numbers of fish attracted to baited cameras and population density: Studies on demersal grenadiers Coryphaenoides (Nematonurus) armatus in the abyssal NE Atlantic Ocean. Fisheries Research. 36:133–137.

Richards BL, Smith, Steven G., Ault JS, DiNardo GT, Kobayashi DR, Domokos R, Anderson J, Taylor JC, Misa W, Giuseffi L, et al. 2016. Design and implementation of a bottomfish fishery-independent survey in the main Hawaiian Islands. Honolulu, HI Report No.: NMFS-PIFSC-53. https://doi.org/10.7289/V5/TM-PIFSC-67.

Rotherham D, Underwood AJ, Chapman MG, Gray CA. 2007. A strategy for developing scientific sampling tools for fishery-independent surveys of estuarine fish in New South Wales, Australia. ICES Journal of Marine Science: Journal du Conseil ICES Journal of Marine Science: Journal du Conseil. 64:1512–1516.

Rueda M. 2001. Spatial distribution of fish species in a tropical estuarine lagoon: a geostatistical appraisal. Marine Ecology Progress Series. 222:217–226.

Smith SG, Ault JS, Bohnsack JA, Harper DE, Luo J, McClellan DB. 2011. Multispecies survey design for assessing reef-fish stocks, spatially explicit management performance, and ecosystem condition. Fisheries Research. 109:25–41. doi:10.1016/j.fishres.2011.01.012.

Stallings CD. 2009. Fishery-independent data reveal negative effect of human population density on Caribbean predatory fish communities. PLOS ONE. 4:e5333. doi:10.1371/journal.pone.0005333.

United States Department of Commerce. 2016. Pacific Island Fisheries; Hawaii Bottomfish and Seamount Groundfish; Revised Essential Fish Habitat and Habitat Areas of Particular Concern.

University of Hawaii. 2019a. Pacific Islands Benthic Habitat Mapping Center. [accessed 2019 May 28]. http://www.soest.hawaii.edu/pibhmc/cms/data-by-location/main-hawaiian-islands/.

University of Hawaii. 2019b. Hawaii Mapping Research Group. [accessed 2019 May 28]. http://www.soest.hawaii.edu/HMRG/cms/.

Western Pacific Regional Fishery Management Council. 2010. Bottomfish Fisheries in the Hawaii Archipelago.

Yoklavich MM, Love MS, Forney KA. 2007. A fishery-independent assessment of an overfished rockfish stock, cowcod (Sebastes levis), using direct observations from an occupied submersible. Canadian Journal Of Fisheries And Aquatic Sciences. 64:1795–1804. doi:10.1139/f07-145.

Appendix A: Step-by-step Guide to Creating a Multibeam Backscatter Synthesis Mosaic

1) Process and grid the multibeam backscatter data in one or more of several open source or commercial software packages including, but not limited to, MB-System (Caress and Chayes, 1996; Paduan, 2016) or Fledermaus Geocoder Toolbox (FMGT) (Fonseca and Calder, n.d.). Both of these packages were used in this study.

    a. Example gridding and filtering commands (mbgrid):

        i. mbgrid -E5.0/5.0m! -F1 -A2 -C10 -I$datalist1 -N -O$outgrid -R$range1 -V -JU -S7
            1. grdfilter $outgrid -G$filtgrid -D0 -Fm50
        ii. mbgrid -E5.0/5.0m! -F1 -A4 -JU -C5 -U5 -N –I$datalist2 –O$outgrid -V –R$range2
            1. mbm_grd2arc -V –I$outgrid –O$arcgrid

The following steps were carried out using ESRI's ArcGIS, specifically the ArcMap module with the Spatial Analyst extension. The original project used ArcGIS version 10.0 with later testing and modification to document the procedure using version 10.5.1.

2) Convert the ASCII grids to rasters using the ASCII to Raster tool,
3) Select Output data type > floating point and define the projection (UTM zone).
4) For each data layer added, select Properties > Symbology > Stretch Type > Min-Max.
5) Select Custom to force histogram creation, if necessary, and then click on the Histograms button to activate histograms to generate and display a histogram of values.
6) Edit the low and high values in the histogram box by manually trimming the upper and lower (left and right) tails by adjusting the sliders (Figure 5),
7) Apply, then compare the results if other grid layers have already been imported and reclassified.
    a) It is recommended to choose the highest quality data set, with broad coverage, as the reference layer from which adjustments to other layers will be made.
    b) Repeat this process on each successive layer.
        i) You may need to go back and adjust the histograms of previous layers.
        ii) Note that this step is only held in the display and will revert if you quit without reclassifying.
        iii) It is also recommended to keep a detailed spreadsheet log of the various steps, how the histograms were changed, and which layers were stacked, or mosaicked, and in what order.
8) When satisfied with the look of the backscatter match between grid layers in the display, the RasterTools > Reclassify tool was used in version 10.0 to apply the reclassification values to the grid itself, rather than just to the display (Figure 6).

    a) Browse, select, and add "r" to output filename. Be sure to load the input raster from the pull-down menu, not the folder icon.

    b) Under layer properties, switch from Classified to Stretched and change stretch type to Min-Max and Reclassify.

i) One caveat must be noted here. The reclassification back to grid feature was changed by ESRI after ArcGIS version 10.0. In newer versions, one must use the Image Renderer tool to create a Format TIFF file (32 to 8 bit), and then generate a grid from the TIFF (8 to 32 bit).

c) Right click on the data layer > Data > Export Data > Use Renderer, enter zero [0] for "NoData as," adjust location of output file and name, select TIFF, and Save.

d) Change the order of the layers displayed here, top to bottom, if necessary.

e) Now, one may proceed with the Mosaic tool as described below for either version 10.0 or 10.5.1.

9) After properly ordering the stack of grid layers (generally higher resolution data on top, lower resolution data on bottom), the final step in the process is to generate a combined mosaic by using the Mosaic to New Raster tool (Figure 7).

a) Go to RasterTools > Mosaic to New Raster.
b) From the drop down menu choose the first layer that will remain on top, then successively choose layers stepping down until the final one is on the bottom of the list.
c) Assign a name for the new, combined mosaic.
d) Add spatial reference information if different, pixel type (the default is 8 bit unsigned, which supports values of 0 to 255), leave cell size blank, number of bands = 1.
   i) The Mosaic Operator choice is not optional, nor is the choice for Mosaic Colormap Mode optional.
      (1) Best results have been achieved with choosing FIRST/FIRST and when the Colormap layer has the widest range of Z values in order to encompass all layers (e.g., 0 to 255 instead of 3 to 255) with no pixels "lost."
e) Now, carry out the Mosaic to New Raster operation by selecting OK.
   i) The default display may come up darker as Standard Deviation. Switch it back to Min-Max after checking the resultant histogram with the Custom setting to ensure the operation was carried out correctly and to completion.
   ii) You should see a composite of histograms as in Figure 7, representing a complete, synthesized mosaic.

As previously mentioned, a detailed log will be useful later if you wish to makes changes, add or subtract layers, and/or prepare metadata.