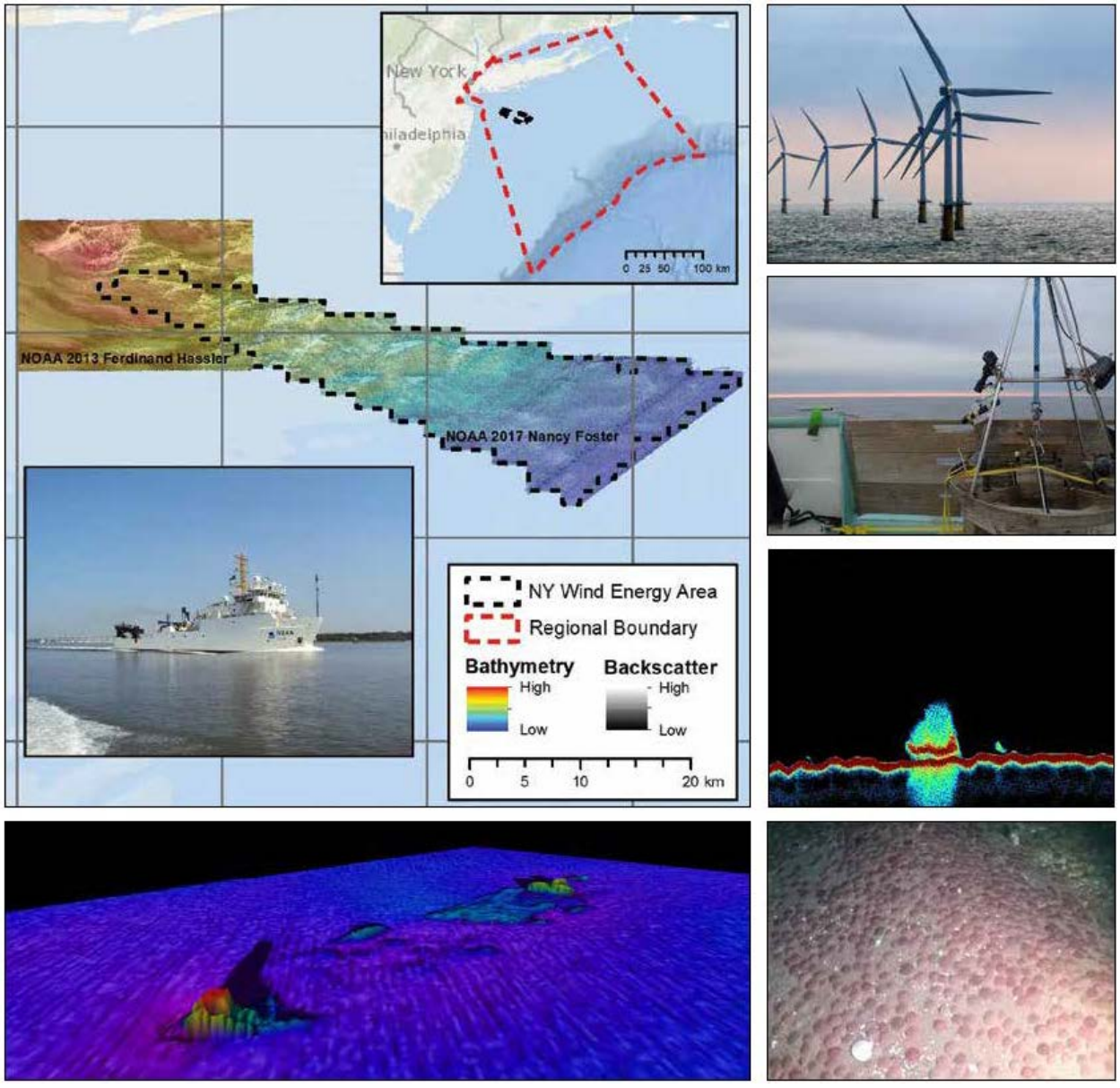
OCS Study
BOEM 2019-069

Comprehensive Seafloor Substrate Mapping and Model Validation in the New York Bight



US Department of the Interior
Bureau of Ocean Energy Management
Office of Renewable Energy Programs

BOEM
Bureau of Ocean Energy Management

OCS Study
BOEM 2019-069

Comprehensive Seafloor Substrate Mapping and Model Validation in the New York Bight

Final Report

Authors

Tim Battista
Will Sautter
Matthew Poti
Erik Ebert
Laura Kracker
Jennifer Kraus, LTjg
Ayman Mabrouk
Bethany Williams
Daniel S. Dorfman
Rachel Husted
Chris J. Jenkins

Prepared under NCCOS IAA MOA-2016-040-10277, BOEM OCS Study 2019-069, and NCCOS BOEM IAA M16PG00015

by

U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Ocean Service
National Centers for Coastal Ocean Science
Marine Spatial Ecology Division
1305 East-West Hwy, SSMC-4
Silver Spring, MD 20910

OCS Study BOEM 2019-069 and NOAA Technical Memorandum NOS NCCOS 255

US Department of the Interior
Bureau of Ocean Energy Management
Office of Renewable Energy Programs

This project was funded by the Bureau of Ocean Energy Management (BOEM) through Intra-agency Agreement M16PG0005 with the U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Centers for Coastal Ocean Science (NCCOS). Dan Dorfman, Rachel Husted, Ayman Mabrouk, Matt Poti and Will Sautter were supported by NOAA Contracts No. EA133C-17-BA-0062, and Erik Ebert under EA133C-17-BA-0049 with CSS, Inc. Chris Jenkins is with the Institute of Arctic and Alpine Research, University of Colorado, Boulder.

This work would not have been possible without the numerous contributors who shared their data, time, and expertise to make this project a success. We would like to especially thank Sarah Hile for organizing and formatting the content of this report; Arliss Winshop, Matt Kendall, Chris Jeffrey, Chris Taylor, and Charlie Menza for their reviews and suggestions; Ken Buja for cartographic support; and the crew and officers of the NOAA ship *Nancy Foster* and R/V *Tiki XIV*.

Citation

Battista, T. W. Sautter, M. Poti, E. Ebert, L. Kracker, J. Kraus, A. Mabrouk, B. Williams, D.S. Dorfman, R. Husted, and C.J. Jenkins. 2019. Comprehensive Seafloor Substrate Mapping and Model Validation in the New York Bight. OCS Study BOEM 2019-069 and NOAA Technical Memorandum NOS NCCOS 255. 187 pp. doi:10.25923/0hw8-gz28

Comprehensive Seafloor Substrate Mapping and Model Validation in the New York Bight

Contents

List of Figures

List of Tables

List of Frequently Used Acronyms

AA	accuracy assessment
AUC	area under the curve
BCT	boosted classification tree
BRESS	Bathymetry and Reflectivity-based Estimator for Seafloor Segmentation
BRT	boosted regression tree
CMECS	Coastal and Marine Ecological Classification Standard
CUBE	Combined Uncertainty Bathymetric Estimator
GV	ground validation
HIPS	Hydrographic Information Processing System
HSAA	Hudson South Auxiliary Area
IDW	inverse distance weighted
kCV	k-fold cross validation
MBES	multibeam echosounders
MVV	modified Van Veen Grab Sampler
NYWEA	New York Wind Energy Area
PCA	principal component analysis
PDE	percent deviance explained
RMSE	root mean square (error)
SBES	splitbeam echosounder
SD	standard deviation
SIPS	Sonar Information Processing System

Chapter 1. Executive Summary

1.1 Project Background

In September 2011, BOEM received an unsolicited request from the New York Power Authority (NYPA) for a commercial lease that proposed an offshore wind power project south of Long Island, New York. Subsequently, BOEM determined that competitive interest in the area proposed by NYPA (hereafter referred to in this document as the New York Wind Energy Area, or NYWEA) exists and initiated the competitive leasing process. On June 21, 2016, BOEM and NOAA establish an Interagency Agreement to conduct a three-year study of NYWEA and mid-Atlantic region titled, "Comprehensive Seafloor Substrate Mapping and Model Validation in the New York Bight". The NYWEA (321 km^2), located south of Long Island, was leased to Statoil Wind US LLC (now Equinor Wind US LLC) on December 16, 2016 for commercial wind energy development. The NYWEA offshore wind energy development also furthers New York State's objective to provide 50% of its electricity from renewable energy by 2030, and to implement a comprehensive Offshore Wind Master Plan. A chronology of milestones associated with developing NYWEA are as follows.

Chronology:

2011 September

- New York Power Authority submits unsolicited project application and lease request.

2013 January

- BOEM issues a Request for Interest to determine competitive interest.

2014 May

- BOEM publishes a Call for Information and Nominations seeking additional nominations from companies interested in the area.
- BOEM issues a Notice of Intent to prepare an Environmental Assessment associated with lease issuance and conducting site characterization surveys and site assessment activities.

2016 March

- BOEM announces the identification of a Wind Energy Area offshore New York.

2016 June

- BOEM publishes a Proposed Sale Notice for commercial leasing for wind power on the Outer Continental Shelf offshore New York.
- BOEM publishes Environmental Assessment for public comments.
- BOEM and NOAA establish an Interagency Agreement to conduct a three-year study of NYWEA and mid-Atlantic region "Comprehensive Seafloor Substrate Mapping and Model Validation in the Atlantic".

2016 October

- BOEM publishes a Final Sale Notice for a lease sale offshore New York.

2016 October

- BOEM issues revised Environmental Assessment due to environmental concerns over sensitive habitat on Cholera Bank.

2016 December

- A lease sale is held by BOEM. Statoil Wind US LLC (now Equinor Wind US LLC) wins the auction with a winning bid of $42,469,725.

2018 June

- Equinor Wind US LLC submits a Site Assessment Plan to BOEM.

The Bureau of Ocean Energy Management issues leases, easements and rights-of-way on the Atlantic Outer Continental Shelf (OCS) for activities that produce or support production, transportation, or transmission of energy from renewable energy sources. Prior to BOEM's approval of the siting of a facility, structure, or cable proposed for a renewable energy project on the OCS, an applicant must submit with its plan the results of its site characterization surveys, with supporting data. In order for BOEM to evaluate impacts to biological, social, physical and economic resources, sufficient baseline information on the area of potential effect from the proposed activity is required. BOEM uses the results of site characterization studies, which include data collection and mapping of geophysical features on the seafloor, to evaluate the potential effect of proposed activities. The study conducted by NOAA, Comprehensive Seafloor Substrate Mapping and Model Validation in the Atlantic, and report herein, are provided to BOEM to support these requirements.

Prior to this study, NOAA National Centers for Coastal Ocean Science (NCCOS) has conducted previous efforts to characterize the ocean floor offshore NY. For example, as part of the NY Department of State's renewable energy planning efforts and its Offshore Atlantic Ocean Study, released in July 2013, NOAA's NCCOS developed a biogeographic assessment with maps and spatial information on bathymetry, surficial sediments, deep sea corals, oceanographic habitat variables, and seabirds offshore of New York (Menza et al. 2012). However, NOAA's previously released predictive models of seafloor substrate and sediment composition for the New York region were identified as needing additional ground validation to assess the model predictions and ensure reliability. The New York biogeographic assessment explicitly identified the need for further site-specific baseline habitat and geologic information within NYWEA and the mid-Atlantic region so that managers are better able to evaluate the potential impacts of offshore wind development at this location. As such, this led to the establishment of the subsequent BOEM and NOAA Interagency agreement to conduct further study, data collection and analysis, contained in this report, for NYWEA and the NY Bight region.

This report describes information collected and analyzed by NOAA to meet BOEM's three primary objectives for the study:

Objective 1: Collect data of the seafloor by the use of acoustic sonar survey to characterize the substrate composition of the seafloor in the NYWEA and vicinity.

Objective 2: Incorporate results of the seafloor data collection into regional predictive models characterizing the composition of the seafloor.

Objective 3: Describe the seafloor habitat and other important biological and geological observations.

The data collection and analysis contained within this report focus on two locations: 1) NYWEA and the mid-Atlantic region study area (Figure 1.1). The NYWEA site is located approximately 11.5 nautical miles (nm) south from Jones Beach, NY and extends approximately 24 nm southeast along its longest portion. It is located in water depths ranging from 25 to 45 m deep. Extensive acoustic and sediment sampling, and ground validation studies were conducted by NOAA NCCOS within NYWEA. The regional study area encompasses 50,082 km^2 of coastal and ocean waters off the coast of New York. This area includes a portion of the mid-Atlantic Bight and much of the area characterized as the New York Bight. Analysis for the regional study area included state and federal waters from the southern shores of Long Island to the edge of the continental shelf and from Nantucket Shoals to the shores of New Jersey. This report contains analysis of hard bottom predictions and sediment texture throughout the regional study area (Figure 1.1).

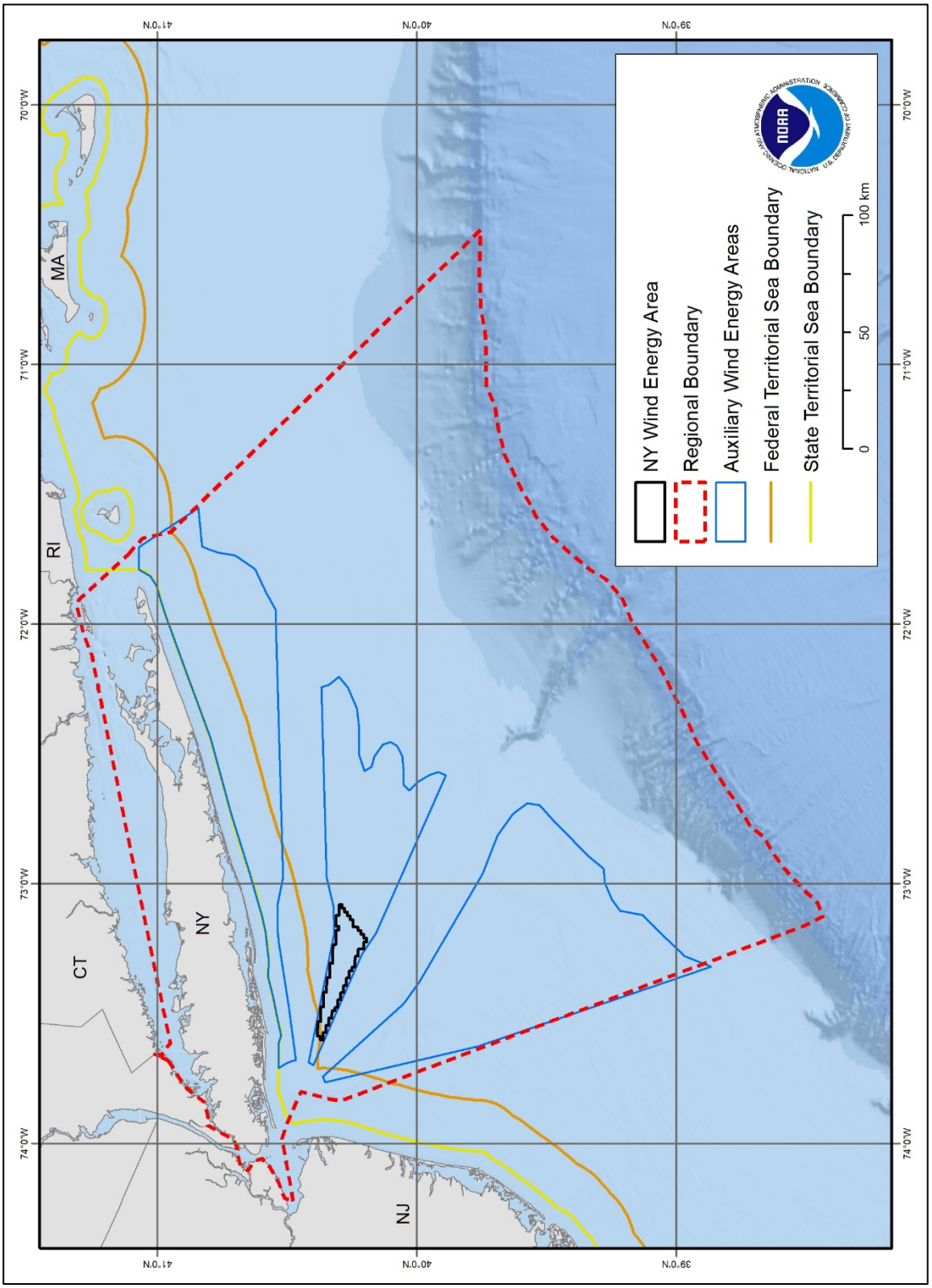


Figure 1.1. A map of the study area used in this report.

1.2 Highlights by Chapter

The following contains a brief summary for each of the subsequent chapters contained in this report.

Chapter 2: Acoustic Survey

Describes ship-based multibeam and fishery acoustic surveys conducted by NOAA within the NYWEA project area. The purpose of these surveys was to collect detailed information in order to better characterize the seafloor depth, roughness, hardness, composition, habitats and topography; and characterize biological use of the region in terms of the fish abundance, distribution, and size. High resolution bathymetry and backscatter data of the seafloor were collected using multibeam sonar, encompassing 283.1 km^2, and concurrently, 1,520 nm of fishery acoustic acquisition.

Chapter 3: Ground Validation and Sediment Survey

Describes the collection of *in situ* seafloor data including underwater photos, videos, and sediment samples. Observations were collected in September 2018 at 400 sites distributed throughout NYWEA, which were subsequently partitioned into 300 ground validation sites, used for model training, and 100 sites used for accuracy assessment. A modified Van Veen Grab Sampler was used to collect video, still photos, and sediment samples to support habitat mapping and sediment composition analysis. The analysis of these data sets is further described in Chapters 7, 9, and 11.

Chapter 4: Seafloor Sonar Processing

Describes the methods used to post-process the raw acoustic multibeam survey data in order to derive accurate, cleaned data sets for subsequent analysis. Industry standard hydrographic processing workflows were applied to the raw data to generate final bathymetry and backscatter surfaces for NYWEA. The application of these surfaces to characterize the morphometric and habitat composition of NYWEA is further described in Chapters 6 and 11.

Chapter 5: Fishery Acoustic Processing

Describes the methods to post-process the raw fishery acoustic sonar data collected for NYWEA. These techniques are employed to quantify the intensity of sound reflected from fish targets to support subsequent analysis of fish size distribution and abundance described in Chapter 12.

Chapter 6: Seafloor Morphometric Analysis

Describes the production of morphometric products from the multibeam sonar data (Chapter 2 and 4) used to characterize the distribution, composition, roughness, shape, and texture of the seafloor and landforms in NYWEA. Results indicate the seafloor in the study area is generally flat (<0.2 degrees slope) with a maximum slope of 4.3 degrees, and 99.9% of the area having slope of less than 1.5 degrees. Furthermore, spatial analysis identified that NYWEA has large areas with consistent topographic complexity and substrate composition commensurate with flat sandy seafloor.

Chapter 7: Ground Validation Video Analysis

Describes the analysis of sediment substrate and cover types from video observations (Chapter 3). *In situ* sampling indicates NYWEA is comprised primarily of sand with broken shell (71.3%), rippled sand waves-geoform (83%), and extensive distribution of the common sand dollar (*Echinarachnius parma*) (90%) and annelids (89%).

Chapter 8: Hard Bottom Predictive Modeling
Describes spatial predictive modeling to predict the extent of hard bottom habitats in the NY Bight region. Environmental predictor variables considered for hard bottom occurrence model included measures of depth and seafloor topography, seafloor substrate, and oceanography. The model of hard bottom occurrence predicted relatively high likelihood of hard bottom in some nearshore areas (e.g., on the edge of Block Channel in the northeast of the study area), on the sides of the Hudson Shelf Valley, and on the slopes of the submarine canyons that incise or partially incise the continental slope. Within the NYWEA, the northwest portion was predicted to have a relatively higher likelihood of hard bottom occurrence, as suggested by the backscatter intensity data collected during this study.

Chapter 9: Sediment Grab Analysis
Describes the results of sediment classification and spatial analysis of *in situ* samples taken in NYWEA (Chapter 3). Analysis revealed that the sediment texture of NYWEA is predominantly well-sorted sand with a conglomeration of pebbles in western portions of the study area.

Chapter 10: Sediment Texture Analysis
Describes updated regional maps of surficial sediment composition (median grain size and percentages of mud, sand, and gravel) at 200 x 200 m resolution for the mid-Atlantic region. Spatial predictions were generated using multivariate (3D) inverse distance weighted interpolation and random forest modeling from surficial sediments point samples. These maps provide additional information about the distribution of surficial sediments in areas outside the NYWEA that may be impacted by activities within it (e.g., cabling associated with offshore wind installations) or that may receive future consideration for activities overseen by BOEM.

Chapter 11: Habitat Maps
Describes the methods, results, and performance of spatial prediction for substrate, geoform, biotic cover, and benthic habitat map. The spatial prediction accuracies ranged from 92.8% (bare sand, no cover) to 11.3% (crustaceans). The habitat map model indicted that "rippled sand with high occurrence of faunal beds" was the most abundant and dominant habitat type, comprising 78.2% of the NYWEA.

Chapter 12: Fish Acoustic Analysis
Describes the methods, results, and patterns of fish distribution in NYWEA. Analysis of the fish acoustic data observed significant variation in the spatial distribution of individual fish, numerous schools of fish with a variety of shapes and dimensions, and plankton layers and patches that were particularly evident during the overnight hours. Small fish were vastly more abundant and broadly distributed throughout the survey area compared to medium and large size class fish, which were strongly associated with the northwest portion of the survey area.

This report concludes with a glossary of technical and statistical terms used throughout the document.

1.3 Data Delivery

In addition to this final report, a compendium Digital Data delivery was delivered to BOEM provided which includes all of the raw geophysical and derived GIS data collected, analyzed, and processed as part of this effort. These datasets can be accessed via BOEM's site and NOAA National Centers for Environmental Information (NCEI).

Chapter 2. Acoustic Survey

Seafloor depth, roughness, hardness and topography are known to be useful predictors for identifying a variety of habitat types such, as sand, hard bottom, coral reefs, seagrass, etc. (Costa et al. 2009; Costa and Battista 2013). Ship-based sonar can be used to conduct acoustic surveys to collect depth data (i.e., bathymetry), which can be analyzed to describe the topography of the seafloor. Most sonars can also collect backscatter data, which can be used to describe the roughness and hardness of the substrate on the seafloor. Additionally, ship-based fishery acoustic sonars can opportunistically be employed while conducting sonar surveys to characterize fish abundance, distribution and size to be able to characterize biological use of a region.

2.1 Chronology

Monday, September 4, 2017 was the original scheduled start date for mission NF-17-09. The project was allotted 25 Days at Sea (DAS) on the NOAA Ship *Nancy Foster* through September 30, 2017. However, due to unforeseen platform and severe weather-related circumstances, several adjustments to the cruise plan were made prior to the successful completion of the survey.

From September 4 at 1000 hr, until September 6 at 1200 hr, the *Nancy Foster* was in a "No Sail" status due to mechanical issues. Repairs to the rescue boat davit were required in order to resume safe operations. Furthermore, the *Nancy Foster* was required to remain at its homeport in Charleston, South Carolina due to the risk of approaching Hurricane Irma in the Atlantic Ocean from September 6-14.

On September 14 at 0800 hr, the *Nancy Foster* departed Charleston and began transiting to Atlantic Highlands Municipal Harbor, New Jersey. The crew were expected to arrive and load visiting scientists via a small boat transfer at roughly 1200 hr on September 16.

The ship proceeded north up the eastern seaboard, approaching the Chesapeake Bay, but was diverted to Norfolk, Virginia on September 15. The NOAA Marine Center had announced a decision to hold all east coast ships in port, as seas from Hurricane Jose were creating conditions deemed unsafe for transit or operations. This status continued until September 29, extended by the formation and uncertain trajectory of Hurricane Maria in the Caribbean Sea.

At 1200 hr on Friday, September 29, 2017 the *Nancy Foster* resumed underway transit for mission NF-17-09 after all scientists had loaded in Norfolk. Mapping team personnel included four NOAA employees and affiliates and one subcontractor with Solmar Hydro, Inc. (Table 2.1).

Table 2.1. NF-17-09 Science Party Personnel.

Name (Last, First)	Title	Date Aboard	Date Disembark
Battista, Tim	Oceanographer	9/29/17	10/09/17
Husted, Rachel*	Physical Scientist	9/29/17	10/09/17
Johnston, Brent††	Hydrographer	9/29/17	10/09/17
Kinney, Juliet^	Physical Scientist	9/28/17	10/09/17
Mabrouk, Ayman*	Physical Scientist	9/29/17	10/09/17

* CSS, Inc. contractor for NOAA ^ ERT, Inc. contractor for NOAA ††Solmar Hydro, Inc.

The ship transited north for approximately 30 hours before reaching the NYWEA. On September 30, starting at approximately 1900 hr, the scientists conducted a patch test calibration of the Reson 7125 and Kongsberg EM710 multibeam echosounders (MBES). The test site was located on the Hudson Canyon and coincided with the site used by the NOAA Ship *Ferdinand R. Hassler*'s calibration in 2013.

Patch tests, in which data is acquired for a series of survey lines run in specific patterns over features with known bathymetry, are used to calibrate multibeam systems. Planned pairs of survey lines allow the measurement of offsets due to any slight misalignment of the sensors regarding four variables: time delay, pitch, roll, and heading. The purpose of the patch test is to correct for systematic errors created by the positioning and mounting angles of the acoustic sensors. A calibrated system will provide accurate and repeatable bathymetric observations of a location regardless of vessel speed, direction, and motion.

Given the project's delayed start due to extenuating weather conditions, NOAA consulted the BOEM Contracting Officer Representative to consult on a strategy to maximize survey coverage of the NYWEA. BOEM approved a strategy to focus NF-17-09 acquisition on areas in NYWEA that lacked complete high-resolution bathymetry coverage. The western corner of the NYWEA project area had been previously surveyed by the NOAA Ship *Ferdinand R. Hassler* in 2013. Evaluation of these previously acquired acoustic data determined that they were of sufficient quality and coverage for further use, thereby reducing duplication of effort. As such, multibeam and fishery acoustic data were not acquired over the western corner of NYWEA during NF-17-09.

Upon arrival to the project site on September 30, 24-hr survey work began at about 2300 hr. The scientists surveyed the project area using a line spacing, line orientation, and survey speed plan optimized for the charted depths and weather conditions. When possible, survey speed increased from seven to approximately 10 knots in an effort to compensate for reduced project time.

Data quality varied by day depending on the sea state conditions. Certain lines were resurveyed to reduce the effect of motion-induced noise detected in the survey data during periods of high sea state. Maritime traffic in the area was minimal, resulting in limited impact on survey operations.

The NF-17-09 research cruise duration ultimately spanned 202 cumulative hours over 10 DAS. In the updated project area, 100% complete seafloor ensonification of the reduced NYWEA project area was achieved with high-resolution multibeam bathymetry and backscatter from both the Kongsberg EM710 and Reson 7125 MBES systems, as well as a full coverage dataset of EK60 fishery acoustics. The scientists disembarked via a small boat transfer to Belmar, New Jersey at 0800 hr on October 9, 2017.

Mission NF-18-07 began on September 4, 2018. The project was allotted 30 DAS on the NOAA Ship *Nancy Foster* through September 29, 2018. Chief Scientist Tim Battista and engineer Ryan Caillouet came aboard the *Nancy Foster* on Thursday, September 6 from Virginia Beach, VA to begin testing the towed camera system while the ship made its transit toward the NYWEA. Upon reaching the project area on September 7, until September 9, scientists also began work to

resurvey a set of lines identified as having reduced data quality due to the several days of bad weather conditions during NF-17-09, and collected EK60 SBES fishery acoustics data.

Additional science party personnel joined the *Nancy Foster* on Sunday, September 9 at the Intrepid Sea, Air & Space Museum pier in Manhattan, New York. The complete mapping team included seven NOAA employees and affiliates (Table 2.2).

Table 2.2. NF-18-07 Science Party Personnel.

Name (Last, First)	Title	Date Aboard	Date Disembark
Battista, Tim	Oceanographer	09/06/18	09/18/18
Caillouet, Ryan	Engineer	09/06/18	09/18/18
Ebert, Erik*	Fishery Scientist	09/06/18	09/29/18
Husted, Rachel*	Physical Scientist	09/09/18	09/29/18
Kraus, Jennifer	Physical Scientist	09/09/18	09/29/18
Mabrouk, Ayman*	Physical Scientist	09/09/18	09/29/18
Williams, Bethany^	Physical Scientist	09/09/18	09/18/18

* CSS, Inc. contractor for NOAA ^ Sea Grant Knauss Fellow

From September 9-15, the *Nancy Foster* was in a "No Sail" status due to the rough weather conditions in the Atlantic Ocean. On Saturday, September 15 at 1000 hr, the *Nancy Foster* began the transit to the NYWEA project area. Unfortunately, roughly 4 hours into the transit critical mechanical issues occurred and the *Nancy Foster* was forced to transit to a dock in Staten Island, New York. From September 16 to September 19, the ship remained in port, hosted a burial at sea, and carried out a planned crew exchange.

Starting on Thursday, September 20 at 1400 hr, the *Nancy Foster* returned to the NYWEA, continued to resurvey selected lines for improved data coverage, and logged the EK60 with 24-hr mapping operations. Data quality varied depending on the sea state. Under poor conditions, lines were at times run in a single direction (to combat sea swell direction) at lower than normal speeds (under 6 knots), which produced better results at a trade-off with survey time efficiency. Certain lines were re-surveyed multiple times to reduce the effect of motion-induced noise. The Survey Technician also recommended powering off the EK60 intermittently due to interference with the multibeam. Maritime traffic in the area had an impact on survey operations, causing some additional delays.

On Monday, September 24, the resurveyed lines in the NYWEA were completed. The team transited south and began surveying in the northwest portion of an area identified by BOEM as the Hudson South Auxiliary Area (HSAA), located south of the Hudson Canyon. Beginning in the southwest corner, the team ran both EM2040 and EM710 sonars, and logged EK60 fishery acoustics data. The line plan followed by the ship was designed to optimize the data quality in terms of survey direction to follow the wave and wind conditions, as well as to maximize the exploration of the new area in the limited time available.

Due to forecasted unworkable weather conditions from September 24-26, the *Nancy Foster* went to shelter in Brooklyn, New York until conditions improved. On Thursday, September 27, the *Nancy Foster* returned to the HSAA to collect as additional multibeam and EK60 coverage as

possible with the remaining Days at Sea for the project. 24-hour mapping operations were completed at approximately 0400 on September 28.

The scientists disembarked via a small boat transfer to Atlantic Highlands, New Jersey at 0900 hr on Saturday, September 29, 2018.

2.2 Platform

The NOAA Ship *Nancy Foster* (R352) is 57 m in length, has a beam of 12 m, and draws approximately 3 m of water (Figure 2.1). Formerly a U.S. Department of the Navy torpedo test vessel, the ship was transferred to NOAA for use as a coastal research vessel in 2001. The *Nancy Foster* now supports seafloor mapping, fish habitat and population studies, physical and chemical oceanography studies, and maritime heritage surveys throughout the U.S. eastern seaboard, Caribbean, and Gulf of Mexico. Multibeam surveys are an integral component of the ship's scientific research support. NOAA's Office of Marine and Aviation Operations has installed a full suite of hydrographic hardware and software to enable hydrographic surveys. The *Nancy Foster* is equipped with wet and dry labs, as well as computers for data acquisition and analysis.

Figure 2.1. NOAA Ship *Nancy Foster* underway. (Photo Credit: Connor Maginn)

2.3 Multibeam Acoustic Systems

Through 2017, the NOAA Ship *Nancy Foster* (R-352) was equipped with a Reson SeaBat 7125-SV2 dual frequency (200 or 400 kHz) shallow-water (5-250 m depth) MBES. The 7125-SV2 was mounted on the vessel hull, port of the keel and forward of the reference point. In February 2018, a Kongsberg EM2040 (200, 300, or 400 kHz) wide band high-resolution shallow water multibeam (6,000 m max depth) MBES was installed, replacing the Reson SeaBat 7125-SV2 system. During the NF-18-07 surveys, EM2040 acquisition was conducted using the 300 kHz frequency to reduce any potential interference with the EK60 fisheries sonar. A deep-water (40-2,500 m depth) Kongsberg EM710 MKII MBES system, with 40-100 kHz range, is also permanently hull-mounted on the *Nancy Foster,* starboard of the keel line. During NF-18-07, EM710 acquisition was conducted using the 40-100 kHz frequency range in Kongsberg's "very shallow" mode. The EM710 sonar was only used for acquisition in the HSAA survey starting on September 24.

For this project, multibeam survey data was collected to International Hydrographic Organization (IHO) Order 1 (<100 m depth) accuracy standards using workflows defined by the NOAA Field Procedures Manual and Hydrographic Survey Specifications and Deliverables Manual (IHO 2008; NOAA NOS 2018). This included maintaining acquisition and preliminary processing logs, real-time quality assurance, and quality control of the incoming data. All sonar range adjustments during acquisition were dictated by changes in water depth. Other sonar parameters remained constant to maximize backscatter quality and consistency throughout the survey.

Hypack/Hysweep software was used for 7125-SV2 bathymetry acquisition and vessel navigation. The system provided precise time tagging of the sensor data and real-time data displays for quality control. Hypack data were output in three formats: .RAW, .HSX, and .7K. Kongsberg Seafloor Information System (SIS) software stored the EM2040 and the EM710 MKII bathymetric data in .ALL format. An Applanix Position and Orientation System for Marine Vessels (POS/MV) was used to measure real-time attitude, heading, heave, and position for the Hypack and SIS systems. Fugro Marinestar Regional GNSS 9G2/VBS satellite-based service was used for positioning corrections. An OceanScience uCTD (underway Conductivity, Temperature, and Depth) measured sound speed profiles while the vessel was underway. uCTDs were cast off the aft deck with a tethered freefall probe and retrieved with a powered level-winding winch. After each uCTD launch, the data were downloaded and processed with UCast software and concatenated using NOAA Pydro Velocipy software.

MBES data were converted and processed with CARIS HIPS (Hydrographic Information Processing System) and SIPS (Sonar Information Processing System). Methodology followed the CARIS HIPS CUBE (Combined Uncertainty Bathymetric Estimator) workflow. These steps include data conversion, filtering, sound velocity correction, tide correction, Total Propagated Uncertainty (TPU) calculation, TrueHeave application, merging, and editing. QPS Fledermaus Geocoder Toolbox (FMGT) was used to process and assess the backscatter intensity data. For NF-18-07 QPS Qimera software was used to process the MBES bathymetry data. Methodology followed the Qimera Dynamic workflow, which is similar to the CARIS HIPS CUBE workflow.

2.4 Fishery Acoustic System

Fish acoustic surveys were also conducted using a splitbeam echosounder (SBES), coincident with MBES acquisition. SBESs are a tool that give researcher a look at the amount biological biomass within a survey area at a given time. This tool is non-invasive and non-lethal way of sampling the fish community compared to other traditional fish sampling methods (trawls or hook and line) as well as less labor intensive. Data collection was continuous and provide researcher with the means to detect at pelagic fish movement patterns with limited disturbance to the natural distribution (Mitson 1998). One drawback to fishery acoustic survey is the lack of individual species information but SBES could aid in follow up research mission to focus those efforts in areas of high biomass. SBES data collection was well-suited for this research mission because of the limited ship time for the scope of the project and data collected provided view of the biological hotspots within the large survey area.

The NOAA Ship *Nancy Foster* is equipped with a Simrad EK60 SBES operated at 120 kHz frequency. The transducer is mounted to the hull and surveyed to a common reference point to

provide precise offsets relative to ship's navigation, multibeam sonars, and other data acquisition systems. The transducer has a nominal beam geometry of 7° and results in a swath or footprint that is about 12% of range from the transducer face (or water depth), or about 3 m swath at the seafloor in 25 m water depth. The pulse transmission (ping) characteristics, data acquisition and data viewing were controlled from a workstation operating Simrad EK80 software (version 1.10.3, Simrad Fisheries) and connected by local area network to the General Purpose Transceivers (GPTs). The ping timing was triggered by and synchronized with a 15 ms delay to the Reson 7125 MBES. Each ping is co-registered with the ship's time server (referenced to Greenwich Mean Time), and the ship's navigation and motion system, which included time, latitude and longitude, pitch, roll, and heave. Output power, pulse length, and other ping transmission properties are provided in Table 2.3. EK60 data files were logged in 100 MB segments.

Table 2.3. Acquisition parameters for the Simrad EK60 SBES on the NOAA Ship *Nancy Foster* used to map fish density distributions in the New York Wind Energy Area (NYWEA).

Parameter	120kHz Frequency
Transducer depth (m)	1.5
Transmit Power (dB-W)	250
Pulse length (μs)	128
Sound velocity (nominal, m s^{-1})	1516.24
Calibration gain (dB)	25.67

The EK60 SBES system was calibrated prior to the cruise (July 11, 2017 at 34° 09.603' N / 76° 38.347' W) using the standard sphere method (Foote et al. 1987) with a 38.1 mm diameter tungsten carbide sphere hung below the transducer. The target sphere has a known theoretical acoustic target strength based on the composition sphere diameter and environmental conditions. The calibration sphere was systematically moved through the beam from forward to aft and port to starboard. The calibration routine calculates the system receiver gain to bring the observed target strength in concordance with the theoretical target strength for the sphere.

2.5 Survey Metrics and Results

In total, 1,520.3 nm were surveyed for 432 track lines distributed throughout the survey area, excluding those areas previously surveyed by NOAA Ship *Ferdinand R. Hassler*. There were 283.1 km^2 of seafloor ensonified with multibeam backscatter acoustic intensity and bathymetry data, ranging from approximately 26 to 43 m depths (Figure. 2.2). In September 2018, 106 km^2 were resurveyed in the NYWEA to enhance data quality (Figure 2.3). In the HSAA, 183 additional nm were also explored, covering an area of 28 km^2.

Fishery acoustics data was logged simultaneously with MBES survey lines during cruises NF-17-09 and NF-18-07, matching the multibeam for 1,520.3 nm of acquisition in the NYWEA and 183 nm in HSAA (Figure 2.4).

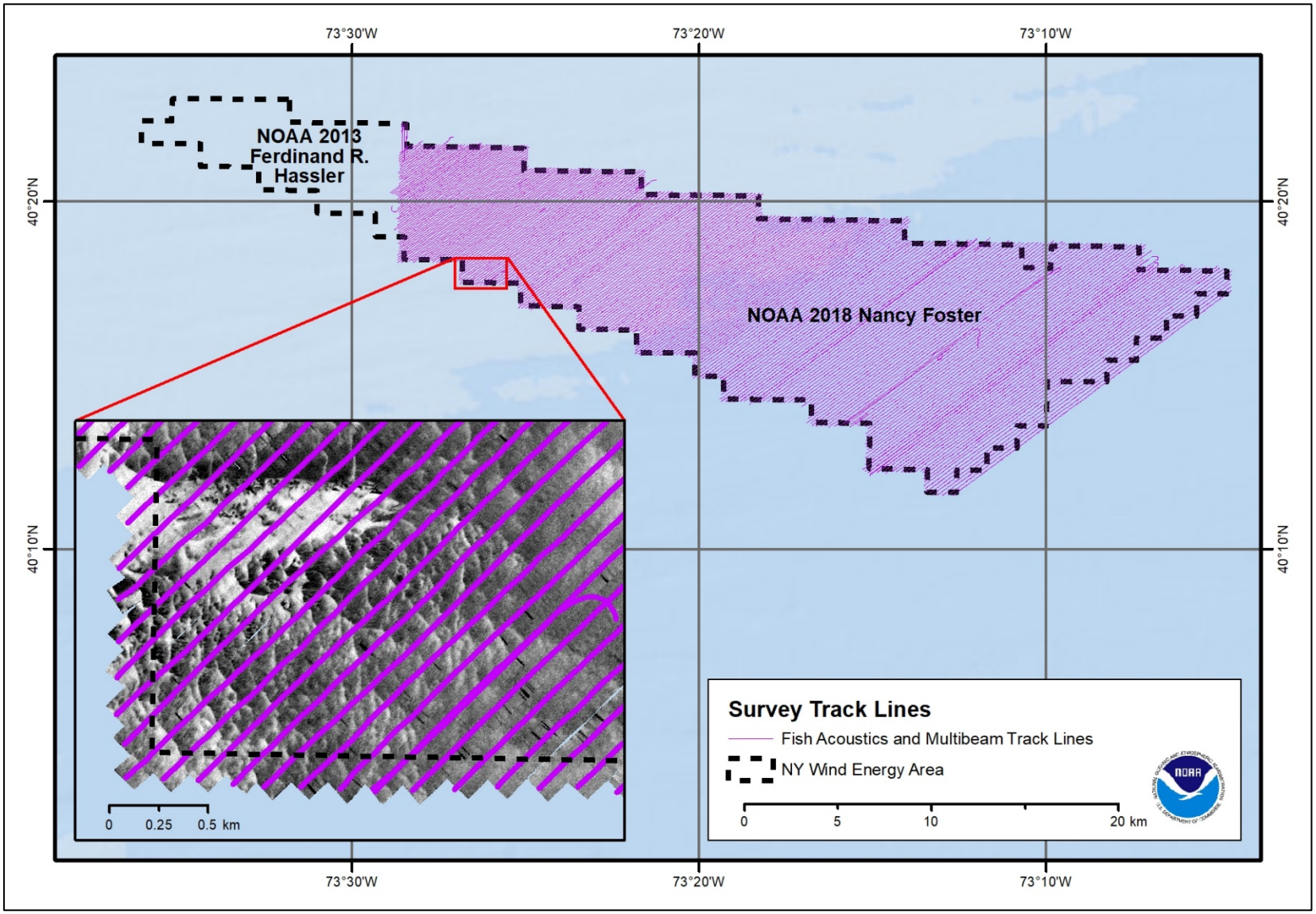


Figure 2.2. NOAA Ship *Nancy Foster* 2017 Survey Track Lines (Fish Acoustics and Multibeam Acquisition) in the New York Wind Energy Area (NYWEA).

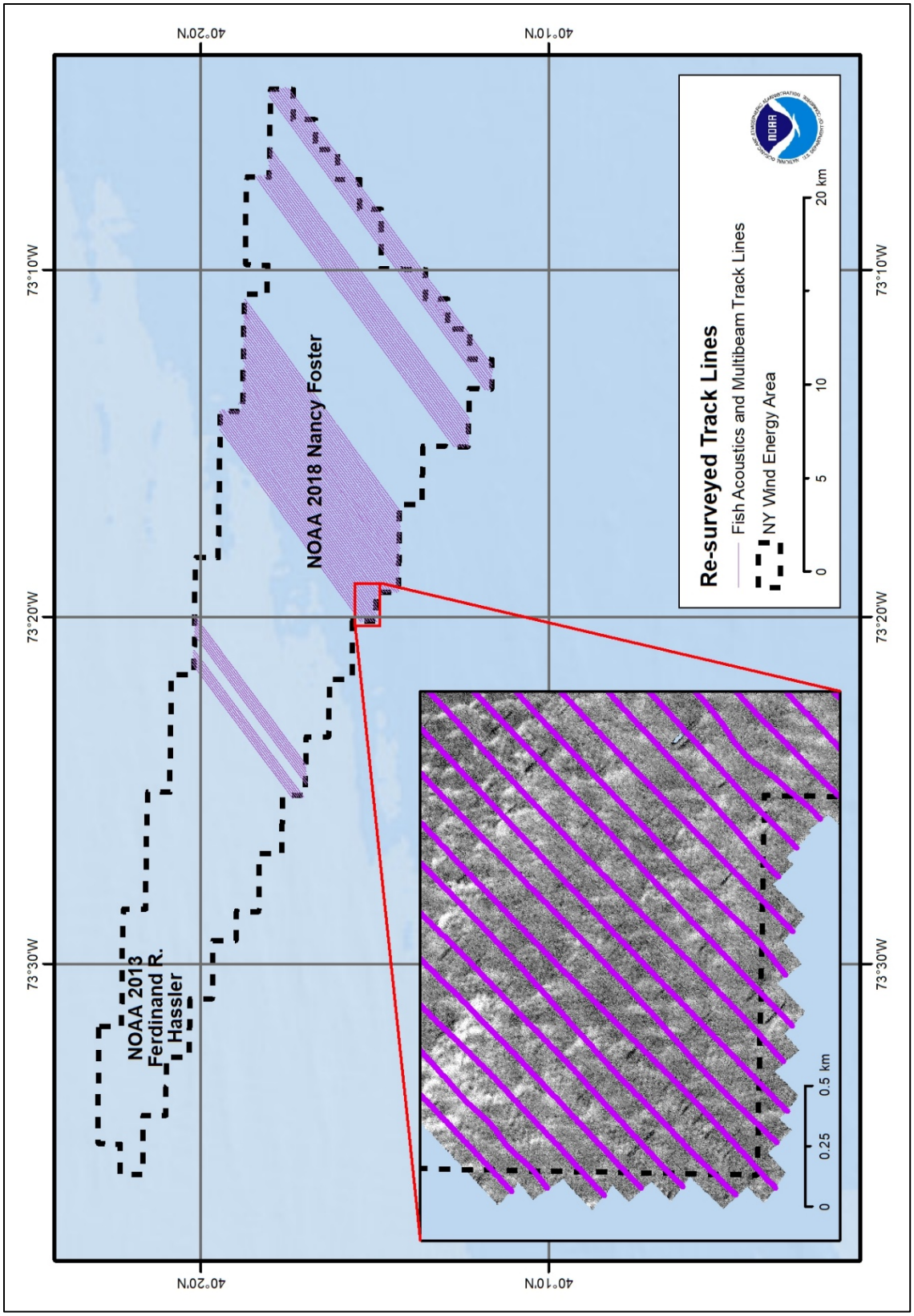


Figure 2.3. NOAA Ship *Nancy Foster* 2018 Re-surveyed Track Lines (Fish Acoustics and Multibeam Acquisition) in the NYWEA.

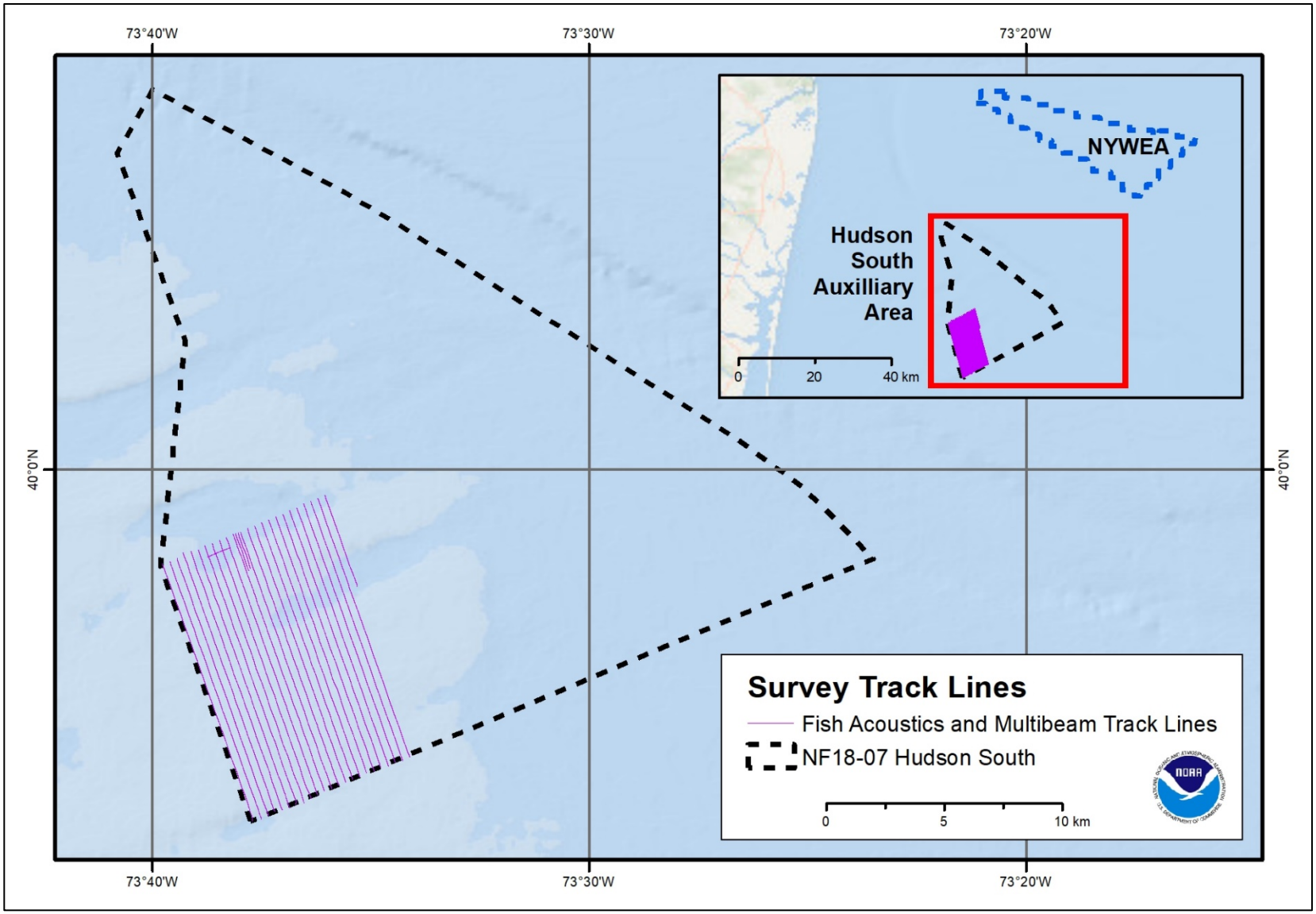


Figure 2.4. NOAA Ship *Nancy Foster* Track Lines (Fish Acoustics and Multibeam Acquisition) in the Hudson South Auxiliary Area (HSAA).

Interesting incidental discoveries may arise due to this new survey in the NYWEA, such as the shipwreck feature clearly visible during preliminary processing of the multibeam bathymetry data in CARIS HIPS and SIPS software (Figure 2.5).

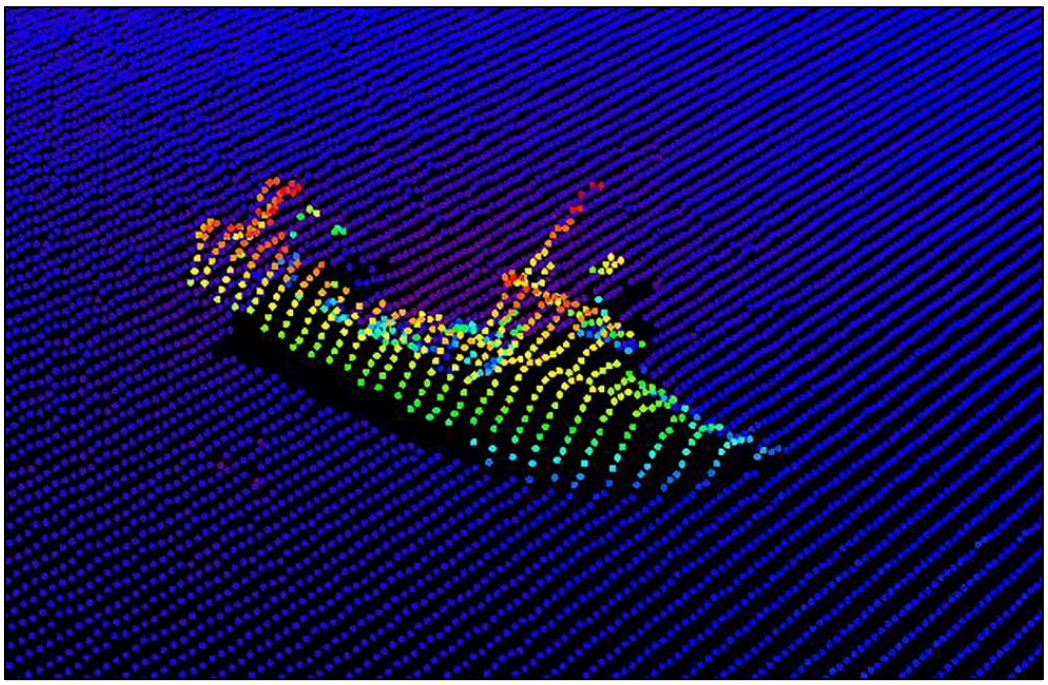


Figure 2.5. Shipwreck feature detected in the NYWEA during the NF-17-09 survey.

2.6 Data Management

On the NOAA Ship *Nancy Foster*, data was transferred daily from the acquisition computers to the ship's server. Back-ups of all of the raw multibeam, fish acoustics, and positioning data acquired during this mission were saved to an external hard-drive and uploaded to the NOAA NCCOS office network data server at NOAA headquarters in Silver Spring, Maryland. A local copy of the fish acoustics data was stored at the Beaufort, North Carolina NCCOS office for processing. An additional complete mission dataset was sent to Solmar Hydro, Inc., contracted to perform final cleaning, processing, and bathymetric product creation (see Chapter 4 Seafloor Sonar Processing) for this project. NCCOS created final multibeam acoustic backscatter mosaic following the receipt of the final processed multibeam bathymetry from Solmar Hydro, Inc. on December 29, 2017.

Chapter 3. Ground Validation and Sediment Survey

Map making or modeling typically incorporates the use of *in situ* sampled data in order to better inform the inference between remotely sensed data and site-specific characteristics. *In situ* data provide the means of detecting the relationships between the observed substrate and cover types and the values in the environmental predictor datasets. A number of techniques and instrument types can be used to collect *in situ* seafloor data including the collection of underwater photos and/or video and sediment sampling. Instrument types include sediment grab samplers, drop cameras, tow sleds, or remotely operated vehicles (ROVs). Observations may be taken at discrete locations or along transects.

In situ sampling is comprised of two categories of data that inform different stages of the map making process - Ground Validation (GV), also known as ground truthing, and Accuracy Assessment (AA) (see Section 11.1.5). For GV, the mapping analyst manually selects GV sites from the draft map based on locations where further information is required. GV data are intended to provide additional information to train and optimize the mathematical models used to predict habitats, and thereby improve model predictions. GV data are needed to create high quality benthic habitat maps because they assist in associating fine scale habitat observations with features detectable in the source remote sensing data or predictor variables. Both sediment sampling and optical data collection (video and photos) were used for *in situ* GV. This chapter describes in detail the equipment and methods used to collect sediment samples used for the GV analysis.

3.1 Survey Design

The sample sites were distributed throughout the project area (Figure 3.1) and included all of the substrate classes identified in the draft habitat map. Of these 400 samples, 300 were used in the GV analysis and 100 were used in the AA analysis. After sampling was completed, the samples were designated as either GV or AA using a random site selection tool in excel.

Draft habitat maps were used to determine the locations of 400 sampling sites collected during the April 2018 cruise aboard the R/V *Tiki XIV* (Figure 3.2). The R/V *Tiki XIV* is an 80 ft steel trawler (23 ft beam and 10 ft draft) operated by Tiki Adventures, Inc. The vessel has a cruising speed of 8 knots and an operational endurance range of 6,700 nm or approximately 14 days, depending on speed and fuel consumption. A team of six scientists, two deck hands, and one captain sailed on the R/V *Tiki XIV* from April 2-13 within the NYWEA project area. Two days were used for the transit to the NYWEA, eight days for sampling, and approximately two days were at anchor or at the pier due to unworkable weather conditions. Typically, between 20-90 sediment samples were collected per 24 hr period, depending on daily sea state working conditions.

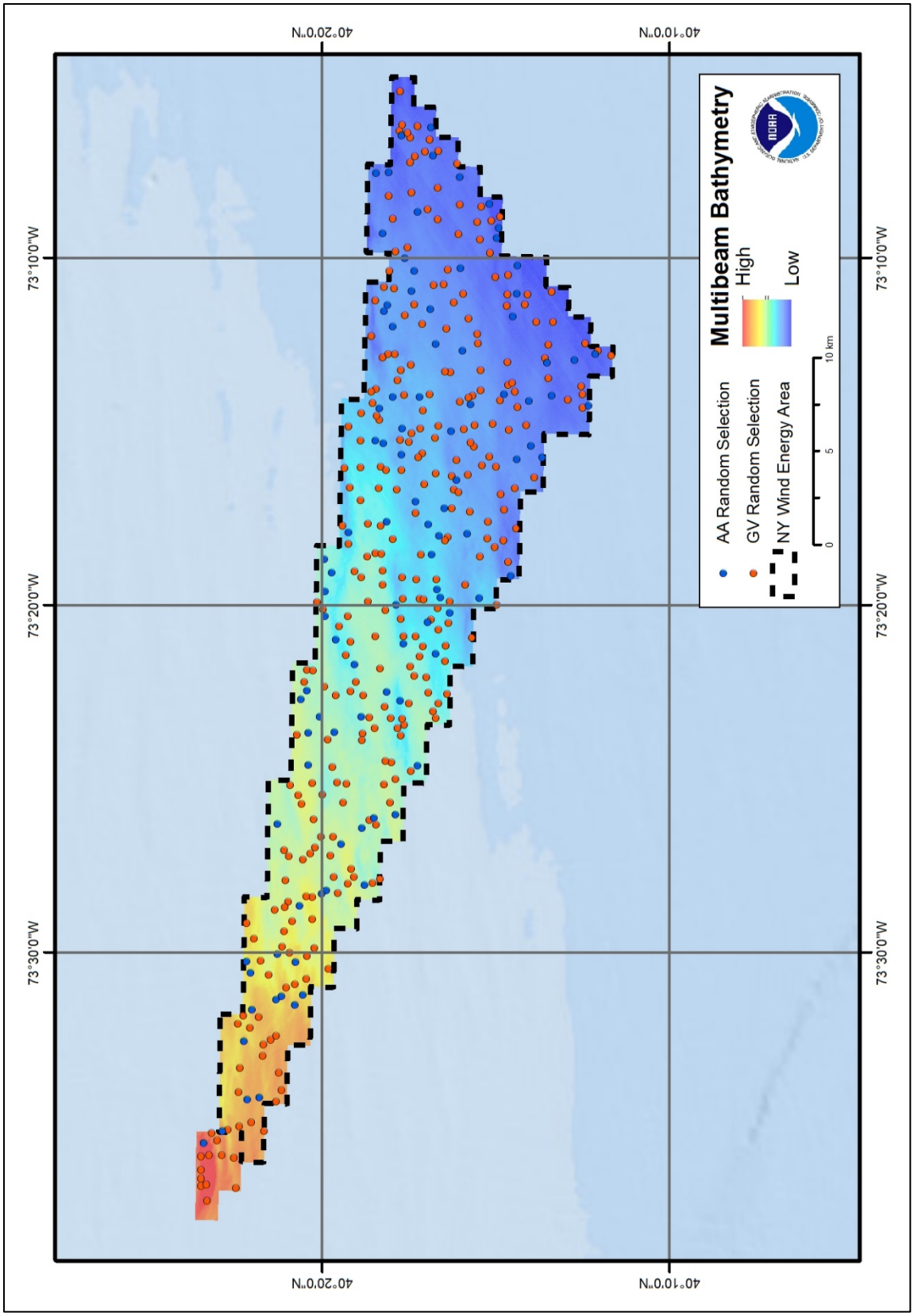


Figure 3.1. Location of Ground Validation (GV) and Accuracy Assessment (AA) sediment sample sites within the NYWEA.

Figure 3.2. The R/V *Tiki XIV* is an 80 ft steel trawler that was contracted by NOAA to conduct GV sampling in April, 2018.

The AA survey planned to use additional seafloor video data collected by using an Underwater Towed Optical Survey System (UTOSS) (Figure 3.3) from September 6-16, 2018. However due to weather and mechanical issues during this period, the collection of this data was not achievable. Collection efforts for the remainder of the trip (September 17-29) aboard the NOAA Ship *Nancy Foster* focused on resurveying multibeam data in the NYWEA to improve data quality and the exploration of the HSAA (see Chapter 2 Acoustic Survey).

Figure 3.3. Front view of the Underwater Towed Optical Survey System (UTOSS) tow camera system.

3.2 Deployment and Collection

Sediment samples were collected using a 250 cm^2 stainless steel, modified Van Veen Grab Sampler (MVV or 'sampler') with a frame (Figure 3.4). The sampler was deployed from the R/V *Tiki XIV* using the vessel's A-frame (12,000 lb weight limit with 10 ft deck clearance) and deck winch (1,700 lb working load). Approximately 200 lb of removable weights were added to the frame to ensure the grabber contacted the seafloor with sufficient force to activate the sediment grabber's closing mechanism (Apeti et al. 2012). Two GoPro Hero4 cameras (one downward facing and one at an oblique angle) were mounted to the sediment grabber's frame to record high definition video. Time calibration of the GoPro cameras to a world GPS clock was performed prior to deployment to synchronize video recording between the cameras. Additionally, a real-time Seaviewer underwater video camera was mounted to the frame to provide the scientists a live view of the sampler as it was lowered to the seafloor. The downward cameras provided a view of the seafloor before the grab sampler disturbed the sediment. The cameras were positioned between two lasers set 20 cm apart for scale. Three external dive lights were also mounted on the frame to improve bottom imaging in low light conditions. The oblique facing cameras allowed the analysts to observe the larger scale geoforms or animals in the water column, as well as record any features or biological cover of the seafloor not directly beneath the sampling instrument.

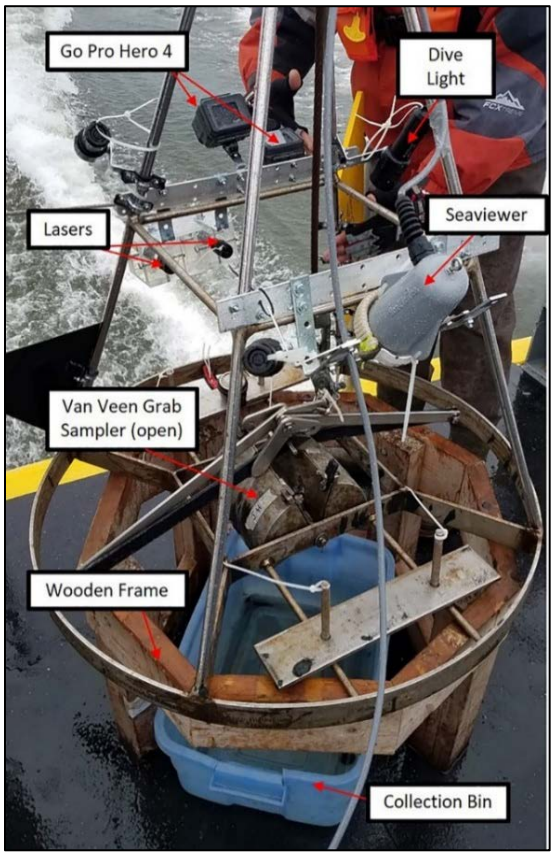


Figure 3.4. The (open) modified Van Veen Grab Sampler (MVV) attached to the frame.

The Seaviewer video was viewed and recorded on a solid state drive using Blackmagic Video Assist Recorder and an external 14 in monitor used for viewing and classifying the habitat real time. The Seaviewer camera was only used for the first two days of sampling due to mechanical issues preventing further use. However, continued use of the Seaviewer camera was noncritical given that video recordings captured by the GoPro cameras were the primary observation device. A mapping grade Trimble Geo7X 6000 handheld GPS receiver was used to record the site location and sediment classifications after each sample was retrieved. The GPS data collected by the Trimble Geo7X 6000 units were converted into a shapefile that contained the location, time, and field observations from each grab sample site. The site location recorded in the GV/AA dataset was the calculated average centroid of the one-second epochs logged in the field during the deployment and retrieval of the sediment grab sampler. A handheld Garmin GPS unit with the site numbers and location preloaded was used to verify the site with the vessel Captain. The GoPro cameras were set to record HD video (1080 x 1920 pixels at 60 frames/second) directly to a 64 GB internal memory card. GoPro video timestamp was calibrated to GPS time using a Trimble unit. A whiteboard plaque was photographed for each site to record the project, date, time, and grab number (Figure 3.5). All videos, photos, and GPS information were backed up daily to two mirrored external hard drives. All underwater videos were reviewed by benthic experts and the presence (1) and absence (0) of each substrate, geoform and biotic cover type listed in the classification scheme were annotated. Multiple substrate, geoform and biotic cover types were present at each site (see Chapter 7 Ground Validation Video Analysis).

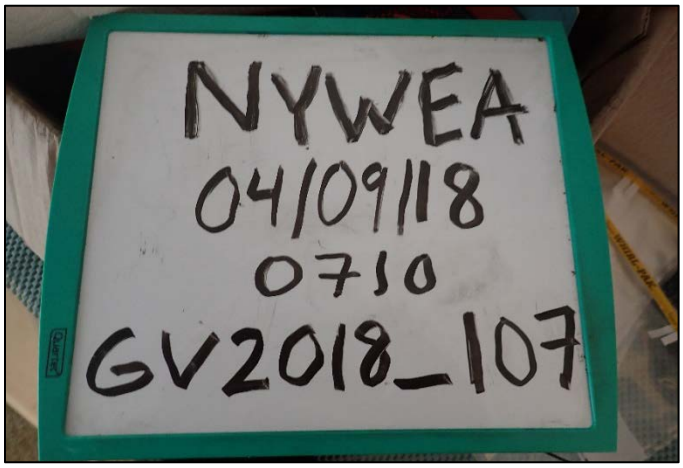


Figure 3.5. Example of the plaque that is photographed before collection of each sample.

3.3 Processing

Upon recovery of the sediment grab sampler onboard the vessel, the MVV was placed on a wooden frame and a surface photo of the sediment grab was taken by opening the top hatch of the sampler and taking a photo of the top of the sample (Figure 3.6). After the photo was taken, the grab sampler jaws were opened and the sample was released into the sediment collection bin. If the sample was washed out, partially filled, or had debris caught in the sampler jaws, the sample would be discarded and the MVV redeployed. If the sample was deemed acceptable (level and intact sediment), the entire sample was then emptied into a sediment collection bin and homogenized using a modified power drill attachment (Figure 3.7). A brief description about the sediment texture and composition was recorded in the Trimble unit. Each sample was

characterized for: grain size, benthos (mollusk, shell, anthropogenic, algae, wood, vegetation), presence/absence of surface oxidation, stiffness (very soft, soft, stiff, or very stiff), color, stratification (coarse to fine or fine to coarse) and the presence/absence of hydrogen sulphide odor. The grain size (clay, silt, sand, granule, pebble, or cobble) and benthos (mollusk, shell, algae, wood, vegetation or anthropogenic contents) was classified as: absent, rare, common and abundant. In the general comments section, particular characteristics of each sample was noted.

Figure 3.6. Surface photo of a sediment grab prior to releasing the sample into a sediment collection bin for further processing.

Figure 3.7. NCCOS scientist opening the MVV sampler to release the sediment sample into the bin below.

The scientists estimated the average grain size of the sample in the field using the Coastal and Marine Ecological Classification Standard (CMECS) Anthropogenic Substrate Scale (Wentworth 1922; CMECS 2017). Upon completion of the visual description, an approximately 250 g sample was taken from the sediment grab, distributed into individual labelled whirl-packs, and stored in a large cooler. After the sampling mission was completed, the sediment samples were shipped to TDI-Brooks International for phi (φ) and grain size analysis (see Chapter 9 Sediment Grab Analysis).

Chapter 4. Seafloor Sonar Processing

Raw acoustic multibeam survey data must undergo extensive post-processing to correct the data for all departures from true depths attributed to the method of sounding or faults in the measuring apparatus. Data correction and cleaning processes are necessary to ensure that the resulting products meet the desired horizontal and vertical accuracy requirements before these outputs are incorporated into subsequent benthic habitat modeling efforts. During data post-processing, all positions and soundings are corrected for sensor and vessel offsets, biases, dynamic attitude, dynamic draft, navigation, sound velocity, tidal variations, and other vessel motion artifacts.

4.1 Processing Methods

For NOAA acoustic survey NF-17-09, the standard CARIS HIPS and SIPS CUBE processing workflow was instituted to process the Kongsberg EM710 and the Reson 7125 multibeam bathymetry (Figure 4.1). The multibeam data were vertically referenced to a Tidal Constituents and Residual Interpolation (TCARI) grid network using seven different mid-Atlantic tidal stations, and was updated with the final observed tides. The real-time GPS navigation and vessel motion from the POS/MV antennas were applied to the processed CARIS multibeam files using custom Smooth Best Estimate of Trajectory (SBETS) for horizontal and vertical accuracy. Sound velocity profiles were applied to the multibeam data to correct for sound speed refraction in the water column. All of these corrections were merged together and TPU was calculated and applied to generate a CUBE surface. Finally, the remaining artifacts and sonar noise were filtered and manually cleaned and the final bathymetric (Figure 4.2) and uncertainty surfaces were exported as a GeoTIFF to IHO to resolution standards by depth.

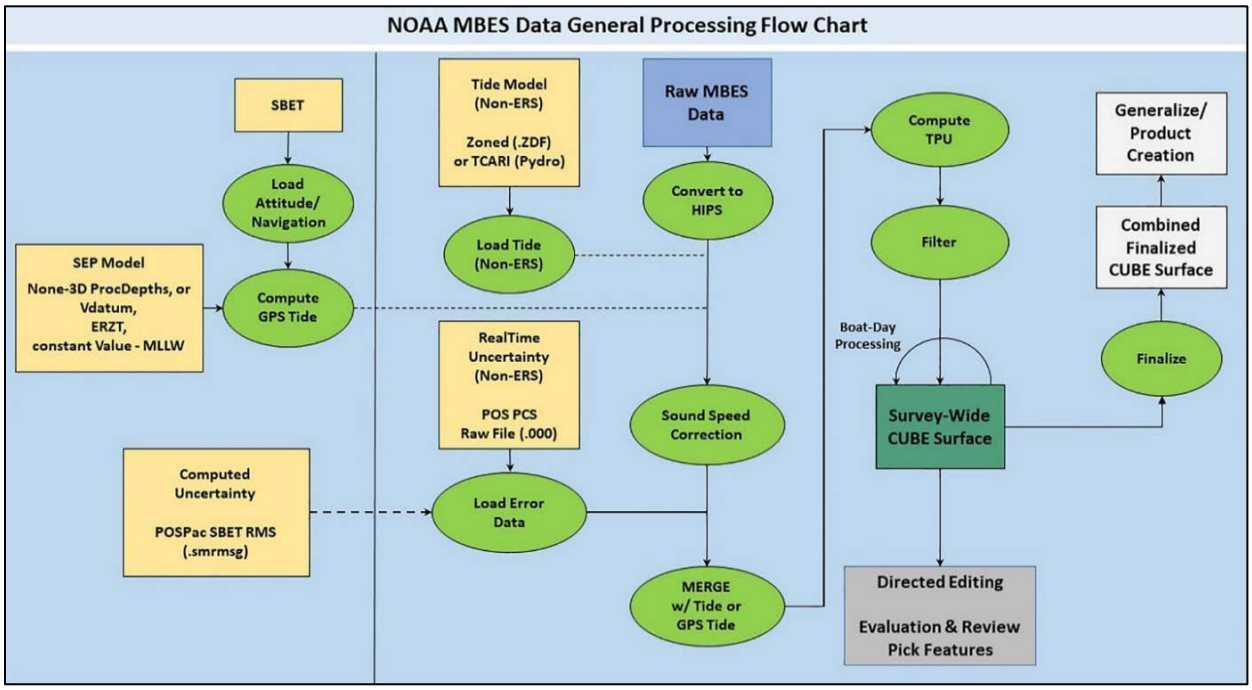


Figure 4.1. Diagram illustrating the general processing workflow from acquisition to delivery.

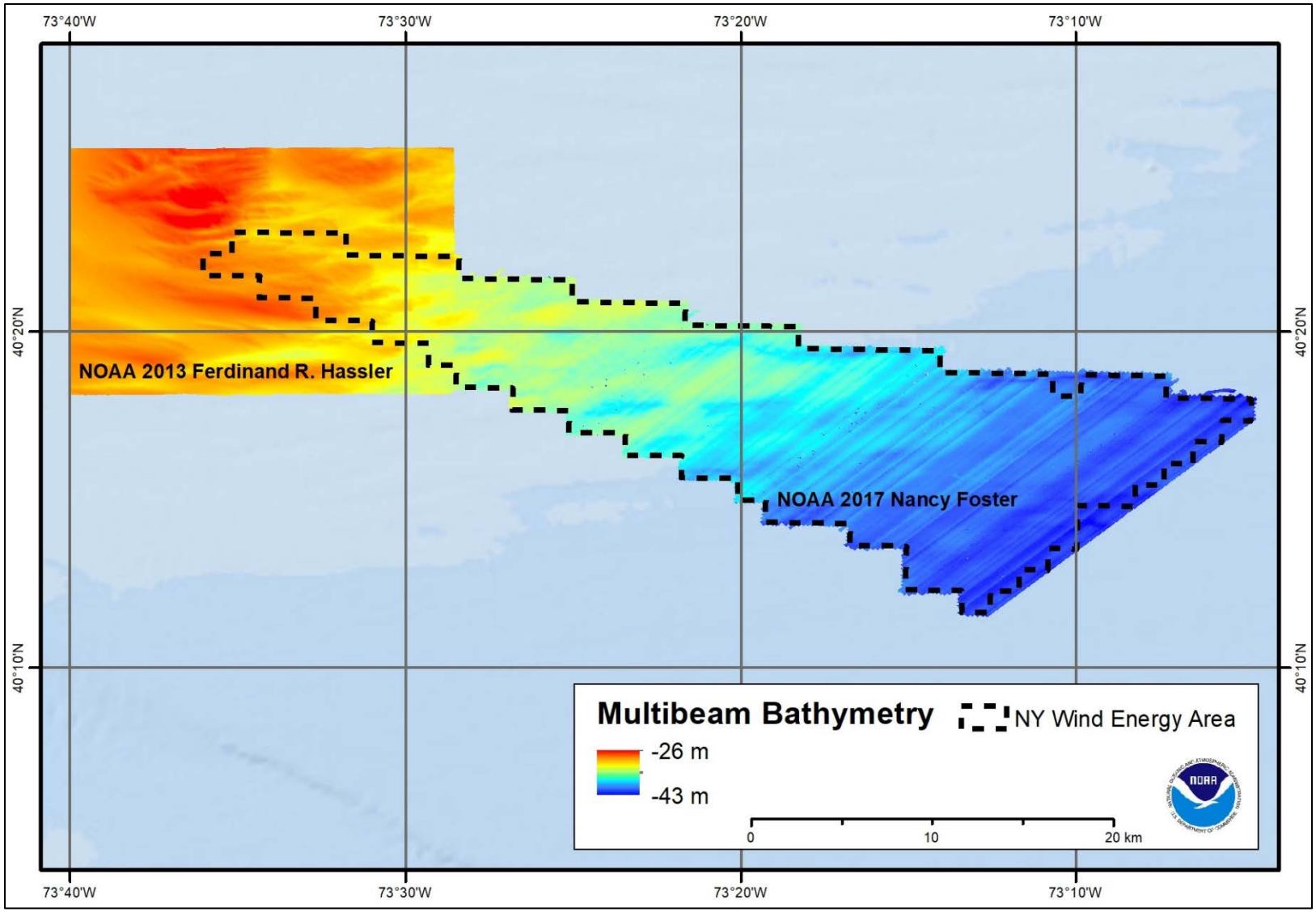


Figure 4.2. Merged multibeam bathymetry coverage in the NYWEA from NOAA Ships *Ferdinand R. Hassler* 2013 and *Nancy Foster* 2017.

For the NF-18-07 survey, the Kongsberg EM2040 and EM710 data were processed using QPS called Qimera bathymetric processing software, which utilizes a similar workflow as CARIS HIPS and SIPS. SBETs were also created using the POSPac Mobile Mapping Suite (MMS) software and applied to all data to correct for any vertical and horizontal offsets. A Vdatum separation model (SEP) with GNSS Vertical Referencing Method was applied to vertically transform the data to the ellipsoid (NAD83). However, due to GPS system anomalies during bad weather, transformation of the data to the ellipsoid using this method was unsuccessful for several lines. This resulted in several lines with major vertical offsets from the rest of the ellipsoidally referenced surface. Therefore, these lines were subsequently processed in CARIS using the same workflow utilized in NF-17-09. Tides were applied using a TCARI grid and the data was vertically referenced to Mean Low Low Water (MLLW) using Vdatum. Sound Velocity was applied and TPU calculated to generate a CUBE surface. Sonar noise artifacts were filtered and the final bathymetric (Figures 4.3) and uncertainty surfaces were exported as GeoTIFFs to IHO to resolution standards by depth and the HSAA (Figure 4.4).

Due to poor weather conditions during NF-17-09 (as mentioned in Chapter 2), the original bathymetry data was somewhat degraded for portions of the survey area, with artifacts occurring especially in the eastern portion of the NYWEA. During the NF-18-07 cruise, bathymetry was collected in the NYWEA to replace the poor quality data from NF-17-09. Unfortunately, the resurveyed data from NF-18-07 was not significantly better due to bad weather, as shown in a side-by-side comparison of the bathymetry data from the two surveys in Figure 4.5. Therefore, the NF-18-07 bathymetry data was not used to replace the NF-17-09 data in the morphometric analysis or the predictive modelling for the benthic habitat map.

4.2 Data Corrections

The mapping team resorted to using spatial analysis techniques instead of resurveyed data to improve the quality of the bathymetry surfaces, which would subsequently be used for deriving the morphometric surfaces (see Chapter 6). The uncertainty layer processed from the NF-17-09 CUBE surface was used to filter and extract vessel motion remove values >1.0 standard deviation, effectively filtering out the artifacts generated from vessel motion. A Matlab routine (inpaint_nans) was then used to fill value back in by interpolation (D'Errico 2006). The final interpolated surface was then exported as an 8 m GeoTIFF in ArcGIS and used as the source for the morphometric analysis and predictor for the benthic habitat modeling.

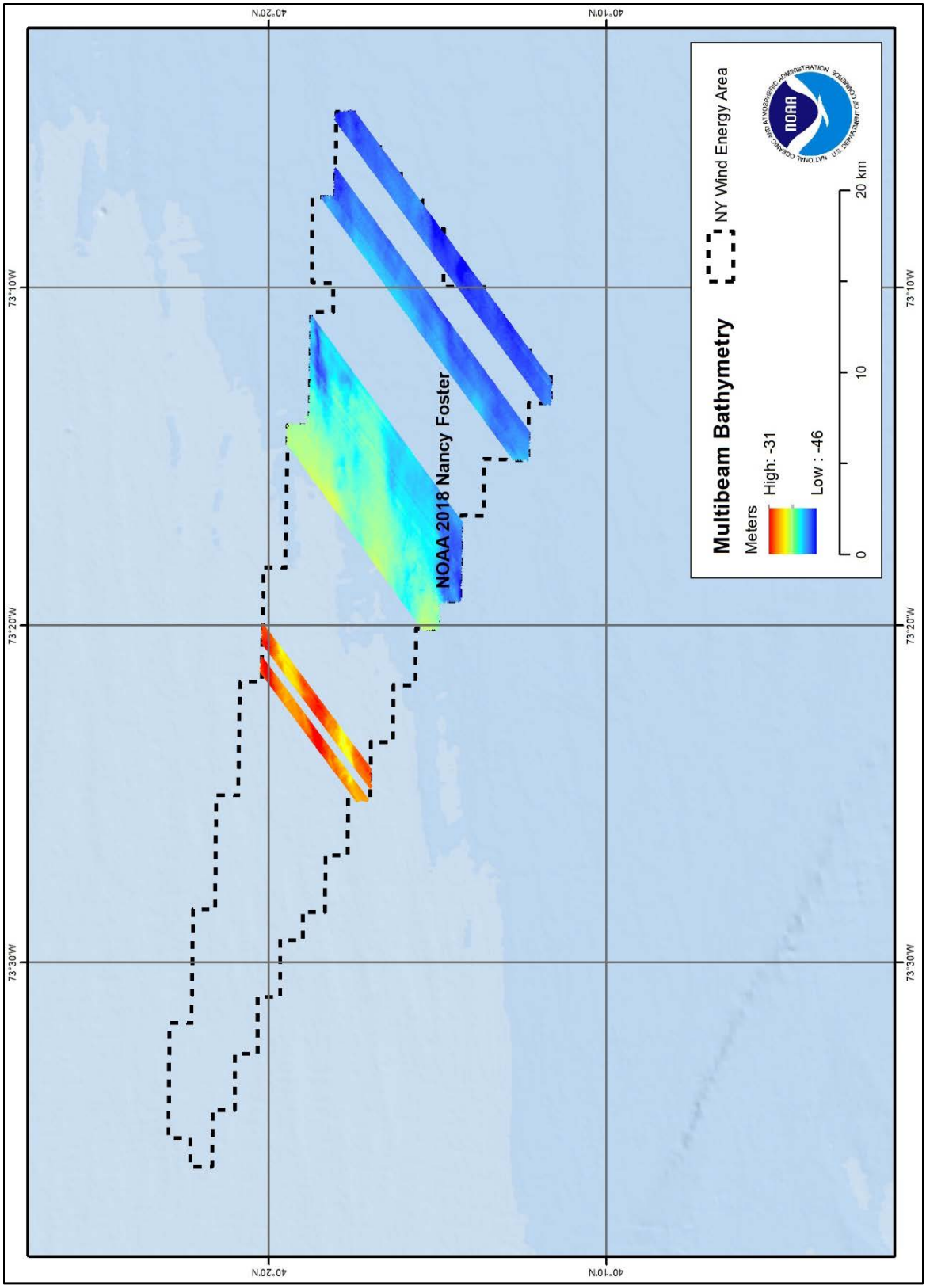


Figure 4.3. Multibeam bathymetry coverage in the NYWEA from the NOAA Ship *Nancy Foster* 2018.

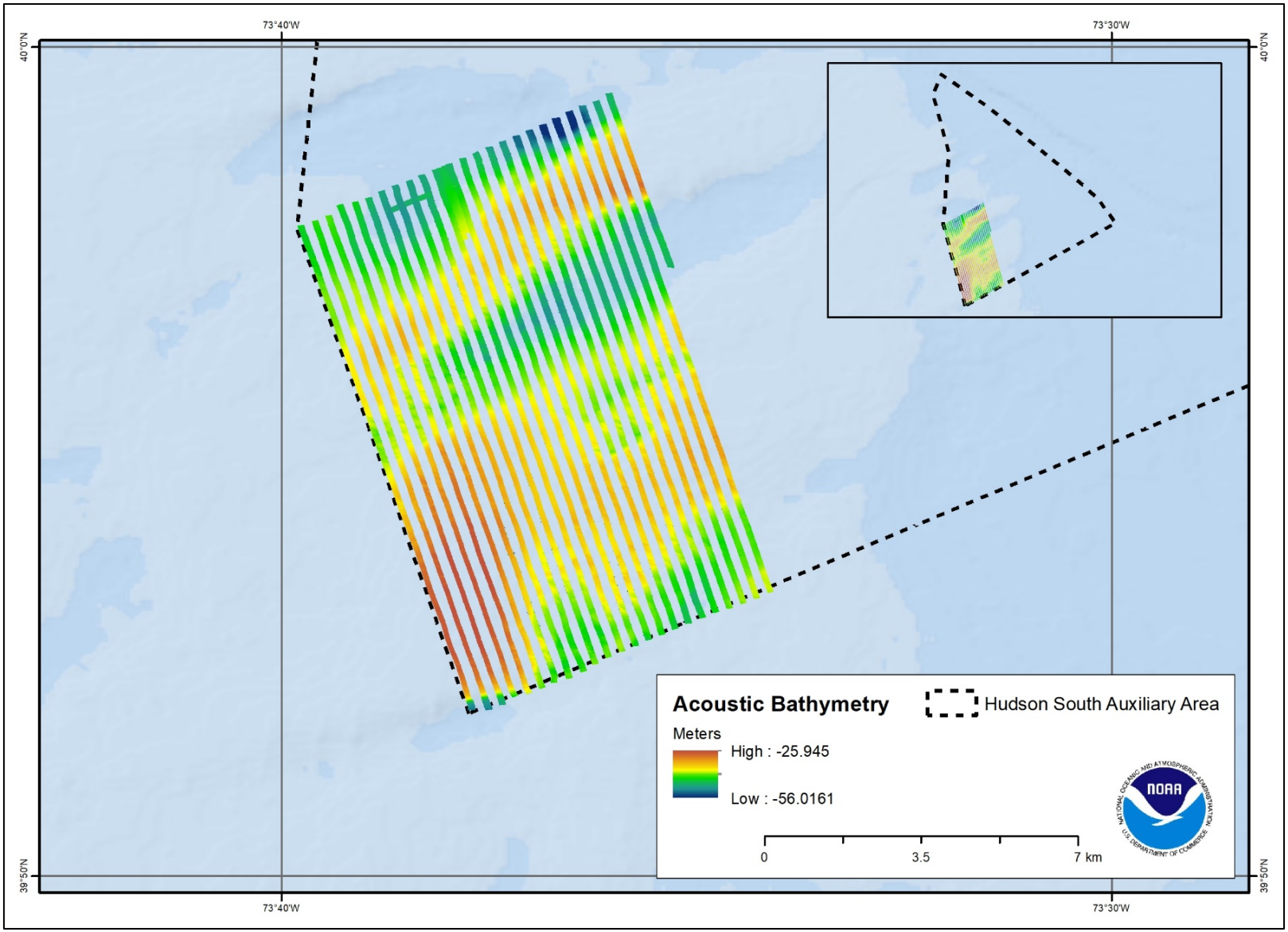


Figure 4.4. Multibeam bathymetry coverage in HSAA.

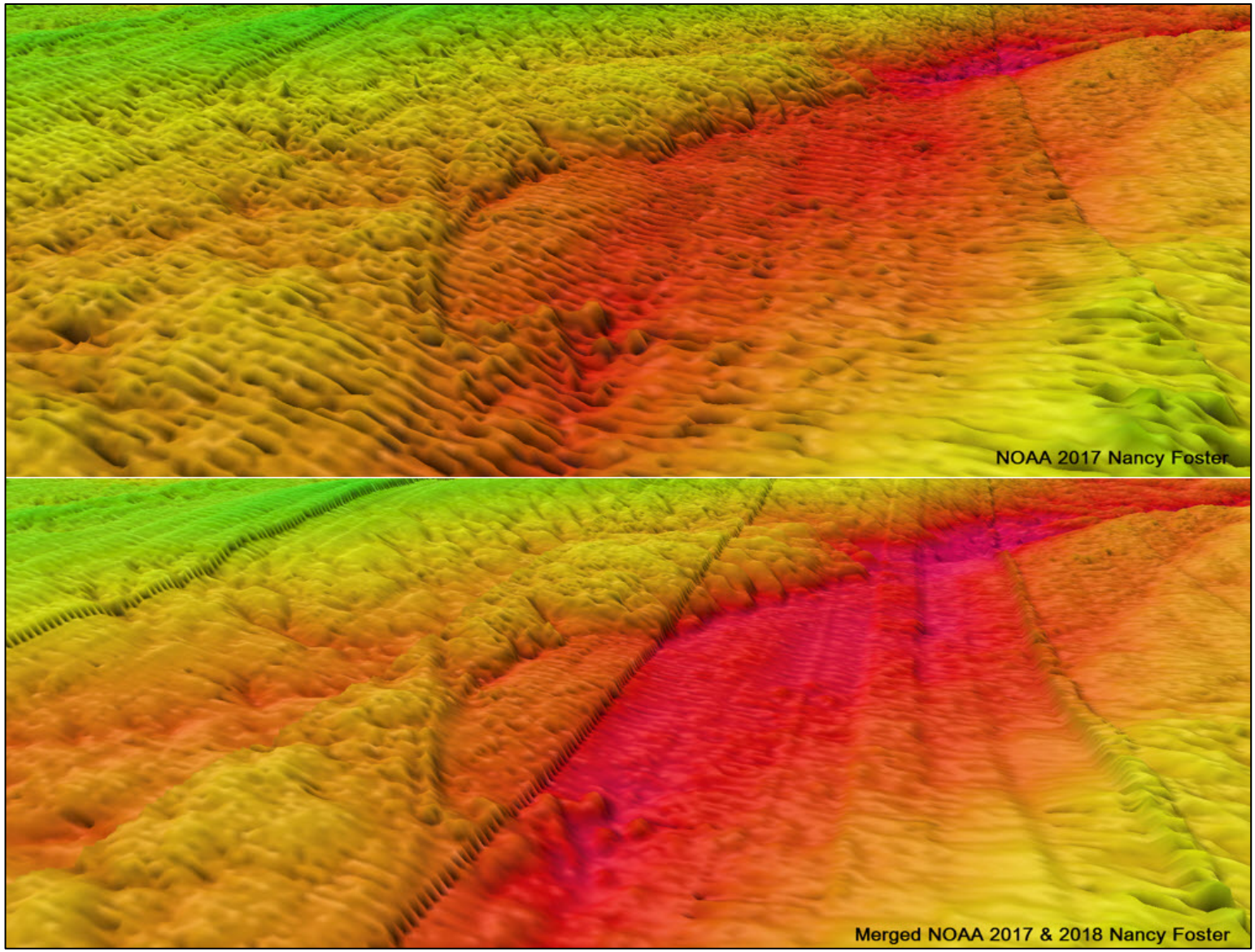


Figure 4.5. Comparison of bathymetric data from collected in 2017 (top) and the merged 2017 and 2018 (bottom) showing degradation of the bathymetry surface when the two surfaces are merged.

4.3 Backscatter Processing

Backscatter information, which can be used to detect seafloor sediment variation from fine to coarser grained sediments and hard bottom features, was recorded during NOAA's multibeam sonar surveys. Backscatter measures the acoustic intensity of sound transmitted by the multibeam pulse and reflected from the seafloor. Backscatter intensity is measured in decibels (dB), with low decibels being finer grained substrate and higher decibels indicating coarser grained substrate or hard bottom (Fonseca and Mayer 2007). This information is extremely useful for detecting sediment composition in relatively flat expanses of seafloor that may be otherwise undetected from bathymetry alone (Fonseca and Mayer 2007). The mapping team used the QPS Fledermaus Geocoder Toolbox software to process the backscatter information and create intensity mosaics. The mosaics were cleared of acoustic noise and refraction artifacts by merging the final cleaned bathymetry generic sonar files (GSF) exported from CARIS (NF-17-09) and Qimera (NF-18-07). The final Reson 7125, EM2040, and EM710 backscatter mosaics were exported as GeoTIFFs with the same IHO depth to resolution standards as the bathymetry surfaces for the NYWEA (Figures 4.6 and 4.7) and the HSAA (Figure 4.8).

The backscatter data from the NF-18-07 survey had improved quality, sharper definition, and finer details than the NF-17-09 survey. The vertical offsets and GPS anomalies did not have the same effect on the intensity measurements as with the previously described bathymetry measurements. However, the NF-18-07 data was collected using a more contemporary multibeam system with different acoustic frequencies, and therefore the dB intensity measurements from the 2018 cruise did not sufficiently align with measurements from the 2017 acquisition. As such, the NF-18-07 backscatter data was also not used to replace the NF-17-09 survey for the morphometric analysis or the predictive modeling.

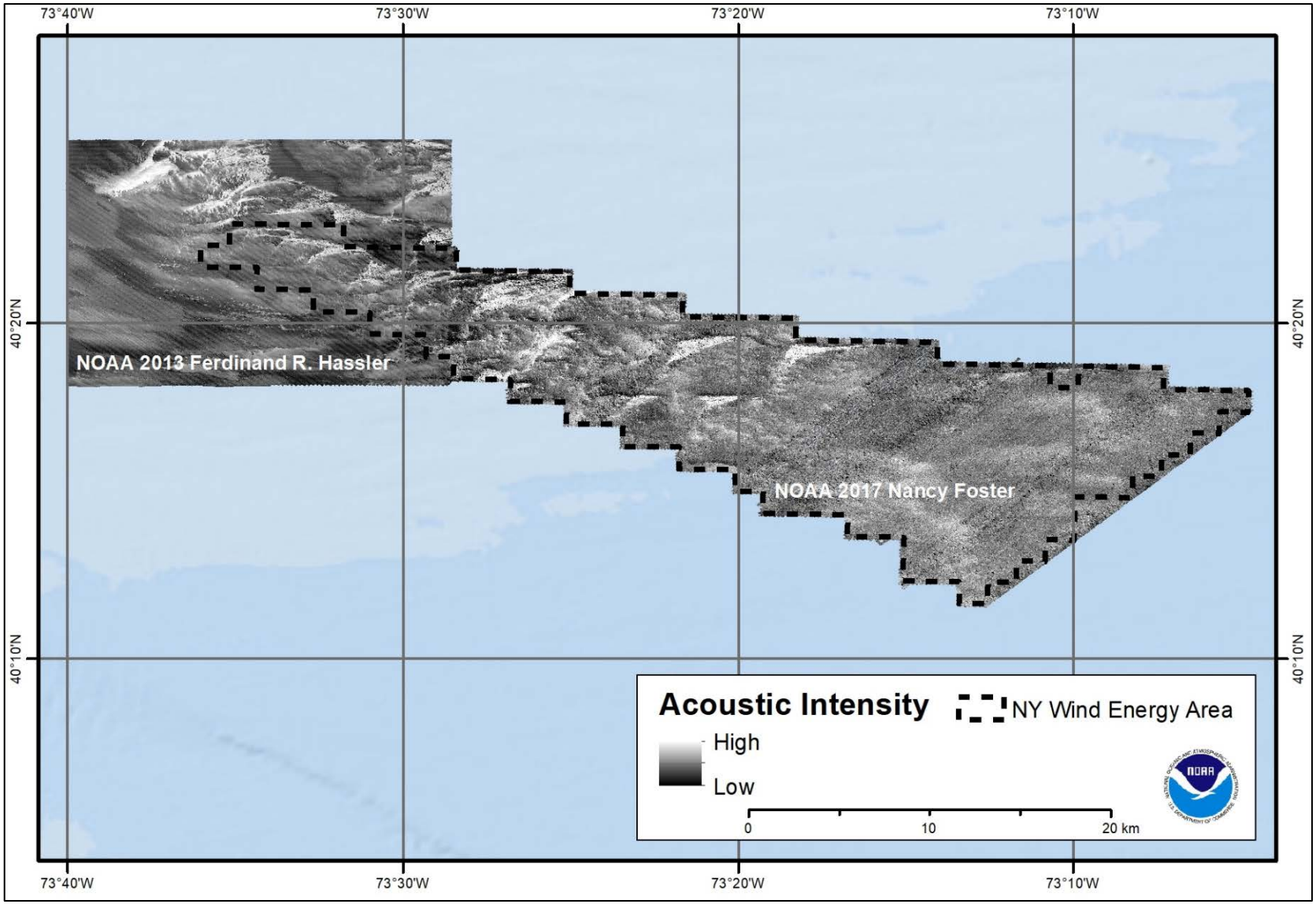


Figure 4.6. Merged multibeam acoustic intensity coverage in the NYWEA from the NOAA Ships *Ferdinand R. Hassler* 2013 and *Nancy Foster* 2017.

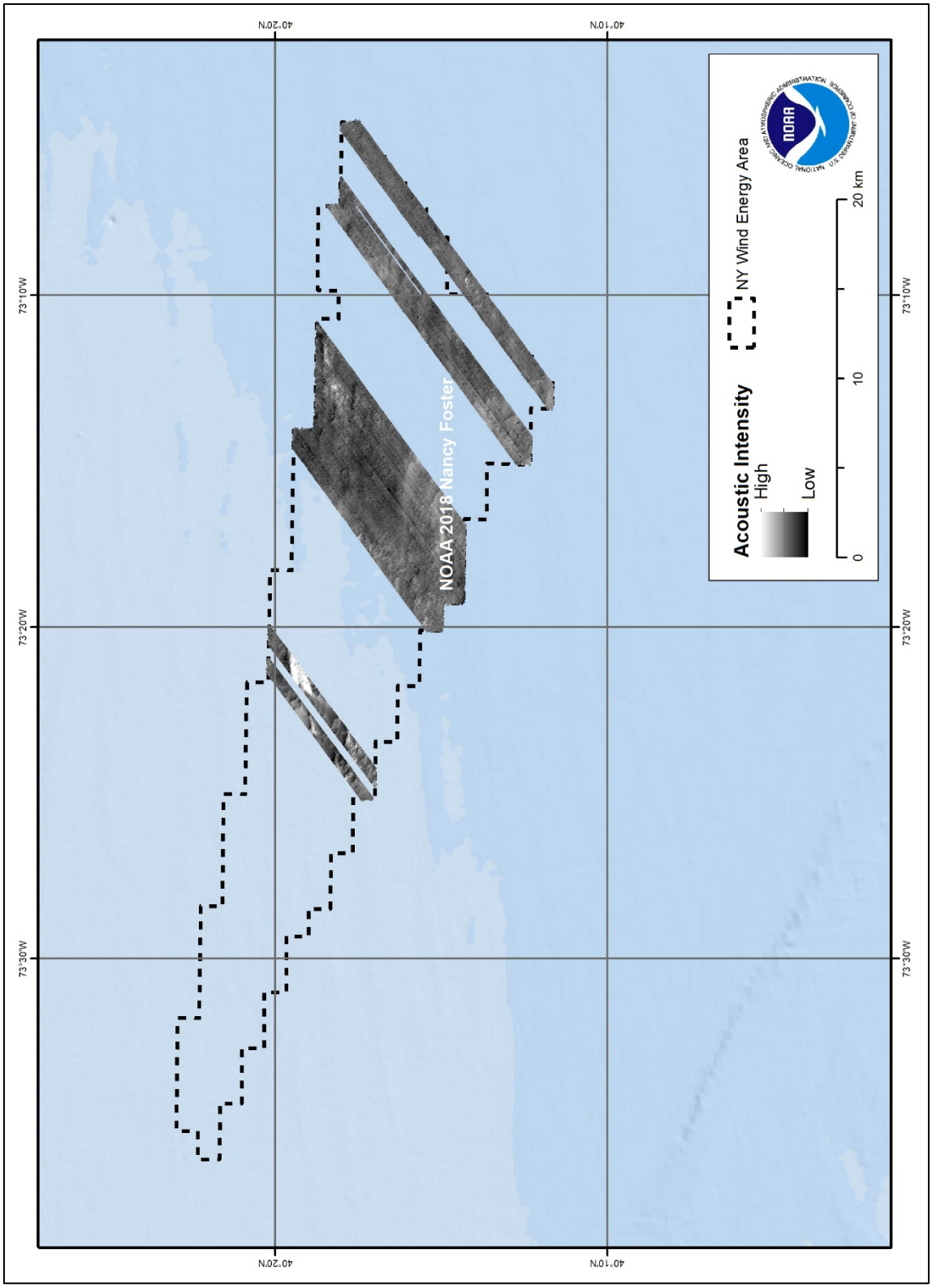


Figure 4.7. Backscatter coverage in the NYWEA from the NOAA Ship *Nancy Foster* 2018.

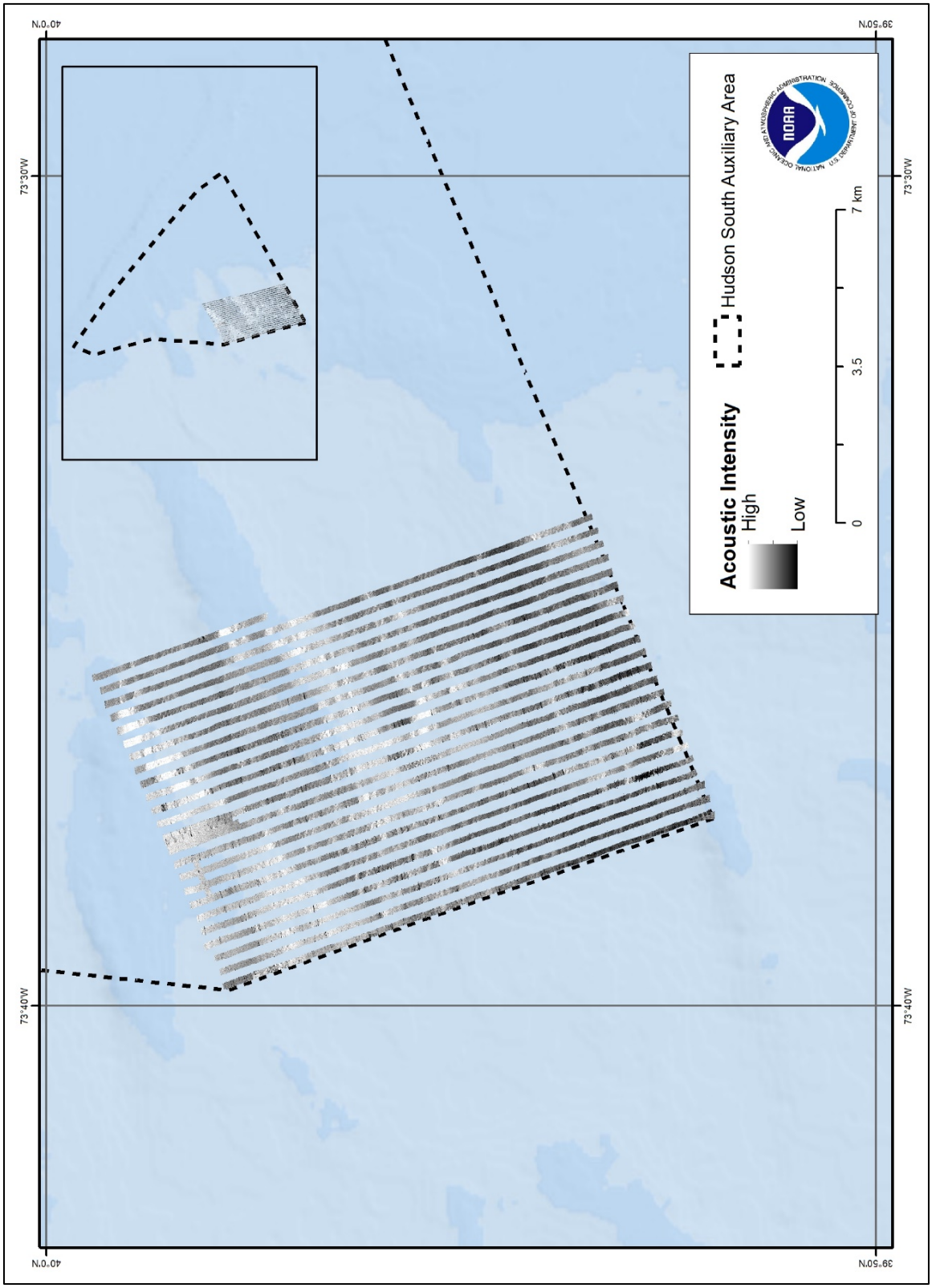


Figure 4.8. NOAA Ship *Nancy Foster* 2018 multibeam acoustic intensity coverage in HSAA.

Chapter 5. Fishery Acoustic Processing

Fishery acoustic sonars or scientific SBES are used to collect data throughout the water column, from the seafloor to the water surface. The SBES emits rapid pulses of high-frequency sound that reflects or echoes off objects (or the seafloor) having differing density than the surrounding water. The fish swim bladder, a gas-filled organ that many fish use to regulate buoyancy, reflects the majority of the sound transmitted by the SBES transducer. The intensity of the reflected sound (target strength, TS, dB ref 1m) is proportional to the size of the swim bladder resulting in an echo positively correlated to fish size. When fish are in close proximity, such as in schools or aggregations, it is not possible to discern individual fish or to characterize the target strength. In this case, the volumetric or area integration of the reflected sound from the school provides an index of its density.

5.1 Processing Methods

The 120 kHz SBES data was processed using Echoview software (version 8.0, Echoview Pty Ltd, Hobart, Tasmania). The data were heave corrected to remove vertical motion caused by swell and waves. The seafloor was delineated and data cleaned (Figure 5.1) to remove interference from other ship systems and air bubbles prior to processing the water column data for fishes. Backscatter returns from plankton and other non-fish targets were excluded using a volume backscattering strength, Sv, threshold of -60 dB. Two different methods, acoustic signal of individual fish versus fish schools, were used for deriving the biological density within the NYWEA survey area.

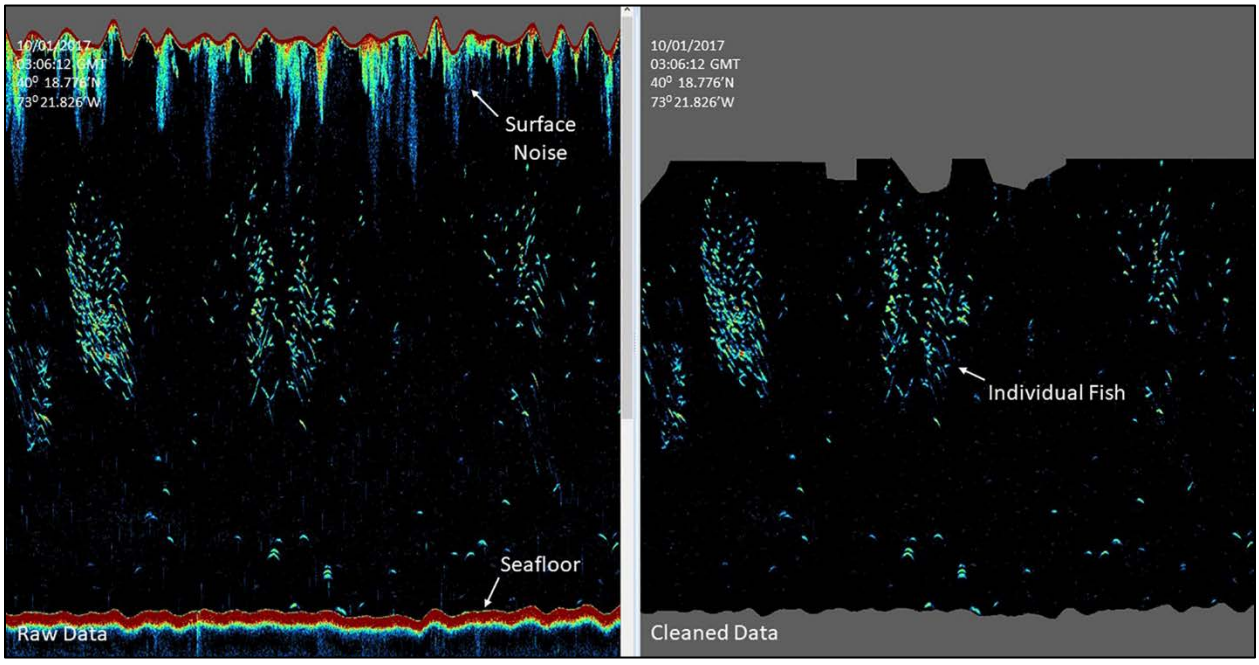


Figure 5.1. Example of raw data (left) with noise from surface bubbles, seafloor reflection and non-fish targets present compared to clean data (right) ready for processing.

5.2 Individual Fish Processing

Using the cleaned SBES data, individual fish were identified using a single target detection and tracking algorithm. The speed of the vessel and rate of ping transmissions resulted in multiple and sequential returns from individual fish. The sequential returns from the fish are referred to as single targets. The split-beam transducer detects the range and horizontal position of the target within the beam at each ping using a phase-differential array. The technique for identifying single targets in SBES data relied upon the data processor's ability to characterize the shape of the return pulse and to specify an acceptable setting that resulted in quality single targets. A 2D single target-tracking algorithm was used to detect sequential echoes generated from individual fish greater than 6 cm total length (Figure 5.2). The 2D algorithm used range and time patterns from the single targets to search for systematic movements of a fish moving through space. The resulting fish identified by the tracking algorithm were stored in a database with a geographic position determined by the ship's GPS, and corrected for relative position of fish within the acoustic beam, depth below the sea surface, and a mean target strength (TS, in dB).

Individual fish and schools data were exported from Echoview in a CSV format. Open-source statistical programming language R (version 3.4.0, R Core Team 2017) was used to summarize and perform all calculations. The acoustic target strength of all single targets within a detected fish was used to calculate fish size (total length, TL) in centimeters using a generalized acoustic size to fish length relationship derived from the:

$$TL = 10^{(TS+64.0035)/19.2}$$

where TS is target strength measured in dB, TL is calculated length in cm (Love 1977). The equation above fits closely with observation of broad classes of fish (Love 1977; Johnston et al. 2006). Individual fish targets were counted and binned into 100 m intervals along survey transects. The density calculation accounted for the increased detection of individual fish as the acoustic beam footprint increases by depth, standardizing the beam width to a 1m swath using the following equation:

$$Cw = 2 \times range \times \tan(0.5BA)$$

where Cw is the weighted count of an individual fish accounting for detection in an increasing beam swath with increasing range, and the tangent of the half beam angle (BA = 7°). Weighted counts were summed for each 100 m interval producing a density with the units fish 100 m^2.

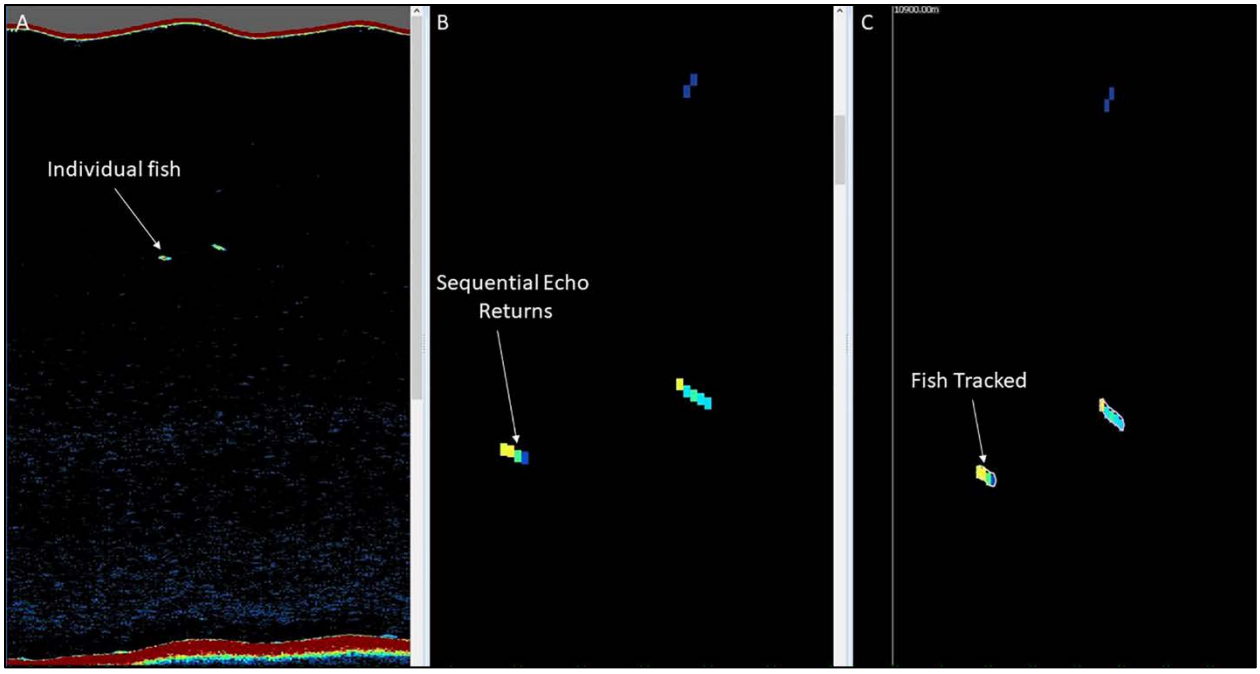


Figure 5.2. Single target detection algorithm processing steps. A) Raw backscatter data (Sv) of an individual fish; B) four sequential echo returns of the same individual fish; C) sequential echo returns identified as an individual fish track with a mean decibel measurement.

5.3 Fish School Processing

When fish are aggregated in acoustic schools (e.g., less than 20 cm vertical spacing between individuals) (Figure 5.3), individual targets cannot be discerned or enumerated. Echo integration was used to provide a measure of volume backscatter as an indicator of fish density. Fish schools were delineated using a visual edge detection algorithm that identified the acoustic backscatter reflected from a single aggregation versus a background signal. The school detection algorithm identified points on the edge of a school that the processor had determined meets the relevant criteria and linked all points together to create a school candidate. The backscatter of the schools was integrated by a distance interval (100 m^2) and expressed as Area Backscattering Coefficient (ABC), which is the area acoustic density of a standardized area to 1 m length and the water column depth as height. To calculate the school fish density, the school acoustic backscatter was scaled to the size of the average fish in the school. When there were no identifiable individual fish within the school for the density calculation, the mean TS of single targets in the outer margins of the school was used as a proxy for the average size of the individual fish inside the school.

Acoustic backscatter (ABC, m^2/m^2) from fish schools were divided by the average backscatter of an individual fish (TS, dB) in order to obtain a density that has units of fish m^2. Density values were then multiplied by 100 to achieve similar magnitude of values as in the density estimates of area swept for individual fish (number of fish per 100 m^{-2}).

All binned data were summed to achieve combined fish densities for all fish and fish schools and converted to ArcGIS point shapefiles (version 10.5, ESRI 2016). The resulting 100 m^2 point

shapefile used a centroid for each interval (100 m^2) surveyed as the positional location for combined individual fish and fish school densities. Spatial Analyst tools in ESRI ArcGIS (version 10.5) were used to interpolate the point shapefile into a raster, using the point to raster function with a 150 m cell size to encompass the complete survey area.

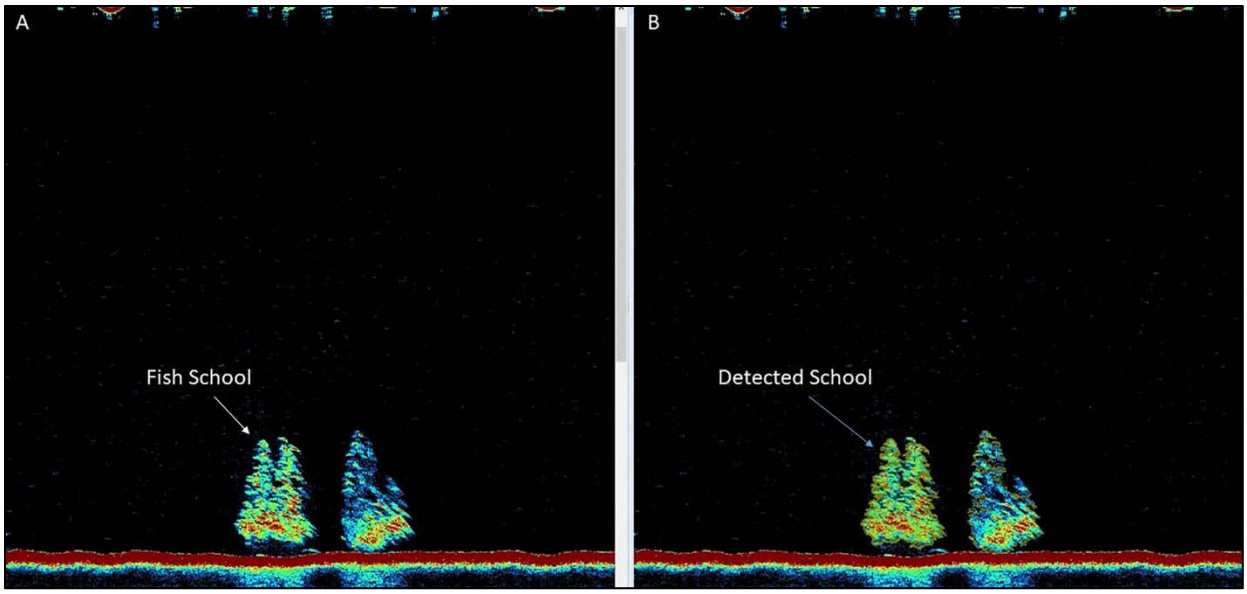


Figure 5.3. A) Example of a large fish aggregation. B) Same Large fish aggregation detected with tracking algorithm with a backscattering measurement.

Chapter 6. Seafloor Morphometric Analysis

Seafloor geomorphology is an important attribute of marine habitats that is inextricably linked with biogeography, biodiversity, and the distribution of marine life (Harris and Baker 2011). Acoustic sonar data can provide a highly resolved, accurate measurement of water depth (i.e., bathymetry). In addition to information on depth, high-resolution bathymetric data can be used to model changes in the shape of the seafloor (i.e., topography) and to highlight variation in seascape (terrain) at a range of spatial scales. In turn, detailed topography, terrain, and other highly resolved seafloor characteristics often correlate with spatial variation in physical, geological, and ecological processes, as well as biological habitats (Zevenbergen and Thorne 1987; Moore et al. 1991). As such, characterizing the shape and identifying unique features of the seafloor can help identify habitat or biodiversity patches that could be affected by human activities that disturb the seafloor. Furthermore, backscatter, the sound intensity reflected off the seafloor and detected by acoustic sonar can be interpreted to describe the sediment composition of the seafloor.

Morphometric products derived from acoustic sonar bathymetry provide a valuable means of synoptically characterizing the composition, roughness, shape, and texture of the seafloor to map and identify the distribution of benthic habitats. In this chapter, 22 morphometrics are evaluated that describe the composition, roughness, shape, and texture of the seafloor and are useful for mapping and characterizing the spatial distribution of distinct benthic features in the NYWEA study area that could be affected by wind-energy development. Similar to previous studies, high-resolution bathymetry data is used to analyze changes in the seafloor topography at varying spatial scales (Costa et al. 2009; Costa and Battista 2013; Costa et al. 2013; Costa et al. 2014). Previous research has shown the utility of these metrics 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).

Bathymetry, backscatter, and the derived morphometric data were employed as predictors in the Boosted Regression Trees (BRT) model to create a detailed habitat map within the NYWEA (see Chapter 11).

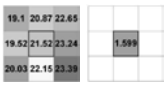
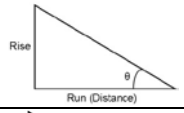
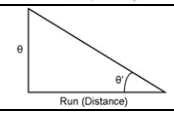
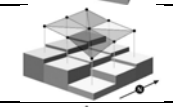
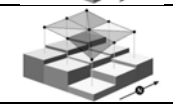
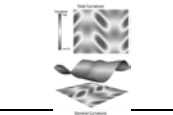
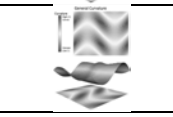
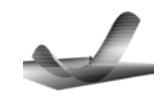
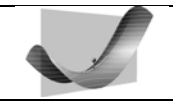
6.1 Methods

Two approaches were used to generate a suite of 22 data layers depicting various measures of seafloor geomorphology from the 8 x 8 m resolution bathymetry and backscatter data collected by acoustic sonar for the NYWEA (see Chapter 4 for description of acoustic sonar data). First, focal statistics were applied to the final 8 m resolution bathymetry data (described in Chapter 4, Section 4.1) to produce 13 morphometric surfaces (Table 6.1). Focal analyses were performed within a neighborhood operation that computed an output raster in which the value for each output cell was derived from the closest eight adjacent cells. Second, nine additional morphometric surfaces were calculated with Bathymetry and Reflectivity-based Estimator for Seafloor Segmentation (BRESS) software (Masetti et al. 2018). The BRESS methodology applied a “field of view” with varying extents based on polygon size to calculate each metric of seafloor geomorphology.

6.1.1 Seafloor Morphometrics Derived Using Focal Statistics

Data layers depicting thirteen unique measures of seafloor geomorphology were calculated from the bathymetry data layer using focal statistics (Table 6.1). In this approach, values for each grid cell were calculated from the values of the eight surrounding cells in a 3 x 3 grid cell (576 m^2) focal neighborhood. Calculations were performed using the 'raster' package in R (Hijmans 2016; R Core Team 2016).

Table 6.1. Descriptions of the 13 measures of seafloor geomorphology derived from focal statistics. All calculations were performed using the 'raster' package in R (Hijmans 2016; R Core Team 2016).

Morphometric Surface		Unit	Description	Reference
Depth Standard Deviation		Meters	Measure of variation in depth	Costa et al. (2013, 2018)
Slope		Degrees	Gradient in the direction of maximum slope	Horn (1981); Jenness (2013)
Slope of the slope		Degrees of degrees	Gradient in the direction of maximum slope of slope	Horn (1981); Jenness (2013)
Aspect (East-West)		Unitless	The sine of the direction of maximum slope	Horn (1981); Jenness (2013)
Aspect (North-South)		Unitless	The cosine of the direction of maximum slope	Horn (1981); Jenness (2013)
Rugosity		Unitless	Ratio of surface area to horizontal planar area	Horn (1981); Jenness (2013)
Rugosity (Arc-chord Method)		Unitless	Ratio of surface area to planar area of best-fitted slope	Du Preez (2015); Jenness (2013)
Total Curvature		Radians/100m^2	Measure of the roughness or ruggedness of the seafloor, with higher values indicating that an area is more rugged; values always ≥0	Evans (1979); Jenness (2013)
General Curvature		Radians/100m – = concave + = convex	Measure of the extent to which the seafloor is convex (e.g., ridges) or concave (e.g., depressions)	Zevenbergen and Thorne (1987); Jenness (2013)
Plan Curvature		Radians/100m – = concave + = convex	Curvature of the surface perpendicular to the slope direction	Evans (1979); Zevenbergen and Thorne (1987); Jenness (2013)
Profile Curvature		Radians/100m – = convex + = concave	The curvature along the line of maximum slope, where the surface is intersected with the plane formed by the aspect and Z-axis	Zevenbergen and Thorne (1987); Jenness (2013)

Note, two methods of plan and profile curvature metrics were analyzed: (1) methods described by Evans (1979), hereafter Evans method, and (2) methods described by Zevenbergen and Thorne (1987), hereafter Zevenbergen and Thorne method. They are discussed with the other metrics in the next section.

6.1.2 Seafloor morphometrics derived from Bathymetry and Reflectivity-based Estimator for Seafloor Segmentation (BRESS)

Data layers depicting nine additional measures of seafloor geomorphology were created using the (BRESS) software (Masetti et al. 2018).The software used the bathymetry and the acoustic backscatter data to identify seafloor segments, measure areas with consistent characteristics, and provide statistical layers that characterized the segments. The algorithms used in BRESS included an initial segmentation of the bathymetric surface into landforms (geomorphologic phonotype), that served as an archetype for a particular terrain morphology (Jasiewicz and Stepinski 2013). BRESS software was designed by Masetti et al. (2018) to mimic the positive aspects of the segmentation process as performed by a skilled analyst, while avoiding the inherent deficiencies of human subjectivity, processing time, and lack of reproducibility.

Initial segmentation of the study area in BRESS was based on characterizing landforms or patterns within the bathymetry values that were based on slope and depth and then grouping areas with consistent values into segments. Landforms types were assigned to groups of grid cells or kernels with similar values for depth and slope. Landform types were: pit, valley, foot slope, concave slope, slope, convex slope, shoulder, ridge, peak, or flat. The landforms employed in BRESS are defined in Jasiewicz and Stepinski (2013), and based on bathymetry within a cell and its surrounding neighbors. Adjacent kernels with similar landforms were aggregated further to create larger polygons of varying sizes that represented areas of consistent topography.

The geomorphon approach used by BRESS employs a variable-scale solution (10-50 m search radius). In this case, each grid cell is considered in relation to its neighbors using several different search radii. The resulting surface is a contiguous set of segments of varying sizes, where morphology is consistent within each segment. The parameters used in the computation of the geomorphon were: search inner radius = 10 m, search outer radius = 50 m, and flatness angle = 0.2 degree.

Table 6.2. Descriptions of the measures of seafloor geomorphology derived using the Bathymetry and Reflectivity-based Estimator for Seafloor Segmentation (BRESS) software (Jasiewicz and Stepinski 2013; Masetti et al. 2018).

Topographic Surface	Description	Unit
Landforms	Classified bathymetry based on landform type, calculates pattern-based statistics, and creates area kernels (connected grid nodes with the same landform type).	Number of landform types
Area Ratio	The ratio between the area of the segment and its maximum possible extension (based on the outer search radius).	None
Average Azimuth	The average orientation of the segment.	Degrees
Average Height	The average height of the visible neighborhood.	Meters
Elongation Ratio	The ratio between the maximum and the minimum dimensions of the segment.	None
Height Range	The height range of the visible neighborhood.	Meters
Height Variance	The height variance (calculated using the Average Height as mean value) of the visible neighborhood.	Meters
Maximum Delta	The maximum elevation delta (that is, the absolute value of the height) of the visible neighborhood.	Meters
Maximum Width	The maximum dimension (x- vs. y-direction) of the segment.	Meters

6.2 Results

6.2.1 Seafloor Morphometrics Derived From Focal Statistics

Results from the focal statistics indicated minimal spatial variation in depth and topographic complexity within the study area (Table 6.3). While the thirteen focal statistics can each be valuable in the analysis of seafloor topography, there is a great deal of redundant information when considering the dataset as a whole. In our analysis, described in detail in Chapter 11, we found nine of these metrics to be valuable in delineating benthic habitat types. Those nine metrics are reported here.

Table 6.3. Summary descriptive univariate statistics for topographic surfaces derived from focal window analyses.

Topographic Surface	Unit	Min	Max	CV
Depth (Mean)	Meters	19.09	42.11	0.14
Depth (Standard Deviation)	Meters	0	0.79	0.65
Slope	Degrees	0	4.3	0.75
Slope of the slope	Degrees of degrees	0	11.5	0.87
Aspect (Eastness)	Unitless	0	1	
Aspect (Northness)	Unitless	0	1	
Total Curvature	Radians/100m^2	0	0.121	2.39
General Curvature	Radians/100m – = concave + = convex	-5.29	14.26	-4.45
Plan Curvature (Zevenbergen and Thorne)	Radians/100m – = concave + = convex	-2.87	7.04	-96.33
Profile Curvature (Evans)	Radians/100m – = convex + = concave	-2.59	2.29	-41.86

Depth

Bathymetric depth (Mean) ranged from 26-43 m. The study area generally was flat and relatively smooth, indicating minimal variance in physical attributes of the seafloor (Table 6.3).

Standard Deviation

Standard deviation (SD) of depth for the study area ranged from 0 to 0.79 m, indicating a very low variation in seafloor depth (Table 6.3, Figure 6.1). Additionally, 99.9% of the study area had a SD of less than 0.1 m. In general, SD of depth may reveal important patterns in certain areas or habitat types. Results indicated that large areas in the eastern portion of the study had minimal topographic variation.

Slope

Variation in seafloor slope is another important indicator of topographic complexity. The study area featured particularly low values for slope and could be generally regarded as a gently sloping to flat-planed landscape. The maximum slope for any focal area was 4.3 degrees (Table 6.3, Figure 6.2). However, less than one percent (0.13%) of the study area had a slope greater than 1.5 degrees (Figure 6.3). The overall mean slope was 0.283 degrees (SD = 0.212 degrees)

and there was little variation in depth (CV = 0.75, where CV is the coefficient of variation). These derived values suggest that minimal topographic complexity existed within the study area.

Slope of the Slope

Slope of slope often is used as an index of seafloor complexity, with higher values indicating increased structural complexity of the seafloor. Overall, the study area generally had low seafloor rugosity and low structural complexity (Figure 6.4). Slope of slope values ranged from 0 to 11.5 degrees with a mean value of 0.628 degrees and a coefficient of variance of 87% (Table 6.3). Focal cells with slope of slope >1 degrees either identified areas where seafloor complexity occurred or were artifacts of the bathymetry. Grid cells with slope of slope values <1 degree were particularly horizontal. Approximately 85% of the study area had slope of slope <1 degree.

Aspect

Aspect can influence physical variables such as strength and direction of currents. In some cases, aspect can also affect the distribution and composition of biological communities (Lecours et al. 2016). Within the study area, aspect might not be an important local factor affecting currents or community composition because slope was predominantly <1.5 degrees, and the seafloor geomorphology was predominantly supine (Table 6.3, Figure 6.5). The spatial patterns exhibited by observed north-south and east-west gradients seemed random and might not reflect broader-scale patterns of north-facing or east-facing slopes that would correspond with physical or ecological variations observed within the study area.

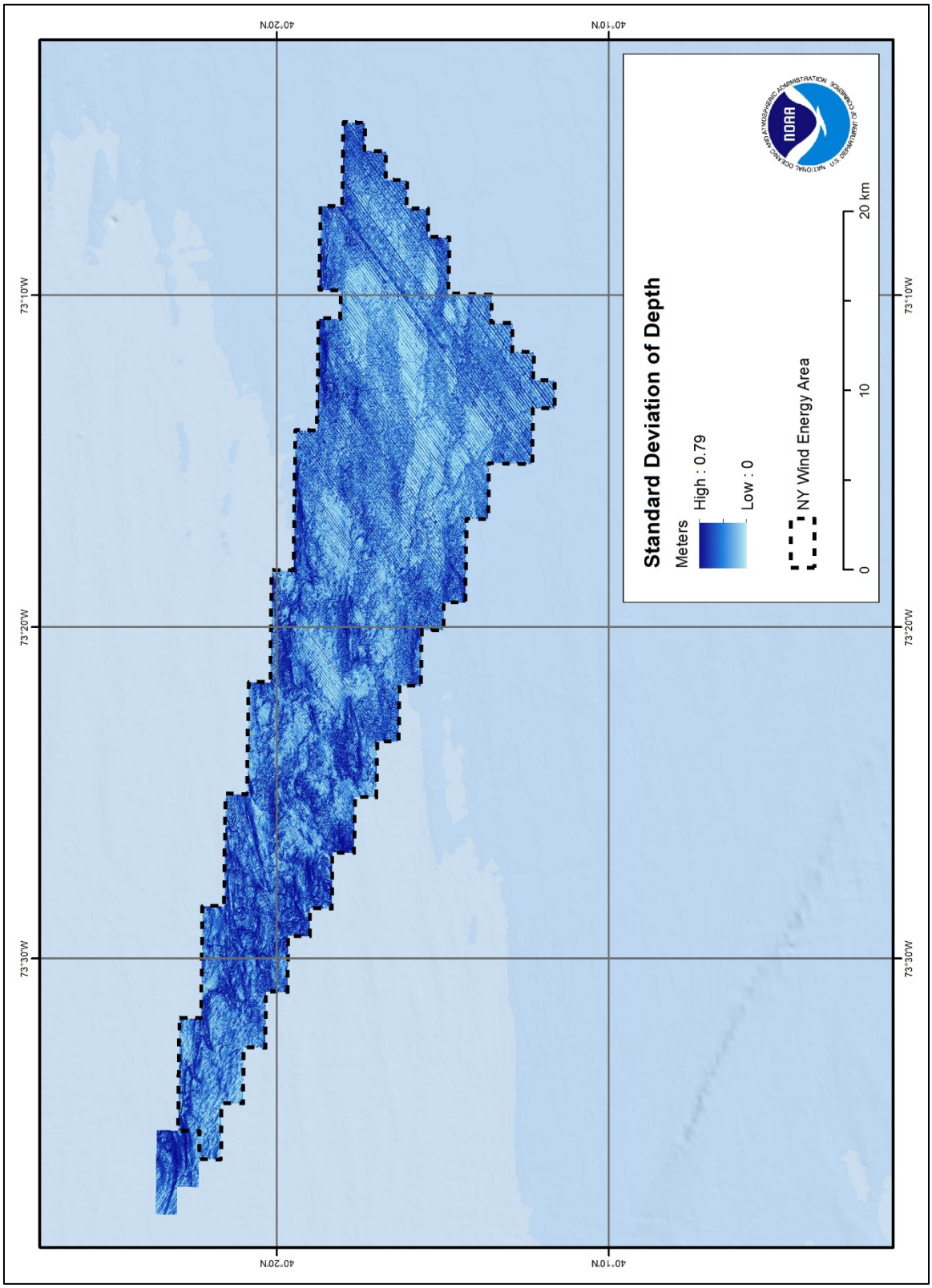


Figure 6.1. Map depicting standard deviation of depth derived from multibeam bathymetry for the NYWEA.

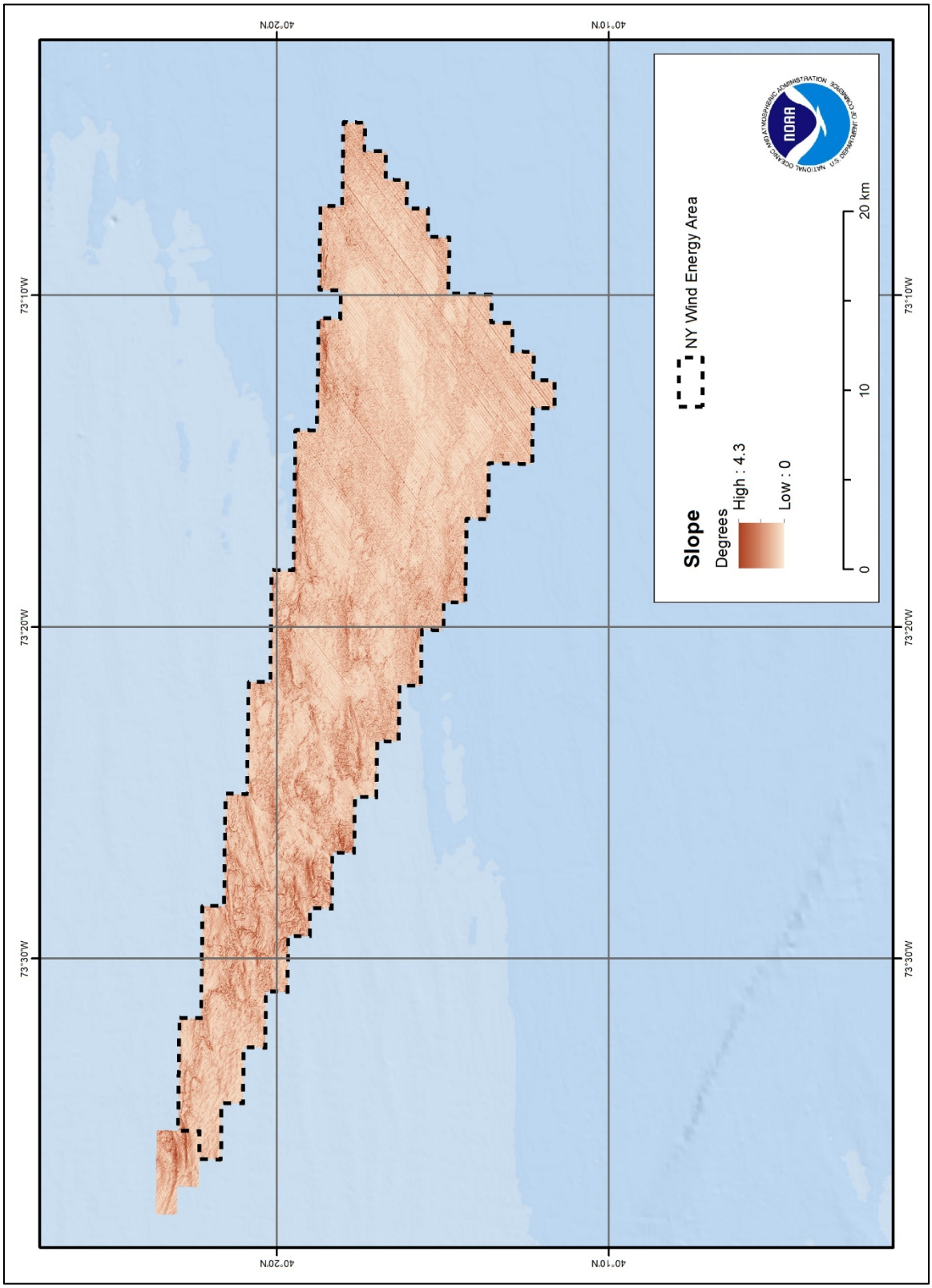


Figure 6.2. Map depicting slope derived from multibeam bathymetry for the NYWEA.

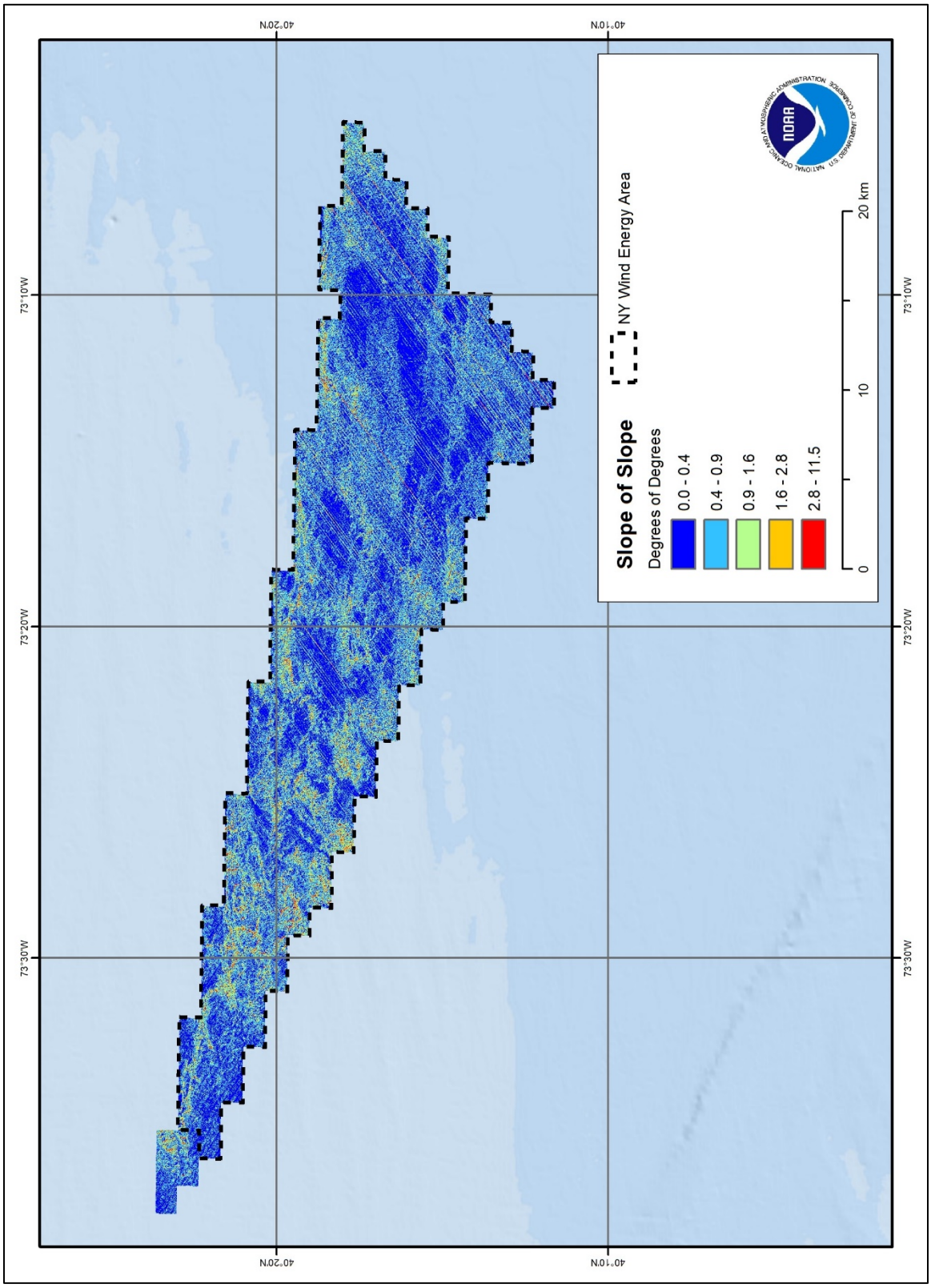


Figure 6.3. Map depicting slope of slope derived from multibeam bathymetry for the NYWEA.

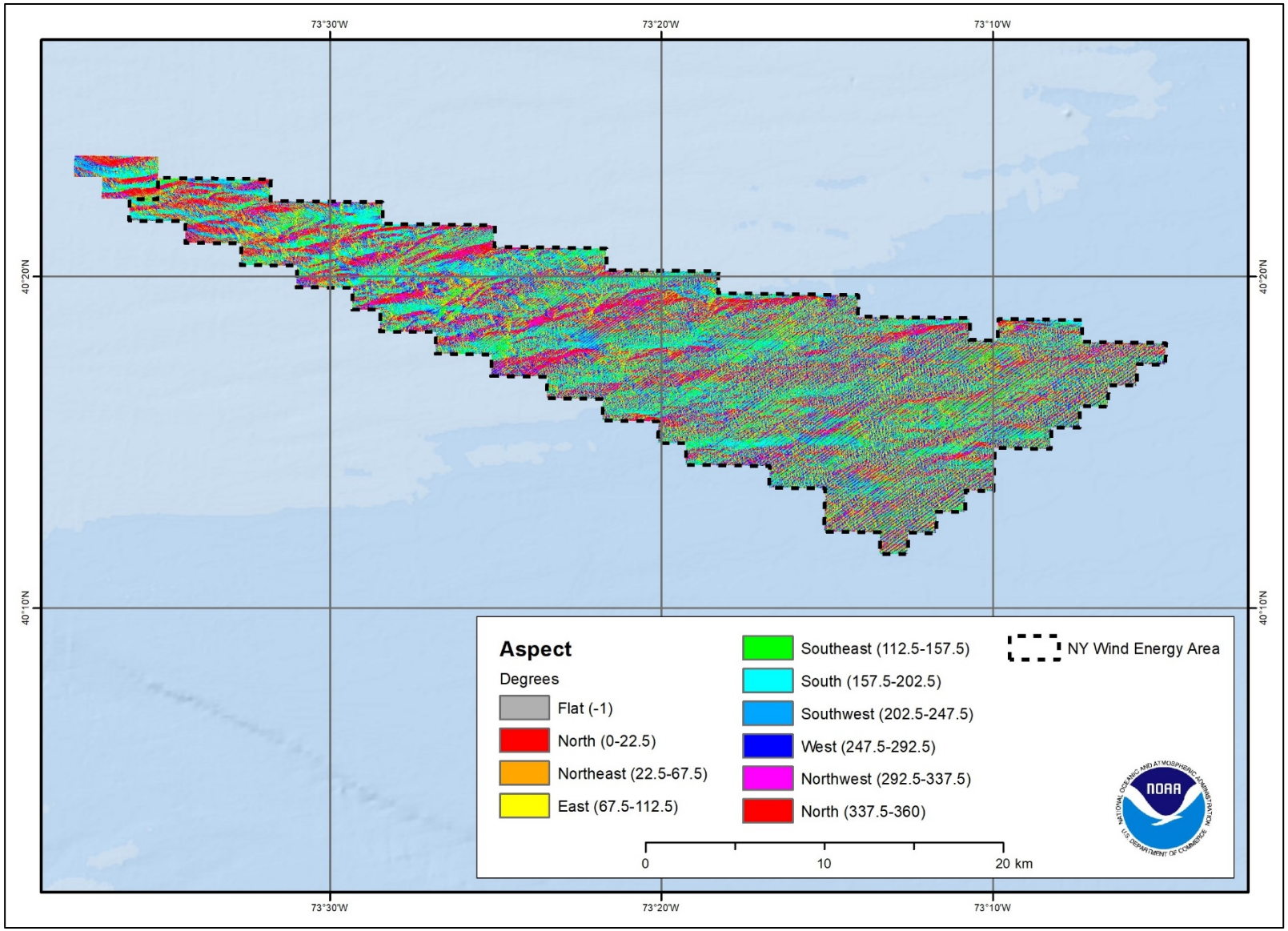


Figure 6.4. Map depicting aspect derived from multibeam bathymetry for the NYWEA.

Total Curvature
Total curvature measured the "surface roughness" to identify places with structural surface complexity or ruggedness within the study area. Although certain areas exhibited surface roughness, total curvature values generally were low within the study area and did not exceed 0.12 radians/100m^2 (Table 6.3, Figure 6.5).

General Curvature
General curvature also measured surface roughness as well as the concavity and convexity of the seafloor to depict ridges and depressions. General curvature ranged in value from -5.29 to 14.26 (Table 6.3) and did not visualize any discernible spatial patterns in concavity or convexity within the study area (Figure 6.6).

Plan Curvature
Plan curvature measured concavity and convexity perpendicular to the horizontal seafloor. Additionally, plan curvature could be used to indicate convergence (- values) or divergence (+ values) of hydrologic flows over a surface. Zevenbergen and Thorne plan curvature measured concavity and convexity to indicate the direction of hydrologic flows over a surface. According to Jenness (2013), the Zevenbergen and Thorne method computes concavity and convexity as well as hydrologic convergence and divergence more accurately than the Evans method, especially for highly accurate and resolved bathymetry data. Zevenbergen and Thorne plan curvature values ranged from -2.8 to 7.0 radians/m^2 (Table 6.3), and similar to Evans plan curvature values, they did not reveal any discernible spatial patterns in the convergence or divergence of currents within the study area (Figure 6.7). Approximately 3% or 10 km^2 of the study area fall outside the range of two SD from the mean (<-0.2 and >0.16).

Profile Curvature
Profile curvature surface roughness measured concavity and convexity perpendicular to the plane of maximum slope, and could be used to depict acceleration (− values) or deceleration (+ values) of hydrologic flows over a surface (Figure 6.8). Evans profile curvature method values ranged from -2.59 to 2.29 (Table 6.3). Similar to the plan curvature metrics, visualization of these profile curvature values did not reveal any discernible spatial patterns in the acceleration or deceleration of currents within the study area.

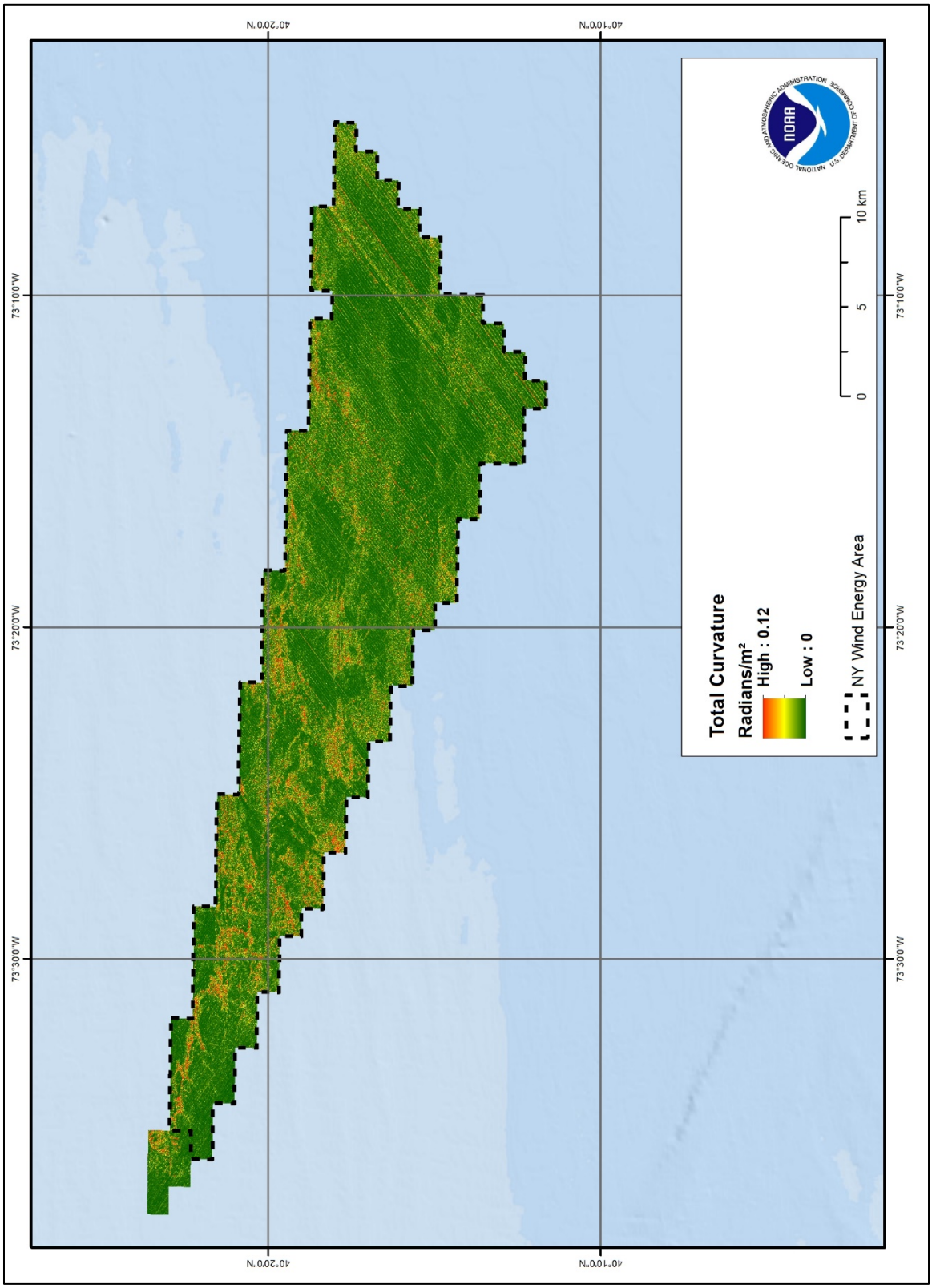


Figure 6.5. Map depicting total curvature derived from multibeam bathymetry for the NYWEA.

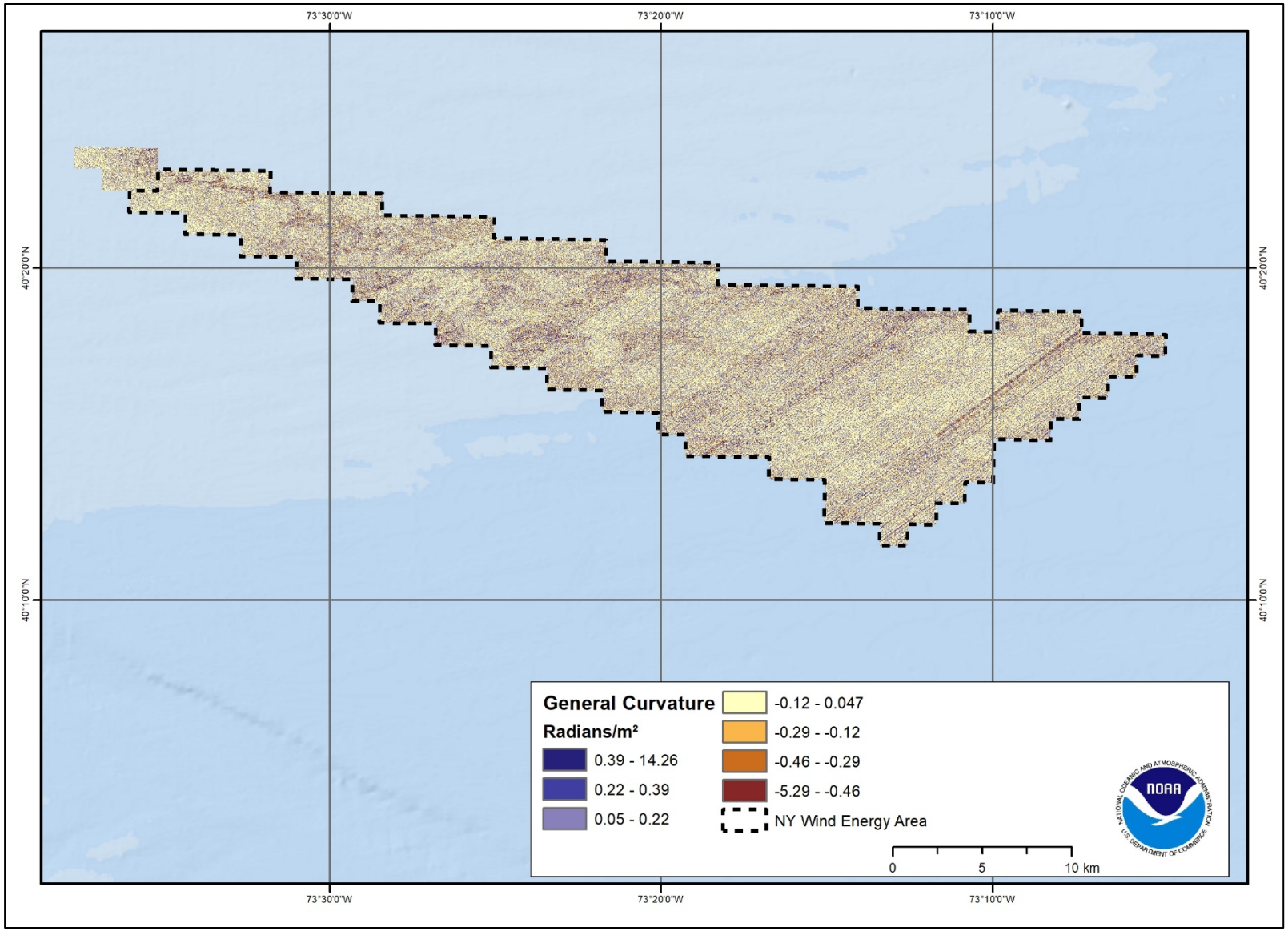


Figure 6.6. Map depicting general curvature derived from multibeam bathymetry for the NYWEA.

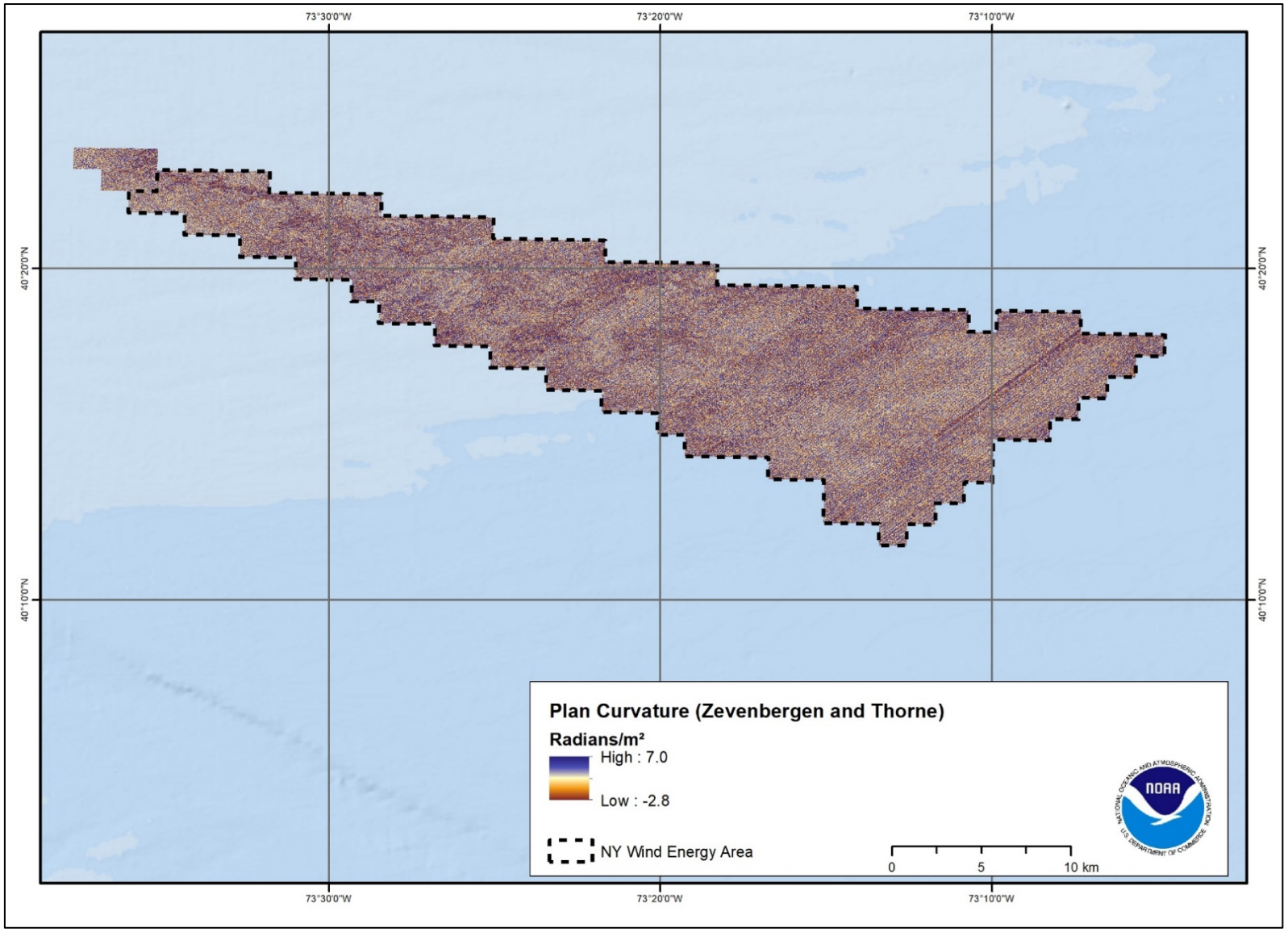


Figure 6.7. Map depicting plan curvature derived from multibeam bathymetry following Zevenbergen and Thorne (1987) for the NYWEA.

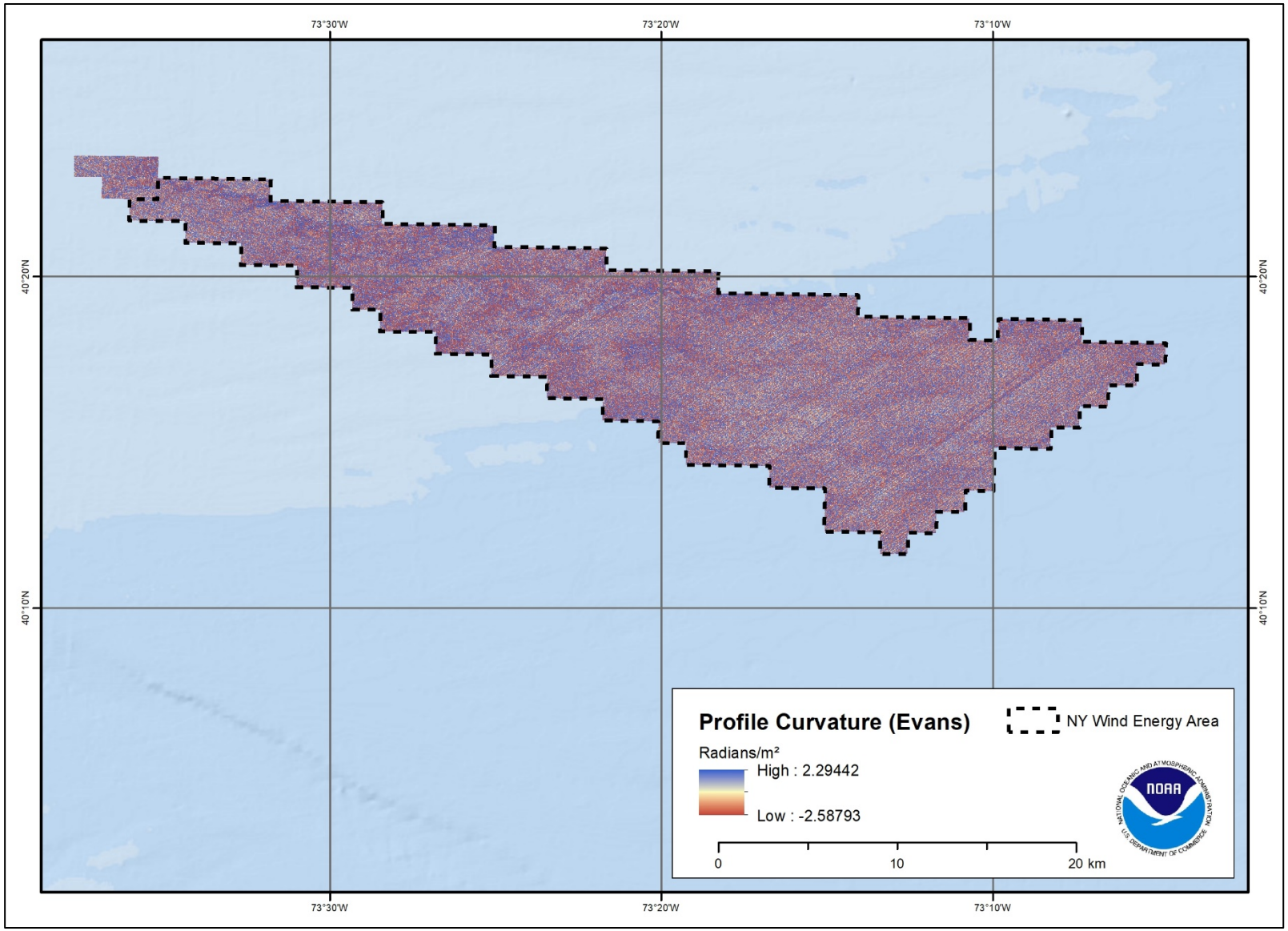


Figure 6.8. Map depicting profile curvature derived from multibeam bathymetry following Evans (1979) for the NYWEA.

6.2.2 Seafloor morphometrics derived from BRESS

The BRESS approach was particularly appropriate for the study area because the landscape featured relatively large areas of grid cells with similar bathymetric and backscatter intensity characteristics. The BRESS segmentation grouped and combined grid cells with similar values for depth and slope and created 175 polygons or landforms with boundaries defined by discontinuities in depth and slope. Since the segments contained kernels with similar values, this approach reduced the number of analysis units from roughly 5 million kernels down to 175 segments that were classified into landform types (Table 6.4).

The BRESS approach includes the calculation of nine metrics. However, our analysis of the information indicated significant information in the landform and average height metrics and not in the metrics of area ratio, average azimuth, elongation ratio, height range, height variance, maximum delta, or maximum width.

Table 6.4 Area and relative proportion of study area occupied by landform types derived from BRESS software (Jasiewicz and Stepinski 2013; Masetti et al. 2018).

Landform	Area (km^2)	Percent of study area
Flat	248.06	75.7
Shoulder	26.08	8.0
Slope	19.70	6.0
Footslope	18.88	5.8
Valley	6.88	2.1
Ridge	4.01	1.2
Convex Slope	1.91	0.6
Concave Slope	1.86	0.6
Pit	0.21	0.0
Peak	0.08	0.0

Landforms

Analysis of the landforms indicated a predominantly flat surface with a slope less than 0.2 degrees for most of the study region. Generally, the eastern areas of the study region was particularly flat whereas some topographic complexity occurred in the western portion of the study area (Figure 6.9). Based on the simplified classification of Masetti et al. (2018), the distribution of landforms were group to produce a more easily interpretable graphic (Figure 6.10).

Average Height

Visualization of average height identified areas where there was depth variability and discriminated them from areas of low depth variability. Average height ranged from -2.13 to 2.25 m, with a mean of 0.216 m (SD = 0.221 m, CV= 1.023). These statistics suggest that localized variation in depth was very small within the study region. Here, the smaller values indicated the flat areas whereas the maximum and minimum values highlighted areas of topographic complexity as well as artifacts within the bathymetry data (Figure 6.11).

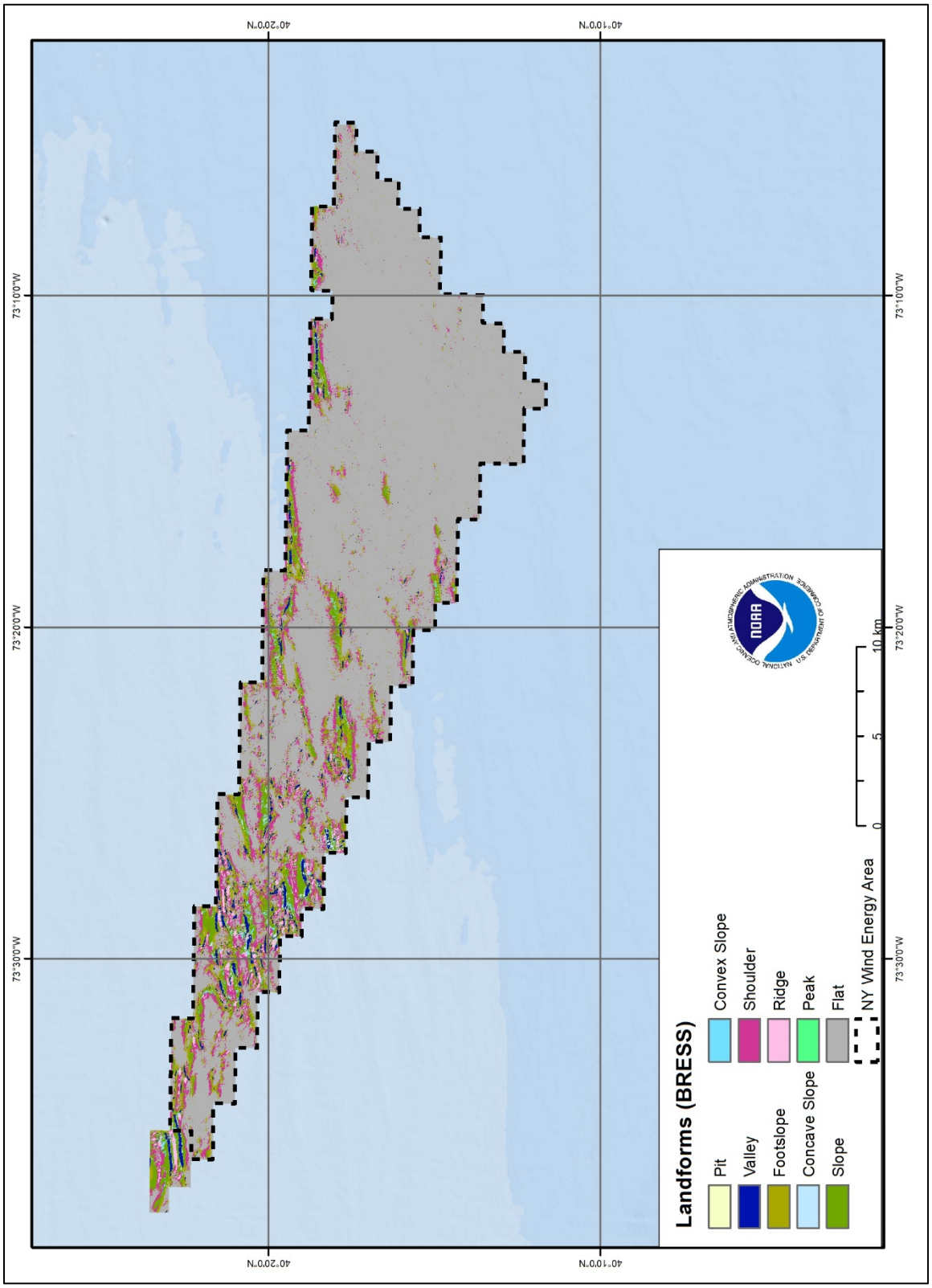


Figure 6.9. Map depicting landforms derived from BRESS for the NYWEA.

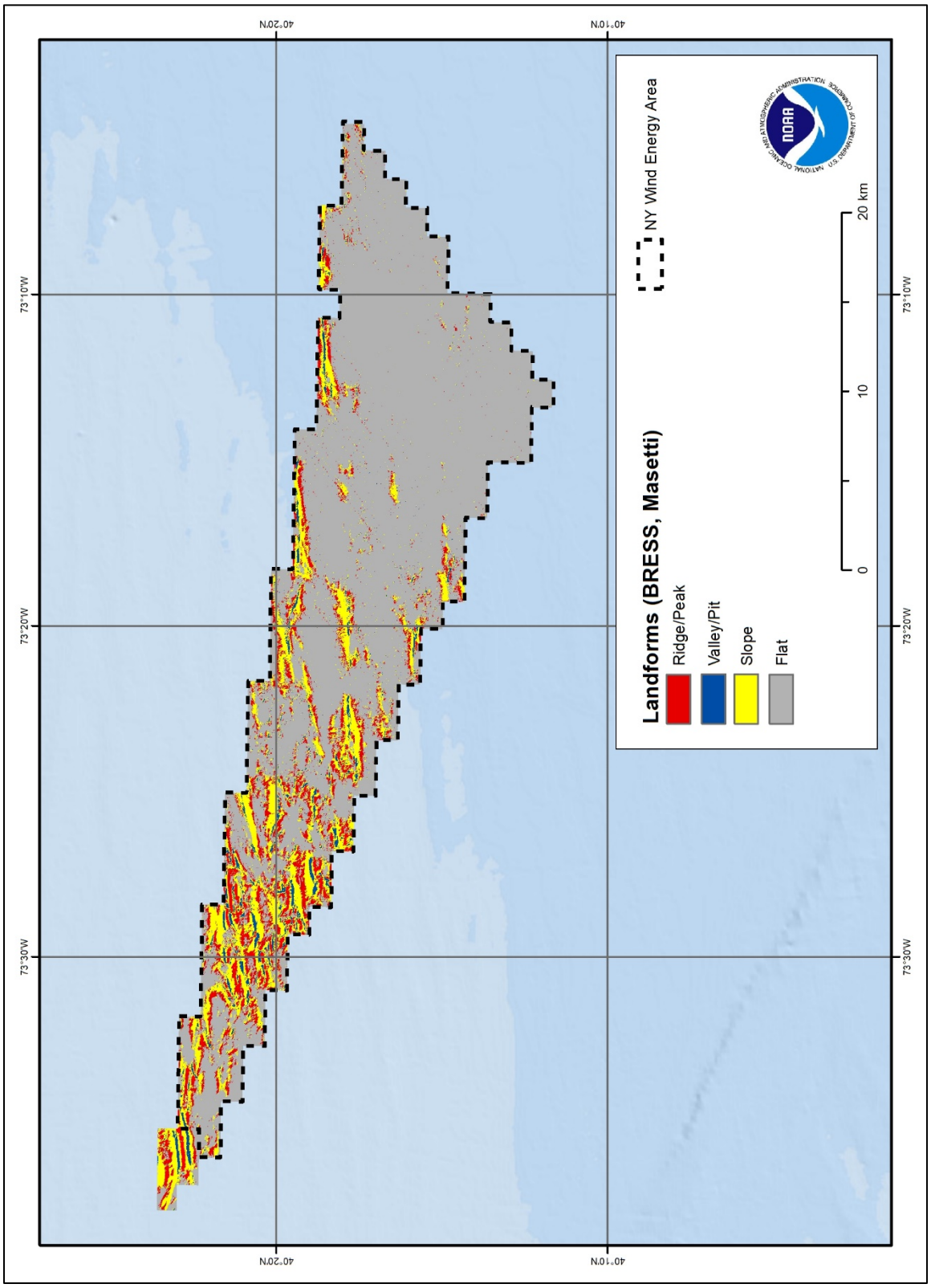


Figure 6.10. Map depicting landforms based on Masetti et al. (2018) for the NYWEA.

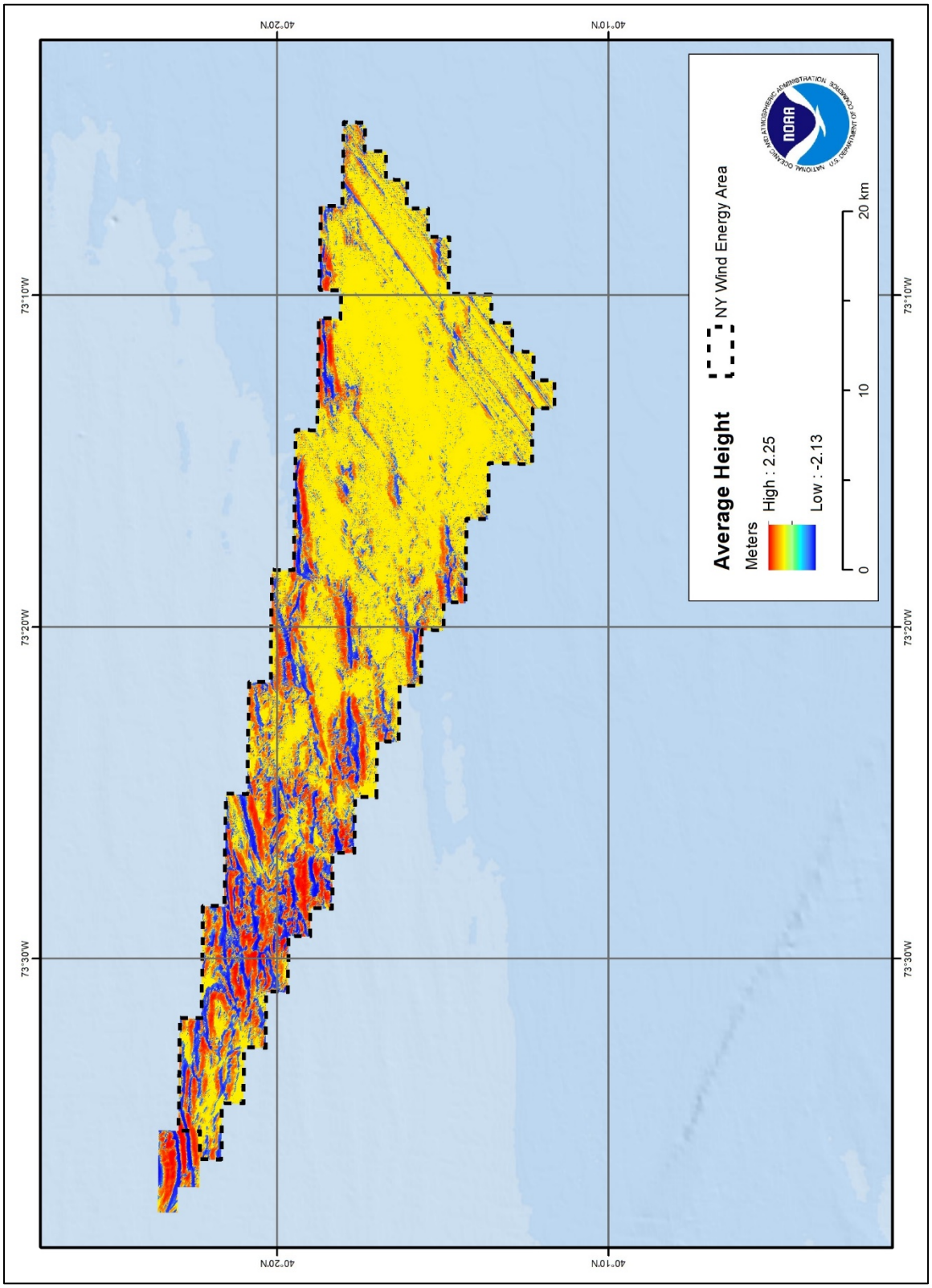


Figure 6.11. Map depicting the average height of BRESS segments for the NYWEA.

6.3 Conclusion

Twenty-two different analyses were used to characterize the geomorphology of the NYWEA study area from the bathymetry and backscatter surfaces. Several metrics were particularly informative from this analysis. The maximum slope within the study area was 4.3 degrees, with 99.9% of the area having slope less than 1.5 degrees. The BRESS analysis identified large swaths of the study area with consistent topographic complexity and substrate composition. Large areas of consistent flat sandy bottom led to large segments particularly in the eastern portion of the study region. Through several rugosity metrics, areas of relatively higher structural complexity along with a few artifacts were identified.

A great deal of the variation of values in seafloor morphometries is due to sea-state induced noise artifacts inherent in the source bathymetric surface. Removing artifacts from the bathymetry could improve results and enable more effective analysis of individual seafloor metrics. Geomorphological metrics along with the bathymetry and backscatter surfaces were used as inputs into boosted regression classification trees to predict habitat types and create a benthic habitat map from bathymetric raster data (see Chapter 11). Metrics with the most discriminatory power appear to be depth, backscatter, SD of depth, slope, slope of slope, north-south and east-west gradients of aspect, total curvature, profile curvature, plan curvature general curvature, and average height.

Chapter 7. Ground Validation Video Analysis

Ground validation (GV) uses direct *in situ* observations to interpret remotely sensed data. In this study, GV data are the basis for correlating observed substrate and cover types with their associated values in the predictor datasets. To make direct observations, underwater videos of the seafloor were collected and analyzed from sites dedicated to GV. Results from the video analysis of the presence and absence of different substrate and geoform types and the percent biotic cover were then used to BRT and enhance the spatial accuracy of habitat maps in Chapter 11. A total of 300 sites were allocated for GV analysis (Figure 7.1). Of these 300 sites, 296 GV sites were analyzed and used for the BRT models, however there were 12 sites that were located at Cholera Bank outside the NYWEA.

7.1 Methods

7.1.1 GPS Processing Method

The planned GV sites were uploaded into the Furuno navigation system of the R/V *Tiki XIV* to guide the Captain to the sampling locations. The GV site planning and the MVV data collection methods and analysis are described in detail in Chapter 3. The Trimble Geo 7X 6000 handheld GPS (global positioning system) receiver was used to record the following data at each GV site: site location, sample time, and *in situ* observations of the grab sample. The CMECS classification schema was uploaded into the Trimble so that the *in situ* presence/absence of substrate types based on the grain size and benthic invertebrate observations could be recorded at each site. The Trimble logged one point each second (epoch) as the grab sampler was lowered into the water and retrieved. Upon completion of the cruise, the Trimble GPS data were post-processed to provide improved, corrected positioning (sub-meter accuracy) using GPS Pathfinder Office (software version 5.81). Through post-processing, an average location (or centroid) from the epochs at each GV site was calculated, and the initial GPS logging time was used as sampling time in the final deliverable. The post-processed Trimble positions and the recorded *in situ* observations were exported as point shapefile using ArcCatalog. A GIS analyst conducted quality control of the data to detect entry mistakes and reviewed the comments in the field for any discrepancies. At sites where multiple grab samples were collected, the final sample was selected to best represent the site location.

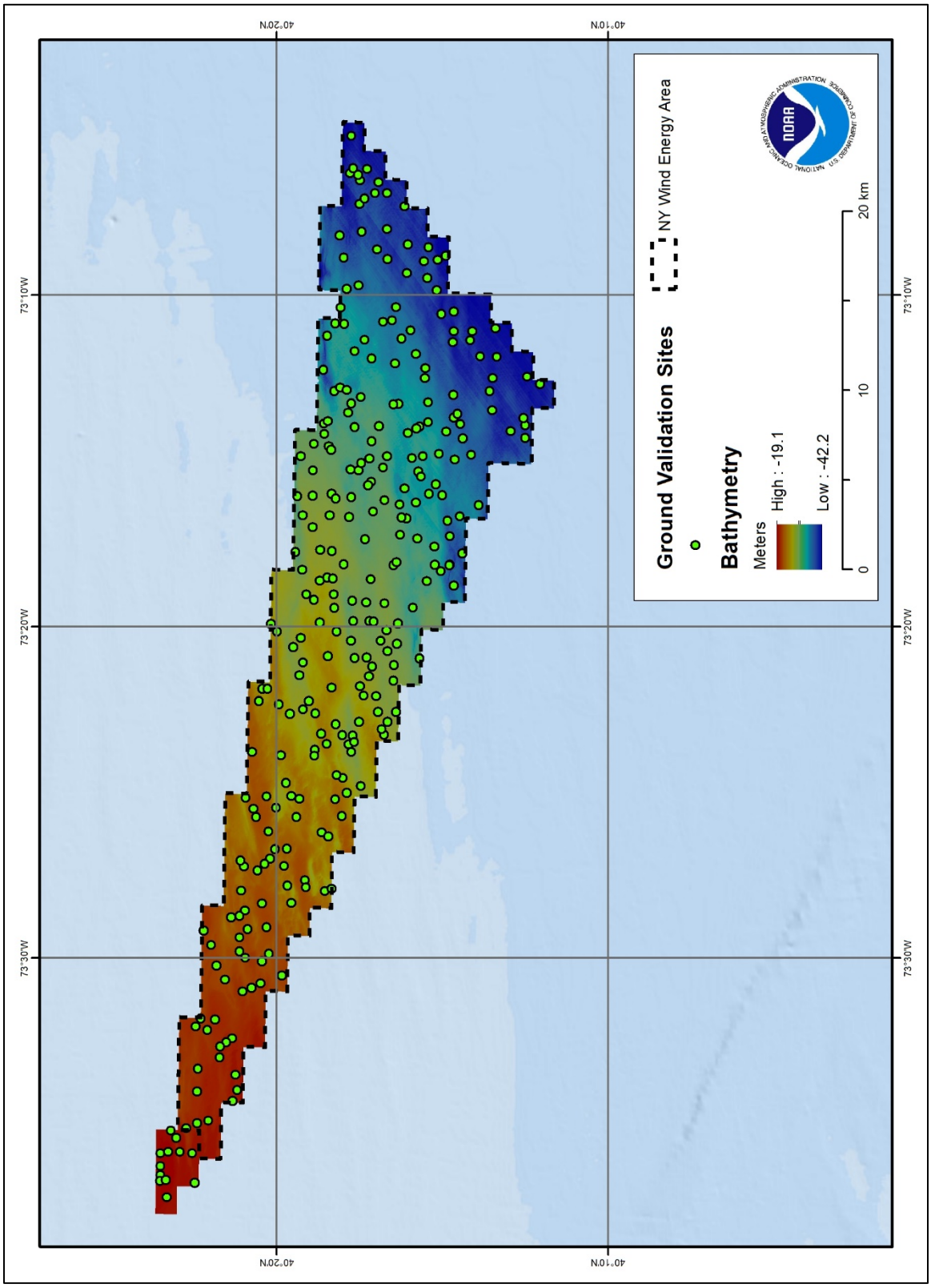


Figure 7.1. The locations of the GV within the NYWEA.

7.1.2 Video Annotation Method

Videos recorded during MVV sampling were reviewed after the cruise to determine the substrate type, geoform and biological cover of the seafloor at each GV site. Three cameras were used to collect video footage: 1) The oblique view GoPro camera with the mounted laser scale, 2) the downward facing GoPro camera, and 3) the HD Seaviewer camera (see Chapter 3 for video specifications). The oblique view GoPro camera provided a larger field of view to interpret the seafloor features and so it was the primary video platform used to analyze the habitat of the seafloor for each GV site. The downward facing GoPro camera was used as a backup to assist in the analysis as the field of view was small especially after sunset and did not provide the analyst with enough data to interpret features of the seafloor. The downward facing camera footage was used at ten sites where the oblique view camera footage was unavailable.

Each GV video was clipped to the first and last frame of useable footage. Videos were then annotated, a process where footage was evaluated to quantify presence and percent cover for substrate, geoform and biota, by a single analyst to prevent variability in results. Underwater video and image annotation is a common method for assessing benthic habitats across ecosystems (Jaffe et al. 2001), and has been used in previous habitat mapping efforts in the mid-Atlantic for offshore wind farm suitability (Guida et al. 2017). While automated annotation is an increasingly popular tool in coral ecosystems (Beijbom et al. 2015), it is less effective in regimes that are dominated by varying combinations of unconsolidated sediment. Thus, annotations in this study were made manually by a single analyst.

Substrates and biota present in each video were recorded using the CMECS habitat classification scheme to characterize benthic assemblages and functional groups. The analyst recorded the presence (1) and absence (0) of substrates and geoforms and visually estimated percentage cover for the biotic component. Additionally modifiers such as anthropogenic features and large megafauna (greater than 3 cm in length or height) not anchored to the bottom were noted during video review (e.g., skates, soles, flounder, dogfish, anglerfish, and jellyfish). Presence/absence annotations from the video analysis were merged with the point shapefile that contained the *in situ* observations of substrates and biota from the MVV sediment grab samples and was used as an input for the BRT models to predict the probability of occurrence for each habitat type and provide a more accurate representation of the seafloor habitat.

7.1.3 CMECS Classifications Scheme

The CMECS system developed by NOAA and the Federal Geographic Data Committee (FGDC) was adopted to record *in situ* and video observations of substrates, geoforms, and biota at GV sites from the MVV survey on the R/V *Tiki XIV*. CMECS is self-described as "a catalog of terms that provides a means for classifying ecological classes using a simple, standard format and common terminology. CMECS offers a way to organize and interpret data about the marine environment, and it provides a common platform for inter-relating data. It builds upon approaches from published national, regional, and local habitat classification procedures, and it offers an umbrella under which a national coastal and marine ecological classification can grow and evolve." (FGDC 2012).

The CMECS classification scheme implemented for this study contains four major components (substrate, geoform, biotic and modifier) used to describe the benthic environment of the NYWEA. Five substrate categories were included: cobbles, pebbles, sand/granules, mud, and

shell. Additionally, to characterize the bedforms of each site, four geoform classes were included: ripples, megaripples, outcrops, and plain. Coverage of the seafloor was estimated for biotic components, and lastly, a modifier component was added to capture other important information from the videos (e.g., mobile fauna, anthropogenic structures). Table 7.1 describes the classification for the components developed for this study.

Table 7.1. Benthic habitat classification scheme used to describe observations for this study.

CMECS Component	Categories/Classes	Definition
Substrate *Unconsolidated Sediments (Presence/Absence)*	Cobbles	64 mm to <256 mm median grain size
	Pebbles	4 mm to <64 mm median grain
	Sand and Granules	0.0625 mm to 4 mm median grain
	Mud (Silt and Clay)	<0.0625 mm grain size
	Shell	broken shell fragments of any size (shell rubble, hash, etc.)
Geoform *Sediment Wave Fields and Outcrops (Presence/Absence)*	Ripples	linear structures of sediment formed by movement of water <1 m wavelength
	Megaripples	structures of sediment formed by movement of water >1 m wavelength
	Outcrops	areas of exposed bedrock
	Plain	relatively flat area of unconsolidated sediment
Biotic Cover *Attached or Bedded Fauna (Presence/Absence and % Cover)*	Molluscs	individual or faunal beds of mussels, clams, oysters, marine snails, or sea slugs
	Echinoderms	individual or beds of sea urchins, sand dollars, and sea stars
	Annelid Worms	sessile or mobile segmented marine worms, including tube, feather, and blood worms (polychaetes)
	Crustaceans	mobile hermit crabs, lobster, shrimp, and amphipods
	Algae	living vegetation attached to the seafloor
	Other	traces of marine fauna, such as egg cases, burrows, track marks
	None (No Cover)	areas with no biotic cover
	Less than 5% cover	*Areas with ≤ 5% of biotic cover, or ≥95% no cover*
Modifier	Anthropogenic	manmade debris or structures from sunken ships or cargo
	Megafauna	demersal or pelagic organisms >3 cm in any dimension
Other/Comments	*Observations from Sediment Grabs*	*Any descriptions or notes from the annotator from the grab sample or video analysis (e.g., color, stiffness, stratification, hydrogen sulfide smell)*

Italics denotes categories not included in CMECS

7.1.4 Spatial Analysis

Geostatistical modeling was conducted to map the spatial distribution of the biotic component by analyzing the percent cover estimates from the GV videos. Interpolations of the different Biotic Cover types across the NYWEA were tested using the Bayesian Kriging method from Geostatistical Analyst tool in ArcGIS (version 10.6). This method was also used for the analysis of grain size distribution for different phi (φ) types and described more in detail in Chapter 9. The sample neighborhood size, shape, and angle for the kriging were tested and the covariations were automatically analyzed with the best fit for each regression to produce the lowest Root Mean Square Error (RMSE) deviations. Locations where Bayesian Kriging predicted a negative percent cover were converted to zero because it is not possible for a negative cover to exist. Upon interpolation, a 32-bit GeoTIFF was created using the NYWEA boundary as a mask and the data were then classified using natural breaks in the data to group cover into percentage categories.

7.2 Results

High definition videos were collected and annotated from 296 GV sites, covering the NYWEA and Cholera Bank. The results from the videos annotations and *in situ* observations were used to train the BRT models in order to predict the habitats throughout the study area (see Chapter 11 for more details). A spatial distribution of the biotic cover components was also generated to compare with the predictive modeling.

7.2.1 Substrate Types

Substrate data presented here was based on *in situ* presence/absence observations from the sediment grab samples rather than estimating cover or composition from video analysis. It was not possible to accurately classify full phi sizes and percent distribution while on the R/V *Tiki XIV* or by observing in the HD underwater video due to inability to confidently measure grain sizes visually. For this reason, an independent grain size analysis was conducted from the sediment grab samples using a sieve test at the TDI-Brooks International laboratories to confidently measure the percent distribution of phi grain size using the Wentworth scale (Wentworth 1922). However, due to time constraints the phi analysis from TDI-Brooks was not used as GV for the BRT models. For the analysis and results of the TDI-Brooks grain size sieve tests, see Chapter 9.

Based on presence/absence observations recorded from the grab samples, substrate types occurred in eight unique combinations (Figures 7.2-7.4). Sands (fine to very coarse) and broken shells (all sizes) were the dominant substrates throughout the NYWEA, present in 100% and 99% of GV sites, respectively. Substrates of 71.3% GV sites were composed exclusively of sandy and broken shell (Figure 7.3; Figure 7.4b).

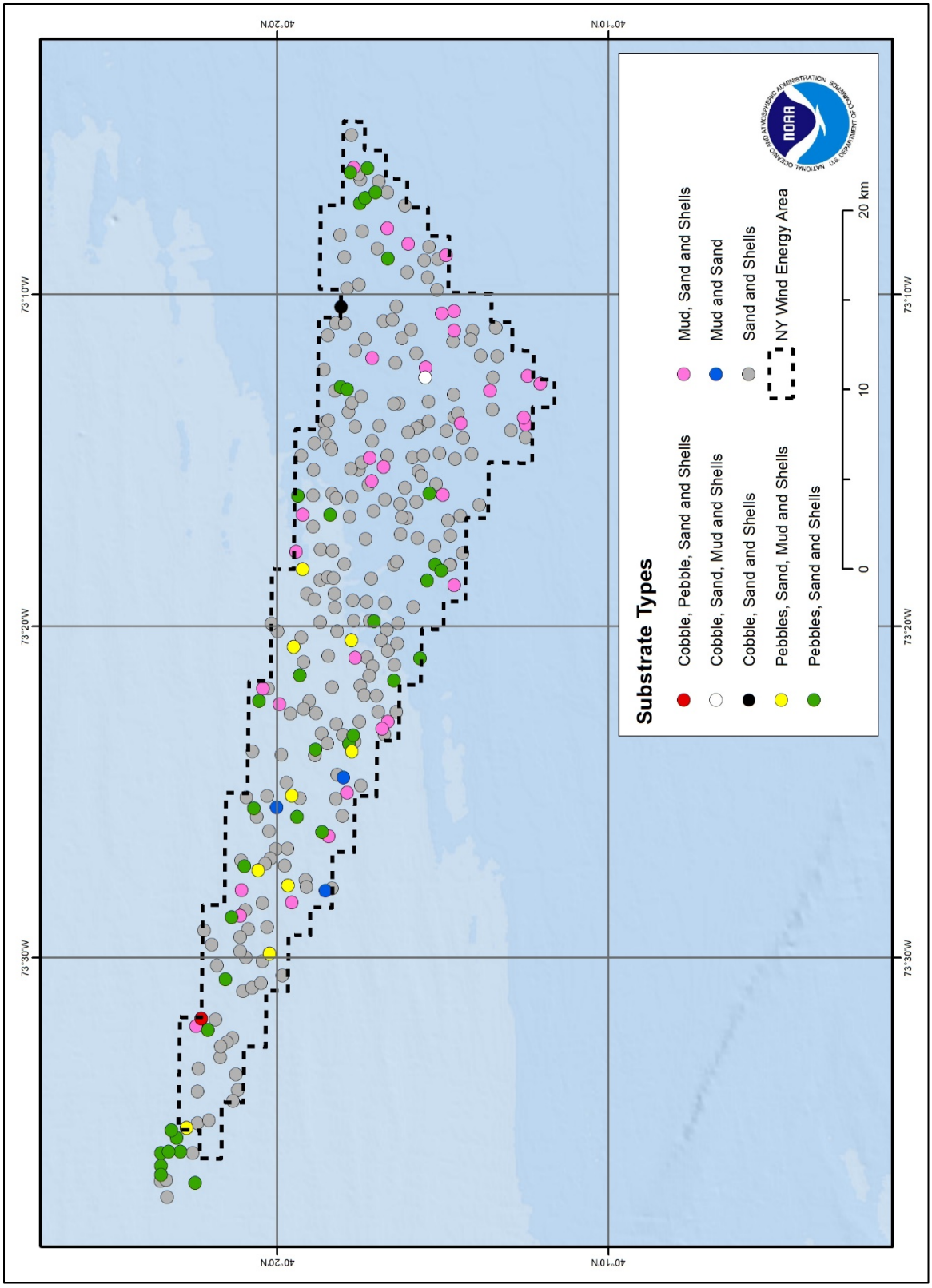


Figure 7.2. Distribution of the different substrate types for NYWEA.

The second most dominant substrate category was the conglomeration of gravelly sediments (granules, pebbles, and cobbles) (Figure 7.4a). Pebbles (small to large) were present in 15.9% of GV sites (Figure 7.3) and distributed irregularly throughout the NYWEA, but commonly occur on Cholera Bank (Figure 7.2). Mud substrates (silts and clays) were detected at only 15.1% of the GV sites (Figure 7.3). Cobbles were rarely observed in the NYWEA and only present at three sample sites, or <1% of GV sites (Figure 7.3). The presences of cobbles appeared erratically and were detected from individual grains collected in the grab samples. No cobbles were sent to the TDI-Laboratory after the homogenization of the sub samples for full phi analysis. The three sites with cobbles do not appear to conform to a spatial pattern and would have greatly skewed the results of the phi distribution analysis, thus they were treated as outliers. Observations were annotated from Predictions of occurrence of each substrate can be found in Chapter 11.

GV methods indicate the NYWEA is dominated by sandy substrate. Previous work by Guida et al. (2017) also found that sand comprised the majority of substrate in this area. More broadly, proposed wind energy areas throughout the northwest Atlantic from Virginia to Massachusetts are also sand dominated (Guida et al. 2017). Thus, the dominance of sand observed throughout the NYWEA in the current study is consistent with previous work within the study area and regionally.

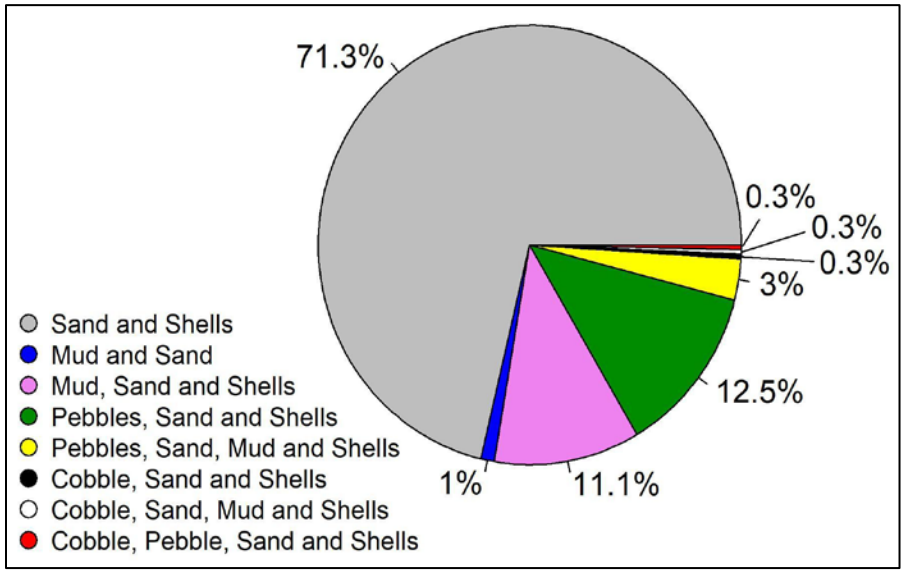


Figure 7.3. Percentage of sites with substrate types present at GV sites in NYWEA.

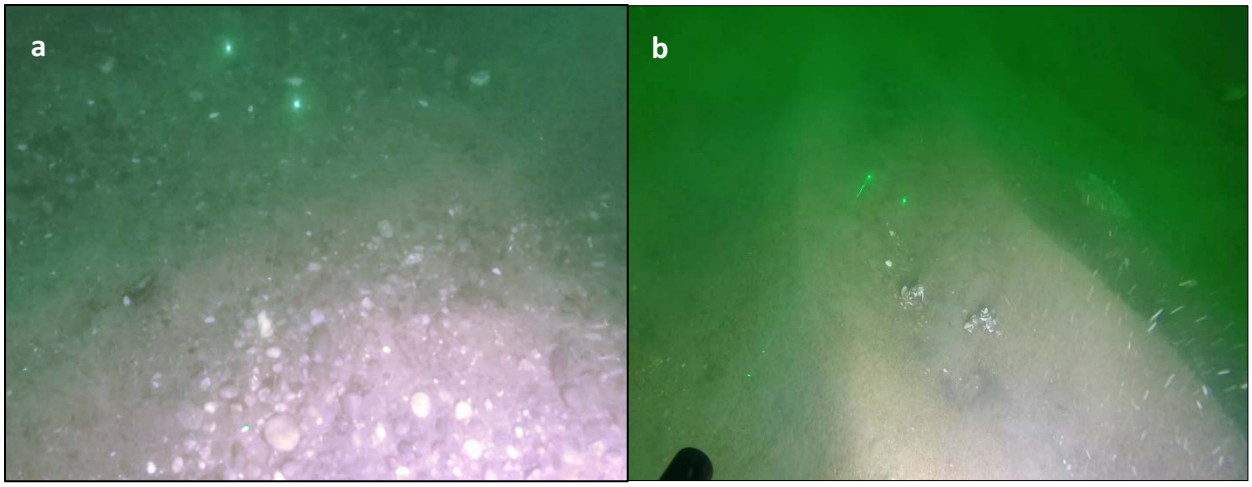


Figure 7.4. (a) Pebble, sand and shells, and (b) sand and shell substrates at GV sites.

7.2.2 Geoform Types

Three distinct geoform types (ripples, megaripples, and plain) were observed at GV sites (Figures 7.5-7.7). No rocky outcrops were observed at any sites. Ripples (Figure 7.5b) were the dominant geoform in the study area, representing 83% of the sites (Figure 7. 7).

Additionally, we observed megaripples at 12% of the GV sites (Figure 7.7). Megaripples are "large, sand waves or ripple-like features having wavelengths greater than 1 meter or a ripple height greater than 10 cm; Megaripples are formed in a subaqueous environment, and they are also known as subaqueous dunes" (Bates and Jackson 1984; Figure 7.5a). Previous work in this area has also noted the presence of megaripples (Guida et al. 2017). In this study, megaripples occurred mainly in the deep southeast zone of the study area (Figure 7.6). Plain (flat) sites were the least common geoform type within the study area and were present at only 5% of GV sites (Figure 7.7, Figure 7.5c). More detailed discussion of spatially predicted distributions of geoforms can be found in Chapter 11.

Figure 7.5. Geoforms: (a) megaripples, (b) ripples, and (c) plain at GV sites within NYWEA.

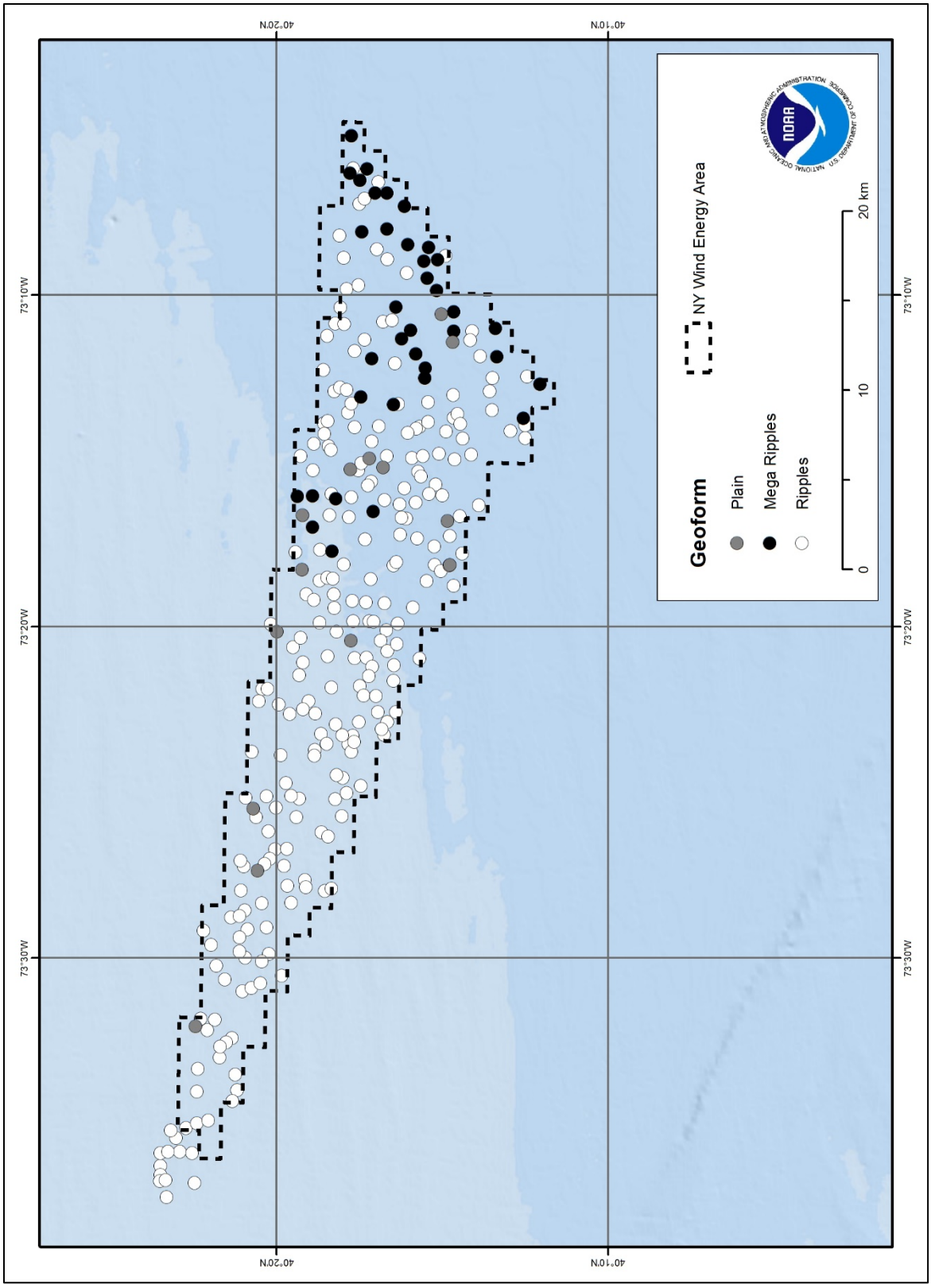


Figure 7.6. Geoform distribution at the GV sites within NYWEA.

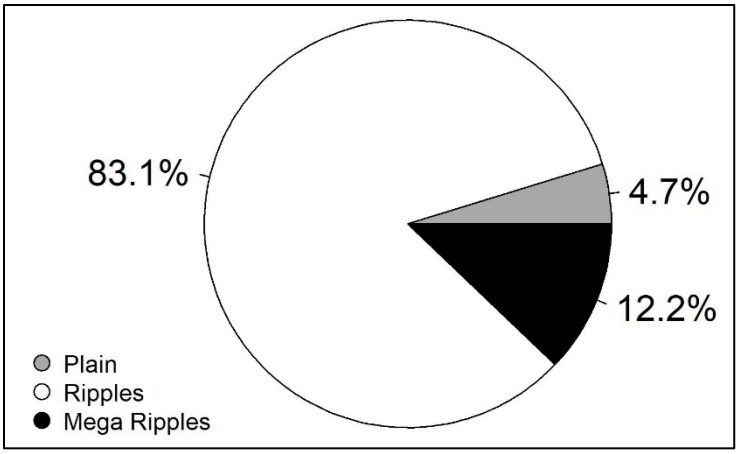


Figure 7.7. Percentage of geoforms present at GV sites within NYWEA.

7.2.3 Biotic Cover and Modifier

A variety of biota were observed at GV sites throughout the NYWEA. Benthic-dwelling epifauna were the primary group observed, with mobile fauna rarely observed. Echinoderms, specifically the common sand dollar (*Echinarachnius parma*), were the dominant biotic component and the only species to represent significant cover of the seafloor at GV sites (Figures 7.8-7.10). They were present at 90% of the GV sites and cover ranged from 1-90% (Figure 7.9).

Figure 7.8. High density of Echinoderms *Echinarachnius parma* (sand dollar) at a GV site within NYWEA.

Other groups observed throughout GV sites were various species of worms, mollusks, crustaceans, algae, and megafauna common to the North Atlantic (Figures 7.11-7.13). While individual organisms in the groups were observed in grab samples and the video analysis, their quantity and distribution throughout NYWEA was not uniform. Annelids, generally polychaetes and oligochaetes, were present at 30% of the GV sites (Figure 7.12). Mollusks, specifically bivalves and gastropods were present at 11.5% of GV sites, with moon snails (species unknown) a primary gastropod observed (Figures 7.12 and 7.13a).

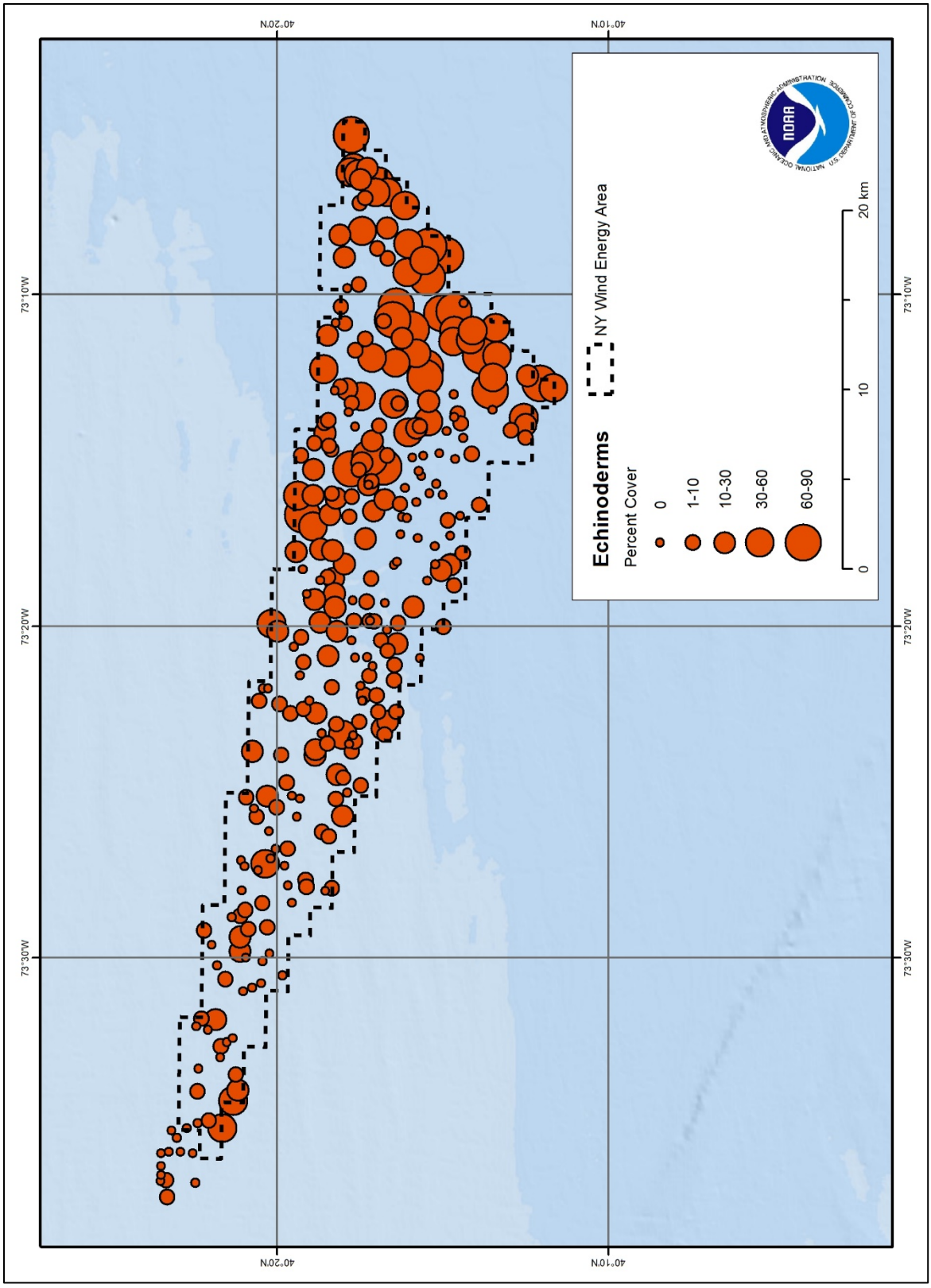


Figure 7.9. Percent cover of common sand dollar (*Echinarachnius parma*) at GV sites within NYWEA.

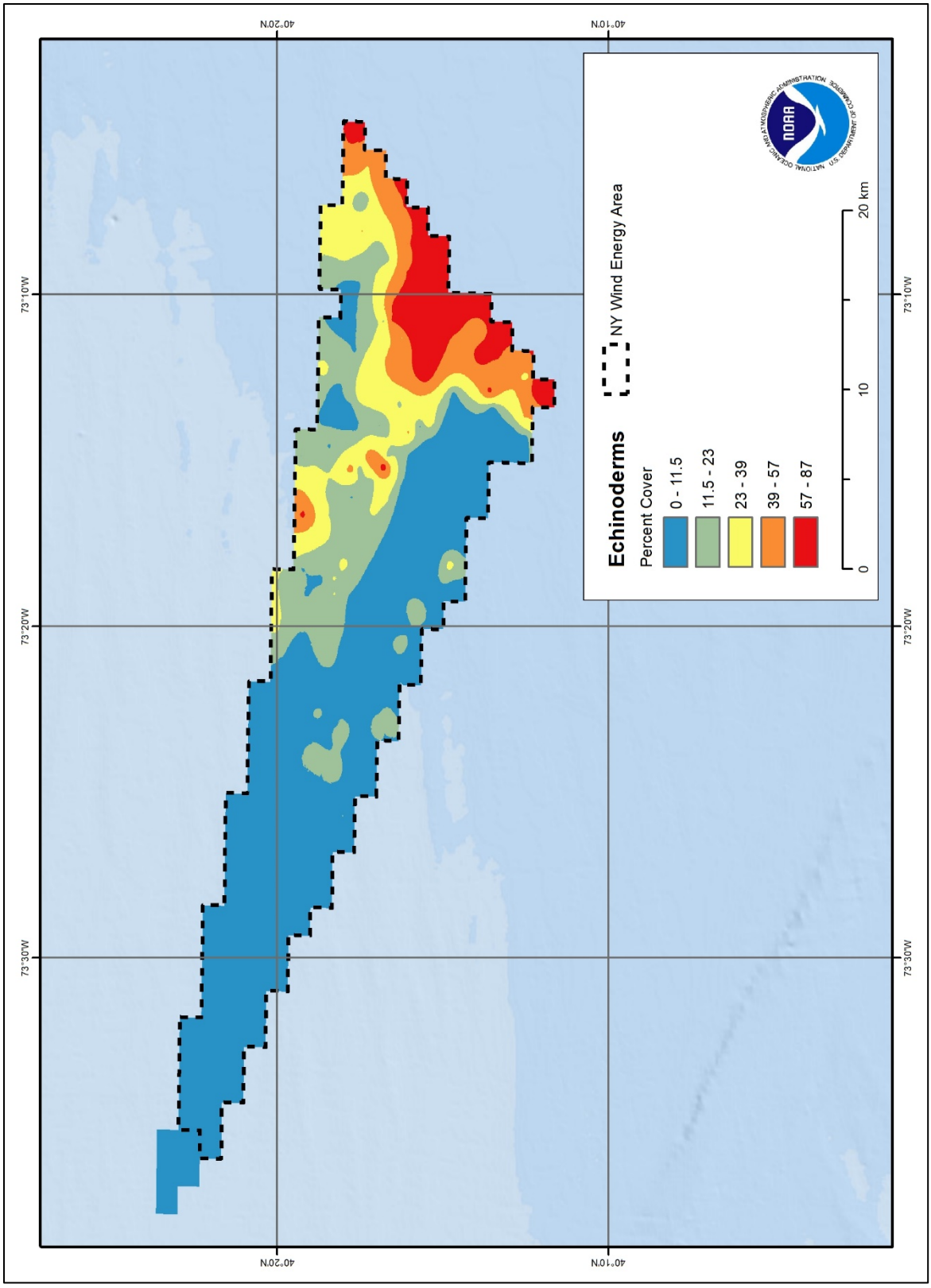


Figure 7.10. Distribution of common sand dollars (*Echinarachnius parma*) using Bayesian Kriging.

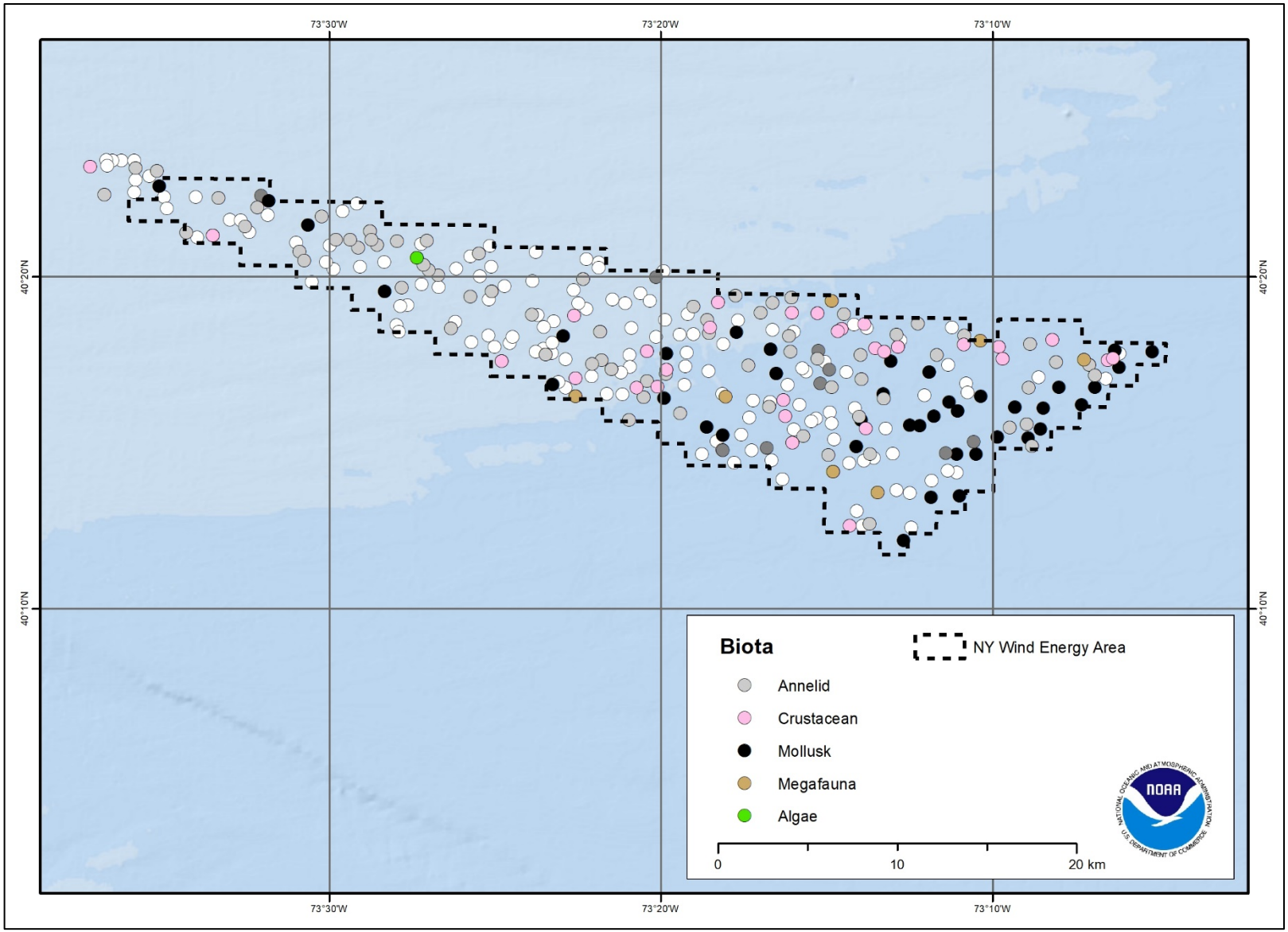


Figure 7.11. Presence of other biota (annelids, crustaceans, mollusks, algae, and megafauna) at GV sites within NYWEA.

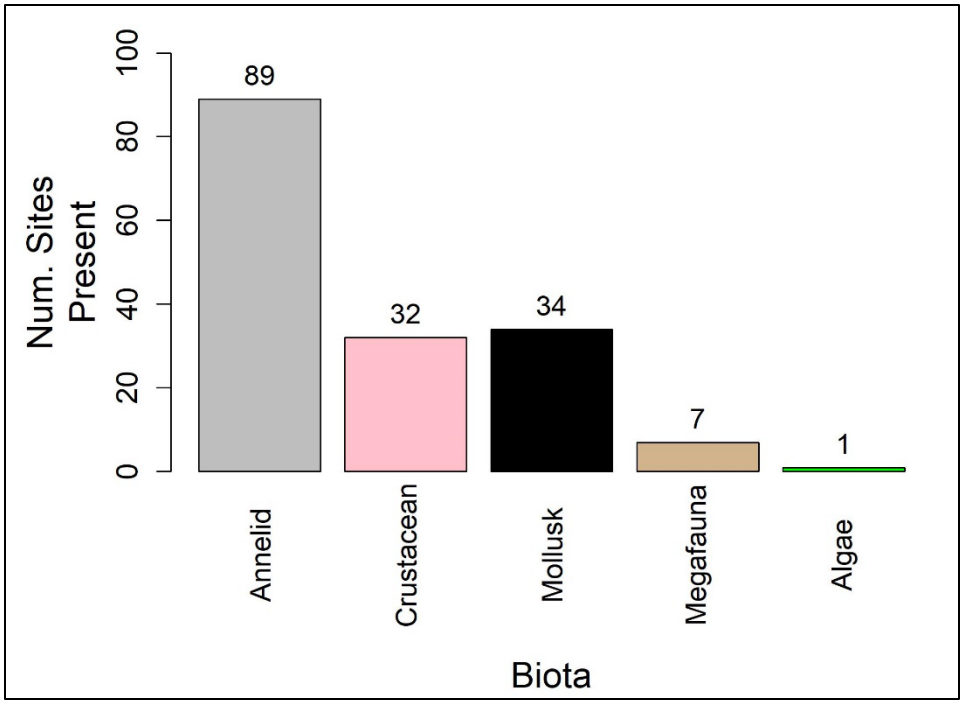


Figure 7.12. Frequency of occurrence of benthic biota at GV sites within NYWEA.

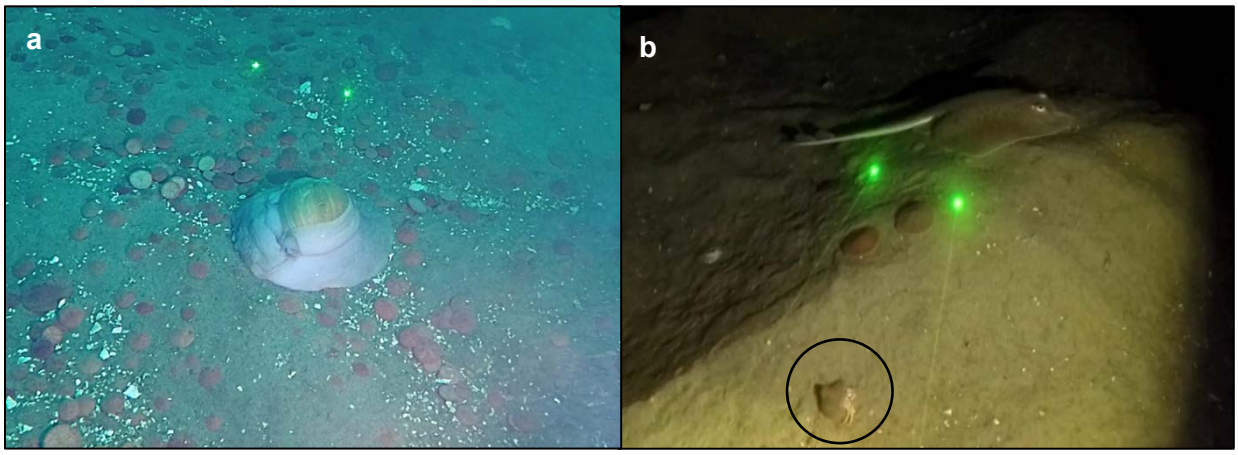


Figure 7.13. (a) Moon snail (species unknown) at a GV site within NYWEA, and (b) a skate (*Leucoraja* spp.) and a hermit crab (species unknown; circled) at another GV site within NYWEA.

Crustaceans, specifically hermit crabs and amphipods, were observed at 10.8% of GV sites (Figure 7.12). Algae was only present at one site (Figure 7.12). Megafauna (consisting entirely of skates, *Leucoraja* spp.) and algae were rarely observed; with megafauna seen at only 2.4% of GV sites and algae present at only one site (Figures 7.11 and 7.12). No anthropogenic structures were found at GV sites, however shipwrecks have been detected in this area including several during the NF17-09 cruise (see Chapter 2). The grab sampler was not used at this location due to the risk of entanglement and harm to the structures of the shipwreck.

Distribution models were only created for sand dollars, as they were the only significantly present biotic component. The distribution of all of the other biotic components modeled very poorly or were sufficiently rare to model. Results from the Bayesian Kriging interpolations showed high uncertainty (RMSE = 14.93), but did conform with the presence/absence distribution (Figure 7.9) and the BRT modeling in Chapter 11. The Bayesian Kriging models illustrate that the sand dollars are highly concentrated at the eastern edge of the NYWEA and spread northwest to the 30 m depth contour (Figure 7.10).

The high abundance of echinoderms (sand dollars) observed in this study was consistent with previous work conducted in the NYWEA and throughout the northwest Atlantic, where surveys revealed dominance of the benthic epifaunal community by sand dollars (Malek et al. 2014; Guida et al. 2017). Sand dollars prefer flat fine sediment habitats (Wigley and Theroux 1981; Malek et al. 2014), which dominate the eastern portion of the NYWEA (Figure 7.2). Therefore as expected, sand dollar cover in areas with homogeneous fine sediment habitats had a high cover compared to the very low cover of the more heterogeneous habitats (e.g., pebbles and cobbles) (Figure 7.9). In this study, sand dollars were also often observed in high densities at sites where megaripples were present (Figure 7.6). The relationship between substrate, geoform, and sand dollar presence is further described in Chapter 11.

As noted, skates were the only megafauna species to be observed at GV sites. Previous trawl surveys throughout the northwest Atlantic, including the NYWEA, have noted skates as the dominant catch year round (Guida et al. 2017). Other biota (e.g., annelids, crustaceans, mollusks) were distributed evenly throughout the area, and therefore were not spatial correlated with the distribution of a particular substrate or geoform type.

Chapter 8. Hard Bottom Predictive Modeling

Maps depicting the distribution of hard bottom habitats can provide important information to guide coastal and marine spatial planning, including the siting of offshore wind installations. In addition to practical considerations about where to conduct offshore construction activities, information about the distribution of hard bottom habitats is critical for making environmentally sound decisions about offshore activities that may impact sensitive biota. Hard bottom substrate can provide a stable point of attachment for sessile invertebrates, such as corals and sponges that are slow growing and particularly vulnerable to disturbance (Freiwald 2002). Hard bottom substrate and associated invertebrate communities can also provide refuge and settlement habitat for various fishes (Guida et al. 2017). More generally, areas of hard bottom substrate that are structurally complex have been associated with higher levels of biodiversity (Steimle and Zetlin 2000).

In this study, the seafloor within the NYWEA was characterized using multibeam acoustic sonar, towed video camera images, and sediment grab samples (see Chapters 9 and 11). This analysis suggested that the NYWEA consists primarily of soft bottom substrate. This chapter describes a regional map of the distribution of hard bottom habitats that was developed for the entire New York Bight to provide additional context about the seafloor in areas outside of the NYWEA that may be impacted by activities within the NYWEA (e.g., cabling associated with offshore wind installations) or that may receive future consideration for activities overseen by BOEM. However, the types of high-resolution information collected in the characterization of the NYWEA are not available for much of the New York Bight. Dunn and Halpin (2009) demonstrated that spatial predictive modeling could be used to predict the extent of hard bottom habitats at a regional scale from a regional bathymetry dataset. Poti et al. (2012) developed a model predicting the occurrence of hard bottom habitats in the New York Bight using point locations of known hard bottom habitats and regional bathymetry data from the NOAA Coastal Relief Model (NGDC 1999). For this report, an updated model predicting the occurrence of hard bottom habitats at 200 x 200 m grid resolution in the New York Bight was created to incorporate the considerable amount of bathymetry data collected in the region over the past decade using multibeam sonar.

8.1 Methods

8.1.1 Hard bottom Occurrence Records

Point locations of hard bottom occurrence were extracted from dbSEABED (Jenkins 2018; see Chapter 10 for description of dbSEABED). Records included in the analysis by Poti et al. (2012) from the U.S. Geological Survey (USGS) usSEABED database (Reid et al. 2005), and from hydrographic survey annotations in NOAA's National Ocean Service (NOS) and U.S. Coast and Geodetic Survey Bottom Type Descriptions database (NOAA NOS 2013) are included in the holdings of dbSEABED.

8.1.2 Environmental Predictor Variables

Environmental predictor variables considered for the model of hard bottom occurrence included measures of depth and seafloor topography, seafloor substrate, and oceanography.
Environmental predictor variables depicting depth and seafloor topography were derived from a synthesis of recent bathymetry data collected using multibeam sonar and older bathymetry data

collected using single beam echosounders. These data were acquired from NOAA NCEI's Bathymetry Data Viewer (NOAA NCEI 2019). To create the bathymetry synthesis, first a gridded data layer at 25 x 25 m resolution was interpolated from the point data collected by single beam echo sounders using the Empirical Bayesian Kriging tool in ArcGIS (Krivoruchko 2012). Kriging is a stochastic interpolation method that uses the assumption that measured values are more similar for neighboring samples than for samples farther away (Tobler 1970) to estimate values at locations that have not been sampled (Cressie 1993). A final gridded bathymetry layer at 25 x 25 m resolution was then created by mosaicking the kriged bathymetry layer with the bathymetry data layers collected by multibeam sonar, with priority given to more recent and higher-resolution multibeam bathymetry data in areas where multiple datasets overlapped. Several environmental predictor variables depicting measures of seafloor topography were calculated from the final bathymetry synthesis (see Table 6.1 in Chapter 6 for detailed descriptions of these variables). The bathymetry (depth) and seafloor topography data layers were each aggregated to the 200 x 200 m resolution of the model grid by calculating the mean of all 25 x 25 m grid cells within each 200 x 200 m grid cell using the Aggregate tool in ArcGIS. By first creating the initial bathymetry synthesis layer and derived seafloor topography layers at 25 x 25 m resolution and then aggregating to the 200 x 200 m resolution of the model grid rather than just creating the bathymetry synthesis at 200 x 200 m resolution, the intent was to capture as much of the fine-scale variability in seafloor topography from the multibeam bathymetry data as possible. Many of the multibeam bathymetry datasets had a native resolution close to 25 x 25 m.

Environmental predictor variables depicting seafloor substrate (see Chapter 10 for a description of these data layers) were initially considered in preliminary models, but were later excluded because model performance declined when these variables were included. This effect on model performance is likely because the data layers depicting surficial sediment composition were interpolated from relatively sparse data, and the scale at which they resolve properties such as sediment grain size was too coarse relative to the locations of hard bottom occurrence. Similarly, environmental predictor variables depicting oceanography (i.e., bottom ocean currents) were initially considered, but later excluded due to the coarse spatial resolution of available ocean circulation models for the New York Bight. A pairwise correlation analysis was used to identify and exclude highly correlated (Spearman rank correlation, $|\rho| > 0.7$) environmental predictor variables. The final set of environmental predictor variables used in the model of hard bottom occurrence included depth, slope, aspect (eastness), aspect (northness), general curvature, plan curvature, and profile curvature.

8.1.3 Maximum Entropy Modeling

Because datasets used to compile point records of hard bottom occurrence did not provide reliable information on the absence of hard bottom (i.e., a sediment grab sample containing unconsolidated sediment does not preclude the presence of adjacent hard substrate), modeling approaches considered were restricted to presence-only methods. Underwater visual surveys (e.g., using ROVs) or data collected using acoustic sonar could provide more reliable absence data for hard bottom, but these data do not exist across much of the New York Bight.

A maximum entropy (MaxEnt) model was used to predict the likelihood of hard bottom occurrence across the New York Bight. MaxEnt is a machine learning approach that estimates functional relationships between occurrence and the environmental predictor variables, with these relationships constrained by the mean value of the environmental predictor variables at the

locations of known occurrences (Phillips et al. 2004, 2006). MaxEnt uses these relationships to predict the relative likelihood of occurrence at all locations in a specified domain. MaxEnt is commonly used in species distribution modeling to create maps of habitat suitability (Elith et al. 2011), and the approach used in this study can be thought of similarly as producing a map depicting the "suitability" of the seafloor for patches of hard bottom substrate. The map of the predicted likelihood of hard bottom occurrence depicts the complementary log-log transform of the raw MaxEnt output (Phillips et al. 2017). It is important to note that it should be considered a measure of relative likelihood rather than a strict measure of the probability of finding hard bottom at a given location.

Model performance was assessed using 10-fold cross validation. Hard bottom occurrence data were divided into 10 subsets, and each subset was used to evaluate a model fit using the data in the other nine subsets. The statistic used to evaluate model performance was the area under the receiver operating characteristic curve (AUC). In addition to calculating the mean prediction of hard bottom occurrence across the 10 cross-validation folds, the coefficient of variation in predictions was also calculated as a measure of variability in predictions.

8.2 Results and Discussion

8.2.1 Hard bottom Occurrence Records

The compilation of point locations of hard bottom occurrence from dbSEABED included 57 records in the New York Bight (Figure 8.1). The relatively low number and the distribution of these records reflect the fact that most of the continental shelf of the New York Bight is covered by sand, with only scattered hard bottom habitats (Poppe et al. 1994; Steimle and Zetlin 2000), but also that there are still many areas in the New York Bight where the seafloor has not been extensively surveyed. While many of the records were located on the continental slope, this does not necessarily suggest that sampling was biased toward the slope. For example, dbSEABED contains numerous survey records on the continental shelf (see Figure 10.1), where few records of hard bottom occurrence existed.

A dataset developed by The Nature Conservancy as part of the Northwest Atlantic Marine Ecoregional Assessment (Greene et al. 2010) was also considered for inclusion in the compilation of hard bottom occurrences. However, based on preliminary model runs, these records were not included. Many of these records were from fisheries trawl surveys, and their recorded spatial locations may represent a >1 km trawl, which is less precise than the spatial resolution of the model. A number of these records are located near the heads of submarine canyons, but the locations of the hard bottom habitats that these records represent may actually be within the submarine canyons.

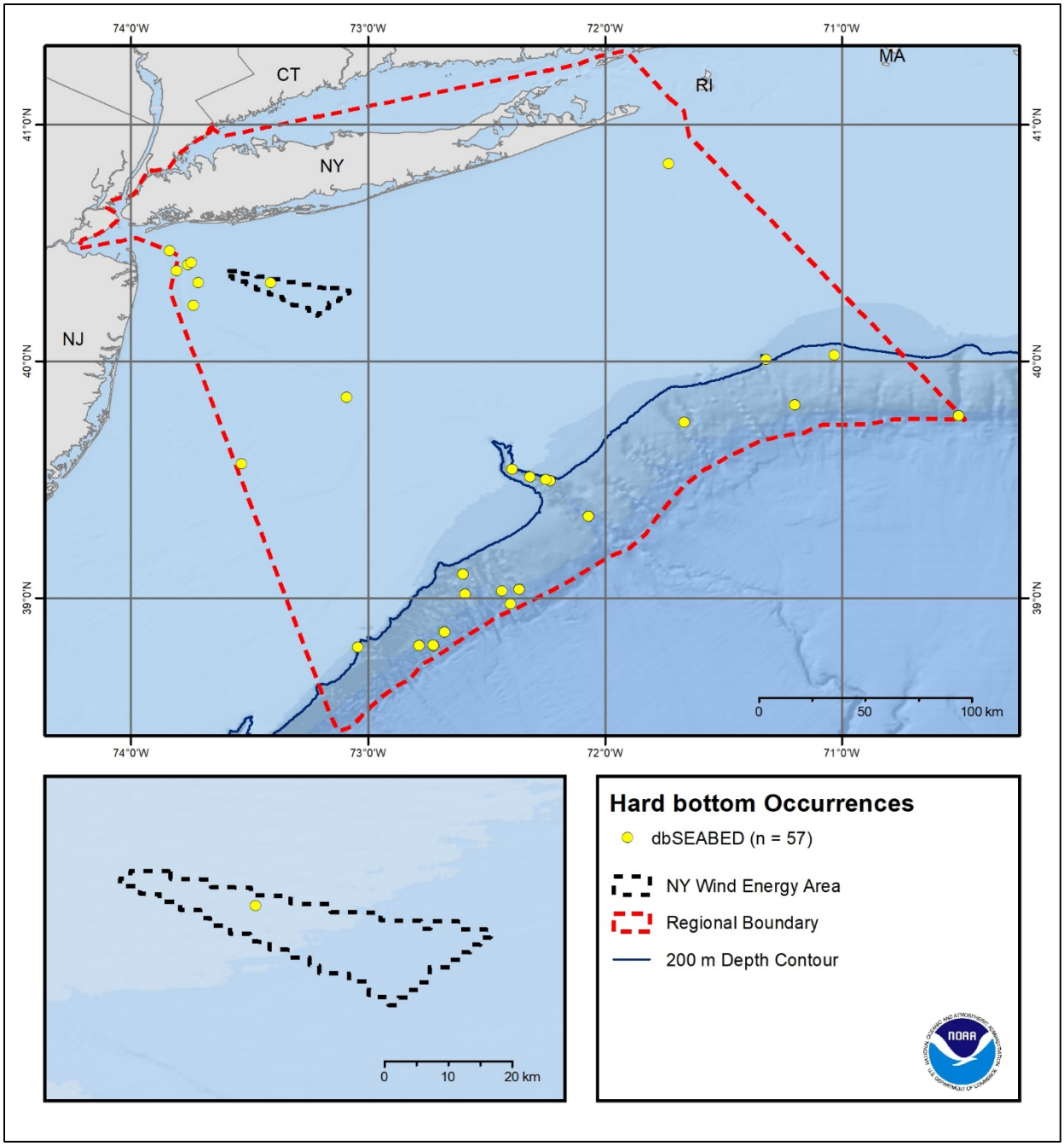


Figure 8.1. Locations of hard bottom occurrences in the New York Bight from dbSEABED.

8.2.2 Predicted Hard Bottom Occurrence

The MaxEnt model of hard bottom occurrence predicted relatively high likelihood of hard bottom in some nearshore areas (e.g., on the edge of Block Channel in the northeast of the study area), on the sides of the Hudson Shelf Valley, and on the slopes of the submarine canyons that incise or partially incise the continental slope (Figure 8.2). Within the NYWEA, the northwest portion was predicted to have a relatively higher likelihood of hard bottom occurrence, as suggested by the backscatter intensity data collected during this study (see Chapter 9). The mean test AUC from cross-validation was 0.84, which indicated that the model performed reasonably well.

It is important to recognize that the model predictions do not provide any estimate or information about the area of predicted hard bottom features or the proportion of hard bottom habitat that may be found in a specific location. Rather, they simply provide a relative measure of how likely hard bottom is to occur in a given grid cell. A grid cell predicted as highly likely to contain hard bottom could conceivably contain mostly sand or other soft bottom, since the patches of hard bottom are likely smaller than the 200 x 200 m resolution of the model. The map of variability in the prediction of hard bottom occurrence suggests that there is relatively higher variability in predictions on the continental shelf as opposed to the continental slope (Figure 8.3), perhaps because hard bottom habitats on the continental shelf are so scattered. The areas with lower prediction variability also correspond to the areas where higher-resolution bathymetry data have been collected with multibeam acoustic surveys.

8.3 Conclusions

With the incorporation of recently collected multibeam bathymetry data, updated maps of predicted hard bottom occurrence were generated at 200 x 200 m resolution, an increase in resolution of approximately16x from the previous maps created by Poti et al. (2012). This improvement, along with the addition of records of hard bottom occurrence in the submarine canyons on the continental slope, allowed for better delineation of features likely to contain hard bottom habitats (e.g., the steep walls of a submarine canyon). However, there are still many areas on the continental shelf that have not been mapped to obtain high-resolution data. It is therefore important to collect additional bathymetry and backscatter intensity data with multibeam sonar, in conjunction with direct sampling of the seafloor, for specific sites where management decisions require information about the distribution of hard bottom habitats. These collections will allow continued improvement in regional models of hard bottom occurrence.

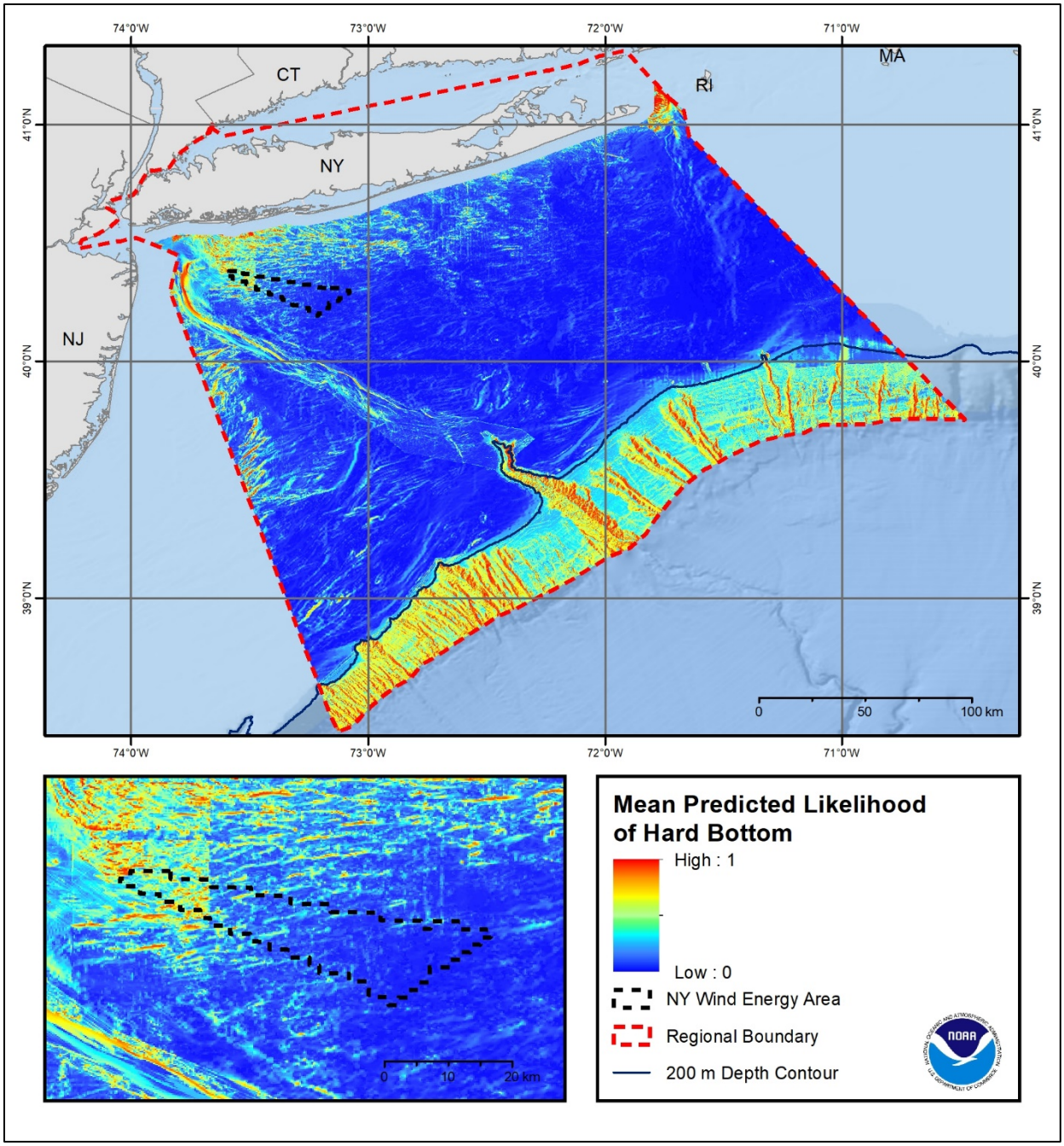


Figure 8.2 Mean predicted likelihood of hard bottom occurrence from a MaxEnt model relating locations of hard bottom habitats to environmental predictor variables derived from a regional synthesis of bathymetry.

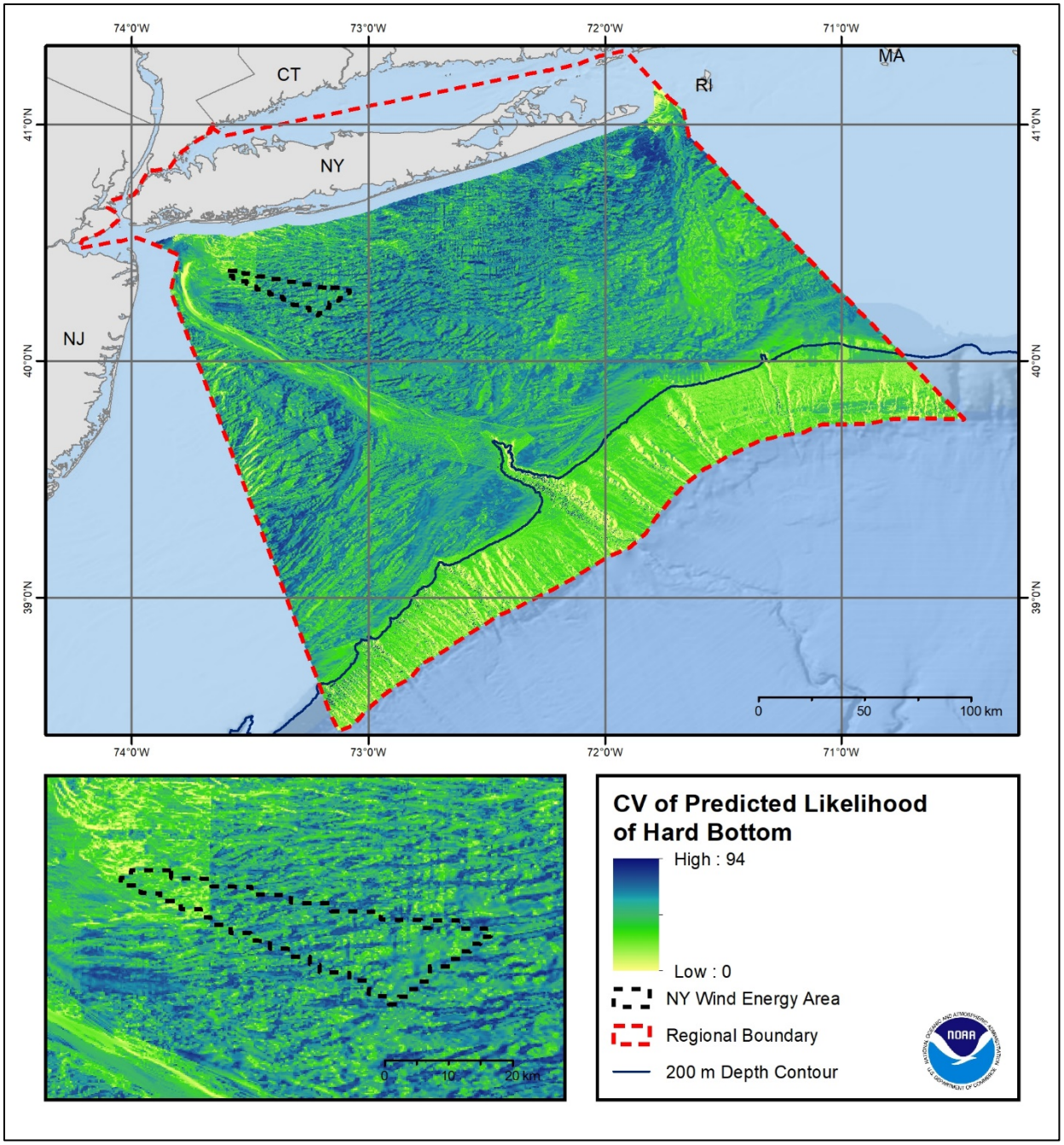


Figure 8.3 Coefficient of variation (CV) of predicted likelihood of hard bottom occurrence from a MaxEnt model relating locations of hard bottom habitats to environmental predictor variables derived from a regional synthesis of bathymetry.

Chapter 9. Sediment Grab Analysis

The substrate of the seafloor is an important component for classifying benthic habitats, which in turn can reveal information about an area's geological and biological origins. For example, the substrate type can influence the shape of geoforms, and the distribution of flora and fauna based on their suitability to substrate types. The compilation of biological and geological components are used to create a unique benthic habitat.

Substrates are typically divided into hard or soft bottom categories. The NYWEA was completely devoid of hard bottom except for the presence of several shipwrecks (see Chapter 2) that are considered anthropogenic substrates (CMECS 2017). The shipwrecks were therefore excluded from the habitat assessment. The backscatter intensity mosaic created from the multibeam sonar surveys (see Chapter 2) was used to map changes and patterns in the soft bottom substrate. Similarly, Goff et al. (2004) used backscatter to classify seafloor substrate types on the continental shelf of New Jersey. However, the intensity of acoustic returns from the seafloor can vary significantly depending on the sensor or platforms used to conduct the surveys, as well as many other physical oceanographic variables (Hughes Clarke et al. 2008). In order to overcome these influences on intensity returns, *in situ* data from sediment grab samples were used to calibrate the intensity values in order to be able to produce a normalized backscatter mosaics.

Comprehensive sediment analysis of NYWEA was conducted through independent, laboratory analysis of sediment grab samples in order to map the grain size distribution (φ, phi) and assess the local and regional substrate components. However, due to the length of time needed to conduct the full laboratory sediment analysis, these results were not available to calibrate the backscatter data or as a predictor dataset for the BRT modeling.

In order to merge backscatter intensity data from different surveys conducted in the NYWEA (i.e., NOAA *Ship Nancy Foster* 2017 and *Ferdinand R. Hassler* 2013), acoustic intensities from the respective surveys were adjusted to provide a normalized backscatter mosaic across the entire study area. The BRT modeling only relied on the presence and absence data and the underwater video analysis from the observations in the field during the R/V *Tiki XIV* survey (see Chapter 7). The results of the comprehensive sediment analysis are the focus of this chapter.

9.1 Methods

9.1.1 Sediment Grab Collection

Sediment analysis was conducted at sample locations spatially coincident with the sites used to provide GV predictors and AA data for the benthic habitat map (see Chapter 3). The sediment grabs conducted on the R/V *Tiki XIV* survey utilized modified methods from the NOAA National Status and Trends Program for the National Benthic Surveillance and Mussel Watch Projects (Lauenstein and Cantillo 1993). Prior to homogenizing the sediment samples, sample stiffness and sorting was observed and recorded as the sample was released from the MVV. Sample "stiffness" was visually assessed and classified by gauging the level of sediment cohesion they retained after being released from the MVV. Sample sediment stiffness was annotated as either "Very Stiff", "Stiff", or "Soft". Sediment stratification (from the surface of the grab sample to the bottom) was also noted for each site and annotated as either "Fine to coarse", "Coarse to Fine", or "None" for a well sorted sample.

Approximately 250 g homogenized sample was taken from each MVV sediment grab site and sent to TDI-Brooks to be sieved for phi analysis. Large shell fragments (>64 mm) and living organisms were carefully removed from the sample prior to placement into a Whirl-pack. The samples were stored in insulated coolers with ice packs to maintain the temperature around 17ºC. The samples were never frozen or dried prior to the analysis as a freeze-thaw cycle of the sample could potentially cause irreversible change in the particle-size distribution due to oxidation and/or agglomeration (Plumb 1981). Once the sediment samples arrived at the TDI-Brooks International laboratory, they were cataloged and refrigerated until they were ready to be analyzed.

9.1.2 Grain Size Analysis

TDI-Brooks International laboratory conducted grain size analysis using sieve and hydrometer methods, per the American Society for Testing and Materials International standards (ASTM 2007). Samples were prepared for analysis by drying overnight in a 105ºC oven, and then disaggregated using a rubber-tipped mortar and pestle to prevent wear down of the sediment. Biogenic substrates present within the sediment, such as broken shell hash and carbonates, were retained in the samples and included in the phi analysis. Each sample was then placed in a stack of gradually tighter sieves, each sieve corresponding to a specific phi size range (Table 9.1). A mechanical sieve shaker was used to shake the stack of sieves. After five minutes of sieve shaking, the sediment retained in each sieve was weighed and compared to the total mass of the sample. The percent distribution of the samples were then calculated from the retained weight for each phi size range.

Table 9.1. Sieve sizes used from TDI-Brooks International Laboratories for grain size analysis.

Sieve Number	Mesh Diameter (mm)	phi Size
NA	64	-6.0
1 1/4 in.	31.5	-5.0
5/8 in.	16	-4.0
5/16 in.	8	-3.0
No. 5	4	-2.0
No. 10	2	-1.0
No. 18	1	0.0
No. 35	0.5	1.0
No. 60	0.25	2.0
No. 120	0.125	3.0
No. 230	0.063	4.0

Grains finer than the No. 230 sieve (0.0625 mm) were collected and transferred to a hydrometer to measure phi sizes 4 to 9.5. A test group of 16 samples was randomly selected to analyze for full phi analysis using the sieve and hydrometer methods to evaluate the overall distribution of coarse and fine grains. Less than 5% of each sample from the pilot group passed through the No. 230 sieve, with most of the pilot samples having less than 2% of the grains analyzed with the hydrometer (Figure 9.1). TDI-Brooks International typically uses the hydrometer test to calculate phi for silt and clays when more than 15% of the sample passes through the No. 230 sieve. However, the content of the silt and clay grains from the pilot test samples did not support the need for further hydrometer testing on the remainder of the samples. The results from the full phi

analysis of the pilot group indicated that the samples were predominantly sand, therefore it was unnecessary to conduct hydrometer tests for all 400 samples. Silts and clays that passed through the dry sieve for each sample were then grouped together and were recorded as "mud" with ≤4 phi.

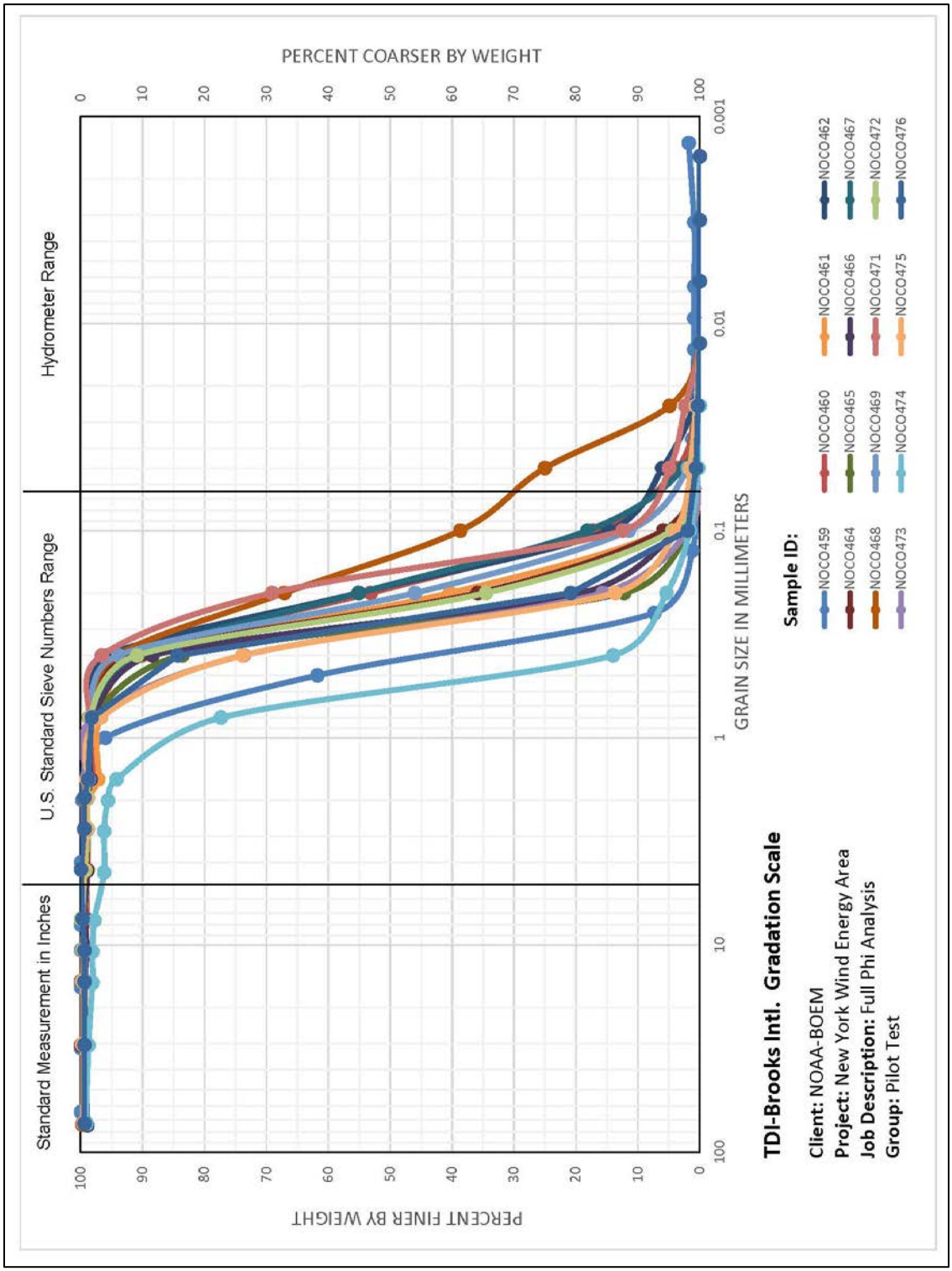


Figure 9.1 Full phi analysis for the pilot test samples using sieve and hydrometer tests.

9.1.3 Sediment Classification and Spatial Analysis

A simple classification method of the substrate type was conducted using the Sedplot software developed by the USGS (Poppe and Eliason 2008). Sedplot classifies sediment samples using the percentages of each phi class to generate ternary diagrams that graphically depict the ratios of muds, sands, and gravels. The program permits the user to select either the Shepard Classification System (Shepard 1954) for fine-grained sediments, or the Folk Classification Scheme (Folk 1954) for more coarsely textured sediments. Since the hydrometer test was not used to get the full percent distribution of silts and clays (>4 phi) for each sample, the Folk Classification System was used to classify the substrate component.

The results from the grain size sieve analysis were then compiled into an ESRI shapefile with the site locations and annotations from the Trimble Geo 7x data during R/V *Tiki XIV* survey and projected into NAD 83 UTM Zone 18N using ArcGIS (version 10.5). The stratification and stiffness of the sediment textures from the observations of the MVV grab samples were also joined to the shapefile (see Section 9.2). The MVV data could then be used to classify the sediment textures, statistically model the grain size distribution, and be visually compared to the backscatter mosaic of the entire NYWEA.

The Geostatistical Analyst toolbox in ArcGIS (version 10.5) was used to produce interpolated surfaces of the phi class distributions from the sediment grab samples using several different methods. These methods were developed to model spatial dependency of continuous data through discrete sample points using semivariogram models (Krivoruchko 2005). The performance of the interpolations can be optimized for the spatial distribution of the data by controlling the size of the search neighborhood and therefore the number of points considered. One interpolation method available is Inverse Distance Weighted (IDW). IDW conducts a direct interpolation of the data while assuming that points closer to each other are more alike than points farther away.

However, IDW is limited in that it does not generate an uncertainty layer for cross validation of model performance. Simple, Ordinary, and Bayesian kriging methods were tested to calculate the statistical models of the point distribution by factoring in the uncertainty and probability of occurrence to create a continuous surface from the sample points. These interpolation techniques also allow the modeler to apply trends in the data, such as the changes in slope and bathymetry as a weighting factor. The summary statistics of each model were analyzed using the Subset Features Tool to choose the kriging method with the lowest RMSE predictions and the lowest standard error. The Bayesian method had the highest correlation of RMSE to standard error curves that validated the uncertainty variability. Bayesian Kriging computes hundreds of semivariograms and automatically chooses the best fit. An IDW and Bayesian Kriging model were compared for each phi class and classified using natural breaks in the data histograms (jenks) to group the grain size distributions into percentage categories. The output from the IDW and Bayesian kriging interpolation methods were compared to the backscatter surface to see if the results made sense and will be discussed in detail in the next section.

9.2 Results

It is generally recognized that sediment grain size distribution is inversely related to distance from shore, wherein gradually finer sediment clasts are transported further away from land and river inputs. However the unconsolidated substrates of the New York bight are relict of past sea

level transgressions and do not conform to modern marine depositional environments (Schlee 1973). Qualitative analysis of the backscatter data collected by the NOAA Ship *Nancy Foster*, the observations from the MVV samples and the quantitative analysis from TDI-Brooks sieve tests provide a better understanding of the distribution, sorting, stratification, and stiffness of these ancient sediments. Once the sediment grab data was projected onto the maps, localized patterns of fine and coarse-grained substrates emerged that correlate with the geoforms detected by the backscatter and morphometrics. The results from these classifications and models are described below.

9.2.1 Substrate classifications from Sieve Test

Sediment sieve results were divided into distinct grain size classes. Folk plots were used to describe the benthic sediment grain-size nomenclature by classifying each sample into one of fifteen major textural groups that were defined by the ratio of gravel, sand, and mud from the sieve test results. This classification scheme provides usable information for determining the final classes for the Habitat Mapping interpretations in Chapter 11. Folk plots of the fine and coarse sediment categories revealed that the sediment texture of the NYWEA is predominantly well-sorted sand with conglomerations of larger clasts in a smaller portion of the sample sites (Figure 9.2a). The Folk Gravel plot (Figure 9.2b) revealed more heterogeneity of sites falling into the "slightly gravelly sand", and six sites in the "sandy gravel" category. Only one site in the northeastern most portion of the NYWEA was determined to be "silty sand" on the Fine Sediment Folk plot (Figure 9.2a).

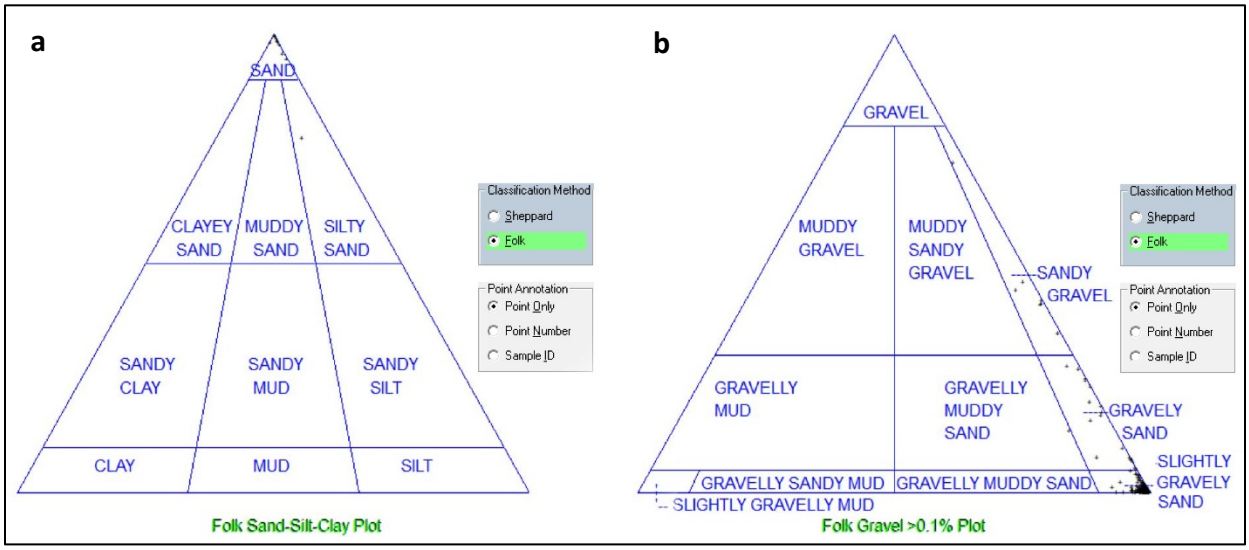


Figure 9.2 Sedplot Folk Diagrams for (a) Sand-Silt-Clay and (b) Gravels.

The Folk Classification system can be useful for categorizing major trends in the data, but it cannot provide any spatial distribution assessment of the distinct sediment categories. Simple classifications based on *in situ* data can mask the real correlation of the substrate to the biotic and geoform components. A more comprehensive approach was implemented to model the percentage of muds, sands, or gravels by the phi class over the study area using geostatistical interpolations as shown in Section 9.2.2.

9.2.2 Grain Size distribution models

Initial spatial distribution of the study area was mapped by creating a pie chart for each sediment sieve result from the grab sample sites in the NYWEA (Figure 9.3). However, geostatistical modeling is a more effective way to illustrate the grain size distribution and reveal spatial trends by interpolating the results of each phi class. It is important to note that these models were chosen to best represent the data spatially, but do not represent actual distinct boundaries for changes in sediment textures. However, due to the sampling strategy used in identifying sample sites within the NYWEA, sampling stations were unequally distributed. This distribution of sites led to under-performance of the IDW model interpolations in areas where sampling was sparse. All of the values that occurred in the sieve test results were displayed in the IDW surface exactly as they were without any uncertainty, which creates “bull’s eye” patterns in the models around sites with relatively higher concentrations of a given class. This was especially comparing the model results with the backscatter base layer imagery showed higher intensity values such as the IDW pebble model (Figure 9.4). The Bayesian Kriging method had better performance with lower RMSE values (Table 9.2) and provided results that coincided with the seafloor signatures from the backscatter. However, the values of the percentage categories for each phi class are relative to the dimensions of the sampling neighborhood, the sample site density, and the uncertainty of the models.

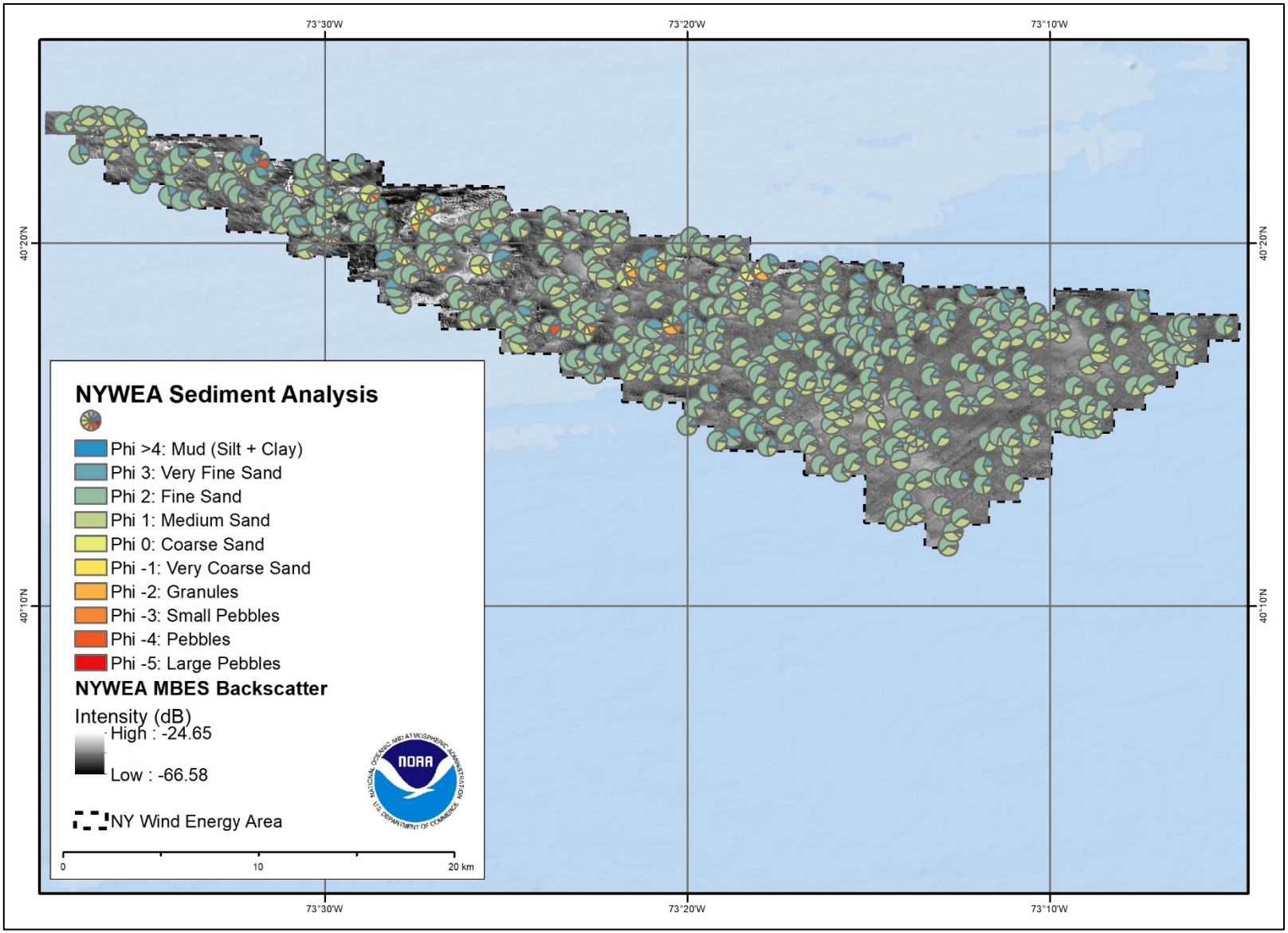


Figure 9.3. Sediment analysis using pie charts for each sample sites.

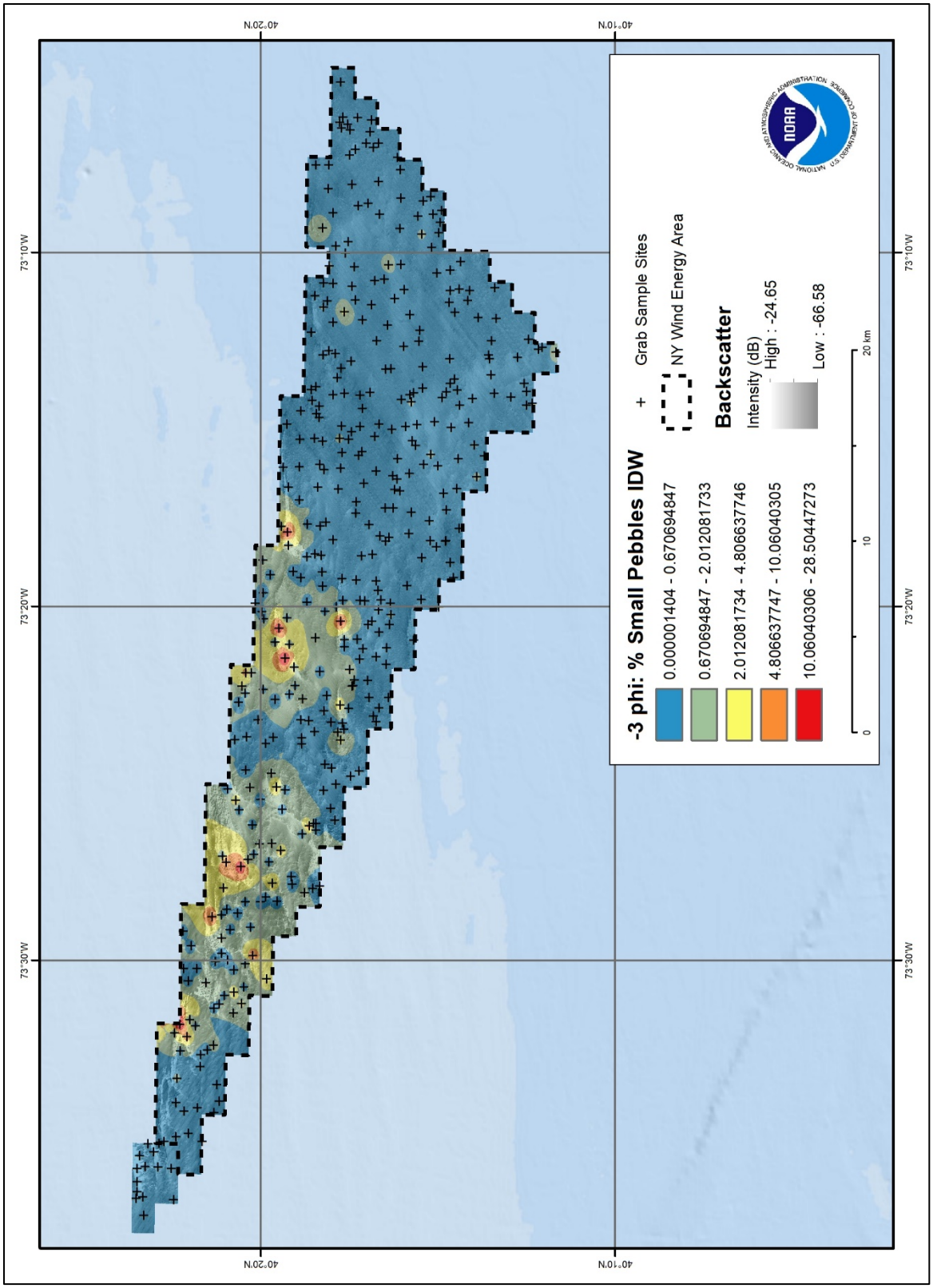


Figure 9.4. Distribution of Small Pebbles (-3 phi) using Inverse Distance Weighting (IDW) from sieve analysis.

Table 9.2 Root Mean Square Error (RMSE) comparisons from IDW and Bayesian Kriging (BayKrig in table) models.

phi	Sediment Class	Max Hood	Min Hood	Sector Type	Angle offset	Radius	IDW RMSE	BayKrig RMSE
4	Mud	15	5	Circle with 4 Sectors	65	5000	1.126854	1.078212
3	Very Fine Sand	15	5	Circle with 4 Sectors	65	5000	10.25833	8.898667
2	Fine Sand	15	5	Circle with 4 Sectors	65	5000	11.71093	14.28582
1	Med Sand	15	5	Circle with 4 Sectors	65	5000	9.528682	9.168033
0	Coarse Sand	15	5	Circle with 4 Sectors	65	5000	3.415909	3.295805
-1	Very Coarse Sand	15	5	Circle with 4 Sectors	65	5000	2.840838	2.700166
-2	Granules	15	5	Circle with 4 Sectors	65	5000	3.689347	3.508178
-3	Small Pebbles	15	5	Circle with 4 Sectors	65	5000	2.708654	2.558305
-4	Pebbles	15	5	Circle with 4 Sectors	65	5000	1.933676	1.893849
-5	Large Pebbles	15	5	Circle with 4 Sectors	65	5000	0.327571	0.3028181

The largest clasts that were analyzed by TDI-Brooks from the R/V *Tiki XIV* survey were Large pebbles (-5 phi), however there were insufficient number of samples to generate an interpolated model. There were also three sites where cobbles (-6 phi) were identified from the MVV samples, however these were not sent to the lab for analysis (see Chapter 7). The Pebble class (-6 phi) was collected in a sufficient number of GV leading to strong model performance. Pebbles were predominantly concentrated in the western-most region of the study area near Cholera Bank, but also were present in a broad belt across the middle of the NYWEA along the 40 m contour (Figure 9.5). The blue area around 40º 20'N was filled in with smaller pebbles (-3 phi) and granules (-2 phi), as seen in Figures 9.6 and 9.7, respectively. Cobbles, pebbles, and granules were grouped together as "gravels" (Folk 1954) and were spatially correlated with the wedge shaped, high intensity areas in the backscatter. These areas had high concentrations of ripple geoforms as were seen in the underwater video in Chapter 7. Gravels and larger pieces of broken shell typically settle in the troughs of ripple formations as finer grain sediments are transported on the ridges (Reineck and Singh 1980). Most gravels were well rounded due to weathering from sediment transport of ancient fluvial channel ways and sorted by shifting currents and major storm events common in the New York Bight (Schlee 1973). The pebbles and other gravelly clasts were dominant in these rippled areas and coincided with high backscatter intensity, but very few samples were collected over these areas due to the unstratified sampling design.

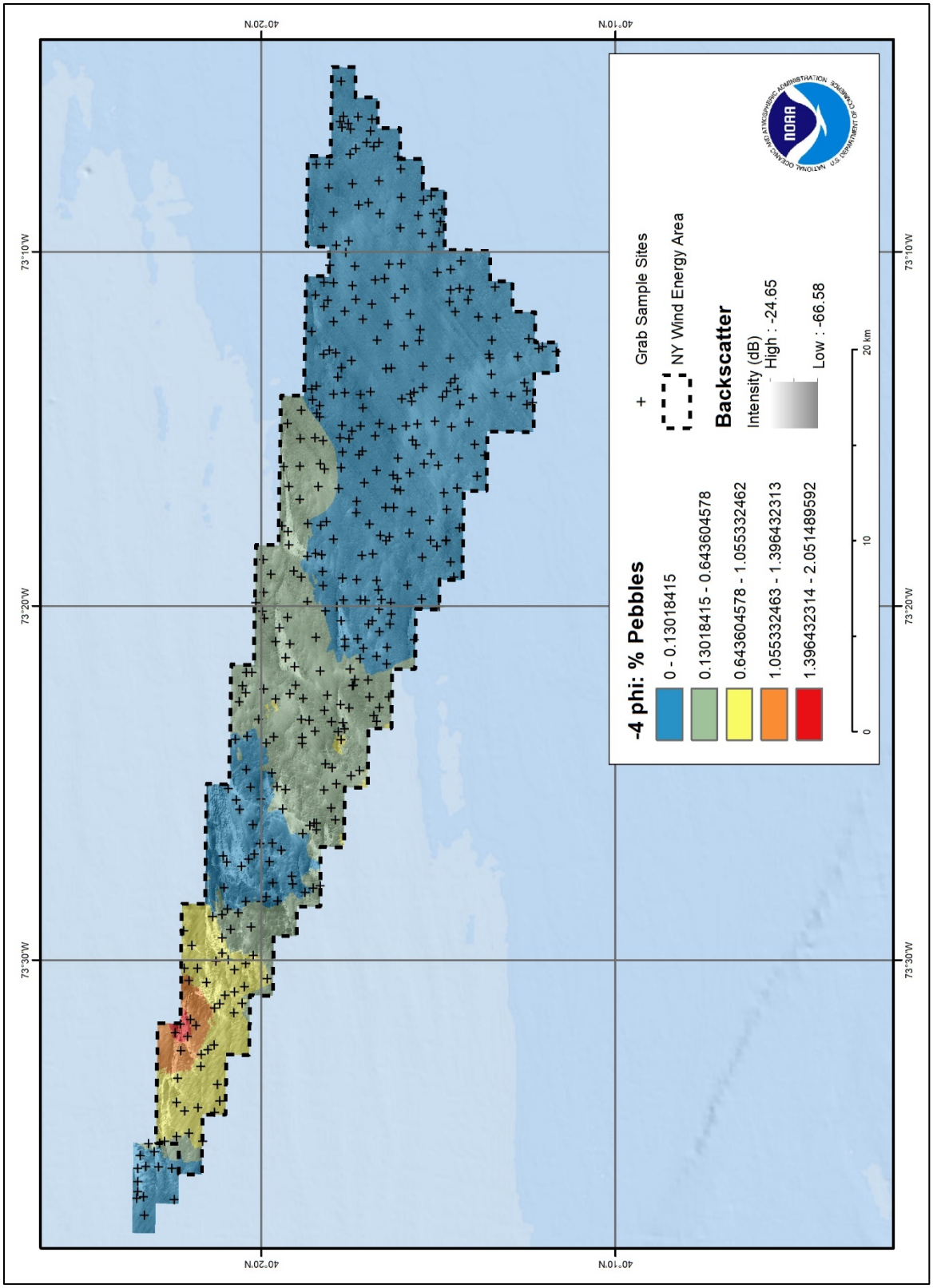


Figure 9.5. Distribution of Pebbles (-4 phi) using Bayesian Kriging from sieve analysis.

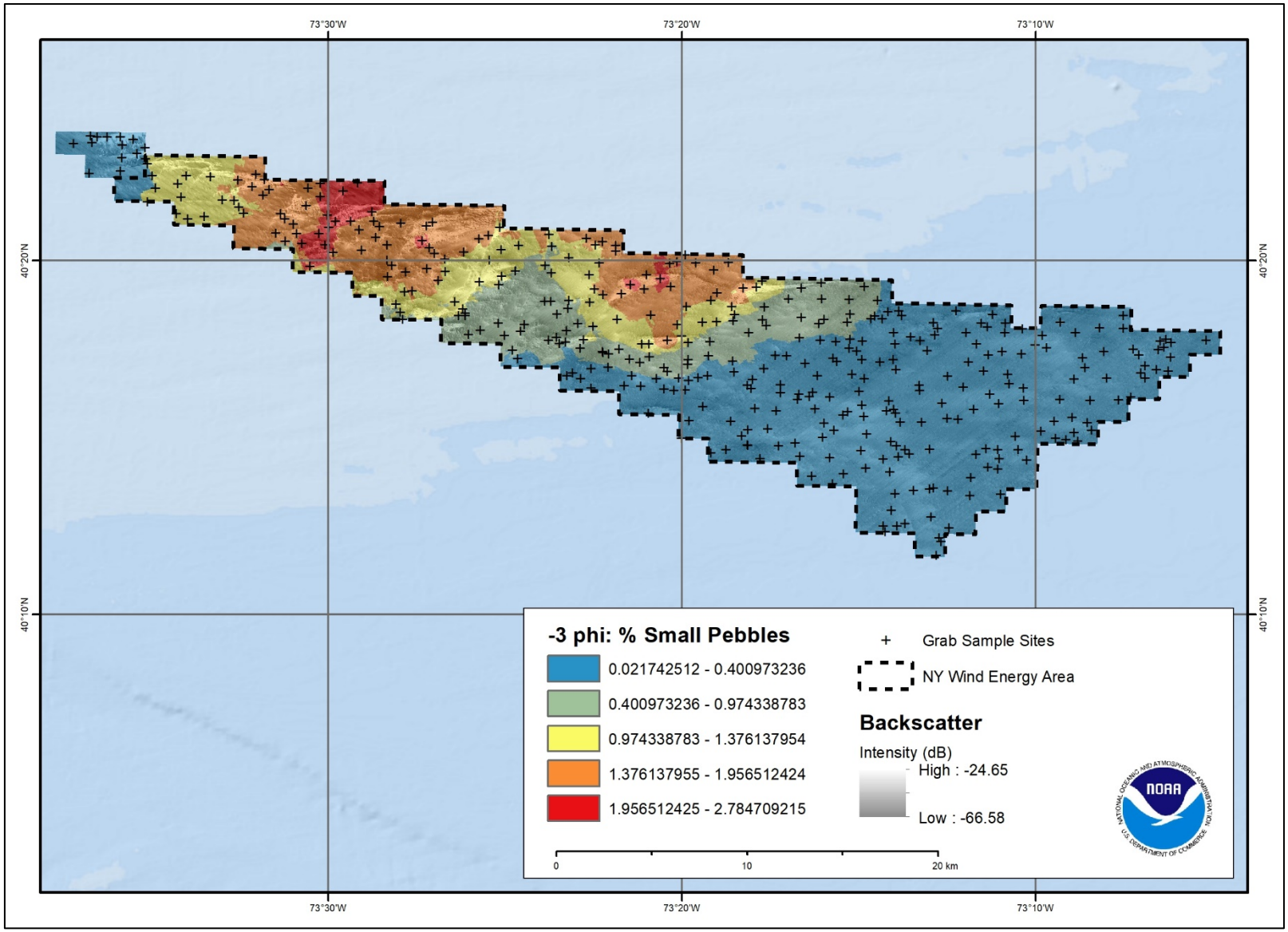


Figure 9.6. Distribution of Small Pebbles (-3 phi) using Bayesian Kriging from sieve analysis.

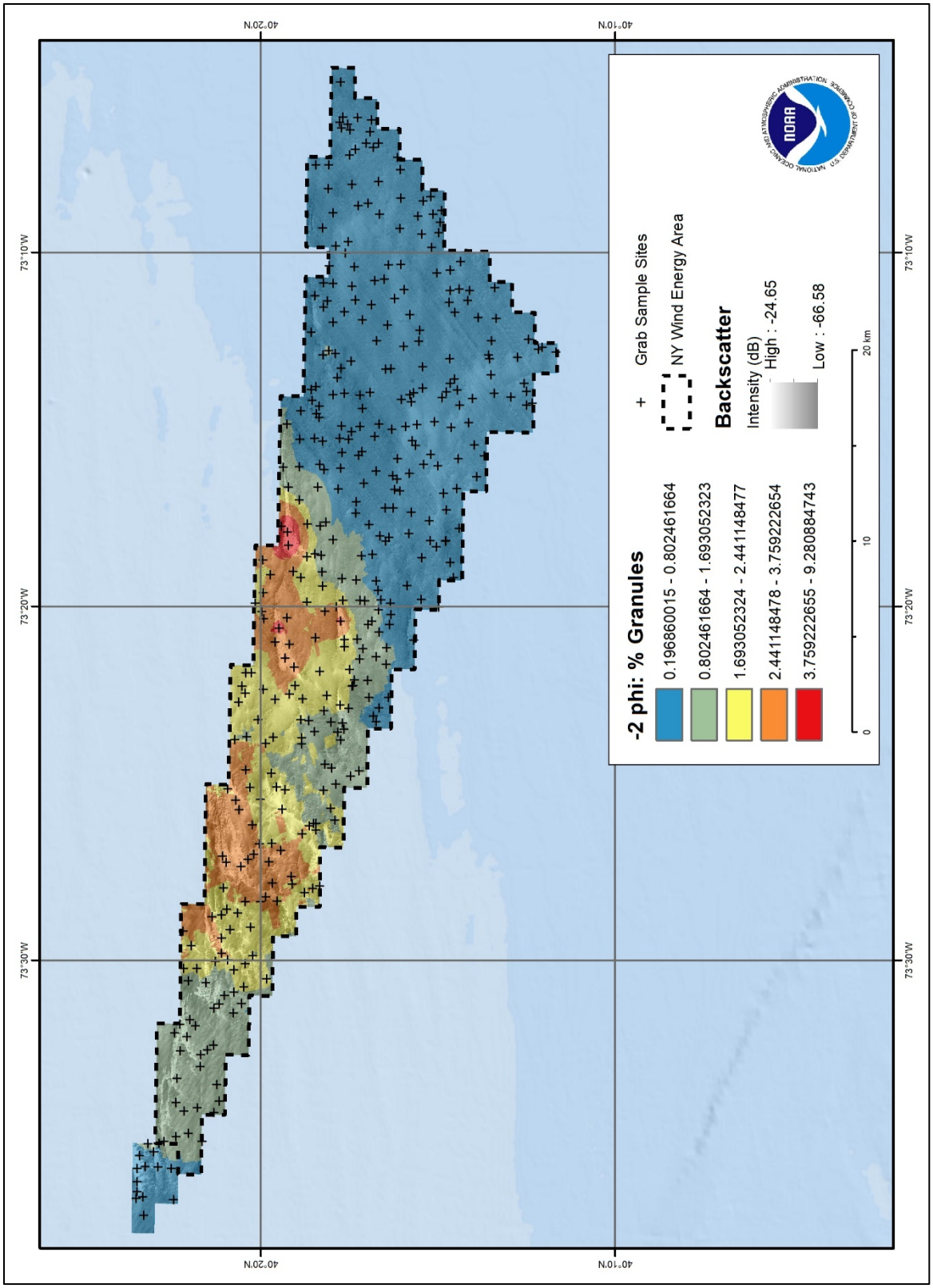


Figure 9.7. Distribution of Granules (-2 phi) using Bayesian Kriging from sieve analysis.

Prior to geostatistical or predictive modeling, the results from sediment sample (i.e., phi analysis, Figure 9.1, and the Folk Plot of the entire dataset, Figure 9.2) indicated that the NYWEA was predominately composed of sandy substrates. The coarse (0 phi) and very coarse (-1 phi) sands were generally concentrated around the wedge-shaped gravel zones at the 40º 20'N latitude (Figure 9.8 and 9.9). However, there were large areas in the eastern portion of the NYWEA where the models indicated strong presence of coarse sand and medium sand (1 phi) (Figure 9.10). These areas were about 10 km away from the gravel zones in the study area, but did show stronger backscatter intensities in the backscatter mosaic than the surrounding fine sands (2 phi) (Figure 9.11). These differences in the intensities were largely due to the interplay of changes in the substrate texture, the shift from highly rippled sands to larger megaripple geoforms, and the reworking of sediments from faunal beds of sand dollars (*Echinarachnius parma)* and other benthic invertebrates. The interpretation of these substrates concur with the GV observations described in Chapter 7.

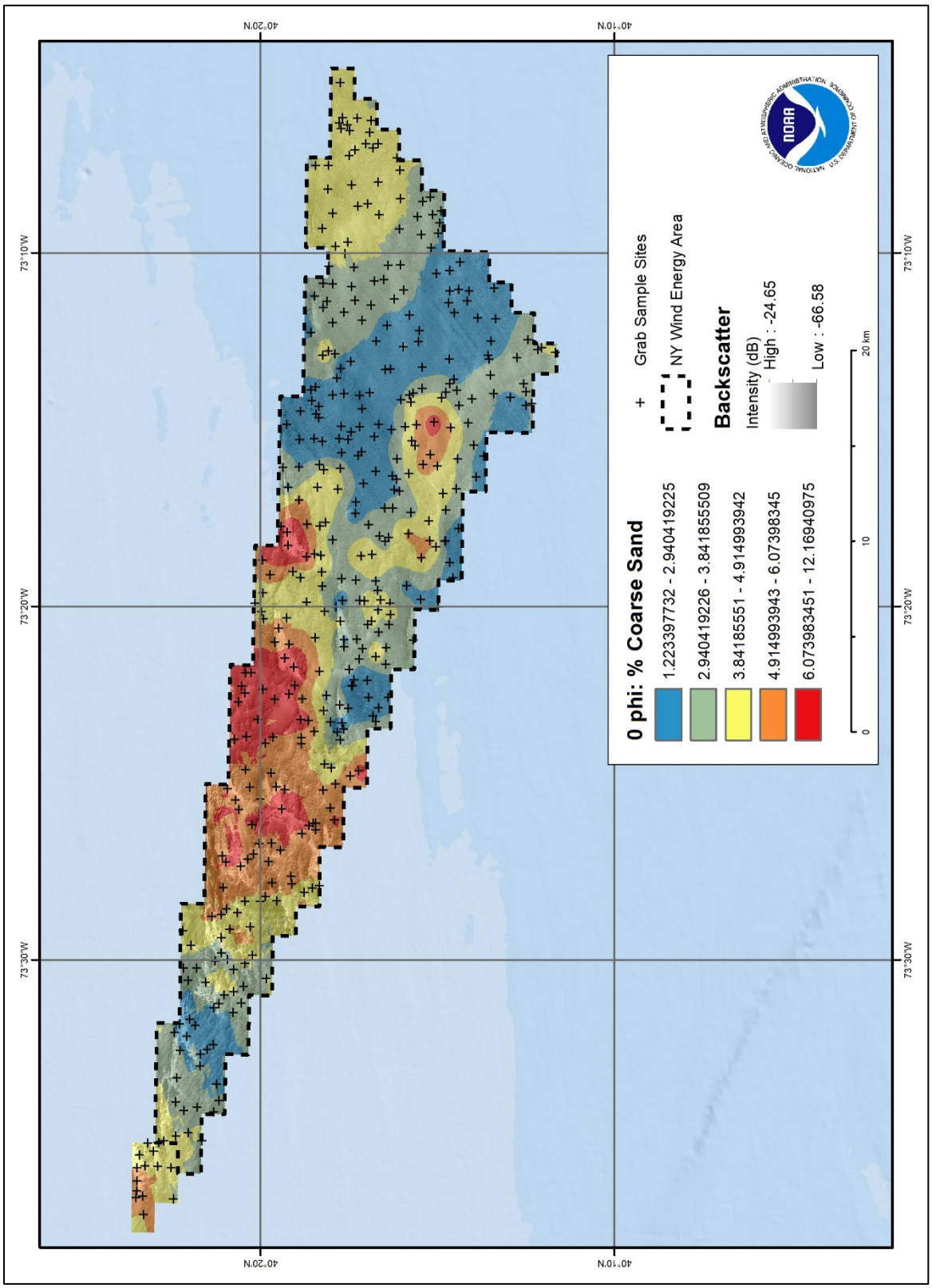


Figure 9.8. Distribution of Coarse Sand (0 phi) using Bayesian Kriging from sieve analysis.

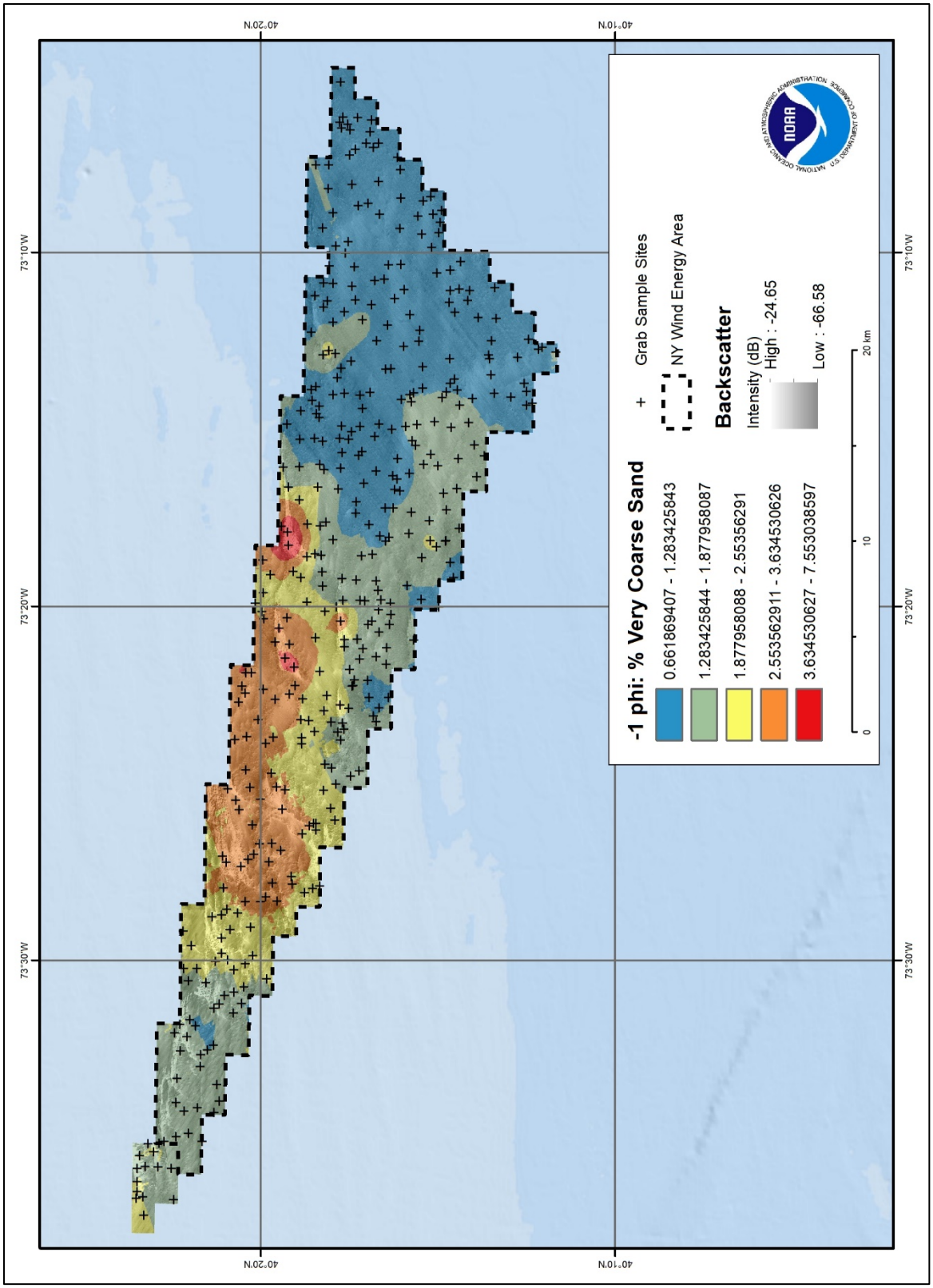


Figure 9.9. Distribution of Very Coarse (-1 phi) using Bayesian Kriging from sieve analysis.

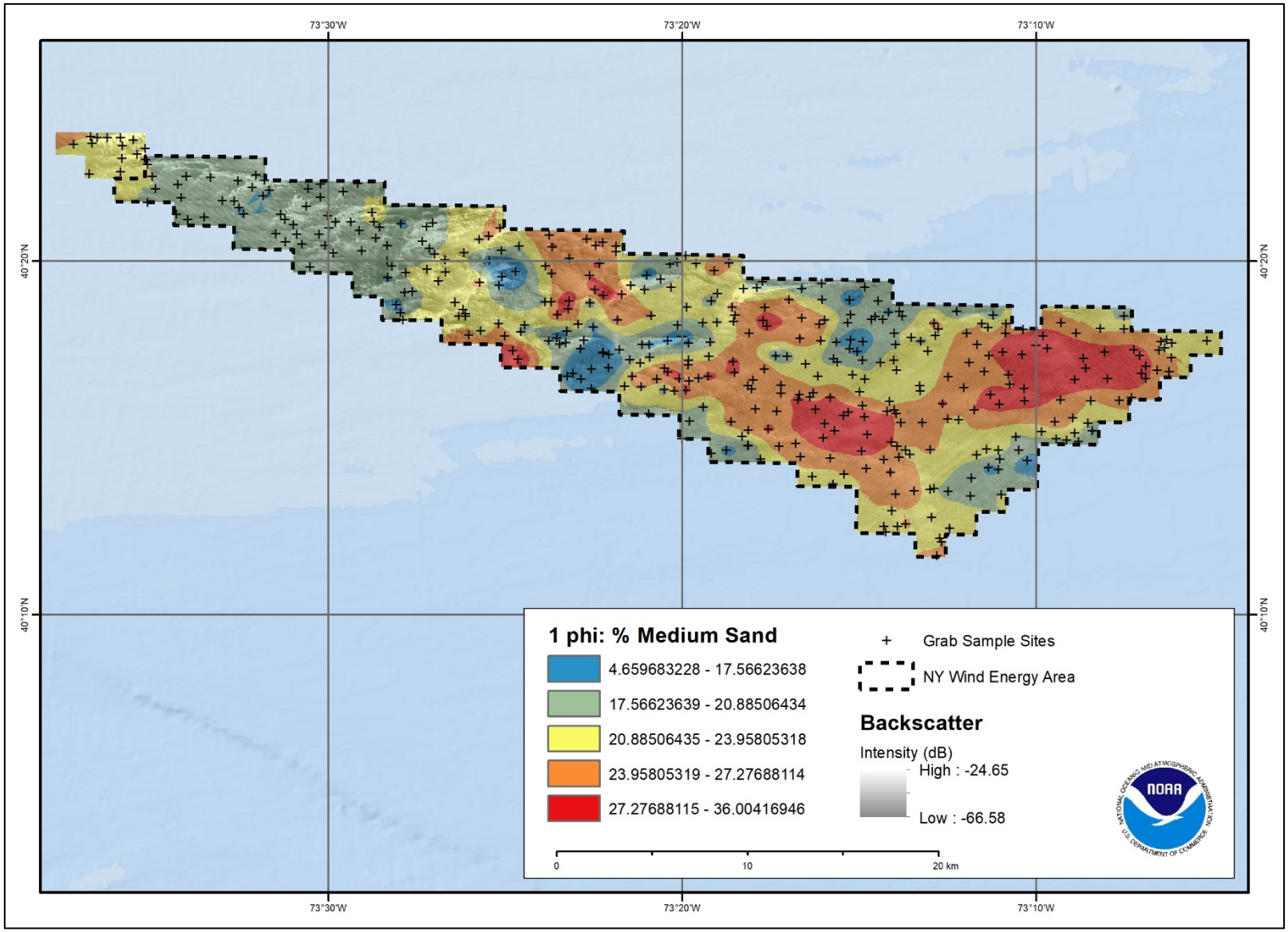


Figure 9.10. Distribution of Medium Sand (1 phi) using Bayesian Kriging from sieve analysis.

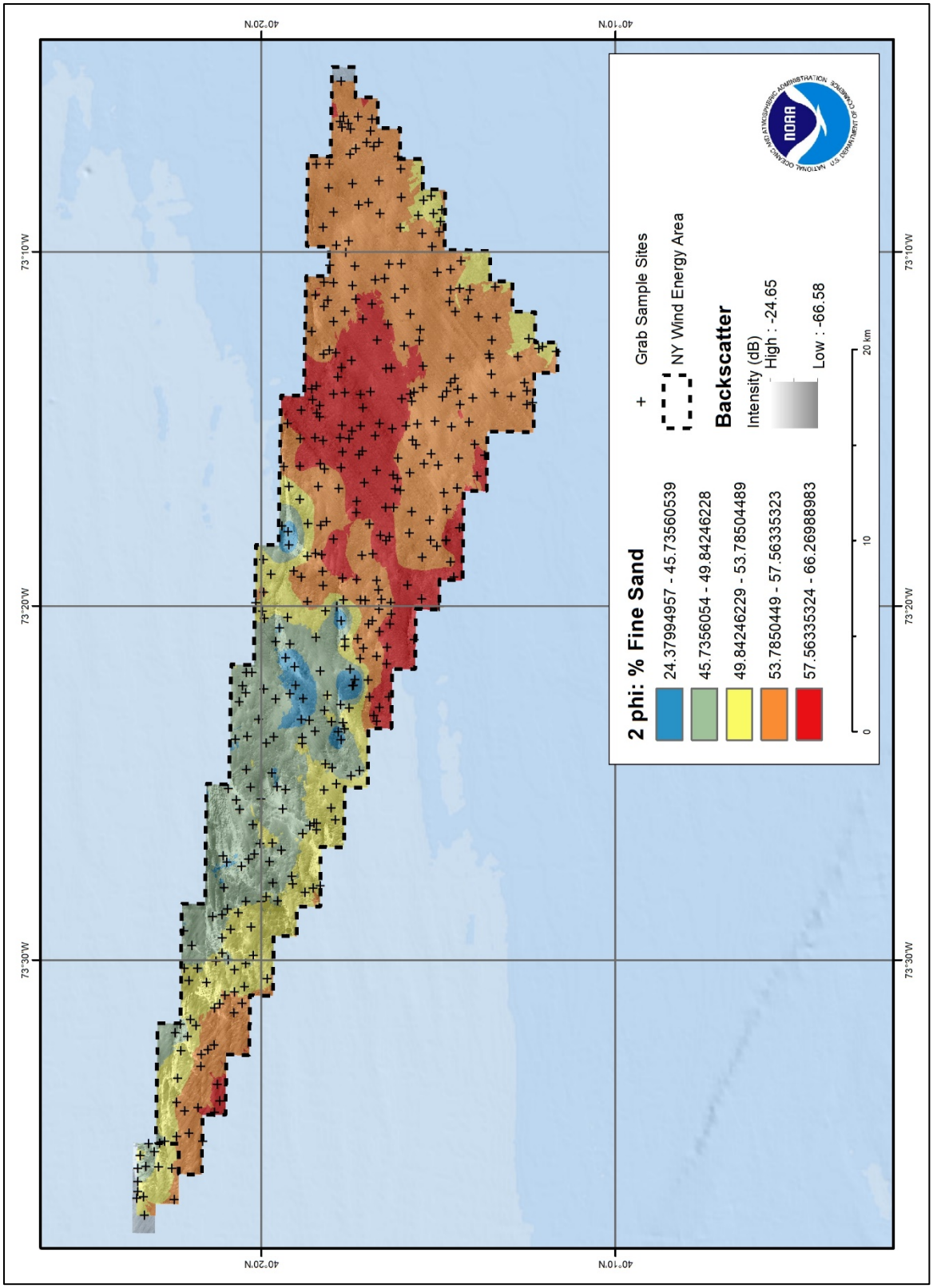


Figure 9.11. Distribution of Fine Sand (2 phi) using Bayesian Kriging from sieve analysis.

The sediment models for very fine sands (3 phi) showed no apparent spatial relationship with distance from shore, geoform patterns, or backscatter intensities (Figure 9.12). It is also important to note that very fine sands grain size class had the second highest RMSE value from the Bayesian Kriging. Very fine sand was concentrated on the western side of the study area near the gravel beds. However, there were insufficient number of sample sites directly on the gravel deposits in this area, so the interpolations may misrepresent the sediment distribution in this particular area. Very fine sand was also prominent on the eastern side of the NYWEA where the geoform is mainly mega ripples. Typically, very fine sediment particles are deposited at slower rates than coarser grain material, but are easily suspended and transported by currents through the saltation process (Reineck and Singh 1980). Thus, the grain size distribution of very fine sediment particles can change in dynamic flow regimes from storm events or shifts in currents. Muds ($\geq$4 phi) had similar distributions as the very fine sand across the study area, however it was less abundant in the sample collections.

The darker shaded regions of the backscatter intensity mosaic coincided with the very fine sands and mud distribution on the eastern side of the study area, as well as surrounding the gravelly areas. This likely resulted from the sonar waveforms being absorbed or scattered by surficial fine sands (Figures 9.12 and 9.13) and muds instead of reflection from coarser grains, giving the imagery lower decibel values and darker shades in the mosaic (Fonseca and Mayer 2007). On average, mud ($\geq$4 phi) had the lowest percent distribution of each sample throughout the sediment sampling survey, but it actually had the highest RMSE of any grain size category. This may be due to the very low percentages of mud in the samples. The highest concentration of samples containing mud were located in the northwestern most corner of the NYWEA, surrounding the gravelly zones (Figure 9.13). The mud deposits around the gravel beds may be a result from the weathering of larger clasts and broken shells in these higher energy areas.

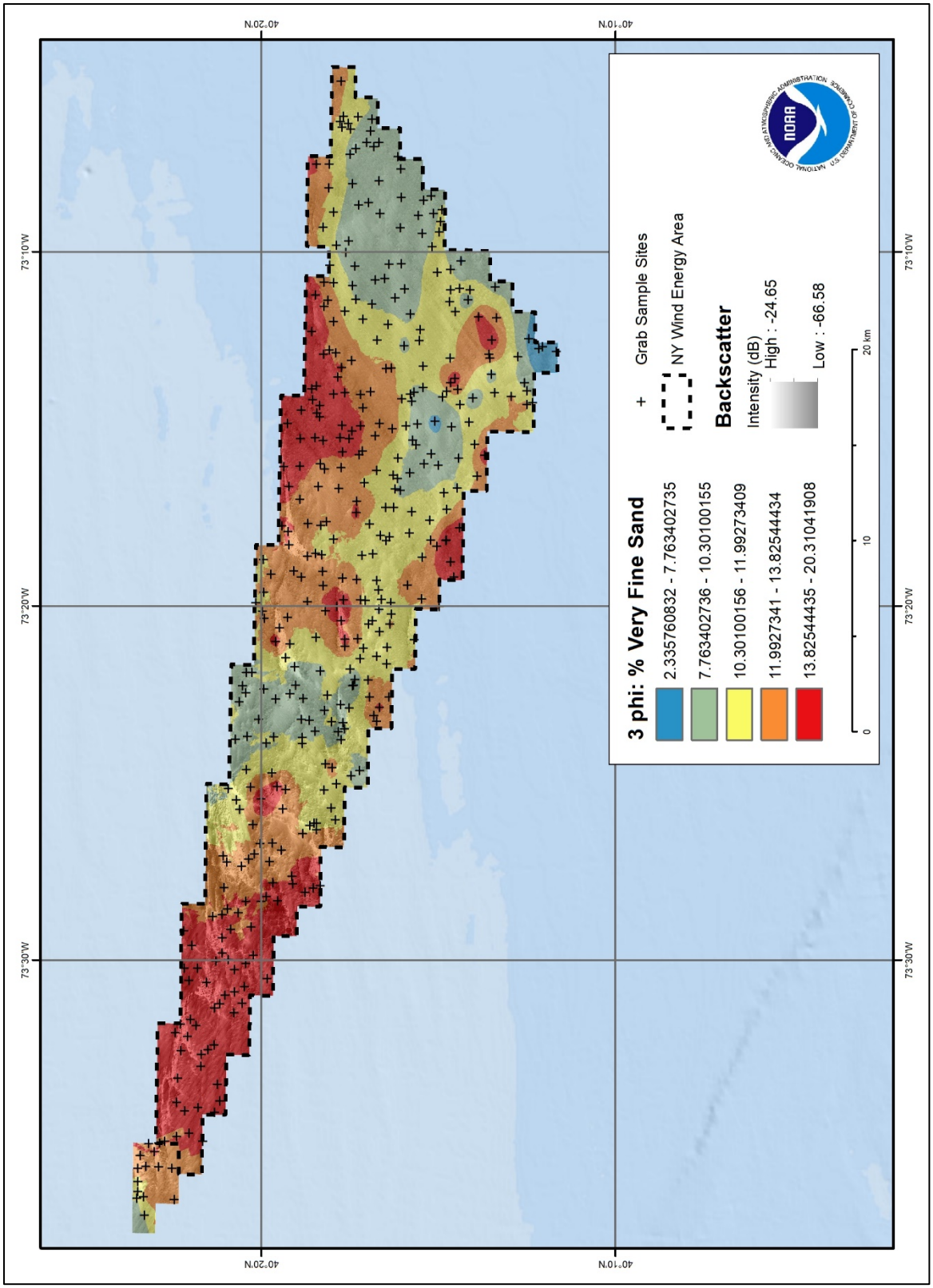


Figure 9.12. Distribution of Very Fine Sand (3 phi) using Bayesian Kriging from sieve analysis.

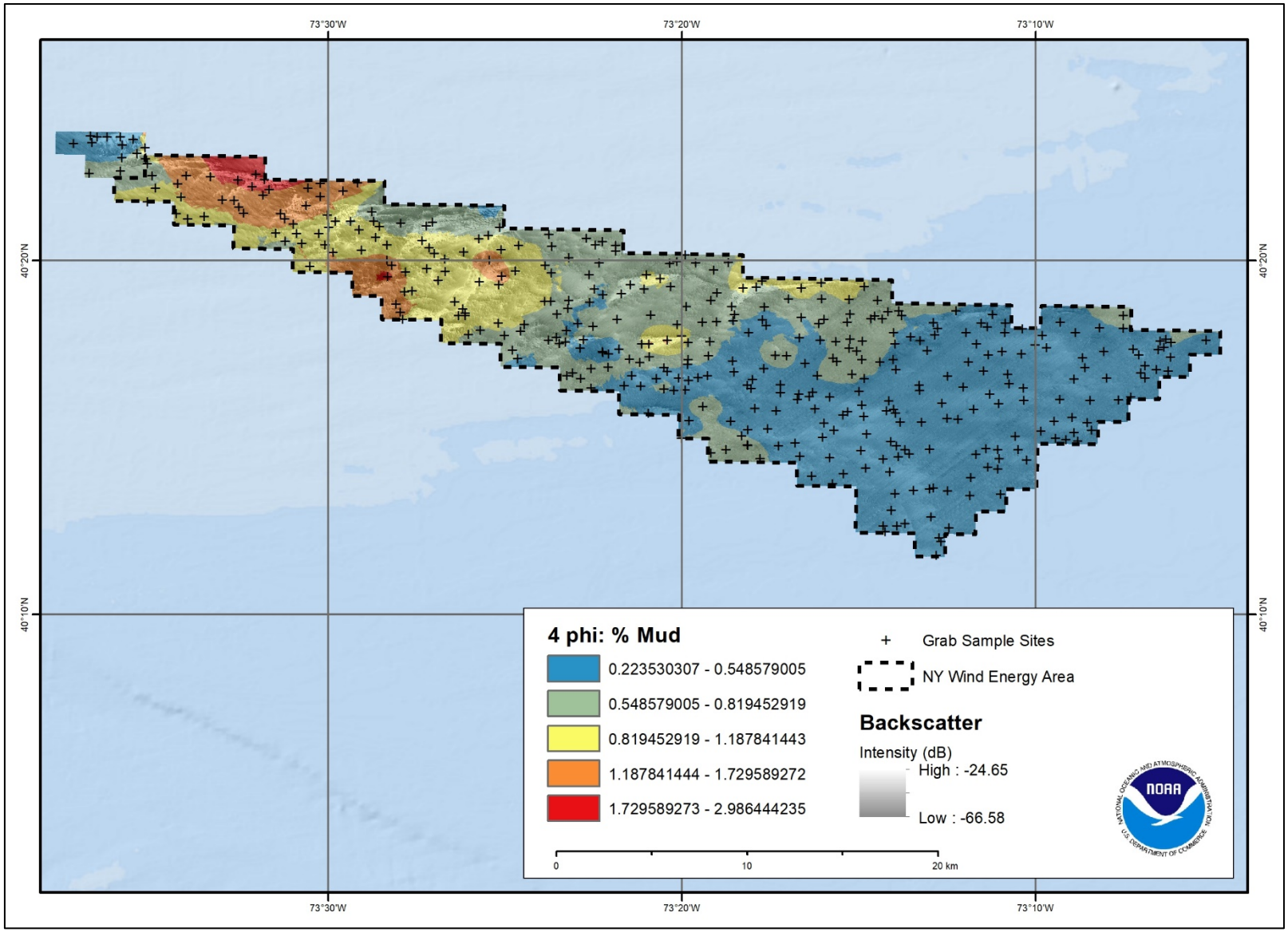


Figure 9.13. Distribution of Muds (4 phi) using Bayesian Kriging from sieve analysis.

9.2.2 Sediment Stiffness and Stratification

Substrate stiffness and stratification can have a direct influence on the absorption or scattering of sound waves from multibeam sonar systems, resulting in different angular responses to the returning echoes (Fonseca and Mayer 2007). Seafloor sediment stiffness and stratification may also have relevance for the engineering, design, and siting of offshore wind turbine platforms and anchoring. The stiffness, or stability of the sediment, is controlled by the cohesive forces between the grains and the pressure of the fluids in the pores (Strout and Tjelta 2005). These properties can have a fundamental influence on the engineering behavior of the substrate. Sediment stiffness was qualitatively measured in the field during the R/V *Tiki XIV* survey by observing the cohesion of the sediment samples as they were released from the MVV. The results from the stiffness test were overlaid on to the backscatter mosaic in Figure 9.14. The sediment stiffness did not appear to have a direct spatial correlation with patterns in the backscatter mosaic or the distribution of different grain sizes. The majority of the samples were characterized as being “stiff”. The lack of geo-physical instruments that could properly measure the stiffness of the sediment samples during the R/V *Tiki XIV* survey may have skewed the stiffness results. Quantitative results of Sediment stiffness can be quantitatively measured using a Free Falling-Cone Penetrometers (FFCP) that can detect *in situ* sleeve resistance, pore pressure, tilt, and even temperature (Stegmann et al. 2006). However, implementation of the FFCP method was beyond the scope of this report.

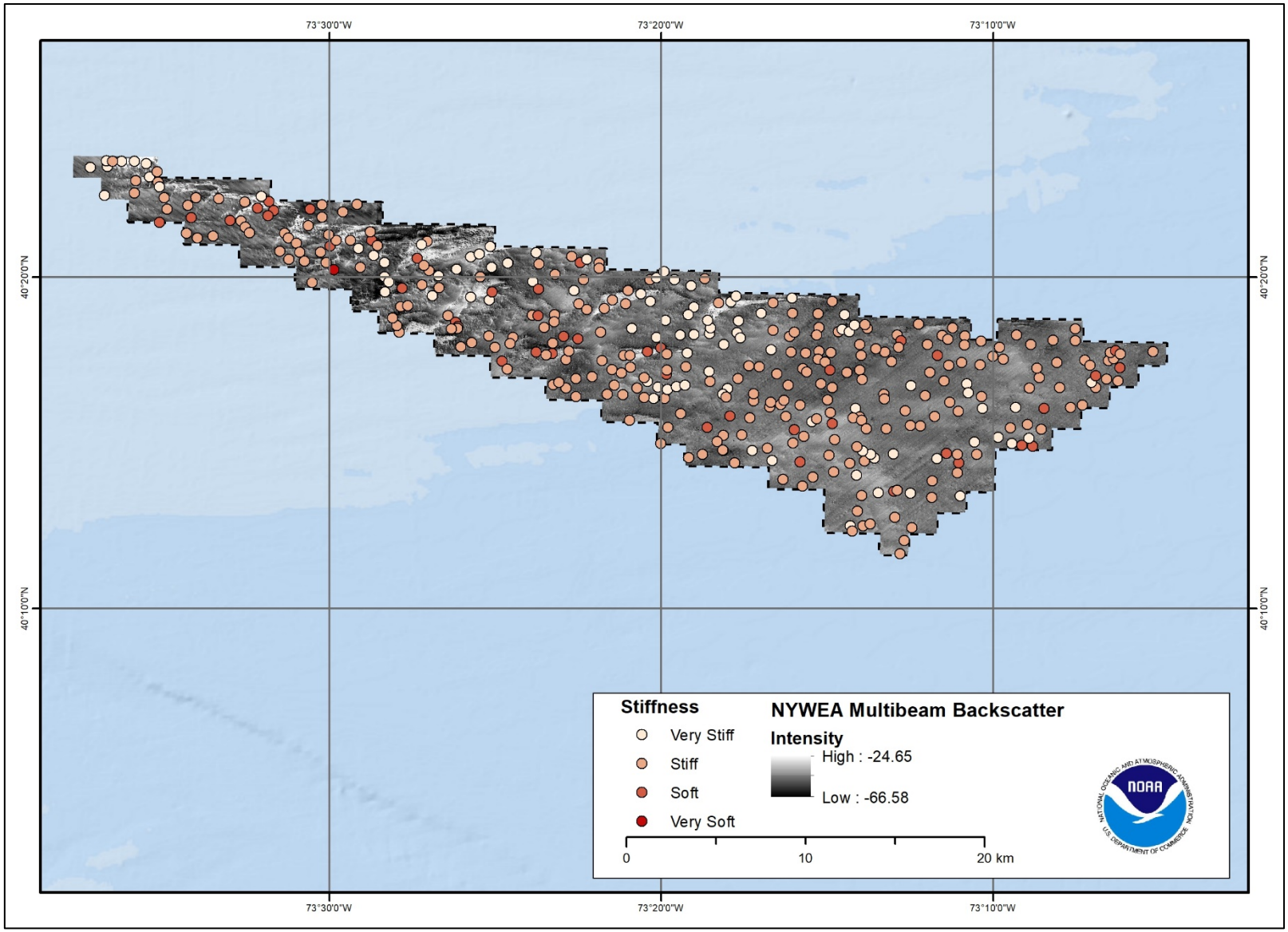


Figure 9.14. The stiffness of sediments measured visually during the R/V *Tiki XIV* survey.

Most of the surficial sediments of the New York Bight are restricted in their sorting range, and typically have very little stratification (Schlee 1973). Sediment samples collected within the NYWEA confirmed that the sediments were very well sorted (Figure 9.15). The findings showed that only 44 of the 390 total sites had any sediment stratification present in the profile (approximately 10 cm) taken by the MVV (11% of total number of samples). A “fine to coarse” profile was only observed at four sites. These samples distinctly occurred at sites on the wedge-shaped gravely deposits identified in the backscatter mosaic. This correlation agrees with the principle that these highly rippled areas have sorted finer clasts on top of coarser grains. Forty sites contained a “Coarse to Fine” sediment profile. These sample sites were mainly located in the eastern half of the NYWEA where there was mostly fine and medium sands. The deposits of coarser sediments on top of the finer materials may be due to reworking of sediments and migrations of megaripples from currents and strong storm systems (Reineck and Singh 1980).

The sedimentary processes and the depositional environment of the NYWEA are generally representative of the New York Bight continental shelf, at large.

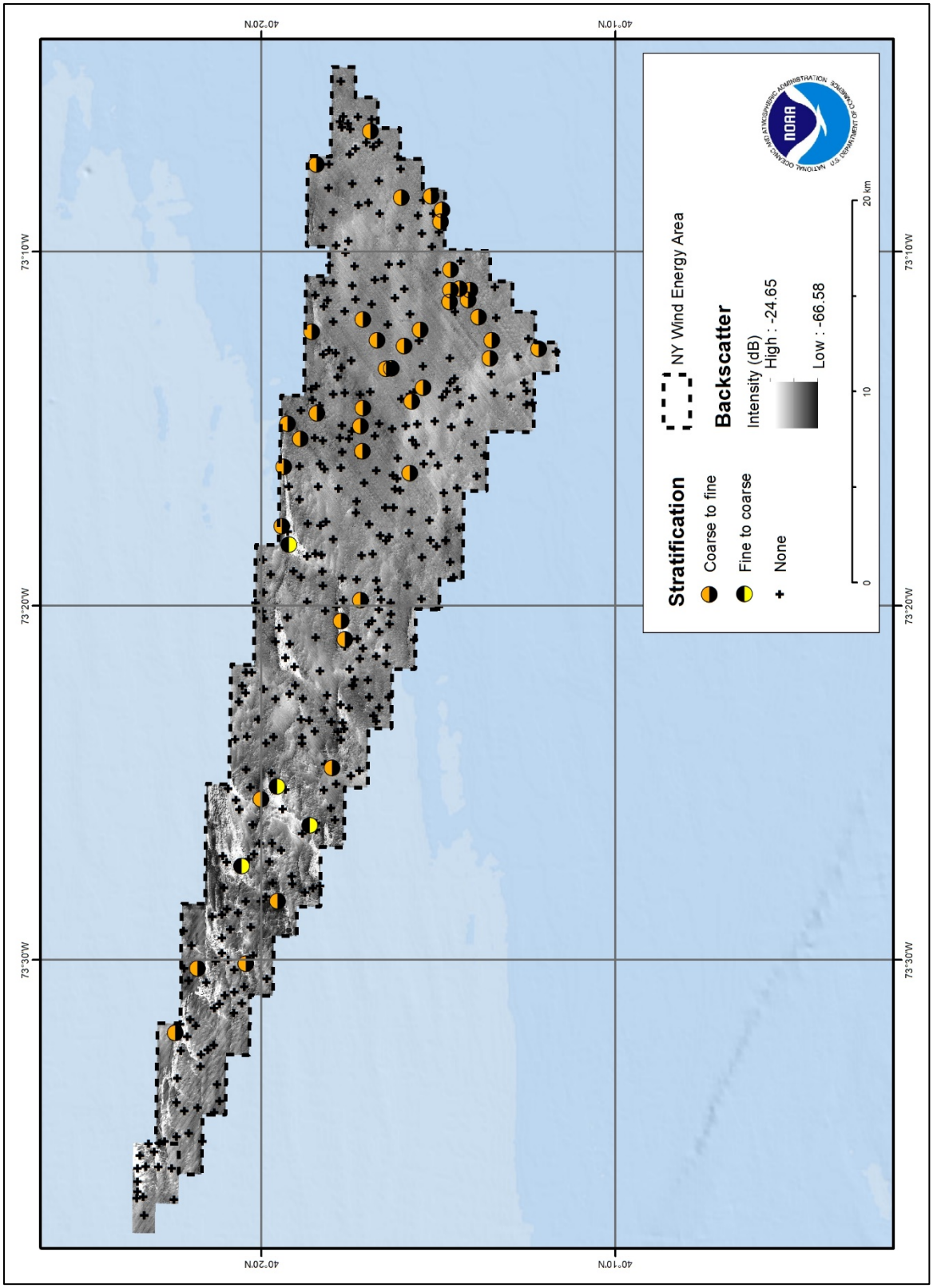


Figure 9.15. The stiffness of sediments measured visually during the R/V *Tiki XIV* survey.

Chapter 10. Sediment Texture Analysis

In addition to maps of hard bottom habitats (Chapter 8) and detailed maps of benthic habitats (Chapter 11), maps of the distribution and characteristics of surficial sediments are also critical for informing the management of marine resources, particularly activities that may affect benthic habitats. Soft bottom substrates can provide habitat for invertebrate communities and demersal fishes, and information on the composition of sediments on the seafloor can be used along with information from fisheries surveys to identify and describe Essential Fish Habitat (Lathrop et al. 2006). Maps of seafloor features are also needed for decisions about offshore engineering activities, such as where to site offshore wind installations.

Similar to other continental shelf regions, the seafloor features of the New York Bight reflect both its geologic history and the influence of dynamic oceanographic and sedimentological processes. Specifically, the actions of those processes on deposits during the >100 m rise in sea level following the end of the last glaciation resulted in the present distribution of surficial sediments (Williams et al. 2006; Goff et al. 2008). While the surficial sediments on the continental shelf of the New York Bight are predominantly sand, transitioning to silt and clay in deeper areas, there are scattered areas of exposed rock, cobble and gravel habitats, and anthropogenic features like dredge disposal sites and artificial reefs (Poppe et al. 1994; Steimle and Zetlin 2000). A prominent seafloor feature in the New York Bight is the Hudson Shelf Valley, extending from the mouth of the Hudson River across the entire continental shelf. It is the only valley on the Atlantic continental shelf not completely filled with sediment following the last glaciation (Freeland et al. 1981; Butman et al. 2003). The Hudson Shelf Valley connects to Hudson Canyon, the largest of a number of submarine canyons that incise or partially incise the continental slope (Butman et al. 2006).

Although some areas of the New York Bight have been surveyed extensively (e.g., Hudson Shelf Valley, Hudson Canyon), for most areas the spatial coverage of sediment sampling data is limited and not sufficient to capture the many spatial scales of variability in surficial sediment characteristics (Goff et al. 2008). To overcome this limitation, seafloor substrate has been mapped and characterized using sidescan sonar (Lathrop et al. 2006) and backscatter intensity from multibeam sonar (Butman et al. 2006; De Falco et al. 2010; Brown et al. 2011). However, Goff et al. (2008) stress that considerable sediment sampling is still needed to calibrate the relationship between acoustic sonar data and sediment properties as this relationship will vary by region.

Several studies have demonstrated that statistical models can be used to create continuous maps of surficial sediment composition from sediment sample data in large, often heterogeneous, legacy databases. For example, Goff et al. (2008) used ordinary kriging to map mean grain size on the Long Island shelf from sediment data in the USGS usSEABED database (Reid et al. 2005). Similarly, Acharya and Panigrahi (2016) used Bayesian kriging to map the distributions of sediment organic carbon and phosphorous on the Eastern Arabian Shelf, and Bockelmann et al. (2018) used kriging to map mud content and grain size in the North Sea. Stephens and Diesing (2015) and Diesing et al. (2017) mapped sediment composition for a large area of continental shelf across parts of the North Sea, English Channel and Celtic Sea, but did so using random forest modeling, a machine learning method. In the New York Bight, Poti et al. (2012)

used geostatistical methods similar to Goff et al. (2008) to develop maps of surficial sediment mean grain size and composition (percentages of mud, sand, and gravel) from sediment data in the USGS Atlantic Coast usSEABED database (Reid et al. 2005).

This chapter presents updated maps of surficial sediment grain size and composition for the New York Bight. The new maps are at higher spatial resolution and extend slightly farther offshore than those created by Poti et al. (2012), and were generated using the data and interpolation methods of dbSEABED, a processing system that integrates seabed datasets to map and analyze seabed properties (Reid et al. 2015; Jenkins 2018). The characterization of surficial sediments for the entire New York Bight provides context for how the seabed within the NYWEA conforms with the broader region. For example, the regional maps provide additional information about the distribution of surficial sediments in areas outside the NYWEA that may be impacted by activities within it (e.g., cabling associated with offshore wind installations) or that may receive future consideration for activities overseen by BOEM. A more detailed characterization of the seabed within the NYWEA is provided in Chapter 9.

10.1 Methods

10.1.1 Surficial Sediment Data

The dbSEABED processing system integrates a multitude of individual datasets, both legacy and modern collections, into a unified database that can be used to map properties of the seafloor (Reid et al. 2005, 2006). The data holdings in the dbSEABED system for the Atlantic continental margin are considerable due to its role in the USGS usSEABED project and include all the records in the Atlantic Coast usSEABED database (Reid et al. 2005). Other notable large collections in the dbSEABED system are the USGS East Coast Sediment Texture Database (Poppe et al. 2014), NOAA's NOS and U.S. Coast and Geodetic Survey (now the National Geodetic Survey) Hydrographic Surveys database (NOAA NOS 2013), and datasets in the Marine Geology archive at NOAA NCEI (NOAA NCEI 2018). As part of this study, new data collections were added to dbSEABED through data mining from published papers, reports and theses, mappings, released datasets, and expedition reports. Because previous data holdings were relatively sparse, finding additional datasets from the continental slope was emphasized. In addition, data from sediment grab samples collected during this study (Chapter 9) were incorporated into dbSEABED. Prior to predicting continuous, gridded data layers for percentages of mud, sand, and gravel from the samples in dbSEABED, values for these variables were transformed using a standard log-ratio approach for compositional data (Aitchison 1986).

10.1.2 Gridded Predictions of Surficial Sediment Composition

Many statistical approaches have been used to generate gridded, spatially continuous data from point samples of environmental data. These include deterministic methods, such as inverse distance weighted interpolation, and stochastic methods, like kriging and machine learning algorithms (Li and Heap 2014; Stephens and Diesing 2015). Li and Heap (2014) reviewed 25 commonly applied methods and noted several important considerations for choosing an appropriate method, including the distribution and quality of sample data and spatial correlation in the data, which may be reduced in areas with higher topographic complexity. For surficial sediment data, these specific considerations are reflected in that many methods do not accurately predict sharp boundaries in sediment distributions, particularly across environmental zonations (e.g., the transition from continental shelf to continental slope).

In this study, continuous gridded maps of sediment composition (median grain size and percentages of mud, sand, and gravel) were created for the New York Bight from the point samples of surficial sediments in dbSEABED using multivariate (3D) IDW interpolation. The 3D IDW approach assumes spatial autocorrelation, that samples closer together are more similar than those farther apart (Tobler 1970), and calculates values at unsampled locations using a weighted combination of the values of neighboring sample locations. The distance between sample points from dbSEABED was measured in 3-dimensional space so that distance weights reflected easting, northing, and depth distances. Maps were generated at 200 x 200 m resolution, the highest resolution supported by the spacing of the sample data. Grid cells containing sample data were assigned the mean value of the sample points within the grid cell. All other grid cells were assigned interpolated values.

To assess the performance of these methods, output maps were examined to determine if the gridded estimates sensibly followed terrain features such as the submarine canyons on the continental slope.

10.2 Results and Discussion

10.2.1 Surficial Sediment Data

Within the dbSEABED holding in the New York Bight, there were 14,755 samples with measurements of median grain size (Figure 10.1). Most of the samples on the continental shelf were comprised of coarse to medium grained sand, with few scattered areas of pebbles and cobbles. Farther offshore, samples transitioned to finer grain size material consisting of silt and clay. The dbSEABED system contained 17,621 samples with measurements of percentage mud (Figure 10.2), percentage sand (Figure 10.3) and percentage gravel (Figure 10.4) in the New York Bight. The highest percentages of mud were found in samples farther offshore, on the continental slope, as well as in the eastern portion of the study area. Conversely, most of the samples on the continental shelf had very high percentages of sand. Few samples had high percentages of gravel.

10.2.2 Gridded Predictions of Surficial Sediment Composition

The interpolated map of median grain size indicated that the majority of the continental shelf in the New York Bight, including the NYWEA, is covered by sand (Figure 10.5). Areas of finer sediments (silt and clay) were predicted in the Hudson Shelf Valley, farther offshore on the continental slope, and in an area to the east. The areas predicted to have finer sediments corresponded to the areas predicted to have the highest percentages of mud, although there were also scattered areas with high-predicted percentages of mud across the continental shelf (Figure 10.6). Very high percentages of sand were predicted across much of the continental shelf, including the NYWEA (Figure 10.7). Areas predicted to have higher percentages of gravel were patchy and scattered across the continental shelf (Figure 10.8). Interestingly, although the NYWEA was predicted as being comprised primarily of sand, there were areas predicted as having relatively a high percentage of gravel just outside the NYWEA.

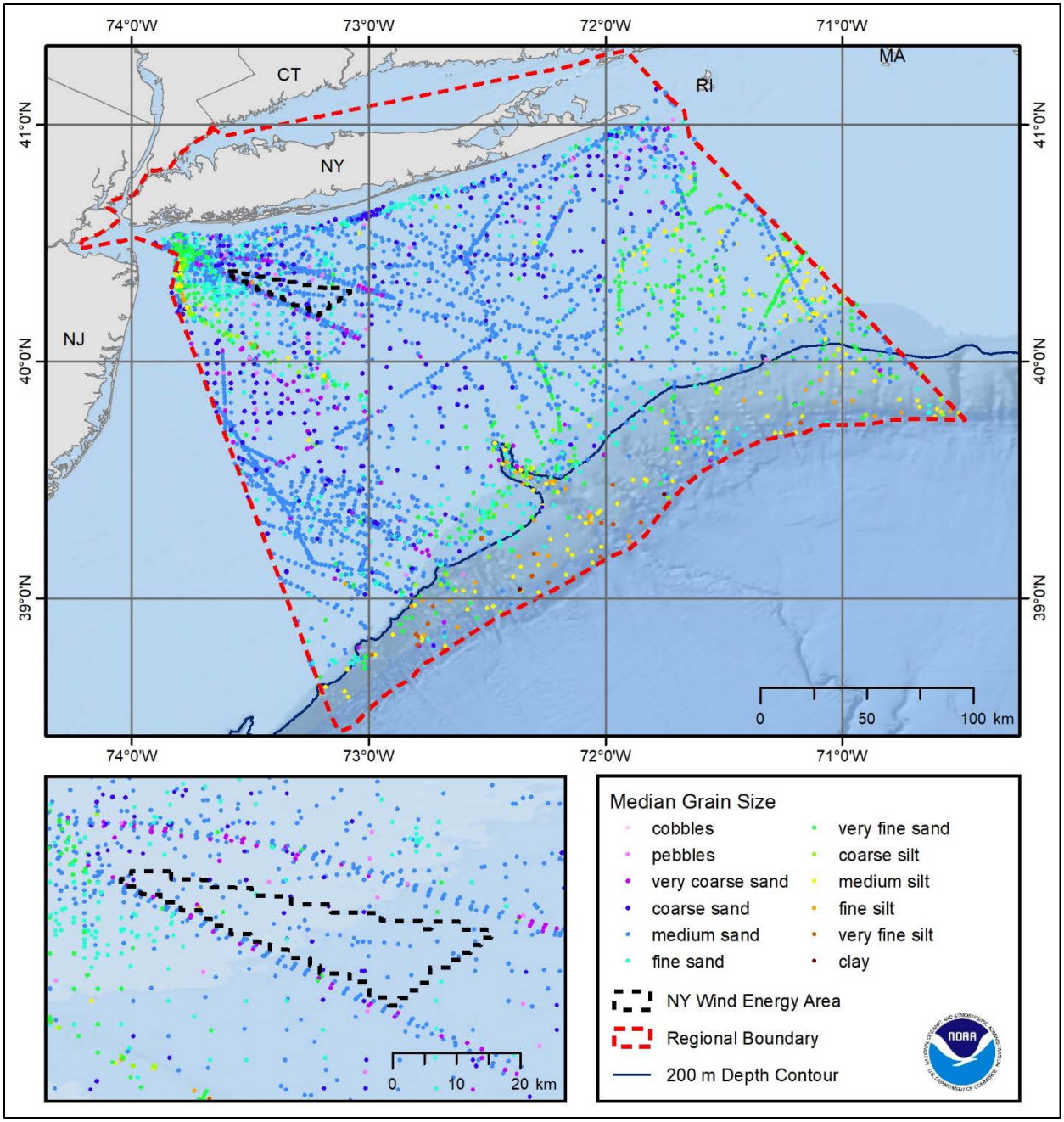


Figure 10.1. Distribution of samples in dbSEABED with measurements of median grain size. Median grain size values are categorized into grain size classes according to the Wentworth scale (Wentworth 1922).

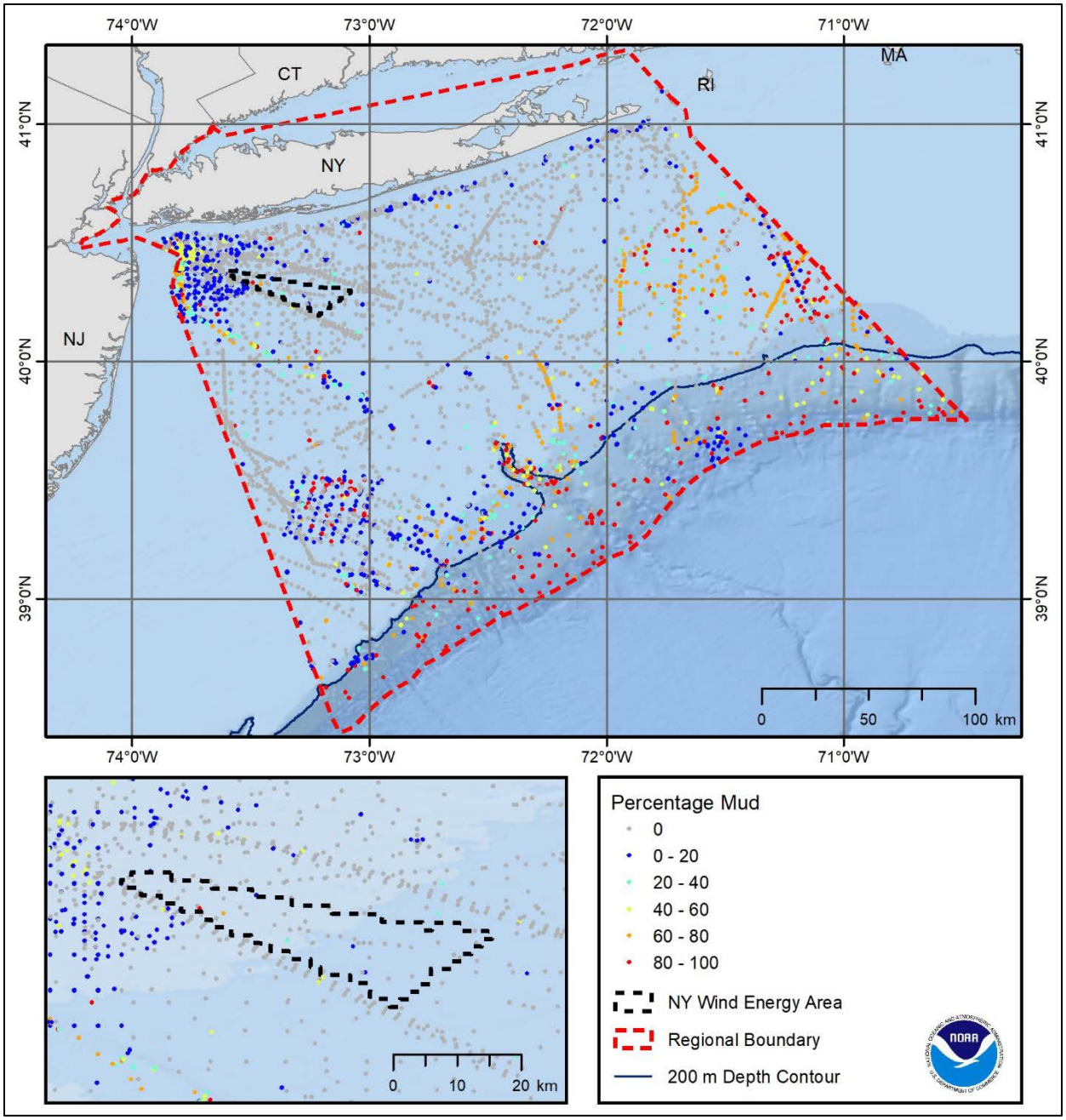


Figure 10.2. Distribution of samples in dbSEABED with measurements of percentage mud.

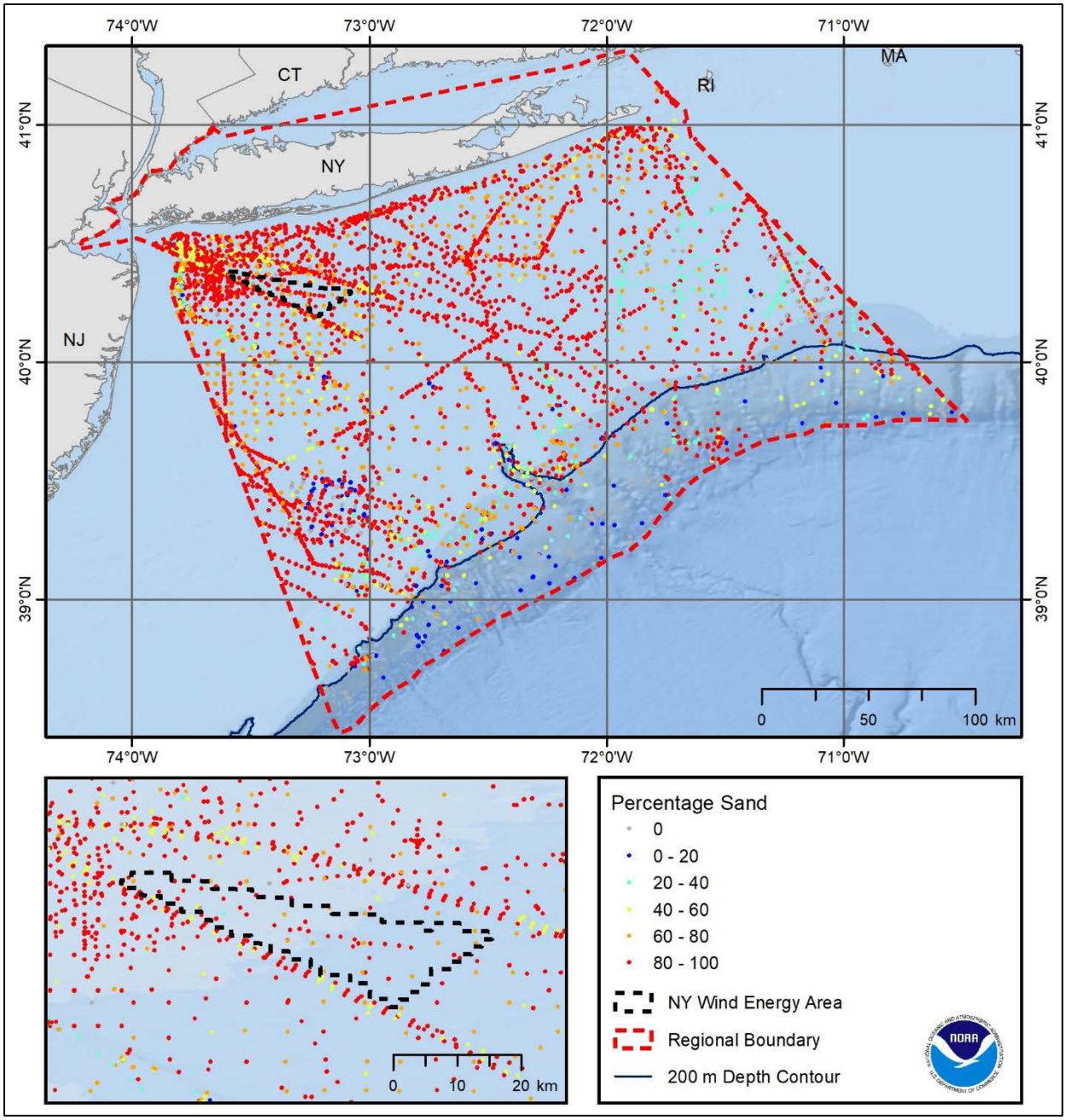


Figure 10.3. Distribution of samples in dbSEABED with measurements of percentage sand.

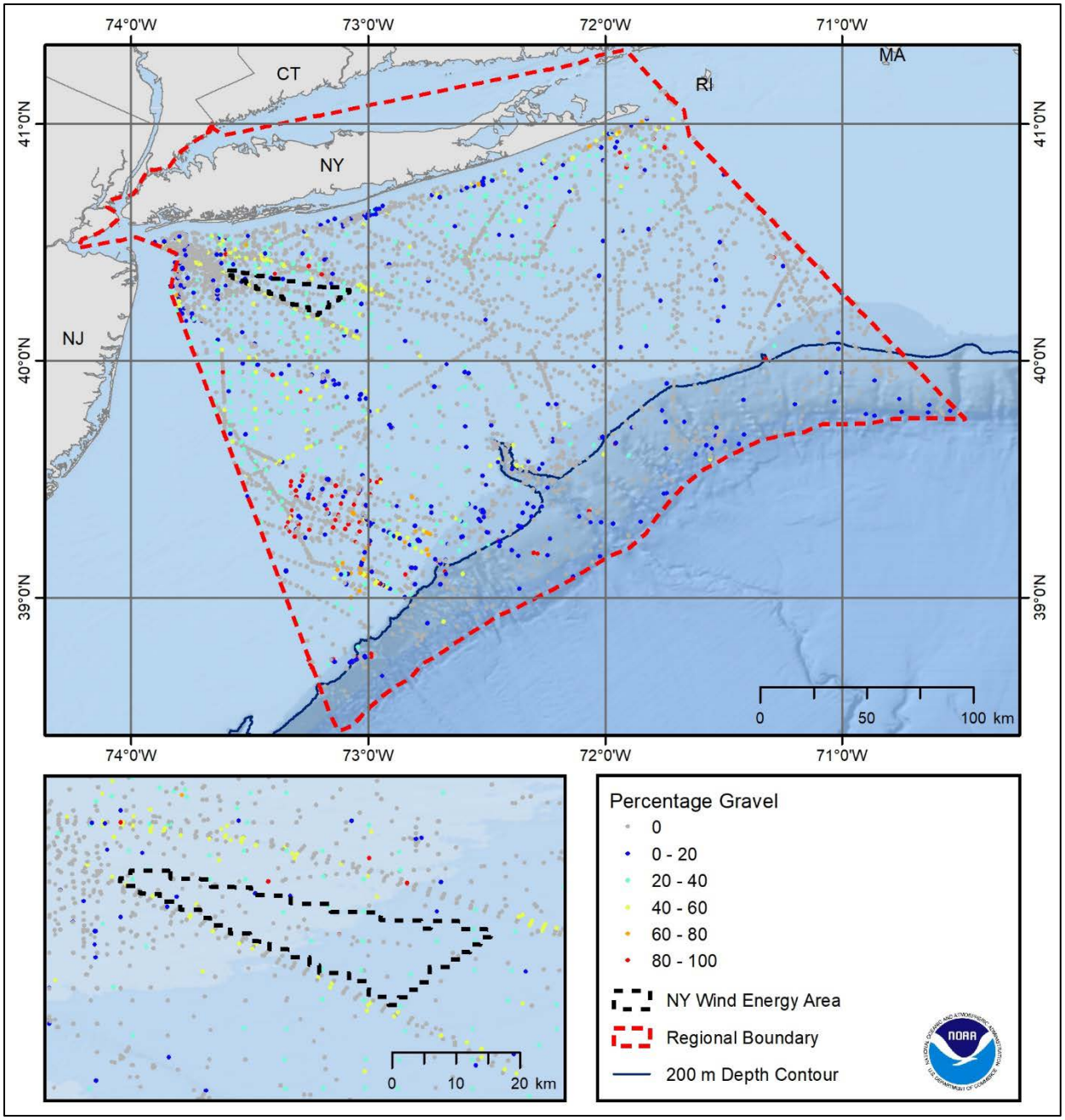


Figure 10.4. Distribution of samples in dbSEABED with measurements of percentage gravel.

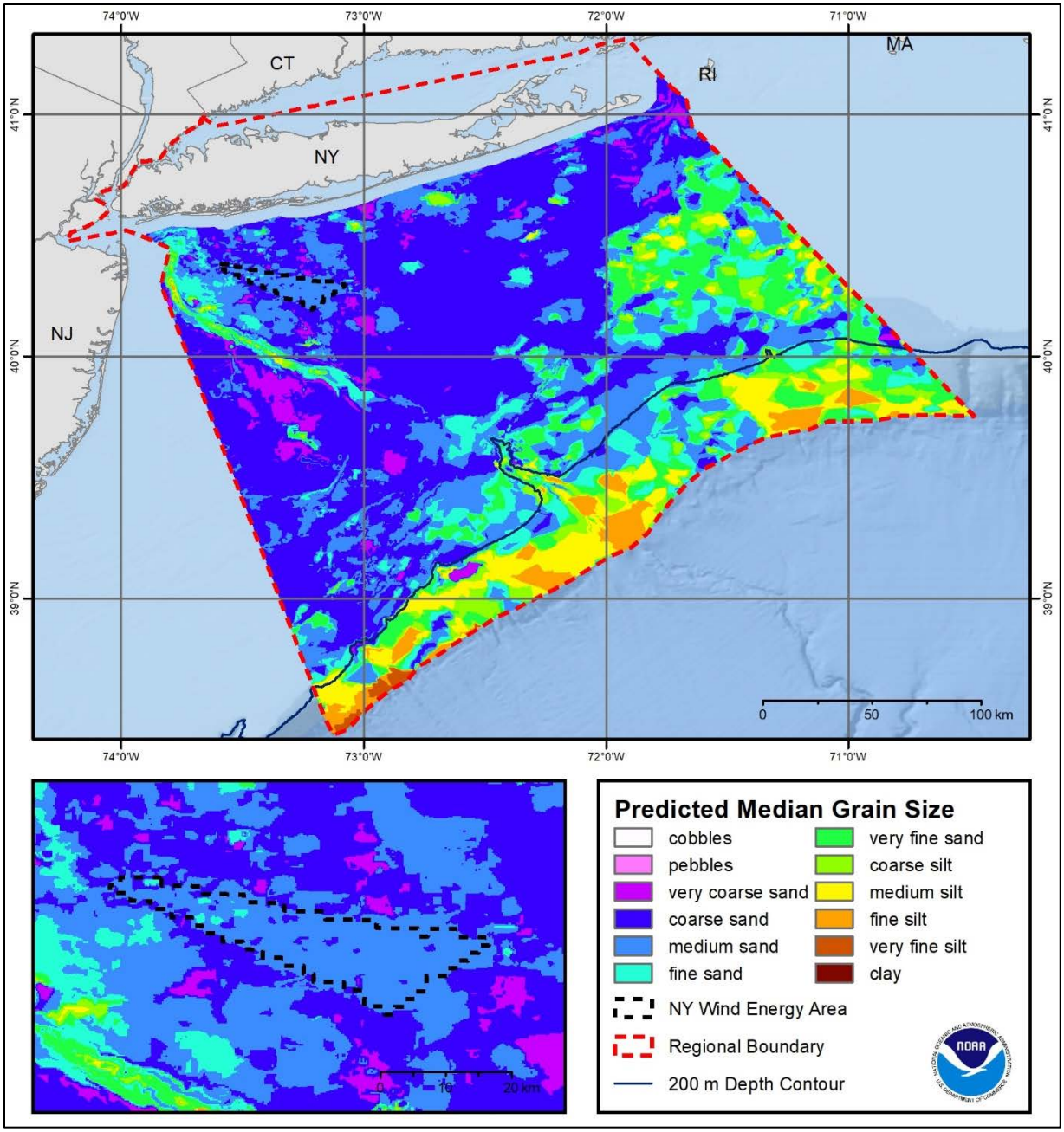


Figure 10.5. Continuous gridded map of median grain size generated using 3D inverse distance weighted interpolation. Median grain size values are categorized into grain size classes according to the Wentworth scale (Wentworth 1922).

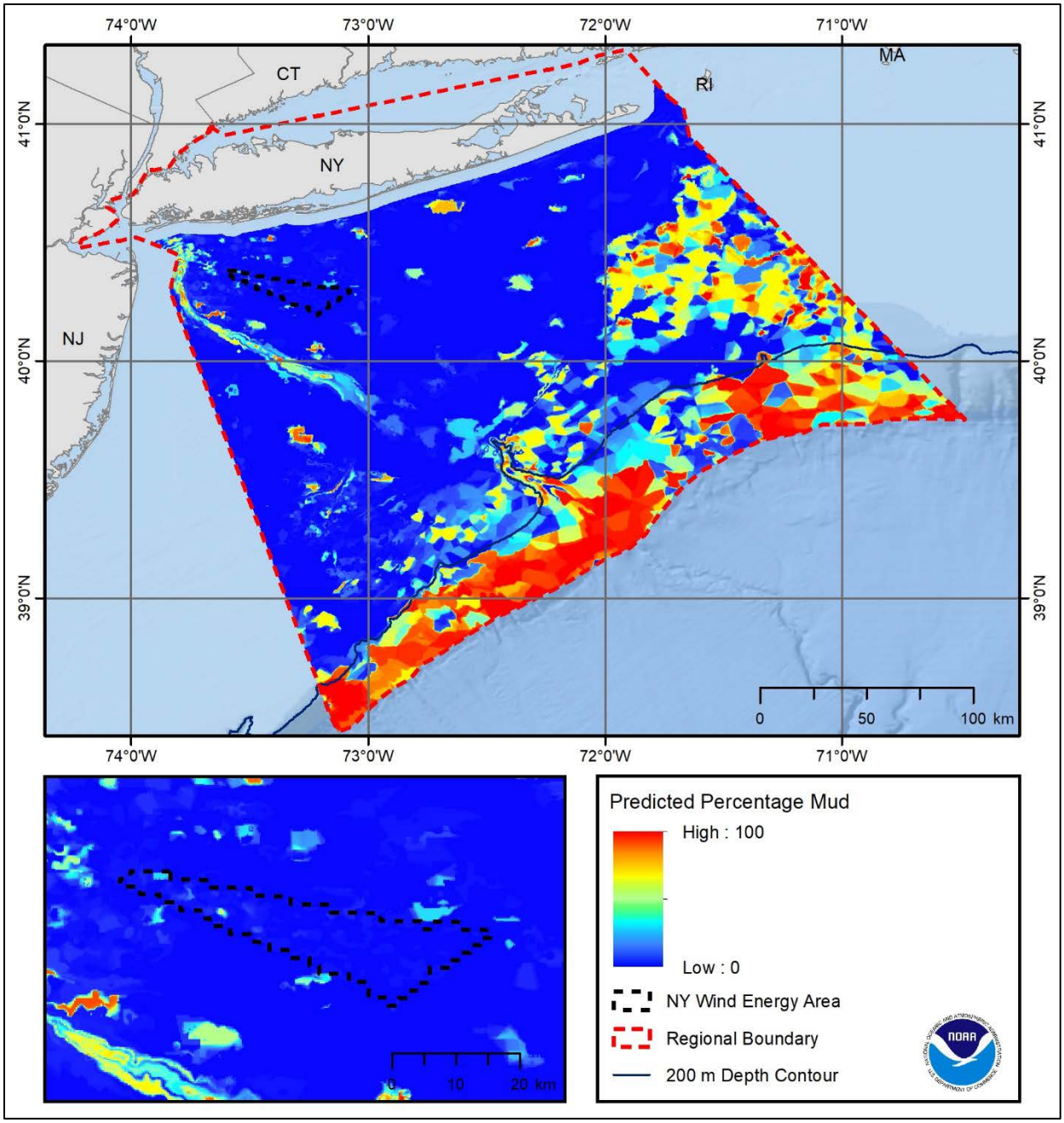


Figure 10.6. Continuous gridded map of percentage mud generated using 3D inverse distance weighted interpolation.

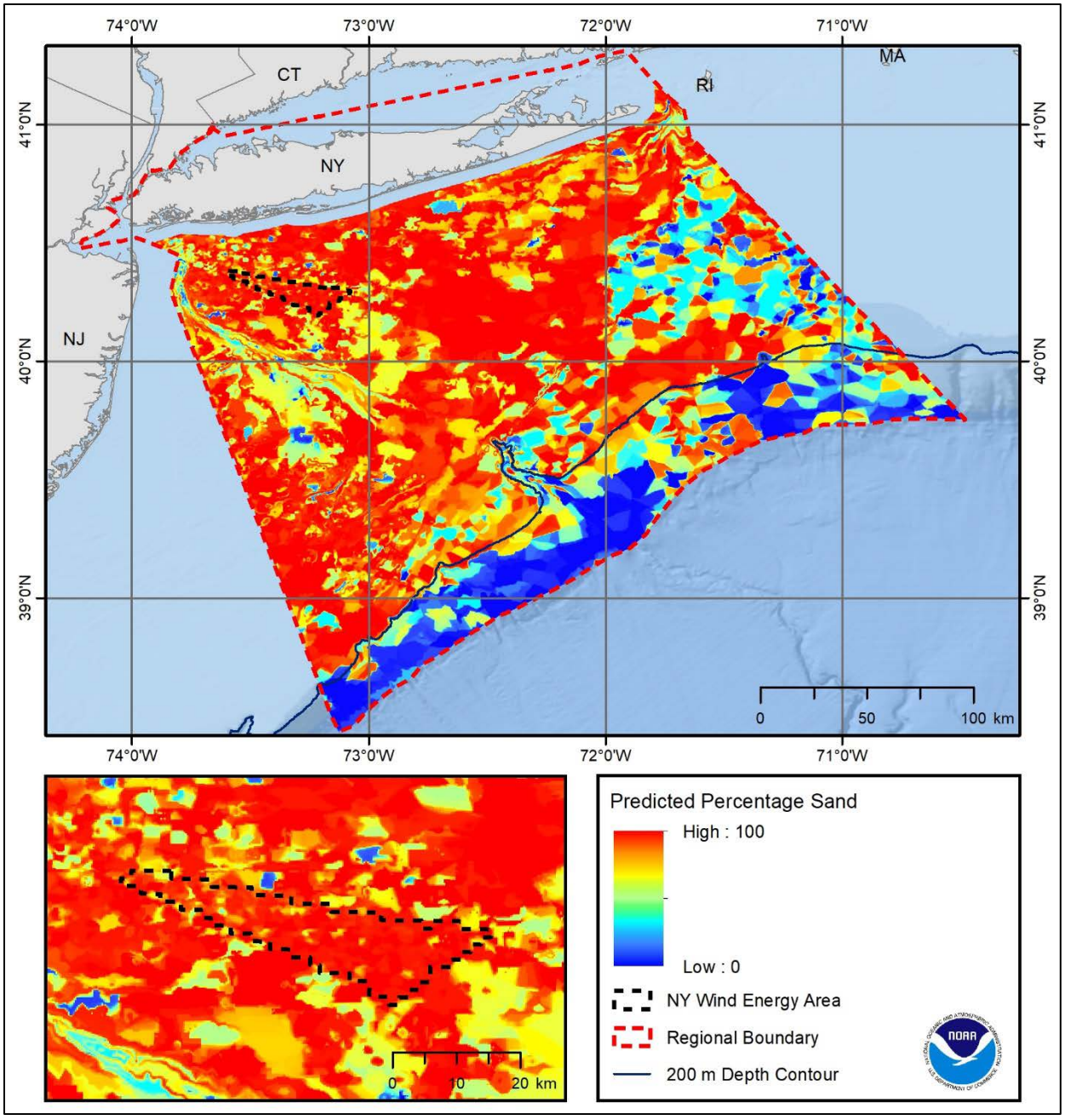


Figure 10.7. Continuous gridded map of percentage sand generated using 3D inverse distance weighted interpolation.

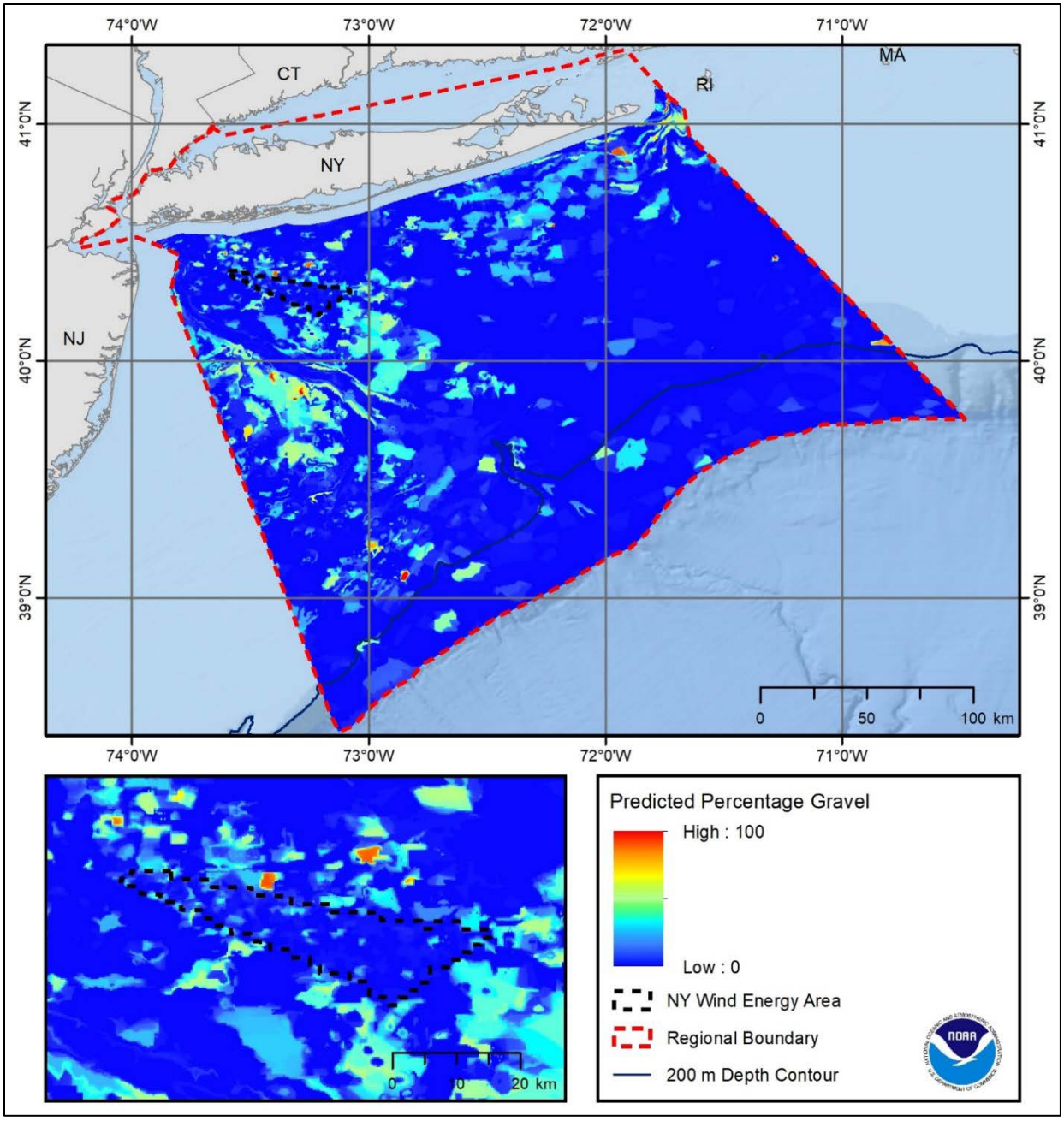


Figure 10.8. Continuous gridded map of percentage gravel generated using 3D inverse distance weighted interpolation.

10.3 Conclusions

Although overall spatial patterns of surficial sediment composition depicted in the updated maps are similar to the maps developed by Poti et al. (2012), the addition of new surficial sediment samples from dbSEABED and the application of the 3D IDW interpolation provided updated maps of surficial sediment composition at 200 x 200 m resolution, an approximately 16x increase in spatial resolution. As a result, the maps delineate the zonation of surficial sediment regimes with greater detail in areas with sufficient sample density. For example, the updated

maps depict more patches of gravelly sediments, which may be too small to have been captured in the previous maps (Poti et al. 2012).

However, it is important to note that there are still many areas in the New York Bight where surficial sediment samples have not been collected. In these areas, interpolations of sediment composition will not be able to capture finer-scale spatial variability and may appear to be overly generalized. One limitation of the 3D IDW interpolation is that it does not provide an estimate of prediction uncertainty. Other methods of spatial interpolation that are stochastic (i.e., they use a statistical model fit to the sample data to make predictions at unsampled locations rather than directly using the values at the sampled locations) do allow for the calculation of a measure of uncertainty at each grid cell location (Li and Heap 2014). Maps of the estimated uncertainty would certainly be useful for identifying where the interpolated maps are most and least reliable.

In this study, random forest models were evaluated as an alternative to the 3D IDW interpolation. Random forest modeling is a machine learning approach in which a statistical model is fit to the sample data using a set of environmental predictor variables (Breiman 2001). Initial maps of sediment composition from the random forest models did not sensibly delineate terrain features such as the submarine canyons on the continental slope. One possible explanation for the poor performance of the random forest models is that data layers representing important environmental predictor variables more directly connected with sedimentation processes (e.g., bottom orbital velocity, median bottom ocean current velocity) were not available at sufficient spatial resolution. In addition, unlike the 3D IDW interpolation, in which neighboring samples had a greater influence on estimates of sediment composition, the random forest models used all the data to estimate the statistical relationships between sediment composition and the environmental predictor variables. The random forest models likely performed poorly because these statistical relationships were not stationary across the New York Bight (e.g., they would not be the same for the continental shelf and continental slope), which spans different terrigenous and biogenic provinces across the continental shelf, slope, and rise.

Since it is unlikely that the entirety of the New York Bight will be sampled or mapped with multibeam sonar, spatial interpolation methods such as the 3D IDW interpolation approach used in this study can be used to provide reasonable estimates of surficial sediment composition. The updated maps presented in this chapter can be used to visualize the spatial patterns in sediment composition in the New York Bight with greater detail than available in previous maps. However, interpolated values should be used with caution in areas where sediment samples are sparser. For specific sites where management decisions require information on sediment composition, it will be important to collect additional sediment samples. These samples, in conjunction with backscatter intensity data from multibeam sonar, would enable a better characterization of the seabed.

Chapter 11. Habitat Maps

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 ecosystem monitoring and assessment programs (Menza et al. 2006), (2) evaluating the efficacy of existing marine protected areas (MPAs) (Pittman et al. 2014), (3) planning for and designing new MPAs (Pittman et al. 2013), (4) assessing the impact of offshore energy facilities on essential fish habitat (BOEM 2016) and (5) targeting research to better understand the socioeconomic, oceanographic and ecological processes affecting ecosystem function and condition (Pittman et al. 2017). Similarly, the benthic habitat maps and products developed under this project provide valuable information to support the evaluation and management of the NYWEA and document baseline conditions of the New York Bight region.

In all, five product types developed in this project are useful in evaluating seafloor characteristics of the NYWEA:

1. Maps of bathymetry and backscatter, along with derived seafloor metrics and a Principle Component Analysis (PCA) based on multibeam surveys across the study area (see Chapters 4 and 6).
2. Delineation of landforms using Bathymetry and Reflectivity-based Estimator for Seafloor Segmentation, or BRESS (Masetti et al. 2018) (see Chapter 6).
3. Sediment grain size and substrate type analysis from the MVV grab samples (see Chapter 9)
4. Predictive surfaces showing the probability of occurrence (and associated error) of substrate, geoform and biotic cover types across the study area (see Section 11.2.3 and 11.2.4).
5. A habitat classification map depicting the co-occurring substrate, geoform and biotic cover types that comprise the final benthic habitat classes (see Section 11.2.5).

11.1 Methods

The approach used here to develop habitat classification maps incorporated predictive models to depict the probability of occurrence of benthic habitats within the NYWEA. The predictive modeling process closely follows that of other recent mapping activities (Kendall et al. 2017; Costa et al. 2018) and included: (1) creating a benthic habitat classification scheme based on CMECS system (CMECS 2017)(refer to Chapter 7, Table 7.1); (2) preparing topographic and geomorphometric datasets to be used as predictors in the habitat models (Chapter 6); (3) collecting underwater videos and sediment grab samples to train the habitat models and evaluate their performance (Chapter 3); and (4) using the habitat models to create spatial predictions of habitat types and a composite habitat classification map of co-occurring habitat types (see Section 11.2). Additionally, output from the BRESS software package delineated benthic landforms (Masetti et al. 2018) in the NYWEA that identifies larger scale geomorphologies that could not be predicted or validated using the underwater video data or grab samples (Chapter 6).

11.1.1 Benthic Habitat Classification Scheme

The classification scheme closely aligns with the CMECS system from the FGDC. The CMECS system is described as "a catalog of terms that provides a means for classifying ecological classes using a simple, standard format and common terminology. CMECS offers a way to organize and interpret data about the marine environment, and it provides a common platform for inter-relating data (CMECS 2017). The CMECS classification system contains three major components that we use to describe the benthic environment of the NYWEA: Substrate Type (e.g., pebbles), Geoform (e.g., sand ripples), and Biotic Cover (e.g., echinoderm beds).

Substrate, geoform and biotic components can be further refined to describe class, subclass and groups, which can be quantified from sediment grabs and underwater video. Given that multibeam data and video constituted the primary source of data used to develop the maps, the classification scheme was constrained to those habitats that were likely to be detected and differentiated by their acoustic properties and could be validated with optical methods.
The initial classification scheme used in NYWEA included five substrate types (cobbles, pebbles, sand, mud, and broken shell), four geoform types (ripples, megaripples, outcrops, and plain), and seven biotic cover types (molluscs, echinoderms, polychaete worms, crustaceans, algae, no cover, less than 5% cover, and other). These were modeled (Table 7.1) to describe benthic habitats and biotic cover. Using this classification scheme, GV and AA sites were annotated to provide a basis for the modeling effort and to validate maps of predicted surfaces across the NYWEA. Final models and predictions were not created for all substrate and cover types documented. In cases where habitat types were very rare (e.g., cobbles had a prevalence = 1.0%), the model failed as these categories were insufficiently present to find a clear mathematical relationship between the habitat type and predictors. Where habitat types were very common (e.g., broken shell had a prevalence = 99%), the model also failed because there were no distinct set of predictors that explained the spatial distribution of the respective habitat types.

The spatial distributions of the remaining substrate, geoform, and biotic cover types were predicted across the NYWEA using mathematical models, namely BRTs. The predicted habitat types included: 'Pebbles', 'Ripples', Megaripples', 'Echinoderms', Crustaceans', 'No Cover'(100% bare), and 'Less Than 5% cover'. The two most prevalent CMECS biotic groups include 'sand dollar beds' and 'mobile crustaceans on soft sediments', at a rate of 90% and 11%, respectively. These two groups were the only biotic components that were considered for the development of the predictive models for the NYWEA.

11.1.2 Predictive Model Inputs

Seafloor depth, backscatter intensity, and morphometric surfaces are known to be excellent predictors for characterizing many different habitat types in tropical environments such as sand, pavement, and coral reefs (Costa et al. 2009; Costa and Battista 2013), but these surfaces can also be very useful in temperate marine environments where benthic habitats are more homogenous. The depth and backscatter data was collected in the field using the multibeam echo sounder by the NOAA Ships *Ferdinand R. Hassler* and the *Nancy Foster* (see Chapter 2). These two surveys were merged to create 8 x 8 m unified bathymetry and backscatter surfaces that covered 100% of the NYWEA. Once the merged multibeam datasets were cleaned of noise and smoothed, they were used as base layers to create the secondary morphometrics using an R script and the BRESS software (see Chapter 6).

In all, 24 different predictors were used as inputs for the BRT to create predictive models for the NYWEA, but only 12 of the predictors were the most influential variables across predictive models (Table 11.1). These surfaces were derived from bathymetry and backscatter using local complexity metrics and statistics from BRESS output. After cross validation, the spatial predictions were then classified into the different substrate, geoform, and biotic cover types using the underwater video and sediment grabs for GV.

Table 11.1. List of the 12 most influential seafloor metrics that were used in the BRT models to create surfaces depicting the probability of occurrence of habitat types.

Source	Predictor	Model label	Reference
MBES survey	Depth	NYB_dtm8m9	this document
MBES survey	Backscatter intensity	nyb_mos_8m	this document
Local complexity metric	Standard deviation of depth	Std_DeDept	Jenness 2013
Local complexity metric	Slope	Slope	Jenness 2013
Local complexity metric	Slope of slope	SlofSl	Jenness 2013
Local complexity metric	Sine of aspect (eastness)	SinAsp	Jenness 2013
Local complexity metric	Cosine of aspect (northness)	Cos_Aspect	Jenness 2013
Local complexity metric	Total curvature	TotCurvEva	Jenness 2013
Local complexity metric	Profile Curvature	ProfCuEvan	Jenness 2013
Local complexity metric	Plan Curvature	PlanCurvZe	Jenness 2013
Local complexity metric	General Curvature	GenCurvZev	Jenness 2013
BRESS software output	Average height	avg_height	Masetti et al. 2018

11.1.3 Predicting and Classifying Benthic Habitats

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. A random subset of data is used to fit a model at each stage. This randomization helps improve model performance (Friedman 2002; Elith et al. 2008). These simple models are then combined linearly to produce one final combined model (Elith et al. 2008). The fitted values in this combined model are more stable than values from an individual model, improving its overall predictive performance (Friedman 2002; Elith et al. 2006; Elith et al. 2008).

Separate spatial predictions were developed using BRTs for each geoform, substrate, and cover type. BCTs were used to combine these individual predictions into a single composite habitat map. Four main steps were used to create habitat predictions and a composite habitat map for the NYWEA (Figure 11.1): (1) preparing the model input data, (2) creating and evaluating substrate, geoform and cover models and spatial predictions, and (3) creating and evaluating a composite habitat map. This work was conducted primarily in ArcGIS (version 10.5, ESRI 2016) and R

software (version 3.3.0, R Core Team 2016) using the dismo (Hijmans et al. 2017), caret (Kuhn 2016), and raster (Hijmans 2016) packages.

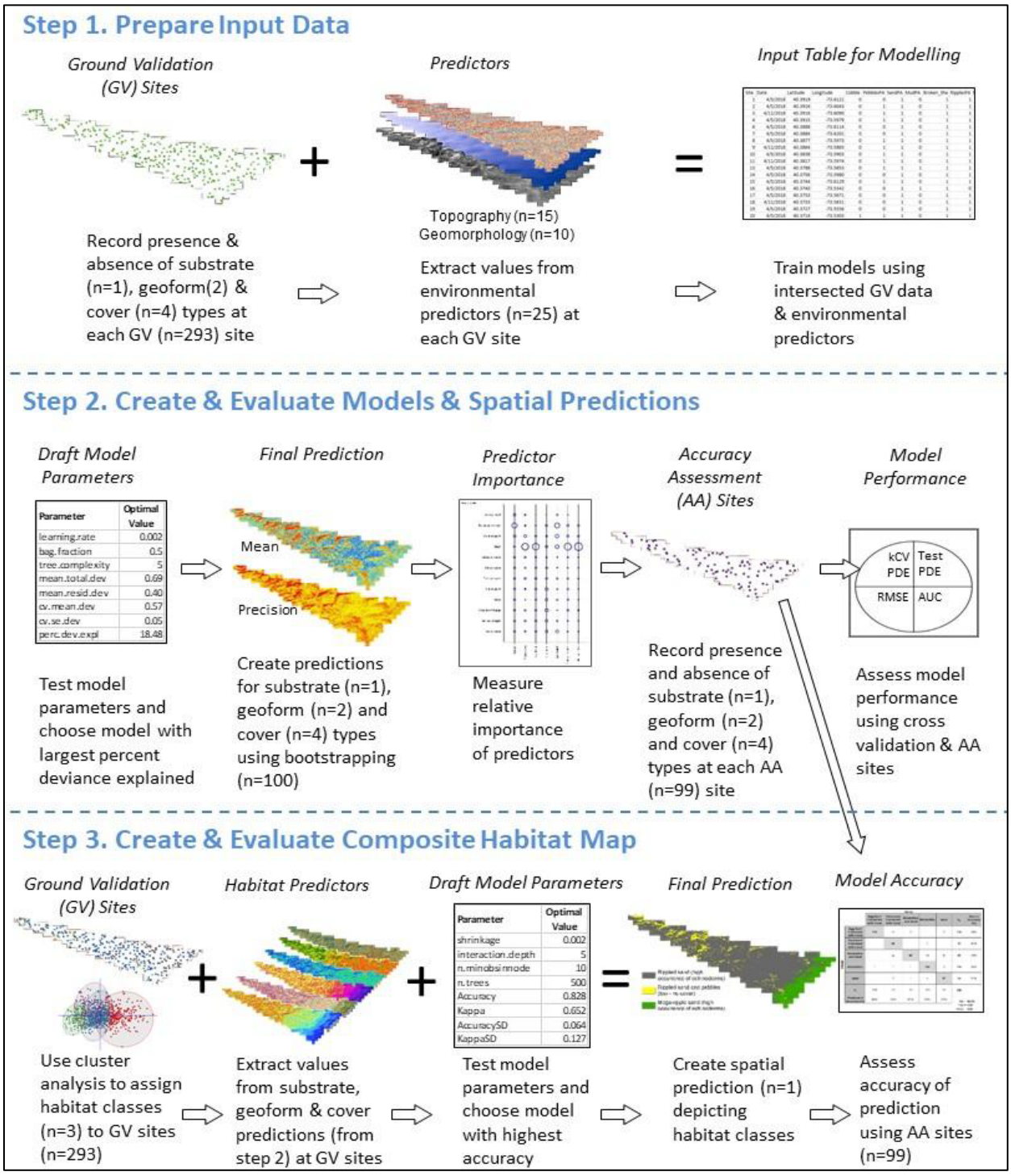


Figure 11.1. Diagram depicting steps in modeling process to predict substrate, geoform, and biotic cover distributions and develop a composite habitat map.

Preparing Input Data (Step 1)

Underwater videos and sediment grab samples from the 293 successful GV sites collected by the R/V *Tiki XIV* (see Chapter 3) were reviewed to determine the presence and absence of five substrate types, four geoform types, and seven biotic cover types at each site (Table 11.2). The video analysis of the GV data were used as a response variable in the BRT modeling process. Response variables were annotated by assigning a presence (1) or absence (0) value for each response variable for every video obtained at each GV site. In addition, presence and absence for each substrate and biotic cover type was observed and recorded for each sediment sample while in the field and attributed as such in the GV table. It was not possible to determine the presence of geoforms (ripples, megaripples, outcrops, plain) from the sediment grab data due to the large scale of those components. The presence of each type of substrate, geoform, and biotic cover were non-exclusive of each other, meaning the co-occurrence of more than one substrate, geoform, or cover type was permitted.

Table 11.2. Prevalence of substrate, geoform, and biotic cover types from sample sites.

Component	Habitat Type	% Prevalence
Substrate	Cobbles	0.7
	Pebbles	16.0
	Sand	100.0
	Mud (Silt and Clay)	15.7
	Shell	99.0
Geoform	Ripples	82.9
	Megaripples	12.3
	Outcrops	0.0
	Plain	4.8
Biotic Cover	Molluscs	11.6
	Echinoderm	90.1
	Polychaete Worms	30.4
	Crustaceans	10.9
	Algae	0.0
	No Cover (bare)	8.9
	Less than 5% Cover	32.8

After reviewing the prevalence of each habitat type or response variable, some of the types were eliminated from subsequent steps of the modeling process. In cases where the prevalence of a substrate, geoform, or cover type was too low (e.g., <0.7% for 'Cobbles') or too high (e.g., 100% for 'Sand') to be successfully modeled using BRTs, these models were dropped. Cobbles were eliminated due to a very low occurrence (n=3) in grab samples. Both the video annotation and grab samples had high occurrences of sand. There was a 98.6% overlap of sand sites across the two data sources (GV video and grab samples) and 100% occurrence of sand when combined. Therefore, the presence of sand ripples and megaripples was used to differentiate sand substrate and subsequently dropped sand as a substrate type in spite of the fact that is occurs throughout

the NYWEA. Outcrop was eliminated due to zero occurrences. Plain was eliminated due to very low occurrences (n=14). While molluscs and polychaete worms were prevalent, the models failed, likely due to lack of association between the environmental predictors and the biotic response. Only one occurrence of algae was observed in a sediment grab and therefore eliminated from the modeling process. All sites had some proportion of no biotic cover. Locations with 100% no cover (bare), as well as <5% cover, were modeled. All remaining response variables were modeled using a binomial (two groups) distribution. No transformations were applied. All of the predictor variables were numeric. The GV sites were intersected with the 25 predictors to extract their value at each location. This spatial intersection combined the GV (response variables) and predictor datasets into a single table.

Creating and Evaluating Models and Spatial Predictions (Step 2)

The BRT models were fit and optimized in R (version 3.3.0, 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) (Table 11.3). 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.

Table 11.3. Suite of BRT model parameters that were tested.

Regularization Parameters	Parameters Tested	Definition	Impact	Example
Learning Rate (lr)	0.001, 0.002	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, 5, 10	Controls how many predictor interactions are fitted in a tree	Decreasing tc will shrink the size (number of nodes) in a tree	tc = 20 will grow larger trees (with more nodes) than tc = 2
Bag Fraction (bf)	0.1, 0.2, 0.5	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

For each of the remaining habitat types (one substrate, two geoforms, and four cover types), 24 combinations were tested for lr, tc, and bf. K-fold cross validation (kCV) was used 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 24 model parameter combinations x 7 substrate, geoform and cover types (n=1,680 models total). Model performance was measured 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.

The model with the highest kCV PDE was selected for each substrate, geoform, and cover type. The relative importance (Elith et al. 2008) of the top five predictors across all seven models was quantified and presented in a bubble plot (Figure 11.2). Circle size is proportional to a predictor's relative importance averaged across 100 model iterations. The larger the circle, the more important the predictor. The top five most influential predictors across all models are shown here and include depth, backscatter intensity, standard deviation of depth, slope, slope of slope, sine of aspect, cosine of aspect, total curvature, profile curvature, plan curvature, general curvature, and average height

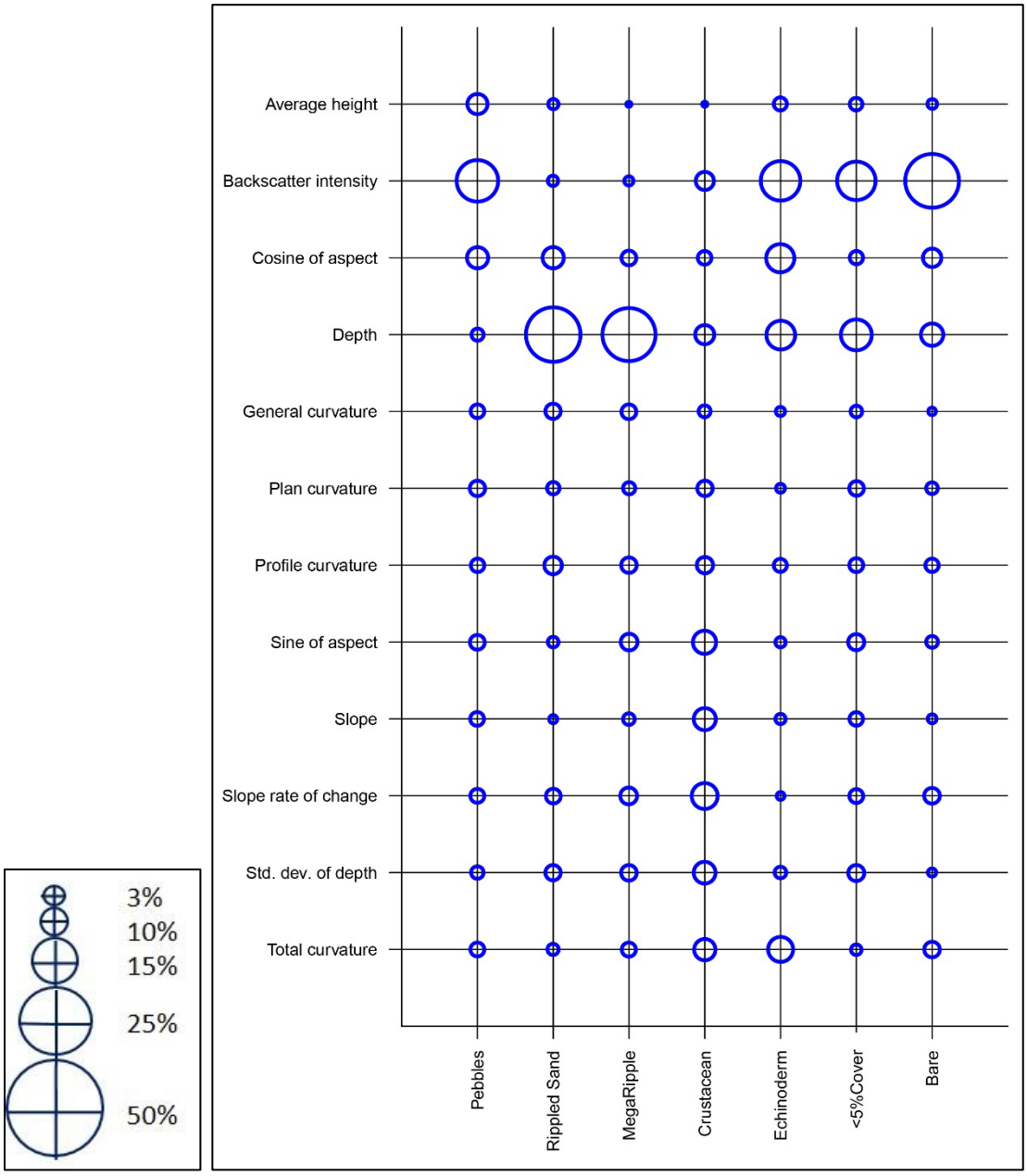


Figure 11.2. Relative importance of the environmental predictors used to develop the six habitat models and spatial predictions.

The best models for the remaining substrate, geoform, and biotic cover types were used to predict their spatial distribution across the NYWEA, performed by 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, geoform 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 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, geoform and cover type, these probabilities and precisions represent the average of, and variation in, 100 model iterations created using bootstrapping (see Technical Glossary).

The performance of each prediction was 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 metrics were calculated because they describe model performance in different ways, and when viewed together, provide a more thorough understanding of the model 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 enough to accurately predict 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 11.1.5). 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." 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.

Creating and Evaluating Composite Habitat Map (Step 3)

BCTs were used to develop a single composite habitat map. The GV data were grouped into three commonly co-occurring geoform, substrate, and cover types using hierarchical cluster analysis. The GV sites with similar substrate, geoform and cover types were grouped into

progressively larger and larger clusters. The resulting geoform, substrate, and cover clusters that best discriminated habitat types into classes were chosen. While there was a fair amount of overlap in characteristics that defined each class (echinoderms and rippled sand were common throughout), aspects that were unique to each class (i.e., pebbles and megaripples) were noted o aid in defining the composite habitat classes.

The primary characteristics of the seafloor that were identified in this analysis include (1) vast expanses of rippled sand with high concentrations of sand dollar beds and scattered molluscs, crustaceans, and polychaete worms, (2) offshore megaripples with similar faunal distributions that, and (3) large wedge shaped congregations of sand ripples with low biological cover on the ridges and mixed substrates of pebbles and broken shell in the troughs. The characteristics that most strongly defined each class were rippled sand and echinoderms, while characteristics that were generally unique to one class included megaripples, pebbles, and no cover; albeit, with a weaker relationship to that class (Table 11.4). An analysis including the occurrence of echinoderms in each class (not shown) was the basis for the class modifiers. Sand dollar beds were a common feature and were used as a modifier to denote classes where echinoderm abundance was high. Sand dollar beds (*Echinarachnius*) have been detected previously throughout other proposed wind energy areas in the Mid-Atlantic region (Guida et al. 2017) using 100 kHz side-scan (Fenstermacher et al. 2001). During the review of GV video, there was evidence of variations in sand dollar bed density with regards to changes in the substrate types. While crustaceans were fairly prevalent (10.9%), they did not occur in great enough densities to be considered to be functioning as habitat and were not included in the composite habitat map.

The BCT model predictors used to create a composite habitat map were: ‘Pebbles, ‘Rippled Sand’, ‘Megaripples’, and ‘No Cover’ probability of occurrence prediction surfaces, with class membership from a cluster analysis used as the response variable. Cluster means of four habitat types into three classes produced with hierarchical clustering (JMP -Ward method) indicates the strength of the association of each habitat type to a given habitat class (Table 11.4). In addition, the composite map was evaluated with and without ‘Echinoderms’ to better understand the role of this primary biotic cover in defining the habitat classes. Next, the 293 GV sites (each of which were assigned one of the three habitat types) were intersected with the probability of occurrence prediction surfaces to extract their value at each location. This spatial intersection combined the GV and probability of occurrence values into a single table.

Table 11.4. Cluster means of four habitat types into three classes produced with hierarchical clustering (JMP-Ward method) indicating the strength of the association of each habitat type to a given habitat class.

	Habitat Type				
Classes	Pebbles	Rippled Sand	Megaripples	No Cover	Class Name (modifier)
1	0.1234	0.8968	0.0620	0.0321	Rippled sand (high occurrence of faunal beds)
2	0.4412	0.8702	0.0690	0.4102	Rippled sand and pebbles (low occurrence of faunal beds)
3	0.1068	0.5917	0.3492	0.0172	Megaripple sand (high occurrence of faunal beds)

Next, model parameters for BCTs were fitted and optimized in R (version 3.3.0, R Core Team 2016) using the caret package (Kuhn 2016). Thirty-six combinations of lr, tc, number of trees (n.trees) were tested and minimum terminal node size (n.minobsinnode) (Table 11.5). 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. 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 for each end point in a classification. 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. kCV PDE were calculated to identify the parameter combination that created the highest performing model. This highest performing model was then spatially applied to create the composite habitat map throughout the region. The original bathymetry and backscatter data, which underlies these models, contained some artifacts that were evident offshore in both the predicted surfaces and the final map product.

Table 11.5. Suite of boosted classification tree (BCT) model parameters and values that were 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, 5, 10	Controls how many predictor interactions are fitted in a tree	Decreasing tc will shrink the size (number of nodes) in a tree	tc = 20 will grow larger trees (with more nodes) than tc = 2
Number of Trees (n.trees)	500, 750, 1000, 1500	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)	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

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 to be sampled at a greater density than common habitats. 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, see Kågesten et al. 2015.

11.1.4 Accuracy Assessment of the Predictive Model Outputs

Analysis of Thematic Accuracy

The thematic accuracy of the final habitat map was quantitatively assessed by NOAA NCCOS using seafloor samples collected during the R/V *Tiki XIV* GV mission (see Chapter 3). The sites were not selected independently of the GV sites because the grab sampling component of the NF18-07 mission was constrained due to severe weather conditions. Thus, 100 AA sites were randomly selected from among the planned 400 sampling locations to evaluate the performance of the predictive models and the accuracy of the composite habitat map. A visual inspection of the AA site locations overlaid with the backscatter indicates the potential seafloor features were fairly well represented by the distribution of site locations (Figure 11.3). The AA grab sample observations, video analysis, and sediment analysis were all conducted using the same methods as the GV datasets by an independent analyst to avoid a sampling bias. Of the 100 AA sites that were collected, the underwater video from 97 sites were successful, and could then be used test the results of the model performance. The three unsuccessful sites were due to the lack of GPS data and poor lighting of the underwater videos.

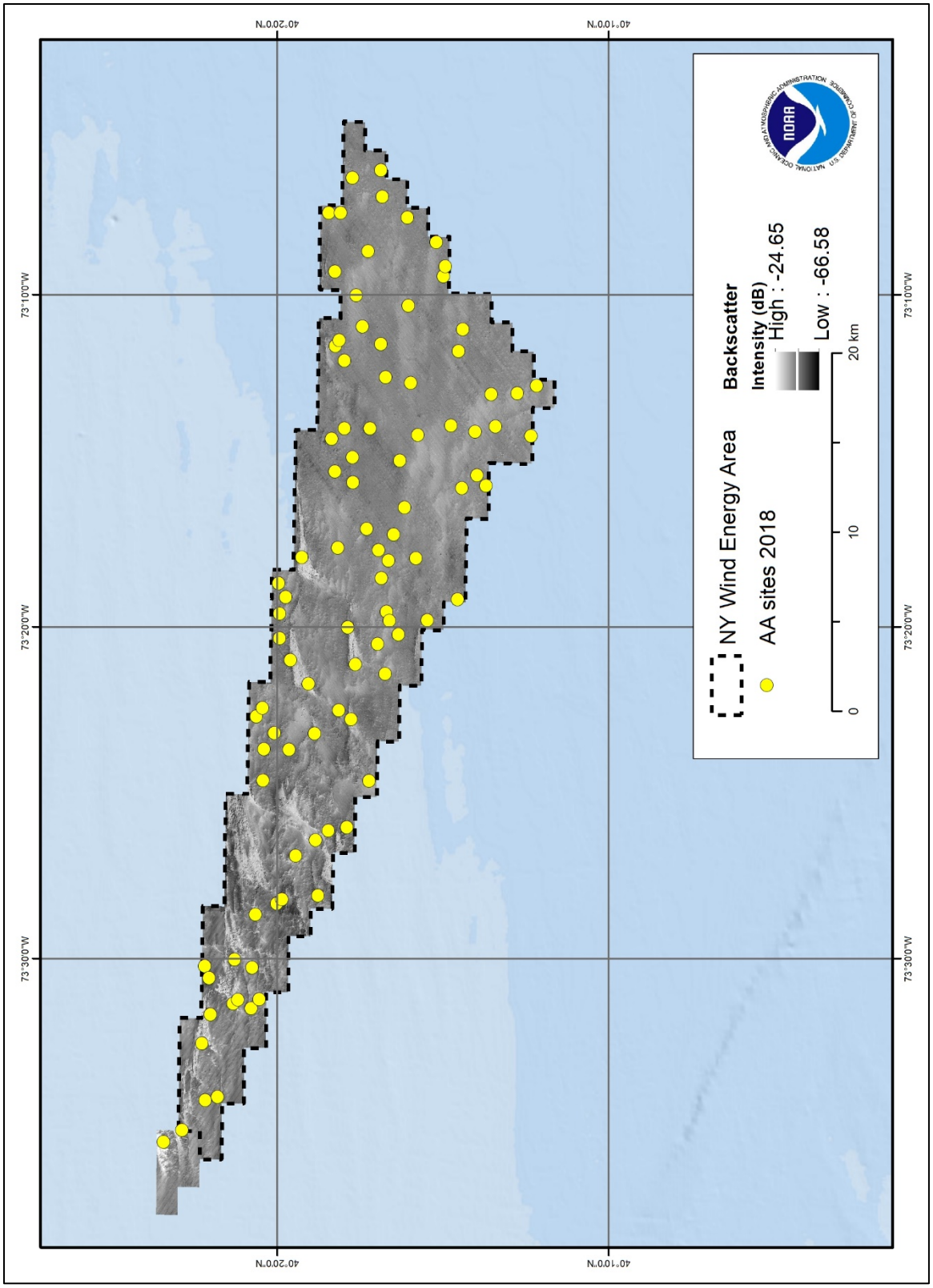


Figure 11.3 Distribution of randomly selected AA sites from the original 400 GV sites overlaid on the backscatter.

The AA sites were grouped into the same three habitats identified by the cluster analysis. Sites were considered correct if the same habitat was present within 8 m (one pixel or raster cell) of the AA site. Analysis for this assessment was conducted by an independent scientist and did not include any interaction with the habitat modelers. A confusion matrix was developed from the 97 AA points describing the composite maps' overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA) (Table 11.6). This matrix was constructed as a square array of numbers arranged in rows (map classification) and columns (AA classification). AA of the composite habitat map also factored in the proportional area of each class to correct for thematic accuracy.

Table 11.6. Confusion matrix for the composite habitat map predicted by the boosted regression tree (BRT) model to calculate the overall accuracy (OA).

		AA (i)				
		Rippled Sand with High Occurrence of Faunal Beds	**Rippled Sand and Pebbles with Low Occurrence of Faunal Beds**	**Mega-Rippled Sand with High Occurrence of Faunal Beds**	**n_{-j}**	**User's Accuracy**
Map (j)	**Rippled Sand with High Occurrence of Faunal Beds**	**57**	2	15	**74**	77%
	Rippled Sand and Pebbles with Low Occurrence of Faunal Beds	4	**3**	1	**8**	38%
	Mega-Rippled Sand with High Occurrence of Faunal Beds	13	0	**2**	**15**	13%
	n_{i-}	74	5	18	**97**	
	Producer's Accuracy (%)	77%	60%	11%	**OA =**	**63.9%**
					T_e =	**0.46**

Pr	Tau_e	V(Tau)	CL(Tau)
0.33	0.46	0.01	0.14

Pr = 1/M where M is number of classes (i.e., 7)

Edited Map		
Habitat	**Area (sq m)**	**Proportion**
1.00	254,510,848.00	0.782
2.00	23,180,480.00	0.071
3.00	47,874,368.00	0.147
Total	325,565,696.00	

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 (Story and Congalton 1986). 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. The Tau coefficient was also calculated 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.

Accuracy Assessment Results and Discussion

In conducting the AA of the habitat map products, it must be duly noted that the presence of 'Sand" substrate was too high to be successfully modeled (see *Preparing Input Data (Step 1)*). In essence, the entire NYWEA area was considered uniformly sand of varying coarseness. As such, an AA could not be conducted on this overarching substrate category. The remaining sub-habitat types that could be modeled, whose prevalence was neither too high nor too low to be included, were assessed for accuracy. The AA results of the sub-habitat types follows.

The three clusters of commonly co-occurring sub-habitat geoform, substrate, and cover types used to create the composite habitat map, not including sand, proved to be difficult to classify and produced a low OA of 63.9% (Table 11.6). Although the predominantly sandy substrate component was easy to analyze in the video and grab samples, and strongly reflected in the models, the geoform and biotic cover components may have underperformed due to several issues. These issues included (1) high uncertainty in distinguishing the geoform components due to scale from the optics of the video analysis, (2) subjectivity in estimating percent cover between the GV and AA video analysts, (3) the influence of smoothing algorithms applied to the sonar data, and (4) the AA sites were a random subset of the original 400 GV sites and not independently selected. It was understood that the models would perform much better with integration of the NF18-07 data to overlay the multibeam data affected by heavy swells from NF17, however time constraints on reprocessing the morphometrics and BRT models precluded the use of this data for this analysis. Thus, the OA of the composite map was lower than expected, but further AAs of the probability of occurrence for individual cover components (Echinoderms, No Cover, and Less than 5% Cover) showed higher model prediction accuracies.

Confusion matrices were helpful in understanding the accuracy of the BRT at predicting the probability of occurrence (or chance) of various biotic cover components. Presence/absence results from the video analysis and grab samples were used to assess the accuracy of the probability of occurrence for different cover types based on percent thresholds that were determined by natural breaks in the statistical outputs. The BRT models were very accurate at predicting areas where there is a high chance of 100% no biotic cover, or bare sand. Only 8.2% percent of the GV sites (24 of 293) and 3% of the AA sites (3 of 97) contained 100% No Cover. The probability of occurrence had to be >21% to be considered present based on the natural breaks in the probability statistics. The predictions of bare sand from the BRT had the highest OA of all the models that were assessed with 92.8% (Table 11.7). The relatively high OA may be due to strong user confidence in determining if a site is desolate of any kind of benthic fauna cover from the underwater video analysis or the relatively low threshold requirement to be considered present or absent.

Table 11.7. Confusion matrix for presence of 'No Cover' by the BRT biotic cover predictions.

		100% No Cover AA (i)			
		Presence 100% NoCov on Video	**Absence 100% NoCov on Video**	**n_{-j}**	**User's Accuracy**
Map (j)	**High Probability of Occurrence for 100% No Cover (>21%)**	**2**	6	**8**	25%
	Low Probability of Occurrence for 100% No Cover (<21%)	1	**88**	**89**	99%
	n_{i-}	3	94	**97**	
	Producer's Accuracy (%)	67%	94%	**OA =**	**92.8%**
				T_e =	**0.89**

The benthic habitats with Less Than 5% biotic cover were observed in 33% of the GV sites and produced moderately strong predictions in the BRT. This type of habitat is mostly comprised of bare sand with a random scattering of individual worms, molluscs, crustaceans, or echinoderms such as sea stars. A probability of occurrence greater than 39% was considered a 'presence' in the model due to the natural breaks in the data. However, the probability of occurrence for less than 5% biotic cover class had a poor OA of 64.9% (Table 11.8). This may be due to the difficulty in estimating the actual percent cover of a few scattered benthic invertebrates on the seafloor from the video analysis.

Table 11.8. Confusion matrix for presence of 'Less Than 5% Cover' by the BRT biotic cover predictions.

		Less Than 5% Cover AA (i)			
		Presence <5% Cover on Video	**Absence <5% Cover on Video**	**n_{-j}**	**User's Accuracy**
Map (j)	**High Probability of Occurrence for Less than 5% Cover (>39%)**	**11**	14	**25**	44%
	Low Probability of Occurrence for Less than 5% Cover (<39%)	20	**52**	**72**	72%
	n_{i-}	31	66	**97**	
	Producer's Accuracy (%)	35%	79%	**OA =**	**64.9%**
				T_e =	**0.89**

Areas that were predicted to have a high probability of occurrence of echinoderms had the best biotic cover of any of the faunal classes. A threshold of >79% chance of echinoderms at each AA site equates to a 'presence' of echinoderms (as determined by natural breaks in the coverage statistics). The BRT model for the probability of occurrence for echinoderms resulted with an OA of 87.6% (Figure 11.9). This accuracy of the probability models for echinoderms was most likely driven by the high correlations of dense sand dollar beds to regions with greater than 95% sand. Habitats that had conglomerations of sand ripples with pebbles and shells were observed to be less suitable for sand dollar beds.

Table 11.9. Confusion matrix for presence of high occurrence of 'Echinoderms' by the BRT biotic cover predictions.

		Echinoderm AA (i)			
		Presence of Echinoderms	**Absence of Echinoderms**	**n_{-j}**	**User's Accuracy**
Map (j)	**High Probability of Occurrence (>79%)**	**81**	4	**85**	95%
	Low Probability of Occurrence (<79%)	8	**4**	**12**	33%
	n_{i-}	89	8	**97**	
	Producer's Accuracy (%)	91%	50%	**OA =**	**87.6%**
				T_e =	**0.81**

The distributions of the crustaceans, molluscs, polychaete worms, and other faunal classes were very poorly modeled in the NYWEA. This is possibly due to low correlations of each biotic cover type to the different substrate or geoforms. Crustaceans had the highest model accuracy of the remaining biotic cover types, however the OA was only 11.3% (Table 11.10). Only one instance of crustacean presence was detected and identified accurately and only 10 instances of crustacean absence were predicted correctly. Polychaete worms, molluscs, algae, and other biotic covers were too sparsely distributed throughout the NYWEA to get an OA greater than 10%, and are not featured as final predictors in this report.

Table 11.10. Confusion matrix for presence of high occurrence of 'Crustaceans' by the BRT biotic cover predictions.

		Crustaceans AA (i)			
		Presence of Crustaceans	**Absence of Crustaceans**	**n_{-j}**	**User's Accuracy**
Map (j)	**High Probability of Occurrence (>14%)**	**1**	78	**79**	1%
	Low Probability of Occurrence (<14%)	8	**10**	**18**	56%
	n_{i-}	9	88	**97**	
	Producer's Accuracy (%)	11%	11%	**OA =**	**11.3%**
				T_e =	**-0.33**

11.2 Habitat Mapping Results and Discussion

BRTs were used to predict the presence of individual substrate, geoform and cover types. BCTs were used to generate a composite habitat map. The results from these models and maps are described below, including the main features of the habitat predictions along with performance and accuracy results.

11.2.1 Model Performance

Six BRT models and resulting spatial predictions describe the probability-of-occurrence for one substrate (Pebbles), two geoform (Rippled Sand, Megaripples) and four biotic cover types (Crustacean, Echinoderm, Less Than 5% Cover, and No Cover (Bare)).

For all models, kCV PDE ranged from 7.1% to 38.2% ($\underline{x}$ =19.6% ±10.5 SE). Model results 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 low amounts of error. In all cases, except Echinoderms and Pebbles, test PDE was negative or low and difficult to interrupt. Low test PDE was most likely a result of some habitat classes being undersampled in the AA data collection.

The best model (100% No Cover) had the highest kCV (38.2%) and AUC (0.94). It also had low bias (-0.028) and RMSE (0.18). Based on AUC criteria alone, the Pebbles, Echinoderms, and Less than 5% Cover models produced good to outstanding performance. For Ripples, Megaripples and Crustaceans, these models did not meet optimum model performance.

The Crustacean model had the lowest kCV PDE (7.13%), and the Megaripple model had the lowest test PDE (-26.9%). Bias was small to moderate for all models, ranging between -0.14 to +0.16 ($\underline{x}$ =-0.001 ±0.1 SE). Bias indicates whether the model under predicted (-) or over predicted (+) the probability-of-occurrence. Lastly, RMSE values ranged from 0.18 to 0.49 ($\underline{x}$ =0.35 ±0.12 SE). The Megaripple model had the largest amount of error (0.49), while the No Cover model had the lowest error (0.18).

11.2.2 Predictor Importance

The relative contribution of each predictor differed among the substrate, geoform, and cover models (Figure 11.10). 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 potential physical drivers that influence the distribution of habitats. Backscatter and/or depth were the primary drivers for all types except Crustaceans. The main contributor to the Substrate (Pebbles) model was backscatter (32.3%) followed by cosine of aspect (8.1%). The geoform models were largely influenced by depth (Ripples = 55.6%; Megaripples = 52.7%), followed by cosine of aspect (Ripples = 8.14%) and slope of slope (Megaripples = 5.15%). Three biotic cover models were largely influenced by backscatter and then depth; specifically Echinoderms = 29.3% backscatter and 14.9% depth; Less than 5% Cover = 27.7% backscatter and 17.2% depth; and No Cover = 54.1% backscatter and 9.0% depth. Slope of slope (12.1%) and sine of aspect (9.7%) were the strongest contributors to the Crustacean model.

11.2.3 Geographic Patterns of Substrate

Substrate: Pebbles

Pebbles (Figure 11.4) were distributed throughout the NYWEA, with a concentration of observations in the northwest section of the NYWEA. Pebbles were observed at 16% of the GV sites (47 of 293) (Figure 11.4b). There were few observations in the southeastern portion of the NYWEA. The predicted surface produced from the 'Pebbles' model reflected this same spatial distribution (Figure 11.4e). There was high probability of pebbles in the western half of the NYWEA which coincided with high CV values in these same locations (Figure 11.4f), indicating higher uncertainty for places where pebbles are more likely to be present.

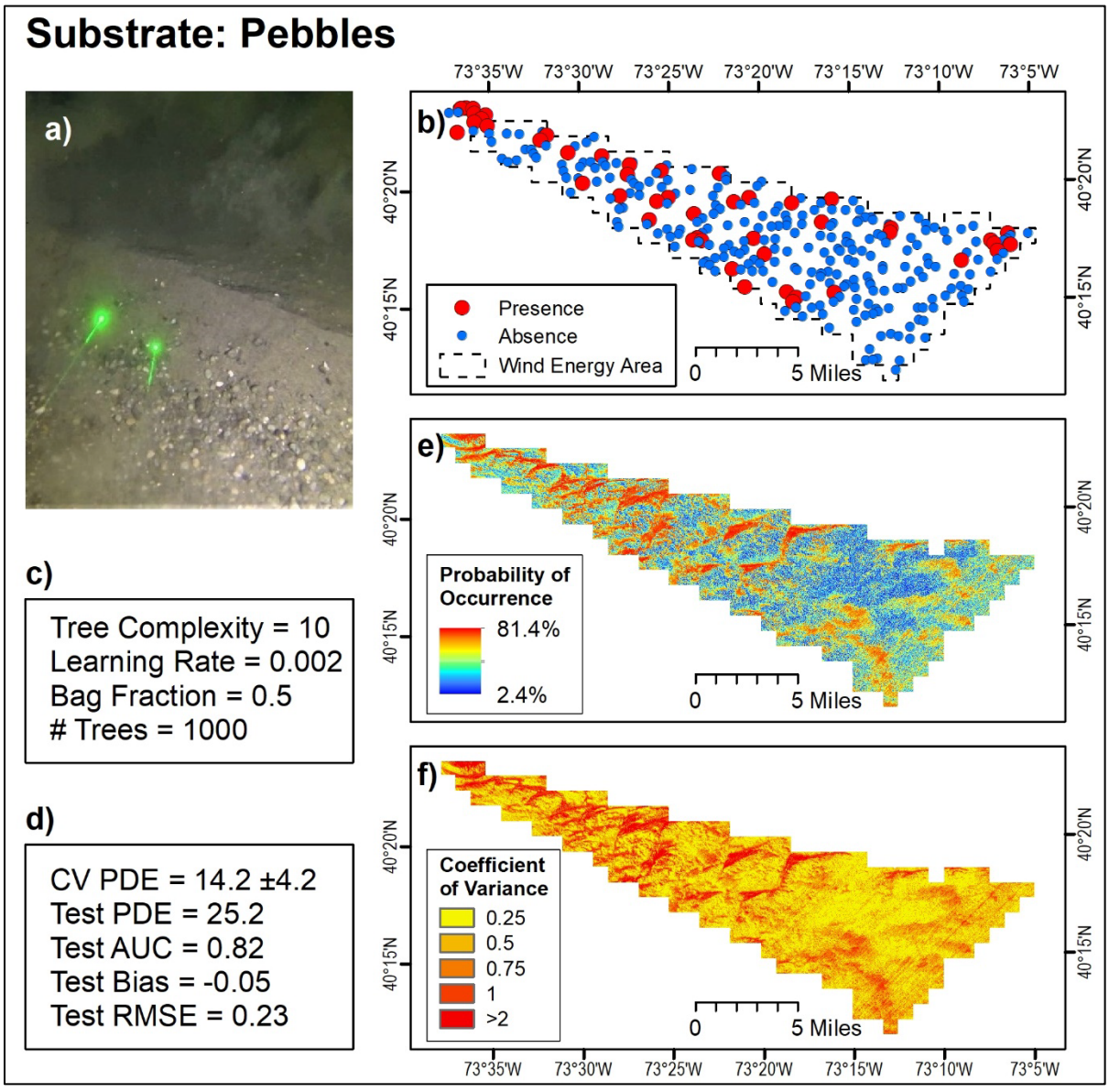


Figure 11.4. Predicted probability of occurrence for 'Pebbles'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

Geoform: Rippled Sand

Rippled Sand (Figure 11.5) was distributed throughout the NYWEA, decreasing slightly toward the offshore. Rippled sand was observed at 83% of the GV sites (243 of 293) (Figure 11.5b). The predicted surface produced from the 'Rippled Sand' model reflected this same spatial distribution with high probability of rippled sand in the western half and lower probability of occurrence moving offshore (Figure 11.5e). CV values were low across the NYWEA (Figure 11.5f), indicating low uncertainty across the entire NYWEA.

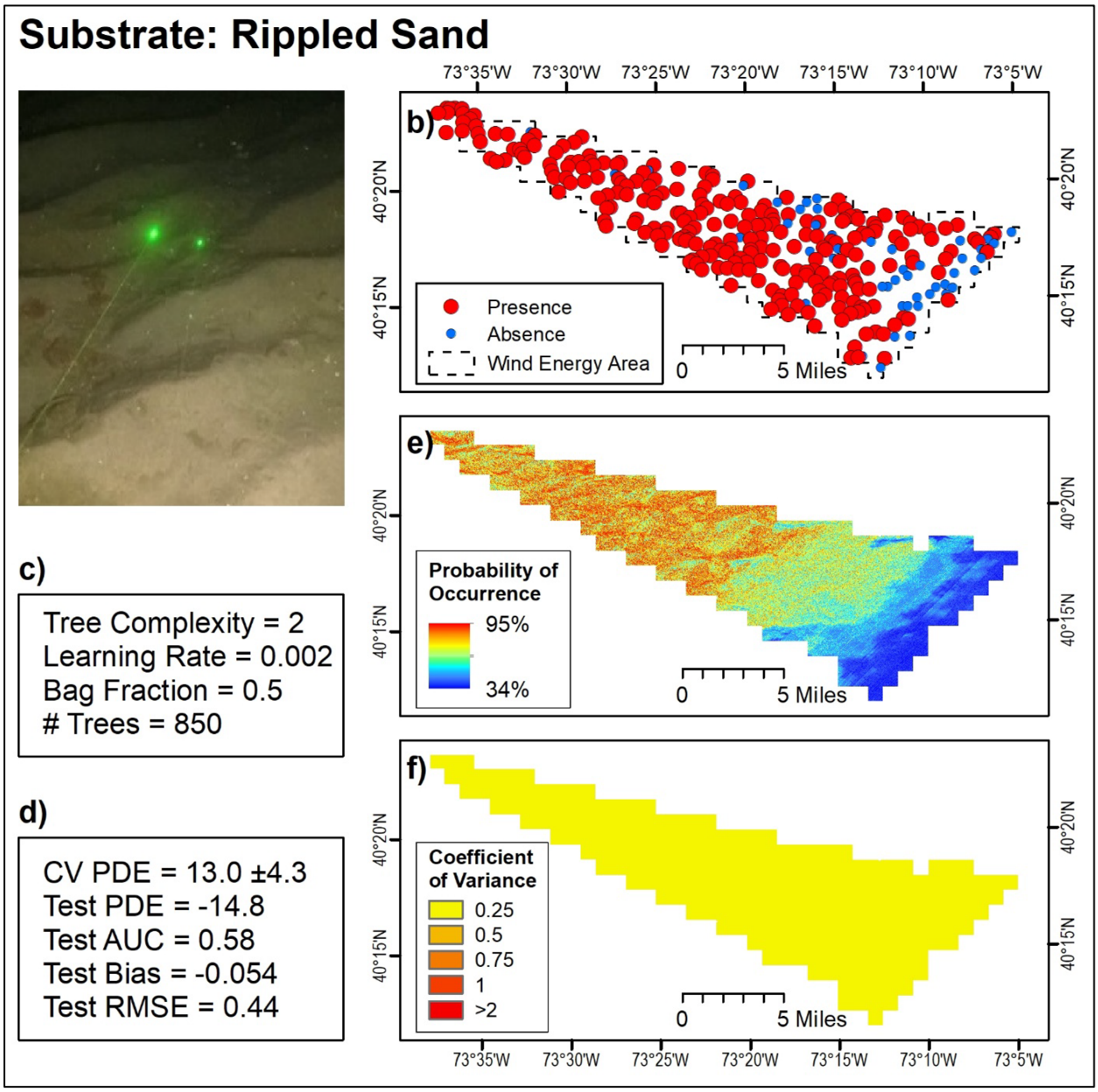


Figure 11.5. Predicted probability of occurrence for 'Rippled Sand'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

Geoform: Megaripples

Megaripples (Figure 11.6) were primarily distributed on the eastern end of the NYWEA and observed at 12.2% of the GV sites (36 of 293) (Figure 11.6b). The predicted surface produced from the 'Megaripples' model reflected this same spatial distribution with high probability of occurrence in the east and lower probability of occurrence closer to shore (Figure 11.6e). CV values were high where the probability of occurrence was high (Figure 11.6f), indicating high uncertainty where the likelihood of occurrence is high.

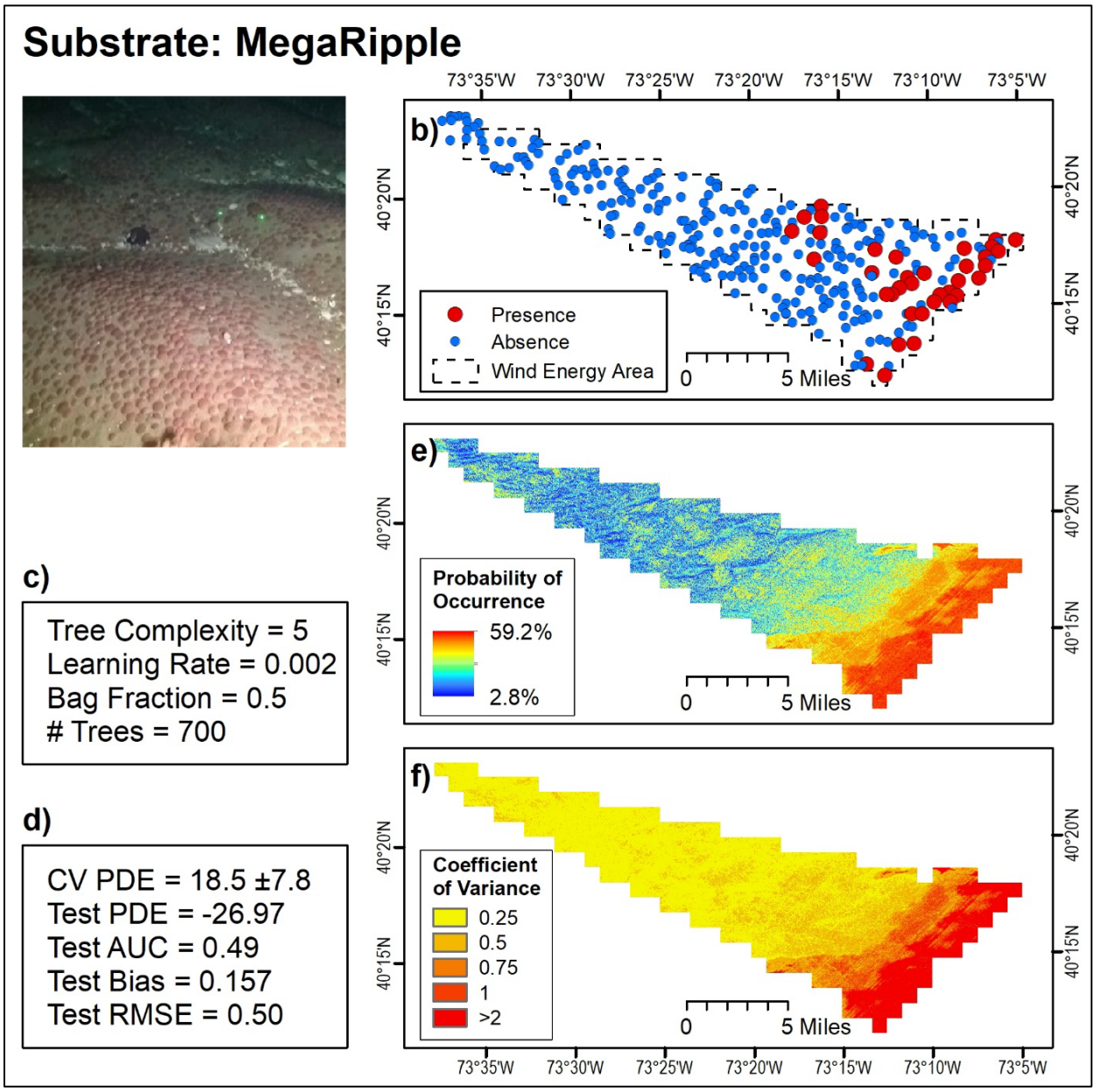


Figure 11.6. Predicted probability of occurrence for 'Megaripples'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

11.2.4 Geographic Patterns of Cover

Cover: Crustacean

Crustaceans (Figure 11.7) were distributed primarily in the central and eastern part of the NYWEA, observed at 10.9% of the GV sites (32 of 293) (Figure 11.7b). The predicted surface produced from the 'Crustacean' model reflected this same spatial distribution of high probability of occurrence, as well as higher uncertainty in the central and eastern section of the NYWEA (Figure 11.7e and 11.7f).

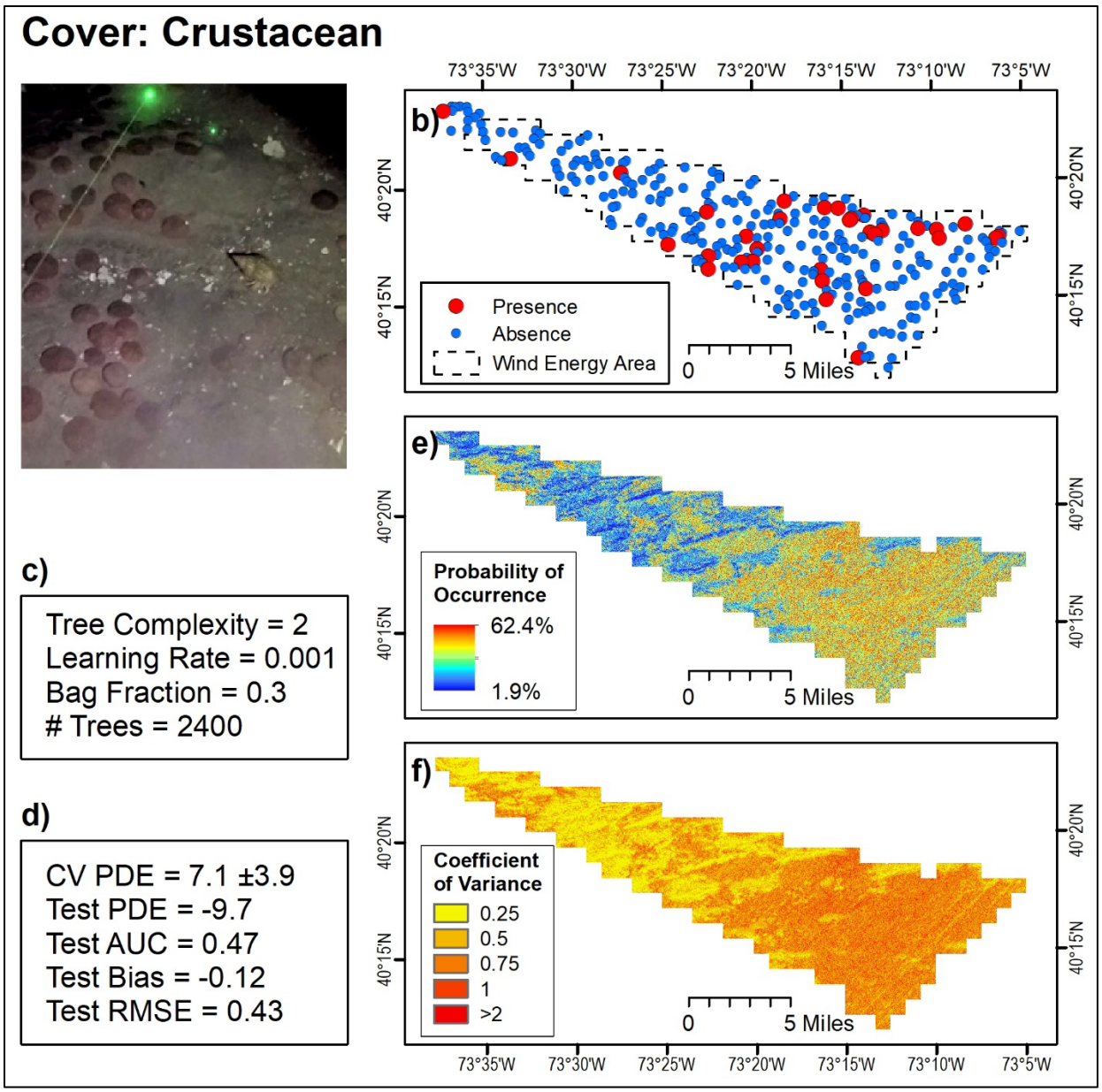


Figure 11.7. Predicted probability of occurrence for 'Crustaceans'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

Cover: Echinoderms

Echinoderms (Figure 11.8) were observed throughout the NYWEA, occurring in 90% of the GV sites (264 of 293) (Figure 11.8b). The predicted surface produced from the 'Echinoderm' model indicated that the probability of occurrence of echinoderms was highest in the eastern half of the NYWEA (Figure 11.8e). CV values were low across the NYWEA (Figure 11.8f), indicating low uncertainty in the model results.

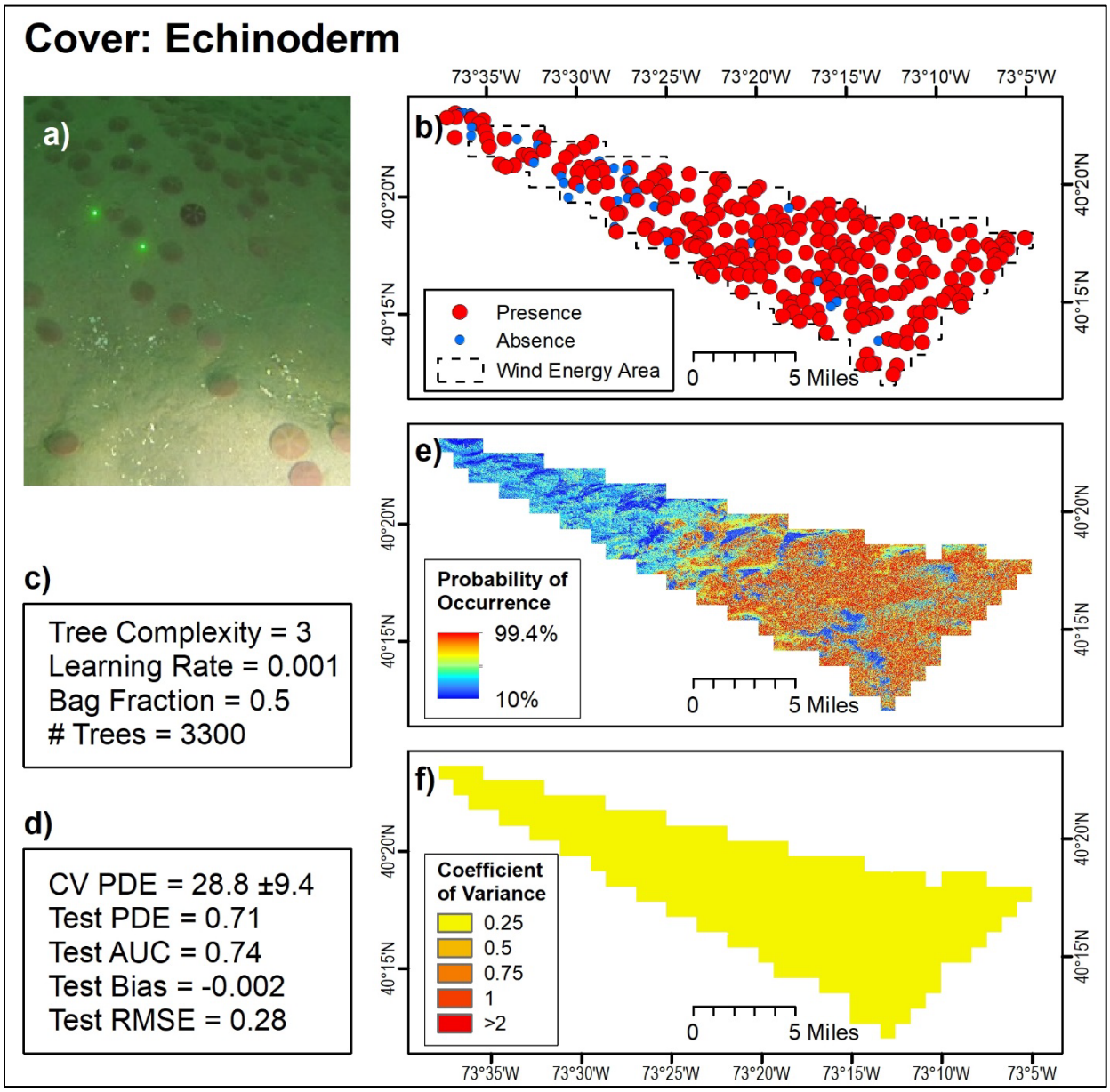


Figure 11.8. Echinoderms. Predicted probability of occurrence for 'Echinoderms'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

Cover: Less than 5% Cover

Less than 5% Cover (Figure 11.9) was distributed throughout the NYWEA, observed in clusters in the western and central/eastern portions of the NYWEA. Less than 5% cover was observed at 33% of the GV sites (96 of 293) (Figure 11.9b). The predicted surface produced from the 'Less than 5% Cover' model reflected a similar spatial distribution with high probability of less than 5% cover in the western half with lower probability of occurrence moving offshore (Figure 11.9e). CV values were low across the NYWEA, especially in the eastern region (Figure 11.9f), indicating low uncertainty in the model.

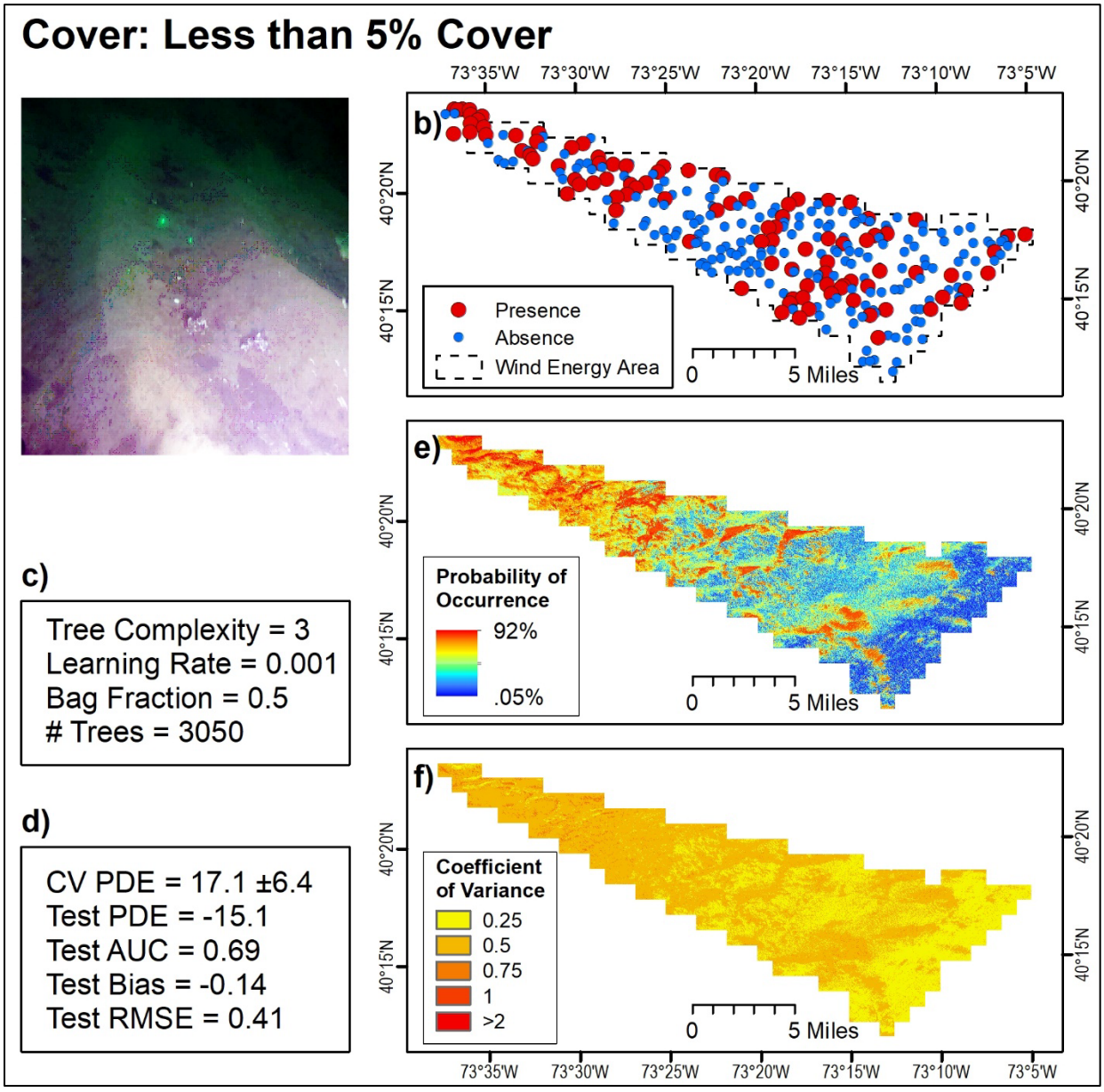


Figure 11.9. Predicted probability of occurrence for 'Less than 5% cover'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

Cover: No Cover

Sites where there was 100% No Cover (Figure 11.10) were scattered, occurring mostly in the northwest portion of the NYWEA. No Cover was observed at 8.9% of the GV sites (25 of 293) (Figure 11.10b). The predicted surface produced from the 'No Cover' model reflected this same spatial distribution of high probability of rippled sand in the western half and lower probability of occurrence moving offshore except for some distinct features in the southern end (Figure 11.10e). CV values were low across the NYWEA (Figure 11.10f), indicating low uncertainty in the model.

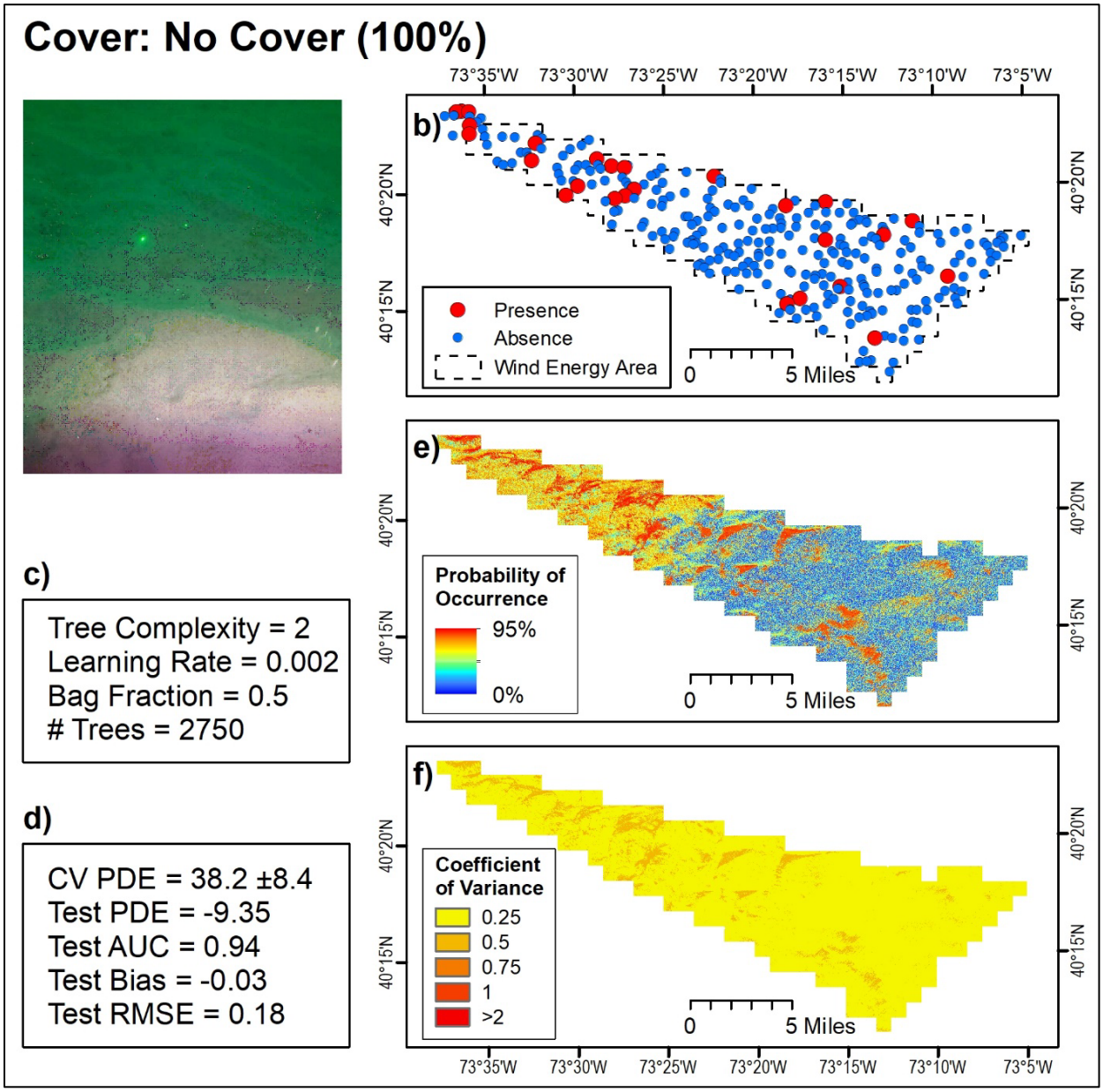


Figure 11.10. Predicted probability of occurrence for 'No Cover'. Figure panels depict: a) reference photo; b) presences and absences in the GV data; c) the input parameters used to create the final model; d) the performance of the final model; e) the predicted average probability-of-occurrence and f) coefficient of variation.

11.2.5 Composite Habitat Map

A total of 325.5 km^2 of the NYWEA was characterized in this project. The composite habitat map displays the predicted spatial distribution of three commonly co-occurring combinations of substrate, geoform and biotic cover types (Figure 11.11). 'Rippled sand with high occurrence of faunal beds' was the most abundant habitat type, comprising 78.2% (254.5 km^2) of the area and was dominant throughout much of the NYWEA. 'Megaripple sand with high occurrence of faunal beds' was the next most abundant habitat mapped, comprising 14.7% (47.9 km^2) of the area. This class dominated the southeast edge of the NYWEA. 'Rippled sand and pebbles with low occurrence of faunal beds' habitats occurs throughout the NYWEA, but especially in the northwest portions of the NYWEA, seemingly distinctive disc-shaped features. This class comprised 7.1% (7.1 km^2) of the NYWEA.

The map products described here take advantage of recent advancements in computing power and model-based mapping techniques to create highly resolved raster maps. A composite habitat map was created describing commonly co-occurring substrate, geoform and cover types. Likewise, the underlying habitat predictions provide a view of the probability of occurrence of several habitat types. Both types of products are pixel based, preserving fine-scale heterogeneity, associated habitat gradients, and smaller benthic features present across the seascape. Compared to a polygon mapping approach, this method diverges from absolute classifications, and instead predicts the likelihood that a specific substrate, geoform or cover type is present at any given location. These predictions also quantified the uncertainty associated with each pixel in the map, helping users understand the precision of these predictions and prioritize where to collect additional information moving forward.

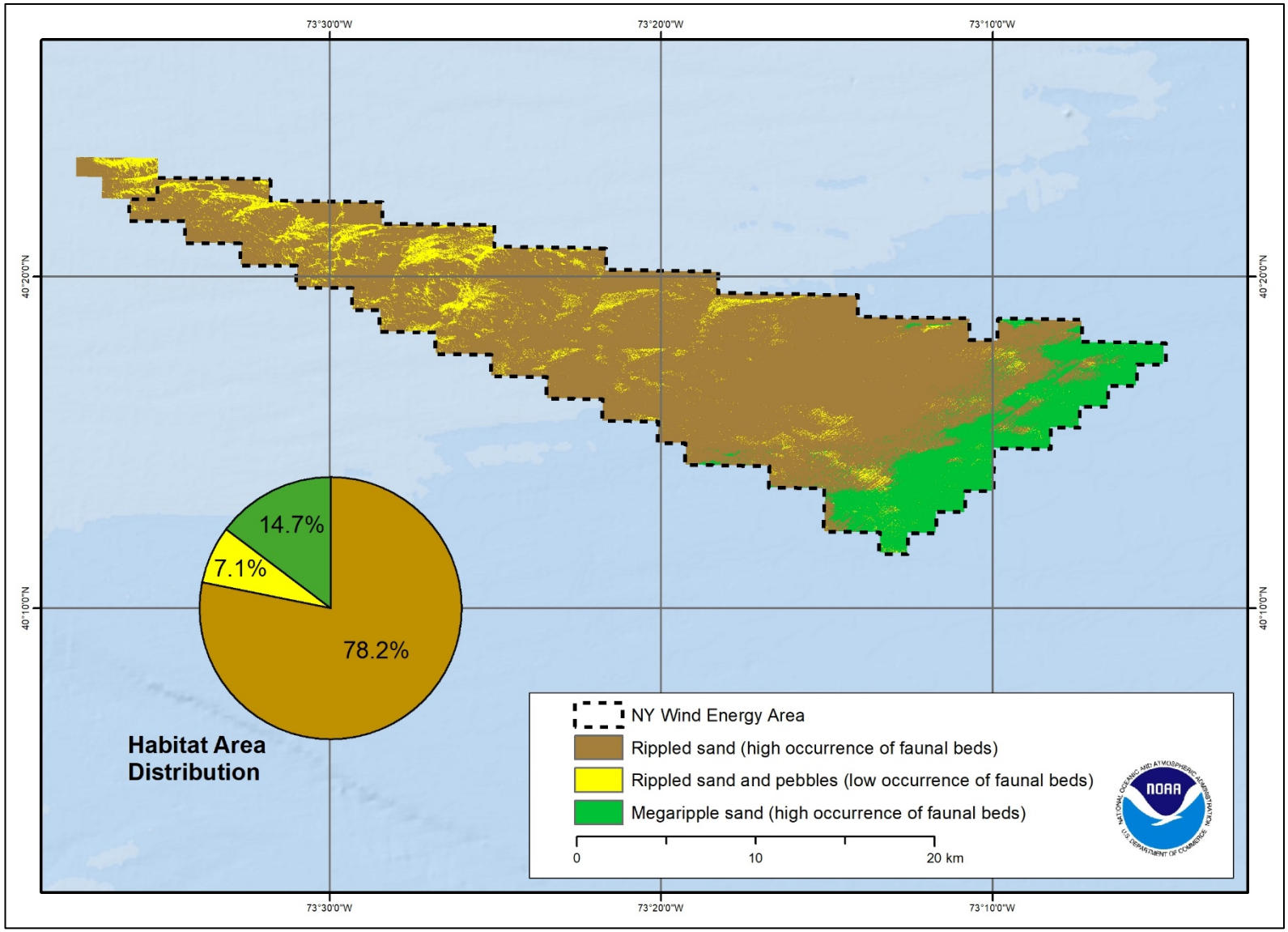


Figure 11.11. Three co-occurring benthic habitats mapped throughout the NYWEA and the overall proportion of area occupied by each habitat class.

Chapter 12. Fish Acoustic Analysis

Data acquired using fisheries acoustics surveys provides valuable information used to understand the inhabitants of marine ecosystems, including predator-prey interaction, diel migration, and broad scale density distribution over a variety of geoforms. Inferring species from the acoustic signatures is not yet possible, but the high spatial resolution and precise estimations of size coupled with seafloor geoform information can inform resource managers regarding physical and biological relationships within that study area (Kracker et al. 2011). The repeatability of large-scale fishery acoustics surveys are ideal for monitoring or assessing the efficacy of marine resources prior to and following implementation of NYWEA. (Foster et al. 2013). Fisheries acoustic products developed under this project document baseline conditions within the NYWEA during the month of September.

12.1 Spatial Distribution

Throughout the survey area, significant variation in the spatial distribution of individual fish, numerous schools of fish with a variety of shapes and dimensions, and plankton layers and patches that were particularly evident during a range of hours (Figure 12.1). The colors within each panel are represented as: the seafloor (dark red), a dense layer of individual fish high in the water column (green-yellow-orange) (Figure 12.1A); plankton throughout the water column (light blue) (Figure 12.1B); a tightly aggregated fish school close to the seafloor (yellow-red) (Figure 12.1C); and loosely aggregated fish schools high in the water column (green-yellow-orange) (Figure 12.1D).

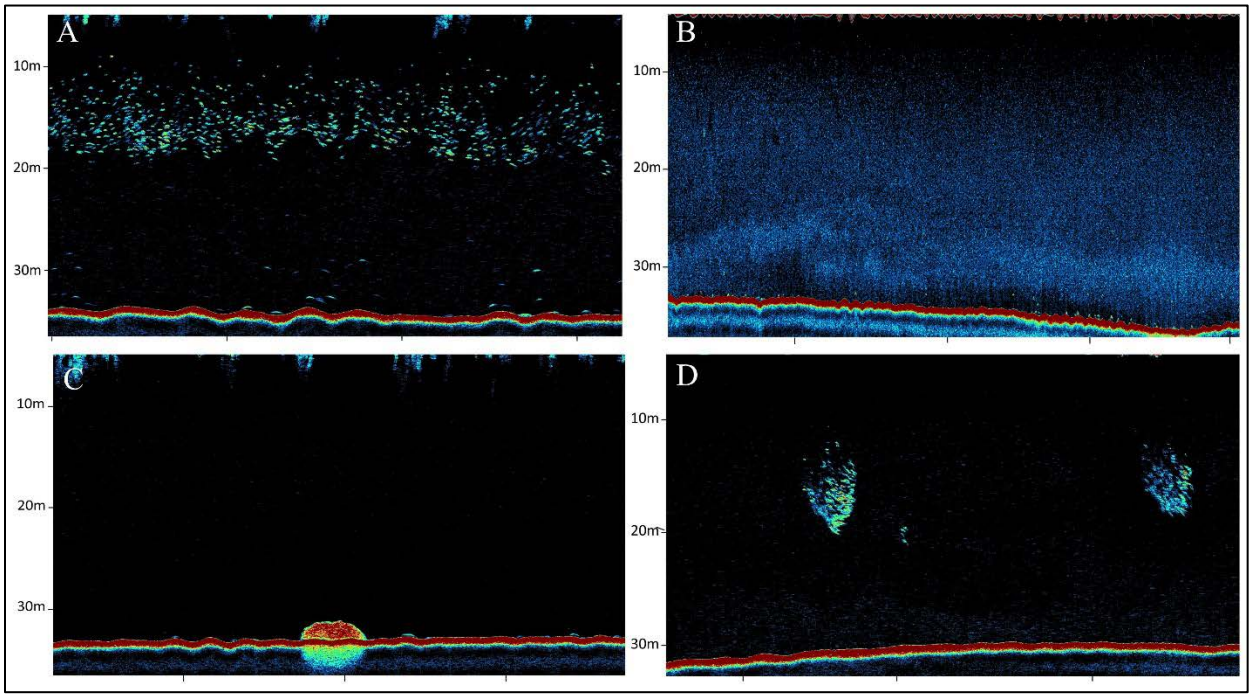


Figure 12.1. Example of splitbeam echosounder (SBES) echograms with depth represented on the y-axis with tick marks on the x-axis representing distance every 100 m. Panels A and B during overnight hours, and panels B and C daylight hours.

The largest concentration of fish were found in the northwest portion of the survey area (Figure 12.2). Spatial distribution of fishes in the survey area varied by size class (total length as estimated from acoustic target strength). Densities were divide into three size classes, small (<11 cm), medium (11-29 cm), and large (>29 cm). Small fish (Figure 12.3) were vastly more abundant and broadly distributed throughout the survey area compared to medium (Figure 12.4) and large (Figure 12.5) size class fish, which were strongly associated with the northwest portion of the survey area. The spatial distribution of fish density in the water column, especially at night, was likely related to fish feeding behavior. In the absence of discernable structured habitat on the seafloor within the study area, the fish density that was observed may be related to seasonal transient use or migration of demersal and pelagic fish.

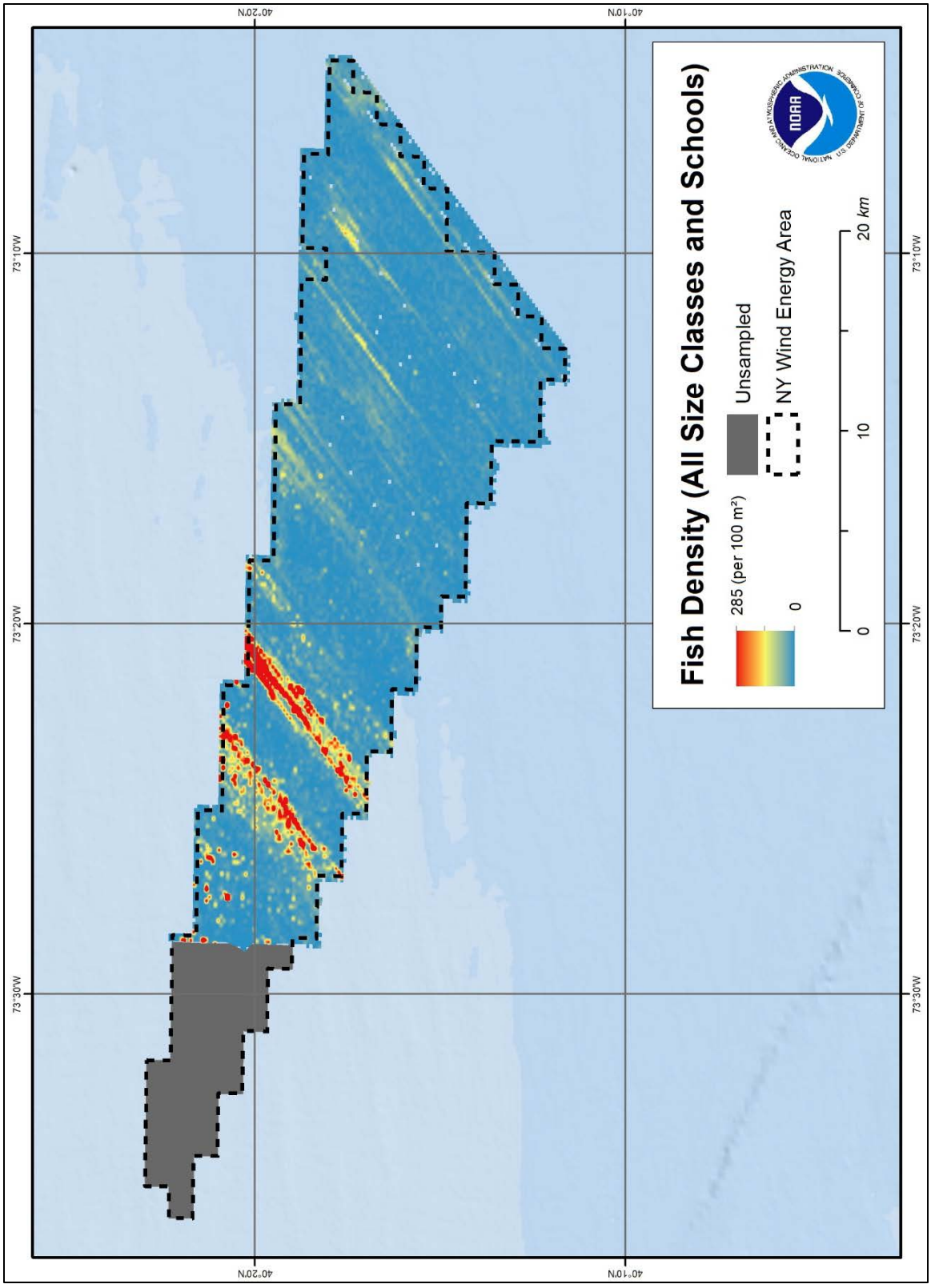


Figure 12.2. Interpolated distribution of total fish density including all size classes and fish schools.

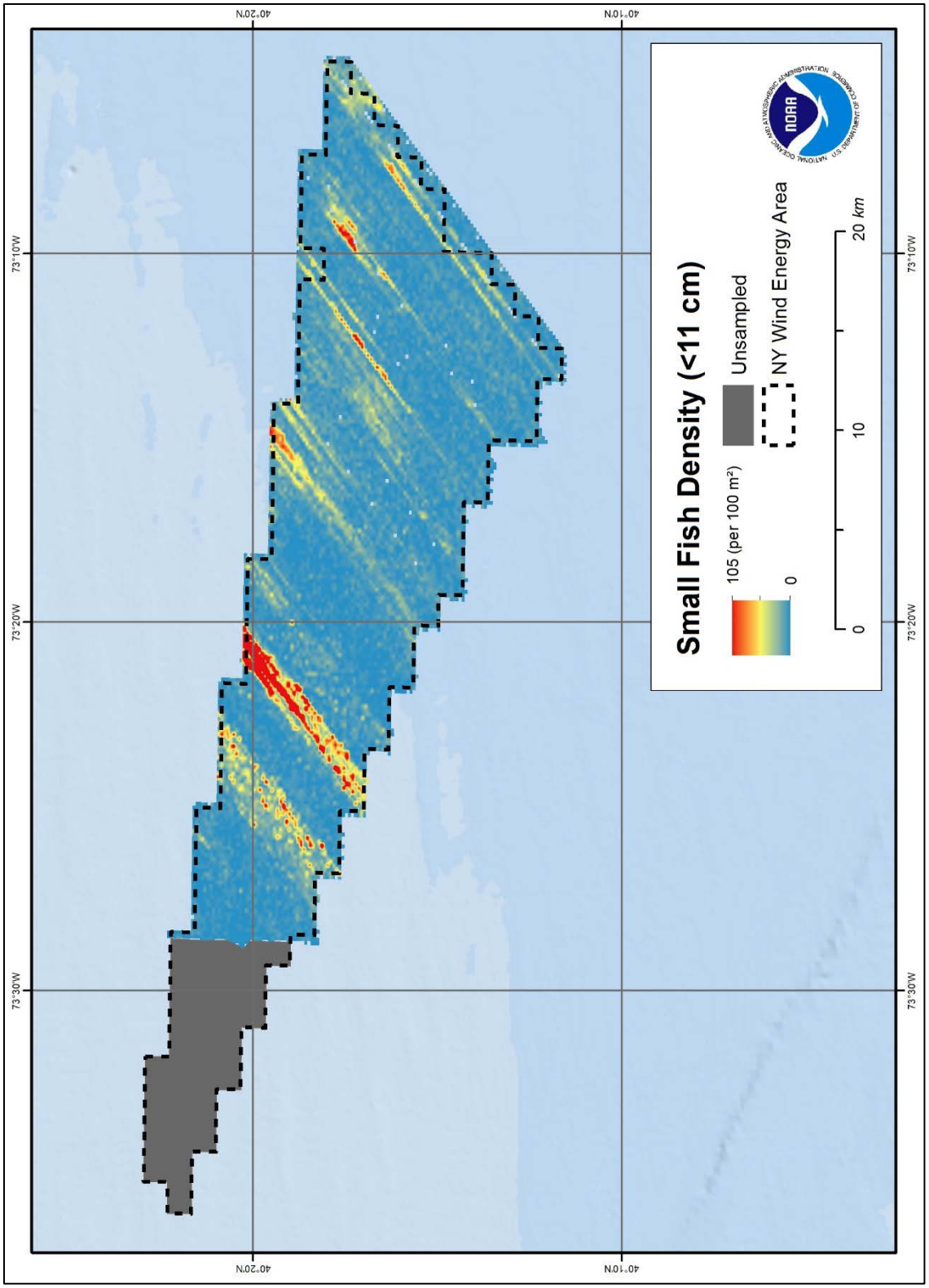


Figure 12.3. Interpolated distribution of small size class (<11 cm) fish density.

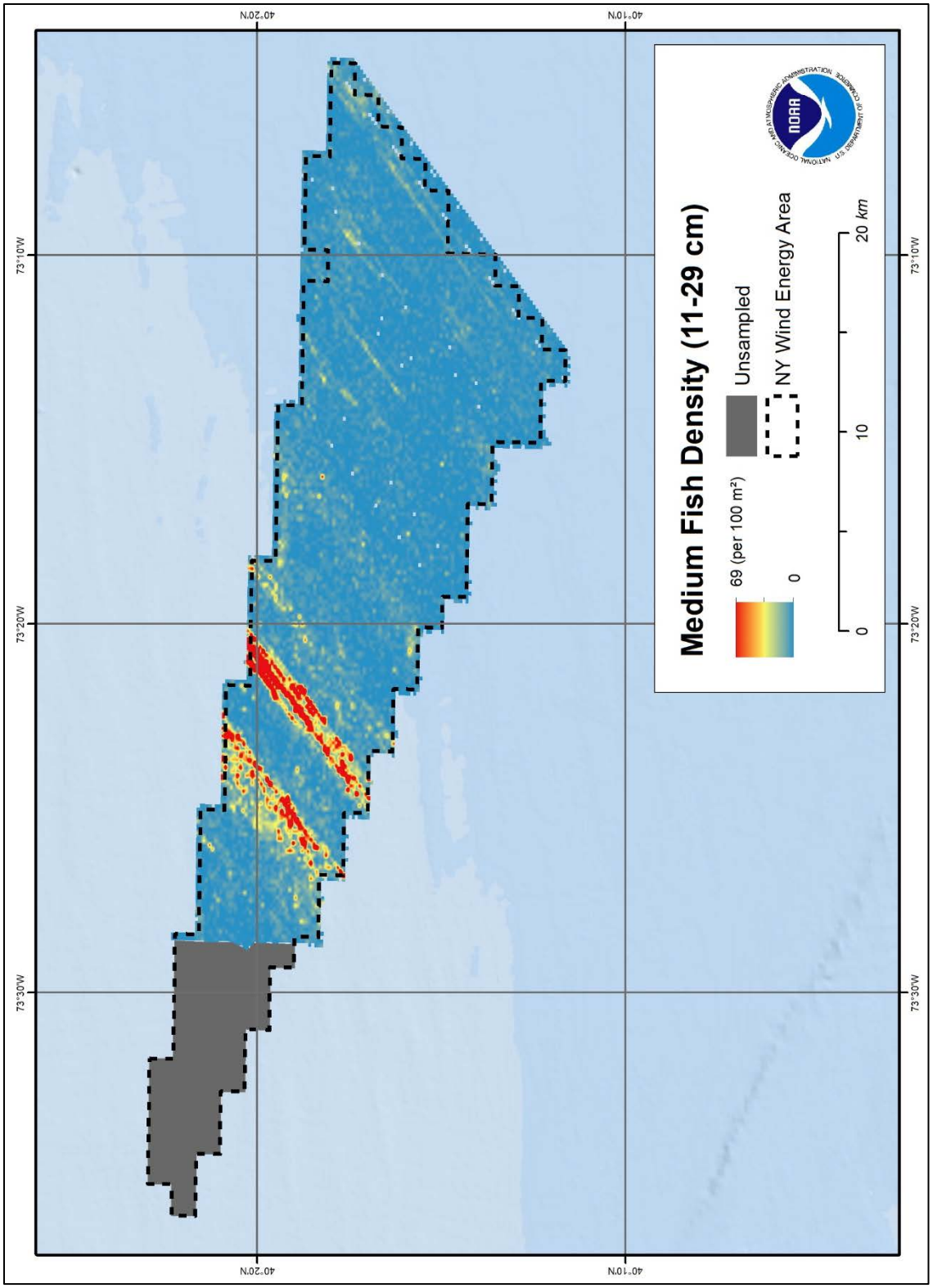


Figure 12.4. Interpolated distribution of medium size class (11–29 cm) fish density.

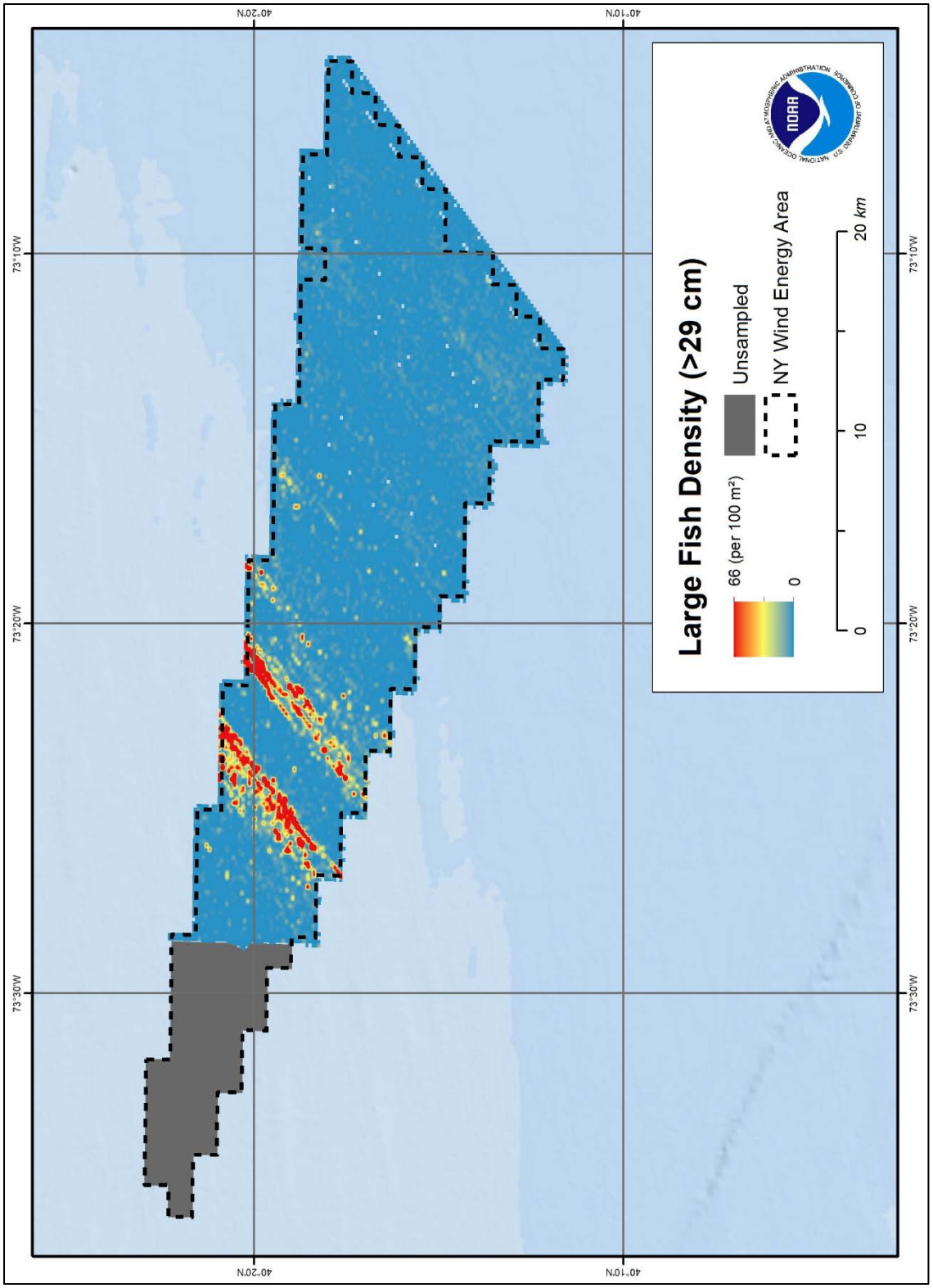


Figure 12.5. Interpolated distribution of large size class (>29 cm) fish density.

12.2 Diel Patterns

Acoustic data collection was conducted over a 24 hour time period. As a result, the data showed a distinct “diel striping” pattern that can be attributed to day-night fish behavior and increased acoustic detection in the water column of certain size classes and feeding guilds from day to night. During the daytime, fish are typically associated with the seafloor or in densely packed schools. At dusk, fish leave the seafloor in more loosely packed schools (Axenrot et al. 2004; Campanella et al. 2019) (Figure 12.6). During daytime hours, acoustic density is typically lower compared to nighttime hours due to the fishes’ association with the seafloor or “deadzone”, as well as underestimating fish within tightly packed schools. The “deadzone” is a region close to the seafloor where fish cannot be acoustically detected due to the spherical nature of the acoustics beam and the strong seafloor echo that masks the weaker fish echoes. The diel striping pattern was evident in all size class maps and coincided with the plankton vertical movement patterns (Figure 12.7). The pattern exhibited clear evidence of diel migration that occurred close to nautical twilight (sun is between 6 degrees and 12 degrees below the horizon), with the vertical migration starting about 2 hours prior to sunset and the descending migration starting about 1 hour prior to sunrise. Most fish were associated with the seafloor during the daytime, but at night, small and medium size fish were seen utilizing the full water column.

Large fish were either associated with the seafloor or in the mid/upper portions of the water column (Figure 12.8). Large, densely packed schools of small to medium size fish were observed throughout the water column during the daytime hours and transitioned to loosely aggregated schools during nighttime hours (Figure 12.9). The daytime schooling behavior of the small to medium size are mechanisms for minimizing predation and did not appear to be associated with features on the seafloor. At dusk, the small to medium size fish transitioned to loosely aggregated schools, which is likely related to feeding behavior. This behavior is coupled with plankton rising into the water column and reduced predation risk from visual predators during the nighttime (Axenrot et. al. 2004; Campanella et al. 2019).

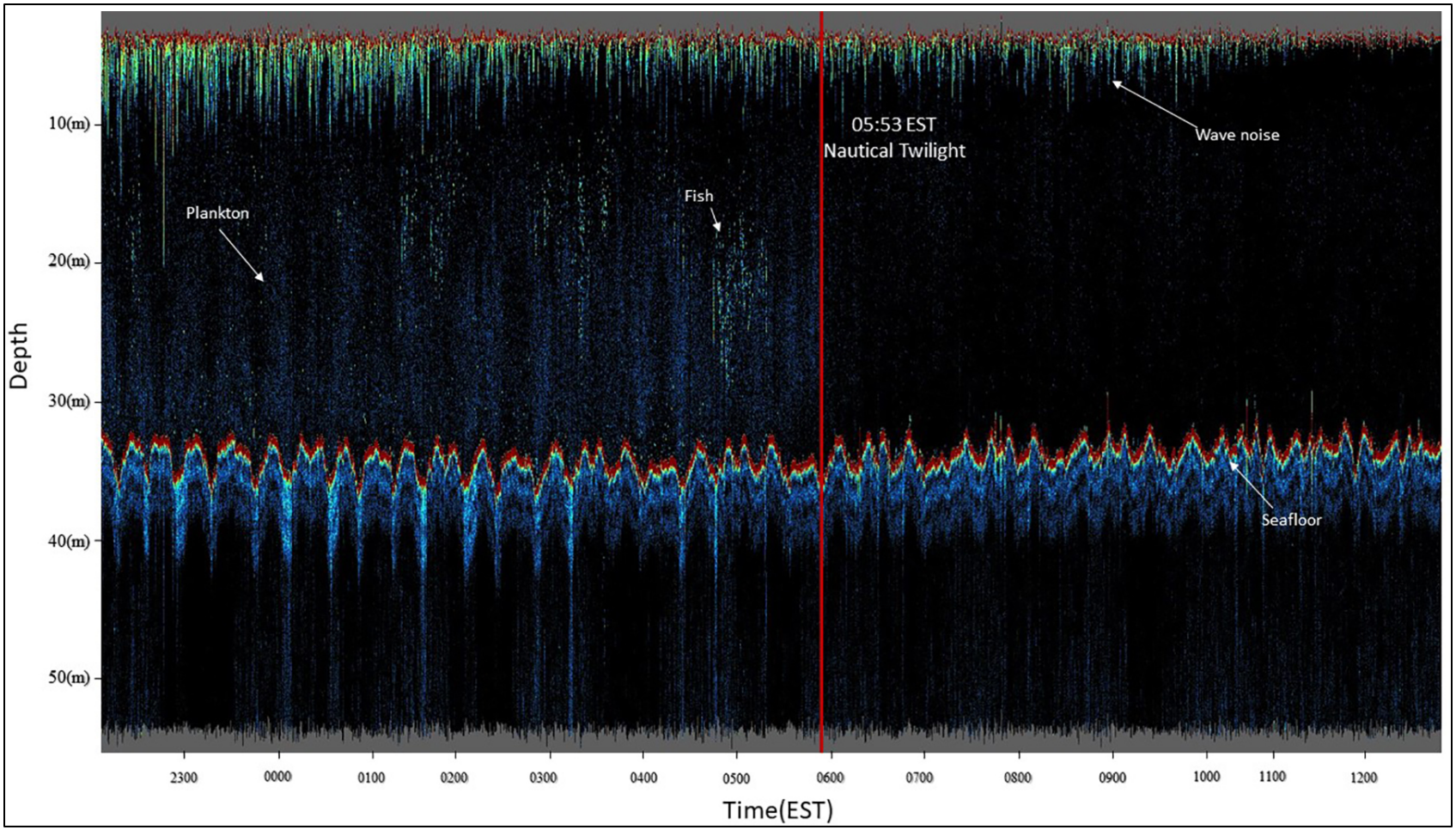


Figure 12.6. SBES echogram of the diel migration of fish and plankton overnight into the mid-afternoon. Red line represents the nautical twilight during the time period of this sample.

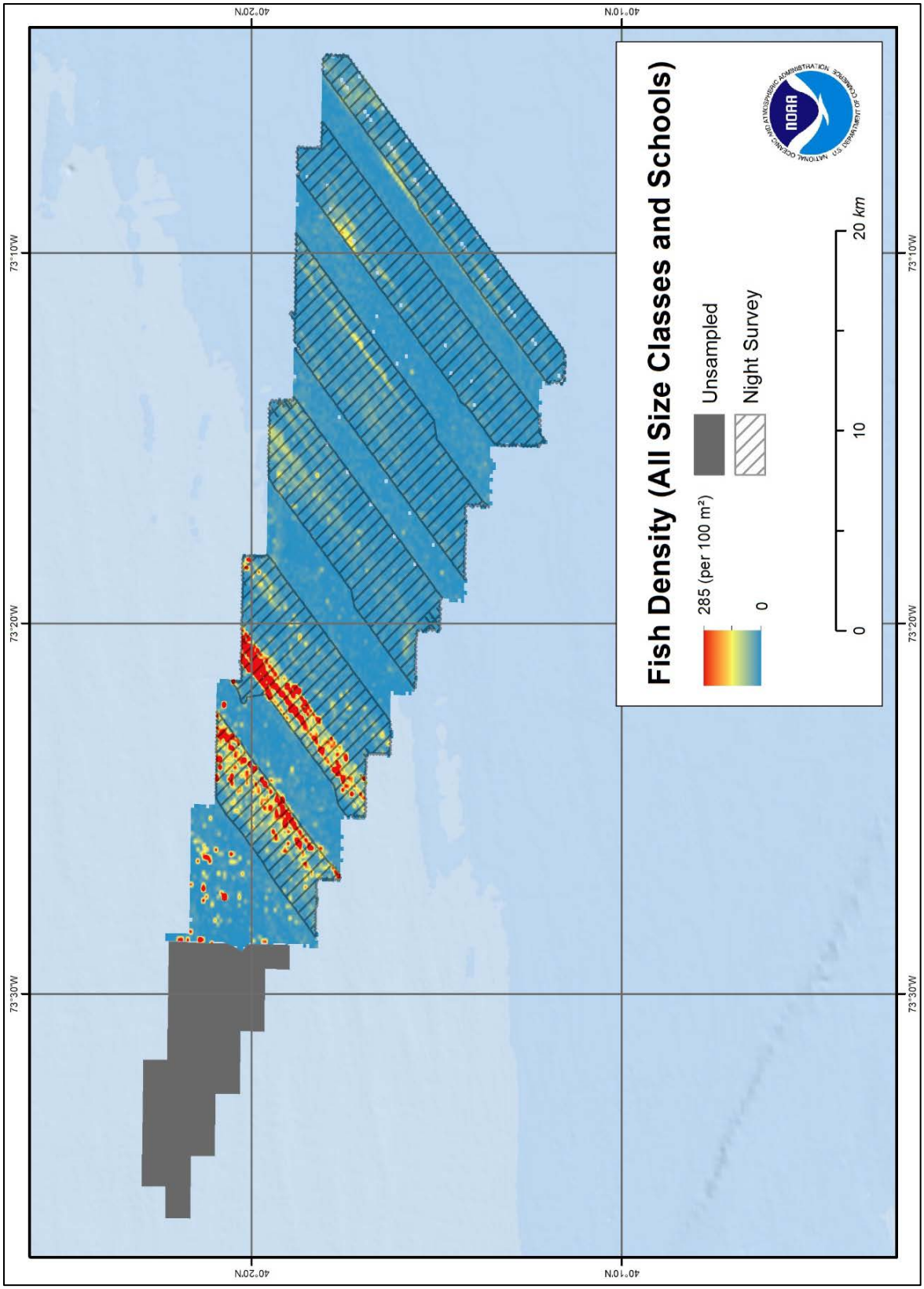


Figure 12.7. Interpolated distribution of total fish density, including all size classes and fish schools.

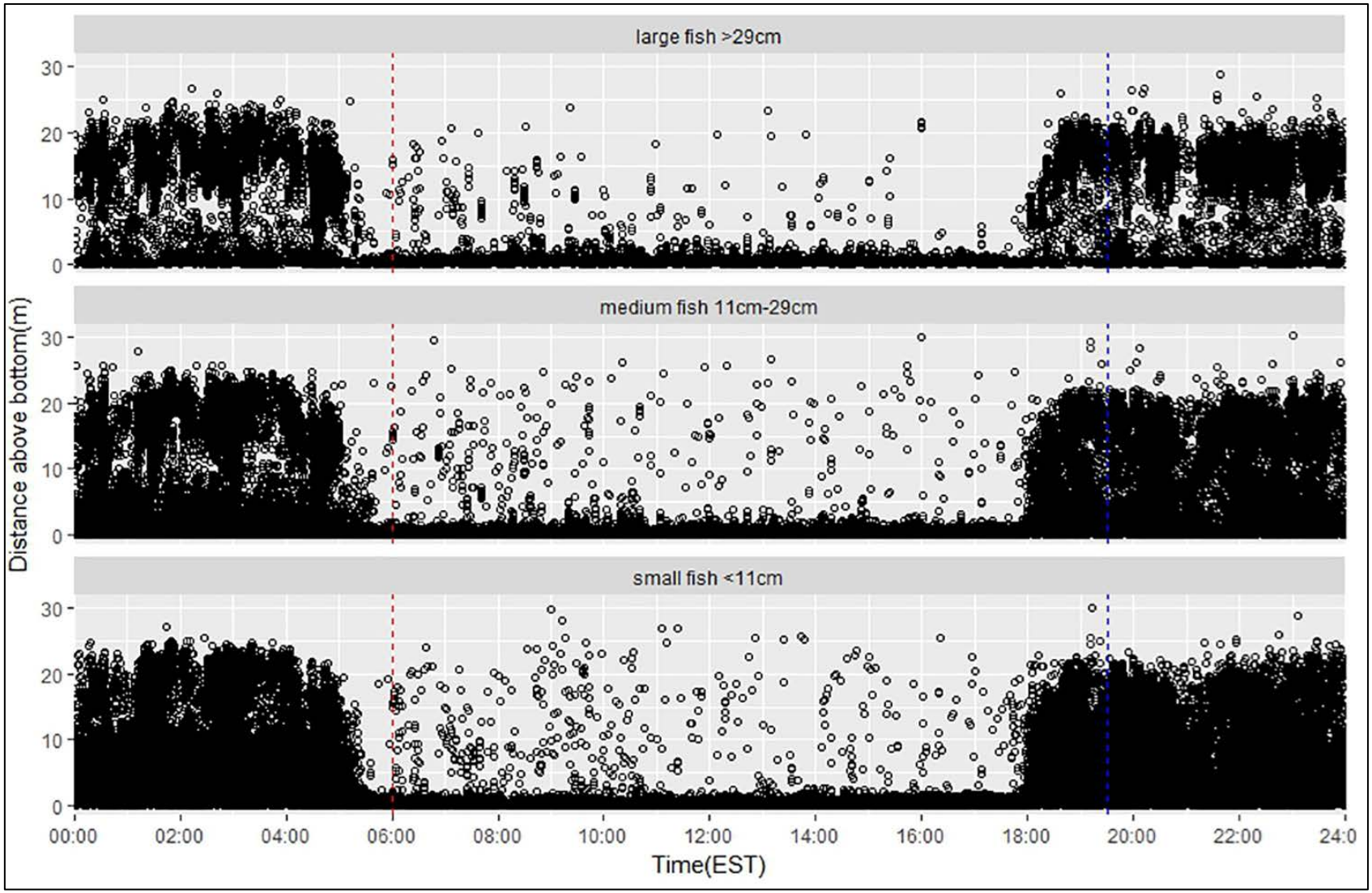


Figure 12.8. Diel pattern of individual fish represented by hollow dots throughout the survey area. The red and blue dashed lines represent nautical twilight for the survey period (October 1-9, 2017).

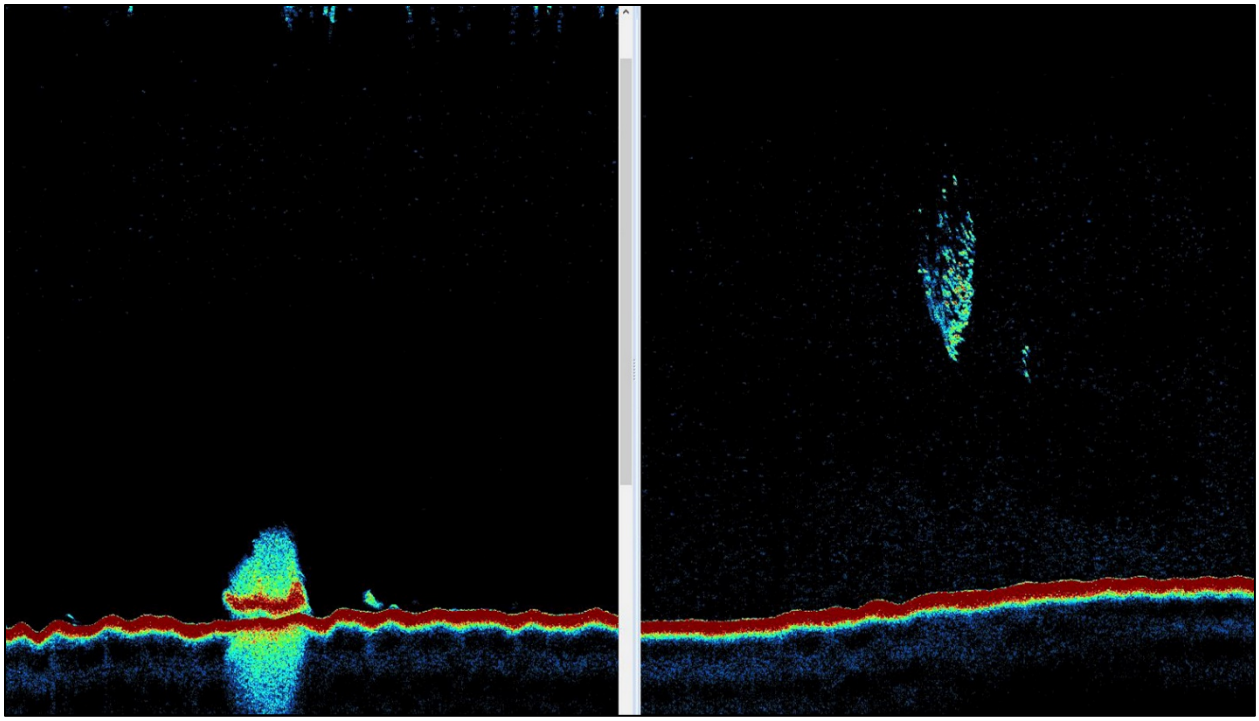


Figure 12.9. SBES echogram demonstrating the difference between a tightly packed school (left) versus a loosely packed school (right).

12.3 Geoform-Density Relationship

Fish density (per 100 m^2) and the composite habitat map (see Chapter 11) were used to analyze the association of fish densities with habitat types within the survey area. In ArcGIS, the three habitat types (i.e., rippled sand, rippled sand/pebbles, and megaripples) were spatially joined with the fish acoustic density points. Densities were averaged to account for the different spatial size of each habitat (Figure 12.10). Fish densities in the rippled sand/pebbles were greater than densities over rippled sand. Fish densities for both rippled sand and rippled sand/pebbles were significantly greater than megaripples. Higher overall density in rippled sand and pebbles was attributed to large pelagic schools of fish that were observed in the water column. The overall uniformity of the habitat types and the lack of hard bottom or structure suggested that fish were not associated with the habitat, but rather with the water column.

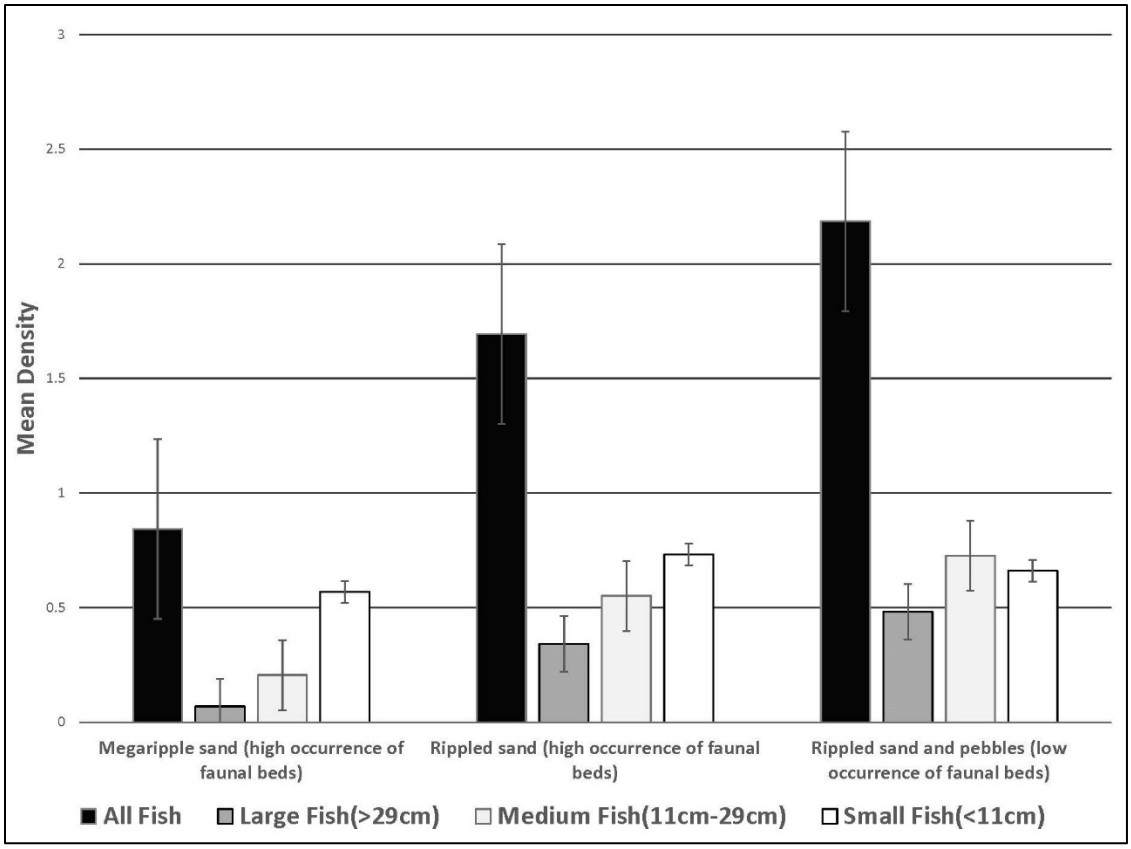


Figure 12.10. Mean acoustic densities (#/100m^2) for fish size classes and over all densities grouped by habitat type.

12.4 Fishery Independent Surveys

SBES data provides a snapshot of the spatial and vertical distribution of fishes during the mapping survey and provides indicators for potential hotspots of biomass associated with seafloor habitats. The acoustic survey methods and results presented here are repeatable and could document changes in the spatial distribution over longer time periods, in response to seasonal or environmental changes, or in response to changes in the seascape, such as the installation of marine energy infrastructure. Without coincident visual or extractive samples, species identification of acoustic fish biomass is not possible. Instead, evidence from fishery-independent surveys from the project area can provide information on what species may have been in the area during the mapping missions. Schooling and feeding behavior patterns paired with fishery trawl surveys provide a family level understanding of the fish within the survey area. Fishery-independent trawl surveys data were acquired from the NOAA Northeast Fisheries Science Center (NEFSC) for survey years 2003-2016 and contain compiled species information (Guida et al. 2017; NOAA NMFS NEFSC 2018). Eight trawl tows were conducted during the cold season (winter/spring) within the restricted spatial footprint of this project area (Figure 12.11). The trawl data showed that Atlantic herring (*Clupea harengus*), sliver hake (*Merluccius bilinearis*), alewife (*Alosa pseudoharengus*) and scup (*Stenotomus chrysops*) were the dominant species present (Figure 12.12). These species schooling behavior and transient migration patterns may comprise the fishes that are seen in acoustics surveys.

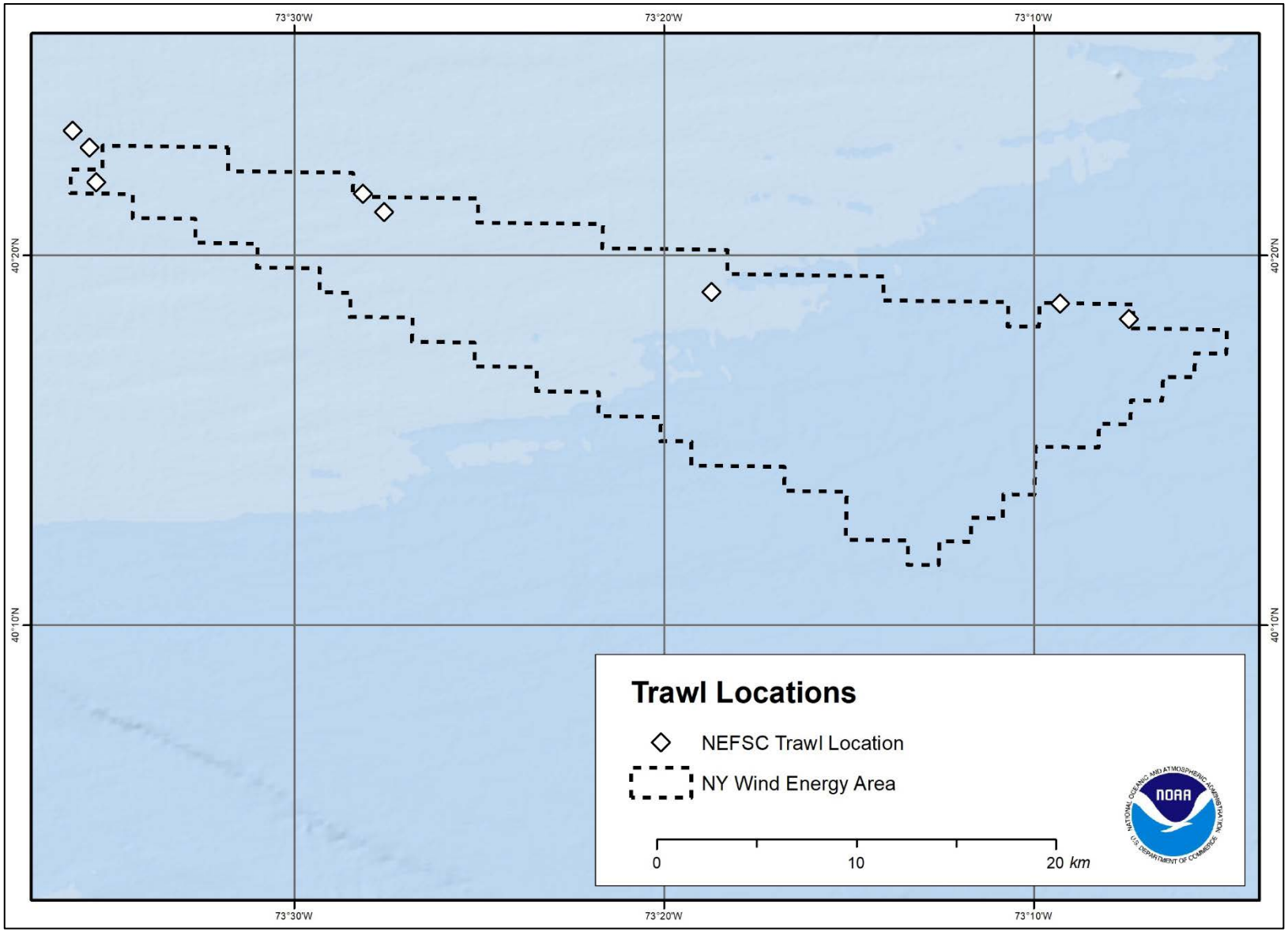


Figure 12.11. Locations of the Northeast Fisheries Science Center (NEFSC) cold (winter/spring) trawl samples. (NOAA NMFS NEFSC 2018).

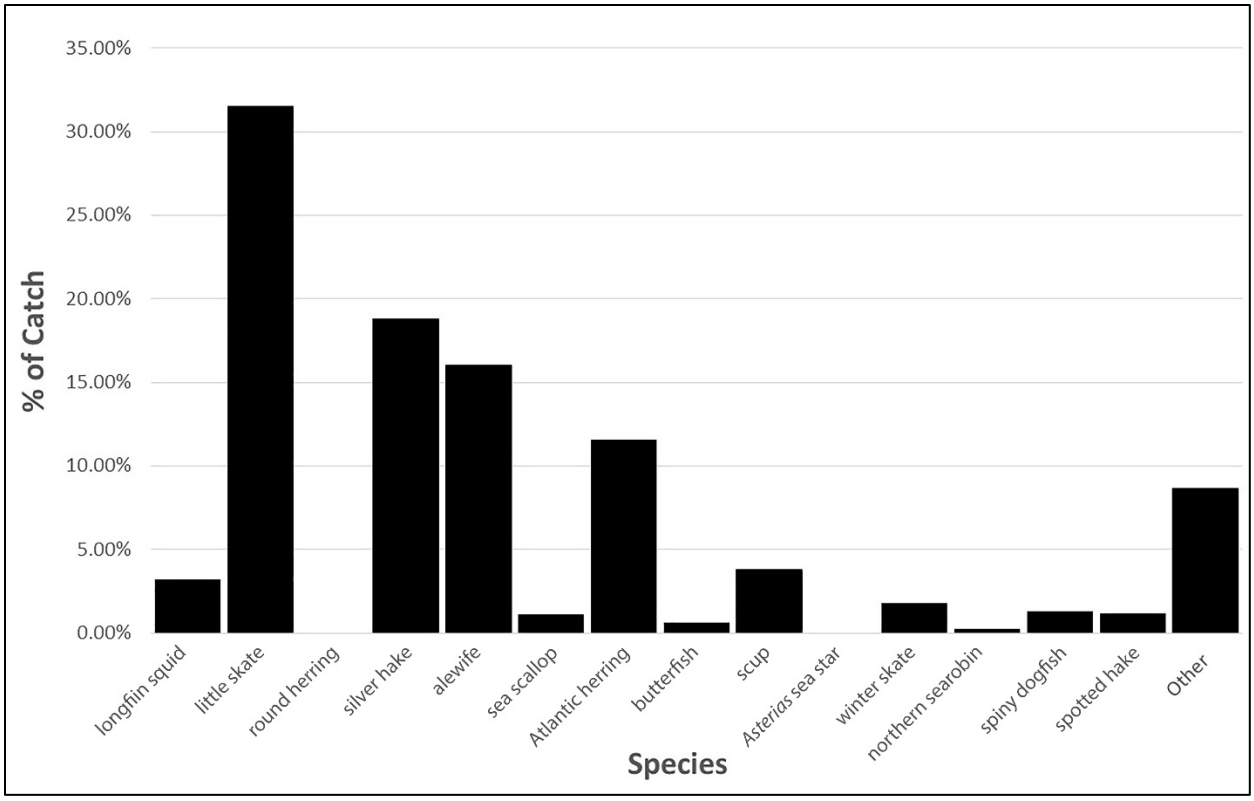


Figure 12.12. Percentage of catch by number from the NEFSC Trawl Survey between 2003 and 2016 cold (winter/spring) season (Guida et al. 2017). Locations of trawls appear in Figure 12.11.

References

Acharya, S.S., and M.K. Panigrahi. 2016. Evaluation of factors controlling the distribution of organic matter and phosphorus in the Eastern Arabian Shelf: a geostatistical reappraisal. Continental Shelf Research 126:79-88. doi: 10.1016/j.csr.2016.08.001

Aitchison, J. 1986. The Statistical Analysis of Compositional Data, Monographs on Statistics and Applied Probability. Chapman & Hall, London. Reprinted 2003 with additional material by The Blackburn Press, Caldwell, NJ, USA.

Apeti, D.A., S.I. Hartwell, W.E. Johnson, and G.G. Lauenstein. 2012. National Status and Trends Bioeffects Program: Field Methods. NOAA National Centers for Coastal Ocean Science, Center for Coastal Monitoring and Assessment. NOAA NCCOS Technical Memorandum 135. Silver Spring, MD. 27 pp.

ASTM. 2007. ASTM D422-63(2007), Standard Test Method for Particle-Size Analysis of Soils. American Society for Testing and Materials. ASTM International, West Conshohocken, PA. 8 pp.

Axenrot, T., T. Didrikas, C. Danielsson, and S. Hansson. 2004. Diel patterns in pelagic fish behaviour and distribution observed from a stationary, bottom-mounted, and upward-facing transducer. ICES Journal of Marine Science 61(7): 1100-1104. doi: 10.1016/j.icesjms.2004.07.006

Bates, R.L., and J.A. Jackson. 1984. Dictionary of geological terms (Rocks, Minerals and Gemstones), Third Edition. Anchor Press/Doubleday, Garden City, NY. 576 pp.

Beijbom, O., P.J. Edmunds, C. Roelfsema, J. Smith, D.I. Kline, B.P. Neal, M.J. Dunlap, V. Moriarty, T. Fan, C. Tan, S. Chan, T. Treibitz, A. Gamst, B.G. Mitchel, and D. Kriegman. 2015. Towards automated annotation of benthic survey images: variability of human experts and operational modes of automation. PLoS ONE 10(7): e0130312. doi: 10.1371/journal.pone.0130312

Bockelmann, F.-D., W. Puls, U. Kleeberg, D. Muller, and K.-C. Emeis. 2018. Mapping mud content and median grain-size of North Sea sediments – a geostatistical approach. Marine Geology 397:60-71. doi: 10.1016/j.margeo.2017.11.003

BOEM. 2016. Essential Fish Habitat Assessment for the Gulf of Mexico. US Department of the Interior, Bureau of Ocean Energy Management, Gulf of Mexico OCS Region. OCS Study BOEM 2016-016. 62 pp.

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

Breiman, L. 2001. Random forests. Machine Learning 45(1): 5-32. doi: 10.1023/A:101093340

Brown, C.J., B.J. Todd, V.E. Kostylev, and R.A. Pickrill. 2011. Image-based classification of multibeam sonar backscatter data for objective surficial sediment mapping of Georges Bank, Canada. Continental Shelf Research 31: S110-119. doi: 10.1016/j.csr.2010.02.009

Butman, B., T.J. Middleton, E.R. Thieler, and W.C. Schwab. 2003. Topography, shaded relief, and backscatter intensity of the Hudson Shelf Valley, offshore of New York. U.S. Geological Survey Open-File Report 03-372.

Butman, B., D.C. Twichell, P.A. Rona, B.E. Tucholke, T.J. Middleton, and J.M. Robb. 2006. Sea floor topography and backscatter intensity of the Hudson Canyon region offshore of New York and New Jersey: U.S. Geological Survey Open-File Report 2004-1441, version 2.0.

Campanella, F., P.J. Auster, J.C. Taylor, R.C. Muñoz. 2019. Dynamics of predator-prey habitat use and behavioral interactions over diel periods at sub-tropical reefs. PLoS ONE 14(2): e0211886. doi: 10.1371/journal.pone.0211886

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.

CMECS. 2017. Coastal and Marine Ecological Classification Standard, Catalog of Units. Online: https://www.cmecscatalog.org/ (Site Accessed 19 February 2019).

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. doi: 10.1080/01431161.2013.800661

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. 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., L. Kracker, T. Battista, W. Sautter, A. Mabrouk, K. Edwards, C. Taylor, and E. Ebert. 2018. 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. doi: 10.7289/V5/TM-NOS-NCCOS-241

Cressie, N.A.C. 1993. Statistics for spatial data (revised ed.). New York: John Wiley & Sons, Inc.

D'Errico, J. 2006. inpaint_nans (version 1.1.0.0), Interpolates (& extrapolates) NaN elements in a 2d array. MathWorks, MATLAB Central. Online: https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans (Site Accessed 19 February 2019)

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. doi: 10.1016/j.csr.2014.05.004

De Falco, G., R. Tonielli, G. Di Martino, S. Innangi, S. Simeone, and I.M. Parnum. 2010. Relationships between multibeam backscatter, sediment grain size and *Posidonia oceanica* seagrass distribution. Continental Shelf Research 30:1941-1950. doi: 10.1016/j.csr.2010.09.006

Diesing, M., S. Kroger, R. Parker, C. Jenkins, C. Mason, and K. Weston. 2017. Predicting the standing stock of organic carbon in surface sediments of the North-West European continental shelf. Biogeochemistry 135:183-200. doi:10.1007/s10533-017-0310-4

Du Preez, C. 2015. A new arc–chord ratio (ACR) rugosity index for quantifying three-dimensional landscape structural complexity. Landscape Ecology 30(1):181-192. doi: 10.1007/s10980-014-0118-8

Dunn, D.C., and P.N. Halpin. 2009. Rugosity-based regional modeling of hard-bottom habitat. Marine Ecology Progress Series 377:1-11. doi: 10.3354/meps07839.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudik, 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. M. Overton, A. Townsend Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberon, S. Williams, M.S. Wisz, and N.E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151. doi: 10.1111/j.2006.0906-7590.04596.x

Elith J., J.R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77:802-81. doi: 10.1111/j.1365-2656.2008.01390.x

Elith, J., S.J. Phillips, T. Hastie, M. Dudik, Y.E. Chee, and C.J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43-57. doi: 10.1111/j.1472-4642.2010.00725.x

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

Evans, I.S. 1979. An integrated system of terrain analysis and slope mapping. Final report on grant DA-ERO-591-73-G0040. Department of Geography, University of Durham, England. 192 pp.

Fenstermacher, L.E., G.B. Crawford, J.C. Borgeld, T. Britt, D.A. George, M.A. Klein, N.W. Driscoll, L.A. Mayer. 2001. Enhanced acoustic backscatter due to high abundance of sand dollars, *Dendraster excentricus*. Marine Georesources & Geotechnology 19(2):135-145. doi: 10.1080/10641190109353808

FGDC. 2012. Coastal and Marine Ecological Classification Standard, Marine and Coastal Spatial Data Subcommittee, Federal Geographic Data Committee. FGDC-STD-018-2012. Online: http://www.natureserve.org/sites/default/files/publications/files/cmecs_version_06-2012_final.pdf (Accessed 21 February 2019)

Folk, R.L. 1954. The Distinction between Grain Size and Mineral Composition in Sedimentary-Rock Nomenclature. The Journal of Geology 62(4):344-359. doi: 10.1086/626171

Fonseca, L., and L. Mayer. 2007. Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. Marine Geophysical Researches 238(2):119-126. doi: 10.1007/s11001-007-9019-4

Foote, K.G., H.P. Knudsen, G. Vestnes, D.N. MacLennan, and E.J. Simmonds.1987. Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Cooperative Research Report Number 144. 63 pp.

Foster G., A. Gleason, B. Costa, T. Battista, and C. Taylor. 2013. Acoustic Applications. pp. 221-251. In: J. Goodman, S. Purkis, and S. Phinn (eds.), Coral Reef Remote Sensing: A Guide for Mapping, Monitoring and Management. Springer Netherlands. 436 pp. doi: 10.1007/978-90-481-9292-2_9

Freeland, G.L., D.J. Stanley, D.J.P. Swift, and D.N. Lambert. 1981. The Hudson Shelf Valley: its role in shelf sediment transport. Marine Geology 42:399-427. doi: 10.1016/0025-3227(81)90172-9

Freiwald, A. 2002. Reef-forming cold-water corals. pp. 365-385. In: G. Wefer, D. Billett, D. Hebbeln, B.B. Jørgensen, M. Schlüter, and T.C.E. van Weering (eds.), Ocean Margin Systems. Springer, Berlin, Heidelberg. 495 pp. doi: 10.1007/978-3-662-05127-6_23

Friedman, J.H. 2002. Stochastic gradient boosting. Computational Statistics and Data Analysis 38(4):367-378. doi: 10.1016/S0167-9473(01)00065-2

Goff, J.A., B.J. Kraft, L.A. Mayer, S.G. Schlock, C.K. Sommerfield, H.C. Olson, S.P.S. Gulick, and S. Nordfjord. 2004. Seabed characterization on the New Jersey middle and outer shelf: correlatability and spatial variability of seafloor sediment properties. Marine Geology 209(4):147-172. doi: 10.1016/j.margeo.2004.05.030

Goff, J.A., C.J. Jenkins, and S.J. Williams. 2008. Seabed mapping and characterization of sediment variability using the usSEABED data base. Continental Shelf Research 30:614-633. doi: 10.1016/j.csr.2007.11.011

Greene, J.K., M.G. Anderson, J. Odell, and N. Steinberg (eds.). 2010. The Northwest Atlantic Marine Ecoregional Assessment: Species, Habitats and Ecosystems. Phase One. The Nature Conservancy, Eastern U.S. Division, Boston, MA.

Guida, V., A. Drohan, H. Welch, J. McHenry, D. Johnson, V. Kentner, J. Brink, D. Timmons, and E. Estela-Gomez. 2017. Habitat Mapping and Assessment of Northeast Wind Energy Areas. US Department of the Interior, Bureau of Ocean Energy Management. OCS Study BOEM 2017-088. Sterling, VA. 312 pp.

Harris, P.T., and E.K. Baker (eds.). 2011. Seafloor Geomorphology as Benthic Habitat: GeoHAB Atlas of Seafloor Geomorphic Features and Benthic Habitats. Elsevier. 936 pp. doi: 10.1016/C2010-0-67010-6

Hasan, R.C., D. Ierodiaconou, L. Laurenson, A. Schimel. 2014. Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. PLoS ONE 9(5): e97339. doi: 10.1371/journal.pone.0097339

Hijmans, R.J. 2016. raster: Geographic Data Analysis and Modeling. R package version 2.5-8. Software Downloaded 25 August 2017. Software Website: https://CRAN.R-project.org/package=raster (Accessed 19 February 2019).

Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2017. dismo: Species Distribution Modeling. R package version 1.1-4. Software Downloaded 25 August 2017. Software Website: https://CRAN.R-project.org/package=dismo (Accessed 19 February 2019)

Horn, B.K.P. 1981. Hill shading and the reflectance map. Proceedings of the IEEE 69:14-47. doi: 10.1109/PROC.1981.11918

Hosmer, D.W., and S. Lemeshow. 2000. Applied logistic regression, 2nd Edition. John Wiley & Sons. New York. 375 pp.

Hughes Clarke, J.E., K. Iwanowska, R. Parrott, G. Duffy, M. Lamplugh, and J. Griffin. 2008. Inter-calibrating multi-source, multi-platform backscatter data sets to assist in compiling regional sediment type maps: Bay of Fundy. Proceedings of the Canadian Hydrographic Conference and National Surveyors Conference 2008. 22 pp.

IHO. 2008. IHO Standards for Hydrographic Surveys (S-44), 5th Edition. Special Publication No. 44. International Hydrographic Bureau, Monaco. 28 pp. Online: https://www.iho.int/iho_pubs/standard/S-44_5E.pdf (Accessed 19 February 2019)

Jaffe, J.S., K.D. Moore, J. McLean, and M.P. Strand. 2001. Underwater optical imaging: status and prospects. Oceanography 14:66-76. doi: 10.5670/oceanog.2001.24

Jasiewicz, J., and T.F. Stepinski. 2013. Geomorphons – a pattern recognition approach to classification and mapping of landforms. Geomorphology. 182:147-156. doi: 10.1016/j.geomorph.2012.11.005

Jenkins, C.J. 2018. dbSEABED: Information Integration System for Marine Substrates. University of Colorado, Boulder, USA. Online: http://instaar.colorado.edu/~jenkinsc/dbseabed//dbseabed.htm (Accessed 23 December 2018)

Jenness, J. 2013. DEM Surface Tools. Jenness Enterprises. Software downloaded November 2015. Online: http://www.jennessent.com/arcgis/surface_area.htm (Accessed 7 March 2019)

Johnston, S.V., J.A. Rivera, A. Rosario, M.A. Timko, P.A. Nealson, and K.K. Kumagai. 2006. Hydroacoustic evaluation of spawning red hind (*Epinephelus guttatus*) aggregations along the coast of Puerto Rico in 2002 and 2003. pp. 10-17. In: J.C. Taylor (ed.), Emerging technologies for reef fisheries research and management. NOAA Professional Paper NMFS 5. 116 pp.

Kågesten, G., W. Sautter, K. Edwards, B. Costa, L. Kracker, and T. Battista. 2015. Shallow-Water Benthic Habitats of Northeast Puerto Rico and Culebra Island. NOAA Technical Memorandum NOS NCCOS 200. Silver Spring, MD. 111 pp. doi: 10.7289/V5Z899FH

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. doi: 10.7289/V5/TM-NOS-NCCOS-229

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.

Krivoruchko, K. 2005. Introduction to Modeling Spatial Processes Using Geostatistical Analyst. Esri. Redlands, CA. 27 pp. Online: https://www.esri.com/library/whitepapers/pdfs/intro-modeling.pdf (Accessed 21 February 2019)

Krivoruchko, K. 2012. Empirical Bayesian Kriging. Fall 2012 Edition. Esri. Redlands, CA. Online: https://www.esri.com/news/arcuser/1012/empirical-byesian-kriging.html (Accessed 19 February 2019)

Kuhn, M. 2016. Caret: Classification and Regression Training. R package version 6.0. Software Downloaded October 2015. Software Website: https://cran.r-project.org/web/packages/caret/index.html (Site Accessed 21 February 2019).

Lathrop, R.G., M. Cole, N. Senyk, and B. Butman. 2006. Seafloor habitat mapping of the New York Bight incorporating sidescan sonar data. Estuarine, Coastal and Shelf Science 68:221-230. doi: 10.1016/j.ecss.2006.01.019

Lauenstein, G.G., and A.Y. Cantillo (eds.). 1993. Sampling and analytical methods of the NS&T program national benthic surveillance and mussel watch projects. Volume II. Comprehensive Descriptions of Complementary Measurements. NOAA Technical Memorandum NOS ORCA 71. Silver Spring, MD. 102 pp.

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. doi:10.3354/meps321267

Lecours, V., M.F.J. Dolan, A. Micallef, and V.L. Lucieer. 2016. A review of marine geomorphometry, the quantitative study of the seafloor. Hydrology and Earth System Sciences 20(8): 3207-3244. doi: 10.5194/hess-20-3207-2016

Li, J., and A.D. Heap. 2014. Spatial interpolation methods applied in the environmental sciences: a review. Environmental Modelling & Software 53:173-189. doi: 10.1016/j.envsoft.2013.12.008

Love, R.H. 1977. Target strength of an individual fish at any aspect. Journal of the Acoustical Society of America 62:1397-1403. doi: 10.1121/1.381672

Ma, Z., and R.L. Redmond. 1995. Tau coefficients for accuracy assessment of classification of remote sensing data. Photogrammetric Engineering and Remote Sensing 61(4):435-439. doi: 10.1016/0924-2716(69)00007-4

Malek, A.J., J.S. Collie, and J. Gartland. 2014. Fine-scale spatial patterns in the demersal fish and invertebrate community in a northwest Atlantic ecosystem. Estuarine, Coastal and Shelf Science 147:1-10. doi: 10.1016/j.ecss.2014.05.028

Masetti, G., L.A. Mayer, and L.G Ward. 2018. A Bathymetry- and Reflectivity-Based Approach for Seafloor Segmentation. Geosciences 8(1): 14 pp. doi: 10.3390/geosciences8010014

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., B.P. Kinlan, D.S. Dorfman, M. Poti, and C. Caldow. 2012. A biogeographic assessment of seabirds, deep sea corals and ocean habitats of the New York Bight: science to support offshore spatial planning. NOAA Technical Memorandum NOS NCCOS 141. Silver Spring, MD. 224 pp.

Mitson, R.B. 1998. Fisheries Acoustics. pp. 137-147. In: J.T. Tanacredi and J. Loret (eds.), Ocean Pulse: A Critical Diagnosis. Springer, Boston, MA. 201 pp. doi: 10.1007/978-1-4899-0136-1

Moore, I.D., R.B. Grayson, and A.R. Ladson. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes 5:3-30. doi: 10.1002/hyp.3360050103

NOAA NCEI. 2018. Marine Geology: Stewardship and services for geosample data from Earth's ocean floor and lakebeds. NOAA National Centers for Environmental Information. Online: https://www.ngdc.noaa.gov/mgg/geology/ (Accessed 25 March 2019)

NOAA NCEI. 2019. Bathymetric Data Viewer. Geospatial Data and Services. NOAA National Centers for Environmental Information. Online: https://maps.ngdc.noaa.gov/viewers/bathymetry/ (Accessed 25 March 2019)

NOAA NGDC. 1999. U.S. Coastal Relief Model - Northeast Atlantic. NOAA National Centers for Environmental Information, formerly National Geophysical Data Center. doi: 10.7289/V5MS3QNZ

NOAA NOS. 2013. NOAA/NOS and USCGS Seabed Descriptions from Hydrographic Surveys. NOAA National Ocean Service, National Centers for Environmental Information. doi: 10.7289/V5BG2KWG.

NOAA NOS. 2018. NOS Hydrographic Surveys Specifications and Deliverables. 159 pp. Online: https://nauticalcharts.noaa.gov/publications/docs/standards-and-requirements/specs/hssd-2018.pdf (Accessed 21 February 2019)

NOAA NMFS NEFSC. 2018. Bottom Trawl Surveys. National Marine Fisheries Service Northeast Fisheries Science Center, and National Centers for Environmental Information, NOAA. Data access: https://data.noaa.gov/dataset/ (Accessed 21 February 2019)

Phillips, S.J., M. Dudik, and R.E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Proceedings of the 21st International Conference on Machine Learning. Banff, Canada. doi: 10.1145/1015330.1015412

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259. doi: 10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., R.P. Anderson, M. Dudik, R.E. Schapire, M.E. Blair. 2017. Opening the black box: an open-source release of MaxEnt. Ecography 40:887-893. doi: 10.1111/ecog.03049

Pittman S.J., B. Costa, and T. Battista. 2009. Using Lidar bathymetry and boosted regression trees to predict the diversity and abundance of fish and corals. Journal of Coastal Research, Special Issue 53:27-38. doi: 10.2112/SI53-004.1

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. doi: 10.1371/journal.pone.0020583

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., 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. doi: 10.7289/V5/TM-NOS-NCCOS-218

Plumb, R.H. 1981. Procedures for Handling and Chemical Analysis of Sediment and Water Samples. Technical Report EPA/CE-81-1. Prepared by Great Lakes Laboratory, State University College at Buffalo for the US Environmental Protection Agency/Corps of Engineers Technical Committee on Criteria for Dredged and Fill Material. U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS. 478 pp.

Poppe, L.J., J.S. Schlee, H.J. Knebel. 1994. Map showing distribution of surficial sediment on the Mid-Atlantic continental margin, Cape Cod to Albemarle Sound. U.S. Geological Survey Miscellaneous Investigations Series Map, IMAP 1987-D. doi: 10.3133/i1987D

Poppe, L.J., and A.E. Eliason. 2008. A Visual Basic program to plot sediment grain-size data on ternary diagrams. Computers and Geosciences 34:561-565. doi: 10.1016/j.cageo.2007.03.019

Poppe, L.J., K.Y. McMullen, S.J. Williams, and V.F. Paskevich. 2014. USGS East-Coast Sediment Analysis: Procedures, Database, and GIS Data. U.S. Geological Survey Open-File Report 2005-1001, version 3.0. Online: https://pubs.usgs.gov/of/2005/1001/ (Accessed 19 February 2019)

Poti, M., B.P. Kinlan, and C. Menza. 2012. Chapter 3: Surficial Sediments. pp. 33-55. In: C. Menza, B.P. Kinlan, D.S. Dorfman, M. Poti, and C. Caldow. (eds.), A Biogeographic Assessment of Seabirds, Deep Sea Corals and Ocean Habitats of the New York Bight: Science to Support Offshore Spatial Planning. NOAA Technical Memorandum NOS NCCOS 141. Silver Spring, MD. 224 p.

R Core Team. 2016. R: A language and environment for statistical computing (Version 3.3.0). R Foundation for Statistical Computing, Vienna, Austria. Online: https://www.r-project.org/ (Accessed 19 February 2019).

R Core Team. 2017. R: A language and environment for statistical computing (version 3.4.0). R Foundation for Statistical Computing, Vienna, Austria. Online: http://www.R-project.org/ (Accessed 19 February 2019)

Reid, J.M., J.A. Reid, C.J. Jenkins, M.E. Hastings, S.J. Williams, and L.J. Poppe. 2005. usSEABED: Atlantic coast offshore surficial sediment data release, version 1.0. U.S. Geological Survey Data Series 118. Online: http://pubs.usgs.gov/ds/2005/118/ (Accessed 19 February 2019)

Reineck, H.-E., and I.B. Singh. 1980. Depositional Sedimentary Environments: With Reference to Terrigenous Clastics. Second, Revised and updated Edition. Springer-Verlag, Berlin Heidelberg, Germany. 551 pp. doi: 10.1007/978-3-642-81498-3

Schlee, J.S. 1973. Atlantic Continental Shelf and Slope of the United States; Sediment Texture of the Northeastern part. Geological Survey Professional Paper 529-L. 63 pp. doi: 10.3133/pp529L

Shepard, F.P. 1954. Nomenclature based on sand-silt-clay ratios. Journal of Sedimentary Research 24(3):151-158. doi: 10.1306/D4269774-2B26-11D7-8648000102C1865D

Stegmann, S., T. Mörz, and A. Kopf. 2006. Initial Results of a new Free Fall-Cone Penetrometer (FF-CPT) for geotechnical *in situ* characterization of soft marine sediments. Norwegian Journal of Geology/Norsk Geologisk Forening 86(3):199-208.

Steimle, F.W., and C. Zetlin. 2000. Reef habitats in the Middle Atlantic Bight: abundance, distribution, associated biological communities, and fishery resource use. Marine Fisheries Review 62:24-42.

Stephens, D., and M. Diesing. 2015. Towards quantitative spatial models of seabed sediment composition. PLos ONE 10(11): e0142502. doi: 10.1371/journal.pone.0142502

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

Strout, J.M., and T.I. Tjelta. 2005. *In situ* pore pressures: What is their significance and how can they be reliably measured? Marine and Petroleum Geology 22(1-2):275-285. doi: 10.1016/j.marpetgeo.2004.10.024

Tobler, 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46:234-240. doi: 10.2307/143141

Wentworth, C.K. 1922. A Scale of Grade and Class Terms for Clastic Sediments. The Journal of Geology 30(5):377-392. doi: 10.1086/622910

Wigley, R.L., and R.B. Theroux. 1981. Atlantic continental shelf and slope of the United States macrobenthic invertebrate fauna of the Middle Atlantic Bight region-faunal composition and quantitative distribution. Geological Survey Professional Paper 529-N. 198 pp. doi: 10.3133/pp529N

Williams, S.J., M.A. Arsenault, L.J. Poppe, J.A. Reid, J.M. Reid, and C.J. Jenkins. 2006. Surficial sediment character of the New York-New Jersey offshore continental shelf region: a GIS compilation. U.S. Geological Survey Open-File Report 2006-1046. Online: http://pubs.usgs.gov/of/2006/1046 (Accessed 19 February 2019)

Zevenbergen, L.W., and C.R. Thorne. 1987. Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12:47-56.

Technical Glossary

Accuracy assessment (AA) – In the context of seafloor mapping, a technique that uses a randomly sampled set of ground truthing videos or sediment grabs to measure thematic accuracy of the habitat map. The score is evaluated by analyzing the User's Accuracy, Producer's Accuracy, and the Overall Accuracy. Proportional Area of each habitat class may also be factored into this assessment to reduce bias from habitats that are more common.

Area backscattering coefficient (ABC) – A measurement of acoustic signature strength over a specified area represented in m^2/m^{-2}.

$$s_a = \int_{z1|} s_v \, dz$$

Area under the receiver operating characteristic curve (AUC) – A receiver operating characteristic (ROC) curve is a graphical representation of how well a model can discriminate between (or predict) two categories of data (e.g., presence and absence). ROC curves can be used to identify the "optimal" thresholds in predictions (e.g., to classify a map of predicted probability of occurrence into presence and absence) for specific management applications. The AUC is the integral of a ROC curve. AUC values range between 0 and 1, with higher AUC values indicating better model performance.

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

Bayesian Kriging – A geostatistical interpolation method that uses data subsetting and simulation to improve the estimation of semivariogram model parameters by accounting for uncertainty in the estimated semivariogram. Other kriging methods calculate a single variogram model without accounting for the uncertainty in semivariogram estimation. Consequently, Bayesian Kriging provides more accurate values for prediction standard error than other kriging methods, which underestimate prediction standard error.

Boosted classification tree (BCT) 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 (BRT) 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 – An iterative technique for fitting predictive models. Models are built in a stagewise 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 resampling technique used to estimate the statistical precision associated with model predictions. Bootstrapping is a technique in which the data are randomly re-sampled with replacement into some number of new datasets. The model is fit to each of these new datasets and used to make a prediction. The precision of these model predictions can then be assessed from the variability across the predictions.

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 in the prediction relative to the mean prediction.

Combined uncertainty bathymetric estimator (CUBE) – An error-model based generator that estimates the depth plus a confidence interval directly to each node point of a bathymetric grid. When the algorithm fails to make a statistically conclusive decision, it will generate multiple hypotheses, attempt to quantify the relative merit of each hypothesis and present them to the operator for a subjective decision.

Cross-validation – A technique for evaluating the performance of a fitted model in which data are divided into subsets of training and test data, the model is fit to the training data, and model performance is assessed based on how well it predicts the values of the test data.

Focal statistics – Calculations in which the value for each grid cell in an output raster is assigned using some function of the values of the input grid cells in a specified neighborhood (e.g., a 3 x 3 grid cell rectangle) around that location.

Ground validation (GV) – In the context of seafloor mapping, a method used to relate image data to real features and materials on the seafloor by manually selecting sites from a draft map based on locations where further information is required. Sampling can be in the form of *in situ* samples or photographs, which ultimately provide additional information to optimize the mathematical models used to predict habitats.

Hydrometer – A device that measures the density of liquids, which could be used to determine the phi and grain size of very fine particles.

Inverse distance weighting (IDW) – An interpolation method that predicts values at unknown locations using a weighted average of the values at known locations, with greater weights given to locations closest to the prediction location. This follows the assumption that locations that are closer to each other are more alike than locations farther apart. IDW is a deterministic method (i.e., it creates predictions directly from the values at known locations), and therefore does not provide an estimate of prediction uncertainty.

k-fold cross-validation (kCV) – A cross-validation technique in which the data are divided into *k* data subsets (i.e., folds). Each fold is used once as the test data to evaluate a model fit using the

data in all the other folds as training data. An advantage of this approach is that all data are used for both training and testing the model.

Kriging – A geostatistical interpolation method that predicts values at unknown locations using a statistical model estimated from the values at known locations through semivariogram analysis. The semivariogram is a plot that is used to assess and quantify spatial autocorrelation in the known locations, with the assumption that locations that are closer to each other are more similar than locations farther apart. A statistical model is fit to the semivariogram and then used to predict values at unknown locations. Since kriging uses a statistical model to create predictions rather than the actual values at known locations, it is a stochastic method in that there is some inherent randomness in the estimation of the semivariogram model parameters. As a result, kriging provides not only predictions but also measures of prediction uncertainty.

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.

Maximum entropy (MaxEnt) model – A modeling algorithm that estimates the functional relationships between occurrence and a set of environmental predictor variables, with the relationships constrained by the mean value of the predictors at known locations of occurrence. These relationships are then used to predict the relative likelihood of occurrence at unknown locations.

Multibeam echosounders (MBES) – A type of sonar used to map the seabed by emitting sound waves in a fan shape beneath the ship's hull. The amount of time it takes for the sound waves to bounce off the seabed and return to the sonar's receiver is used to determine water depth.

Overall accuracy (OA) – In the context of seafloor mapping, the overall score of the accuracy assessment for a benthic habitat map that is calculated by the sum of the major classifications divided by the total number of AA samples.

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.

Phi (φ) Scale – A modification of the Wentworth scale that factors the diameter of a grain (in millimeters) to a logarithmic function, which assigns each grain size class a phi number -6 to 4 (in the NYWEA project).

Producer's accuracy (PA) – In the context of seafloor mapping, a metric to describe the thematic accuracy of the benthic habitat map by measuring how often habitats were incorrectly excluded from their correct habitat class.

Random forest model – A modeling approach that combines a machine learning method, bagging, and random selection of features with traditional tree-based statistical modeling. In this approach, data are randomly sampled with replacement into a large number of subsets. A

regression tree model is fit to each data subset, with a random subset of the features used at each candidate split during training of the decision tree. Models are averaged to generate a final, combined (i.e., "ensemble") model.

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 resampling 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).

Separation (SEP) model – Modeling technique used to translate Global Navigation Satellite System (GNSS) derived vertical positions to chart datum.

Smooth Best Estimate of Trajectory (SBET) – The Smoother is a module that computes the estimates of the inertial navigator by processing data backwards in time and then combining it with estimates from the forward in time filter. This is combined with a Feed-forward Error Control (a POSPac software module) to apply the integrated inertial navigation solution to generate an SBET. This is then applied to multibeam data to enhance the horizontal and vertical accuracy.

Splitbeam echosounder (SBES) – A sonar that is comprised of a transducer that is split into four quadrants allowing the location of targets in three dimensions.

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 survey data.

Sensitivity – A measure of model performance for binary classification models (e.g., presence/absence) that measures the proportion of positives that are correctly predicted as positives. 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 negatives that are correctly predicted as negatives. Like for sensitivity, this measure can be used to identify optimal thresholds for specific management applications. For example, if the goal is 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.

Tidal Constituents and Residual Interpolation (TCARI) – A method of computing water level correctors to reference hydrographic sounding to a tidal datum (e.g. Mean Lower Low Water) using observed water level data.

Total propagated uncertainty (TPU) – A measure for the accuracy to be expected for a point, when all relevant error and uncertainty sources are taken into account.

Training data – Data to which a model is fit in order to estimate model parameter values.

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.

TrueHeave – Applanix technology built into the Positioning and Orientation System for Marine Vessels (POS/MV) that accurately predicts the heave (up/down motion) of the vessel and thus filters out heave artifacts from the survey data.

User's accuracy (UA) – In the context of seafloor mapping, a metric to describe the thematic accuracy of the benthic habitat map using the total number of AA sites that the model correctly predicted for each habitat. The user is the analyst reviewing the AA data conducting the accuracy assessment.

underway Conductivity, Temperature, and Depth (uCTD) – A probe made up of electronic instruments that, when deployed into the water, measures the salinity, temperature, and depth of the water column. An underway CTD can be deployed while the vessel is in motion.

Vdatum – Software designed to vertically transform geospatial data among a variety of tidal, orthometric and ellipsoidal vertical datums.