EXAMINING THE FRONTIER OF AUTONOMOUS UNDERWATER VEHICLE IMAGE ANALYSIS FOR SEA SCALLOP INCIDENTAL MORTALITY

by

Hunter Tipton

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Oceanography

Summer 2021

EXAMINING THE FRONTIER OF AUTONOMOUS UNDERWATER VEHICLE IMAGE ANALYSIS FOR SEA SCALLOP INCIDENTAL MORTALITY

by

Hunter Tipton

Approved: _____

Arthur Trembanis, Ph.D
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Mark Moline, Ph.D.
Director of the College of Earth, Ocean, and Environment

Approved: _____

Estella Atekwana, Ph.D.
Dean of the College of Earth, Ocean, and Environment

Approved: _____

Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to first thank Danielle Ferraro for her contributions to this and previous incidental mortality studies that formed the foundation of my thesis. Additional thanks go out to my advisor, Dr. Arthur Trembanis and committee members David Rudders and Aaron Carlisle for their continuing support for my academic endeavors. Thank you as well to the members of the Robotics Discovery Laboratory for all the great memories and hard work over the span of my time at the university. I would not be where I am today without all your help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Historically, the Atlantic sea scallop (*Placopecten magellanicus*) fishery in the United States has undergone vast changes in management to combat stock volatility and fishing pressure. Today the American sea scallop stock has stabilized and is successfully managed via fishery independent surveys and stock assessment models. One of the factors affecting the stock assessment model for this fishery is incidental fishing mortality, or those individual scallops that experience mortality because of fishing effort but are not retained for use by the fishery. Recent studies sought to enumerate sea scallop incidental mortality on various seabed types. However, most of the scallops examined in these studies were inhabiting sandy bottoms more generally known as soft substrate. This study utilized 38,813 underwater images collected via Autonomous Underwater Vehicle (AUV) and a modified Multiple Before After Control Impact (MBACI) experimental design to examine incidental mortality on gravel and rock, or hard substrates within Closed Area I just north of the great south channel. Of the 31,972 scallops measured from these images 73% were annotated as lying on hard substrate and the highest measured value of incidental mortality was 6.70%. This result suggests that stock assessment values for incidental mortality remain conservative for hard substrate management areas. The results of this study also suggest that both sediment and sea scallop distributions may be more variable than previously understood within the spatial domain of a study site.

Chapter 1

PROJECT INTRODUCTION AND BACKGROUND

Project Introduction

The Atlantic Sea Scallop (*Placopecten magellanicus*) fishery in the Northwestern Atlantic Ocean has witnessed its own collapse and rebound because of fisheries management practices. Beginning in 1982 a fishery management plan (FMP) was enacted to regulate the fishery in light of record low landings reported during the 1970's. Another sharp dip in landings came in the early 1990's, but FMP framework adjustments for stricter management of the resource turned the tide for scallop biomass and landings in 1994 (Hart and Rago, 2006). Adjustments in 1994 included the introduction of rotational management areas, gear regulations, and effort limitations. Recently the fishery has been more stable as landings have stabilized in response to regulation changes. The most recent stock assessment (NEFSC, 2018) reported the 15th consecutive year that the Atlantic Sea Scallop had not been overfished or experienced overfishing. Sustainable exploitation combined with increased biomass levels for the management time series and promising recruitment levels offer hope for the continuation of the Sea Scallop fishery in the Northwest Atlantic.

Fisheries mortality values inform stock assessment models in combination with data from monitoring efforts to establish catch and effort limits for the scallop fishery. Fishery mortality values are an important factor in determining the total mortality for the

Sea Scallop population and the impact of fishing effort on this mortality. Incidental, or non-harvest, mortality is any fishery induced mortality that occurs on individuals not retained by the fishery (Caddy, 1973).

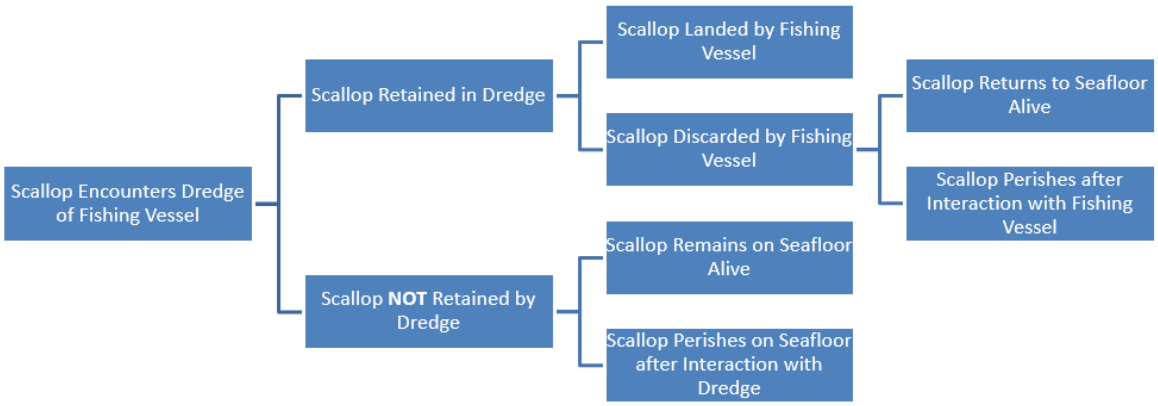


Figure 1.1: The pathways to mortality for sea scallops targeted by the fishery. Figure adapted from Ferraro, 2016.

When scallops interact with the dredge incidental mortality is possible in a variety of ways (Figure 1.1). On the sea floor a scallop may be crushed between the metal frame of a dredge and the substrate, or a scallop could pass through the rings of the dredge as it rides across the bottom and sustain damage as a result. This same culling action is possible on the way to the surface while the dredge is winched from the seafloor to the surface during the process of fishing. The water column also presents issues if thermal

stratification is present, and temperatures change significantly between the surface and the seafloor (Stokesbury, 2011).

At the surface mortality of caught scallops may also occur. Scallops may be crushed on deck while the bag of the dredge is dumped, or they may perish from the environmental conditions outside of the water in times of extreme weather. The sea scallop fishery in the U.S. operates by eviscerating the scallop and discarding the valves and visceral mass back into the sea while retaining the adductor muscle for consumption. Discard mortality is another factor of landed fishing mortality that accounts for scallops retained for some time on the deck of the vessel that are tossed back into the sea. Discard mortality has increased notably within the southernmost management areas of the fishery where the presence of a nematode parasite has caused a greater proportion of shucked scallops to be discarded (Rudders and Roman, 2018). As a result of parasite prevalence, the amount of fishing within the southern reaches of the sea scallop grounds has decreased substantially since 2015 (NEFSC, 2018). Additionally, prolonged fishing activity can heighten predator populations and increase susceptibility to disease due to the fishery practice of discarding the visceral mass of harvested scallops (Jenkins and Brand, 2001; Caddy, 1973). The complexity of the interactions between scallop, dredge, and substrate dictate that incidental mortality values vary across geographic locations, as sediments and environmental conditions change drastically over the span of the Northwest Atlantic fishing grounds.

Most recently incidental mortality values have been enumerated within the Nantucket Lightship Closed Area (NLCA) and Elephant Trunk Closed Area (ETCA) by

utilizing Autonomous Underwater Vehicle (AUV) digital imagery collected following dredging at sea (Ferraro et. al., 2017; Walker et. al., 2016). These studies served as a proof of concept for AUV optical surveying of the dredge scar and surrounding benthos for incidental mortality analysis and offered significant improvement in survey methodology when compared to historical manned submersible survey techniques (Caddy 1973, Murawksi and Serchuk, 1989). Incremental improvements such as these are made possible by the advancement of marine surveying technology. Despite technological improvements in data collection, the post processing of collected imagery remains a bottleneck for optical survey efforts. Manual annotation by human annotation teams remains the primary means of image analysis for a range of optical survey platforms currently used in stock assessment efforts for the scallop fishery today. Recent efforts by Rasmussen et al (2017) sought to address the bottleneck in image processing by implementing an automated image processing workflow. To do so human annotated scallops were used to train and test a YOLOv3 based convolutional neural network platform, named "ScallopScan", for the detection of healthy scallops (Rasmussen et. al. 2017). ScallopScan has detected healthy scallops with an average precision (AP) value of 0.92 for the images of past studies, a result that approaches the 0.95 AP of human annotators. This technological advancement allows for image processing at speeds as high as 12 frames per second once trained and offers another means of comparison with the annotations of humans at minimal additional time investment. Although still in the experimental stages, the insights provided and potential adaptations for a system like

ScallopScan show promise for future studies. Data derived from the most recent ScallopScan experiments are detailed in this manuscript.

In this study I used an inherited the dataset from the 2017 Incidental Mortality cruise in addition to the scallop image annotation interface (SIAS), ScallopScan, and MySQL database of past incidental mortality studies (Ferraro et al, 2017) to examine incidental mortality on an area of harder substrate than previously studied, while also offering incremental improvements to data analysis and visualization techniques made possible through scallop sizing and deep learning image analysis techniques.

The primary research objectives of this thesis are:

- To quantify the incidental mortality of sea scallops remaining on the sea floor post-dredging on areas of hard substrate using AUV optical survey imagery.
- To examine AUV performance for this study as it compares to past studies.
- To investigate variability in substrate composition within and between the study sites of this experiment.
- To quantify the performance of neural network tools in benthic image analysis.
- To characterize the best path forward for future utilizations of this and similar experimental designs.

To achieve these goals, manual annotation of the image set collected at sea in 2017 was completed. Annotation of the 2017 incidental mortality images resulted in 31,972 new scallop annotations from 38,813 images over than span of 11 months' time with a minimal annotation team.

Research Area

This study analyzed images collected from August 15-28, 2017 at three sites within and around Closed Area I (CAI), a rotational management area of the scallop fishery within the Northwestern Atlantic Ocean (Figure 1.2)

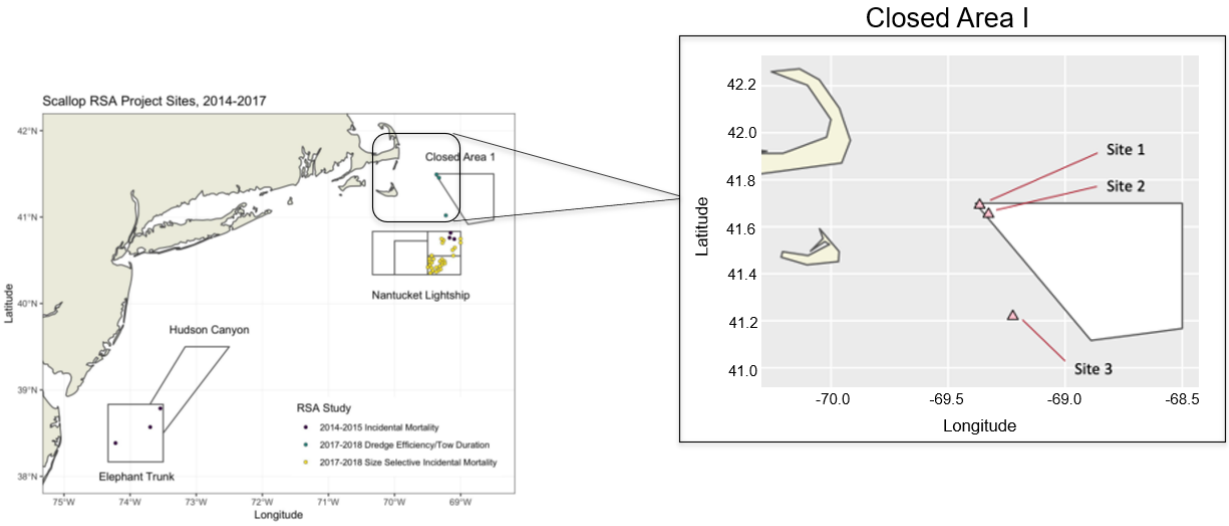


Figure 1.2: A map of Closed Area I (white) with Cape Cod (beige) to its west. Survey stations from this study are signified by red triangles. Longitude and latitude values are displayed on the X and Y axes, respectively. Figure adapted from Trembanis, 2019.

This management area is characterized by rock and gravel substrate and located approximately 28 miles east of the island of Nantucket Massachusetts, and just north of the great south channel. One of the three study sites was moved to outside of the southwest side of the closed area because of lobster pots obstructing the fishing grounds during the time of data collection. The goal of this CAI study site is to establish a value

for incidental mortality on a substrate harder than that of the Nantucket Lightship Closed Area (NLCA) and the Elephant Trunk Closed Area (ETCA) sites of former incidental mortality studies. Although the NLCA is regarded as a hard-bottom site by these studies, approximately 84% of the annotated images from the NLCA were classified as having predominantly "sandy" substrate during annotation. This nuance of the past IM dataset led to the development of another experiment for assessing the literature value of incidental mortality in areas of hard rocky substrate within Closed Area I. Moving to a new study site also increased the spatial scale of incidental mortality experimentation and introduced a new rotational management area for assessment. Finding areas of suitable substrate composition and scallop density proved challenging and is reflected in the scallop sample sizes of this study as they compare to those of the past.

Field Work

The fishing vessel Christian and Alexa was commissioned to complete this incidental mortality study in conjunction with field experiments for a combined Virginia Institute of Marine Science and University of Delaware dredge efficiency and tow duration study (NA17NMF4540044) from August 15-28, 2017. Field experiments for the 2017 incidental mortality survey of Closed Area I followed a similar before-after-control-impact experimental design to earlier incidental mortality studies (Ferraro et al. 2017), but with the purpose of examining incidental mortality on extremely hard substrate. Using the knowledge of the vessel captains, three study sites were chosen after scouting out areas of suspected rocky substrate . The Christian and Alexa is the same vessel that was used in past University of Delaware AUV studies (Walker et al. 2016, Ferraro et al.

2017), eliminating vessel-based biases as a source of variability between this study and those of the past.

Autonomous Underwater Vehicle

The vehicle used in this study was a Teledyne Gavia AUV (Figure 1.3).

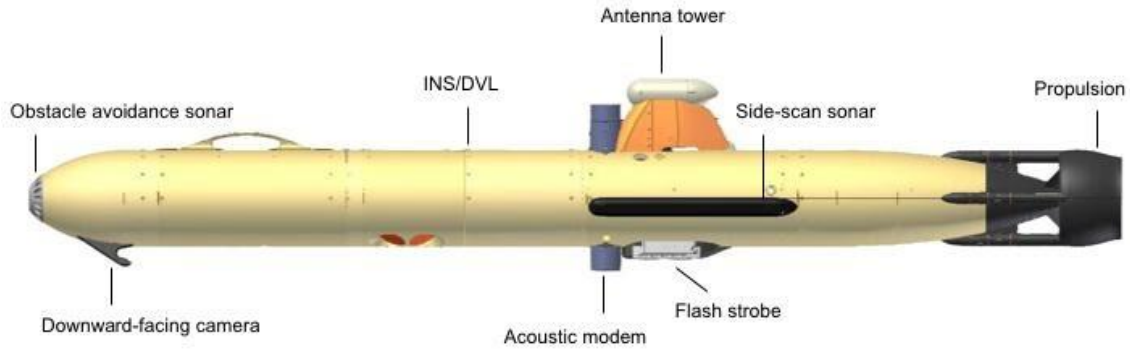


Figure 1.3: The Teledyne Gavia AUV used in this study pictured in survey configuration. Figure adapted from Ferraro, 2016.

Mission plans were standardized across all dredge treatments of the study and consisted of a lawn mower survey pattern comprised of 14 parallel lines 500m long spaced 2m apart. These lines were surveyed at 2.5m vehicle altitude. During missions, the vehicle was equipped with an integrated Point Grey Grasshopper 14S5C/M-C camera module containing a Sony ICZ285AL CCD digital camera shooting at a resolution of 1280x1960 pixels, or 1.2 megapixels. The camera was upgraded in 2016 to install an upgraded flash

strobe and sapphire lens. The horizontal viewing angle of the camera was 41.19 degrees, equating to a 1.88m x 1.41m image footprint at the 2.5m mission altitude. Photos were collected at 1.9 frames per second, resulting in image overlap greater than 40% at a resolution of 2.4mm per pixel. Images were stored in the JPEG format and tagged with metadata including latitude, longitude, altitude, depth, pitch, roll, and heading. In addition to the camera system on board, the AUV was also outfitted with a Marine Sonic side scan sonar capable of 900khz and 1800khz frequencies. The AUV scanned the sea floor with the 1800khz channel at 10m range per side during photo collection for this study to assist in the characterization of dredge scars at survey sites. The vehicle also collected data on temperature, salinity, and dissolved oxygen during mission execution.

Multiple Before After Control Impact Experimental Design

This study implemented the same multiple before after control impact methodology to the previous AUV incidental mortality study (Ferraro et al. 2017), but with the goal of examining substrates harder than those of the NLCA and ECTCA. Three replicate sites were chosen, of which two lay within CAI. Due to a large concentration of fishing gear and lack of suitable scallop density and substrate composition, the third site was moved outside of the closed area to satisfy the targeted study site parameters. The three sites were each designated with treatments including a control (C) Site, a one tow site (A), and a five-tow site (B). The treatments were designed to reflect a closed bottom scenario (C), light fishing pressure (A), and heavy fishing pressure (B). Each site was initiated by a “Before” AUV mission (A1, B1, C1) that was followed with a single or five dredge tows. A replicate “After” AUV mission (A2, B2, C2) was completed as soon as

possible following dredge treatments. AUV missions consisted of 14 parallel lines arranged north to south in a lawnmower pattern (Figure 1.4).

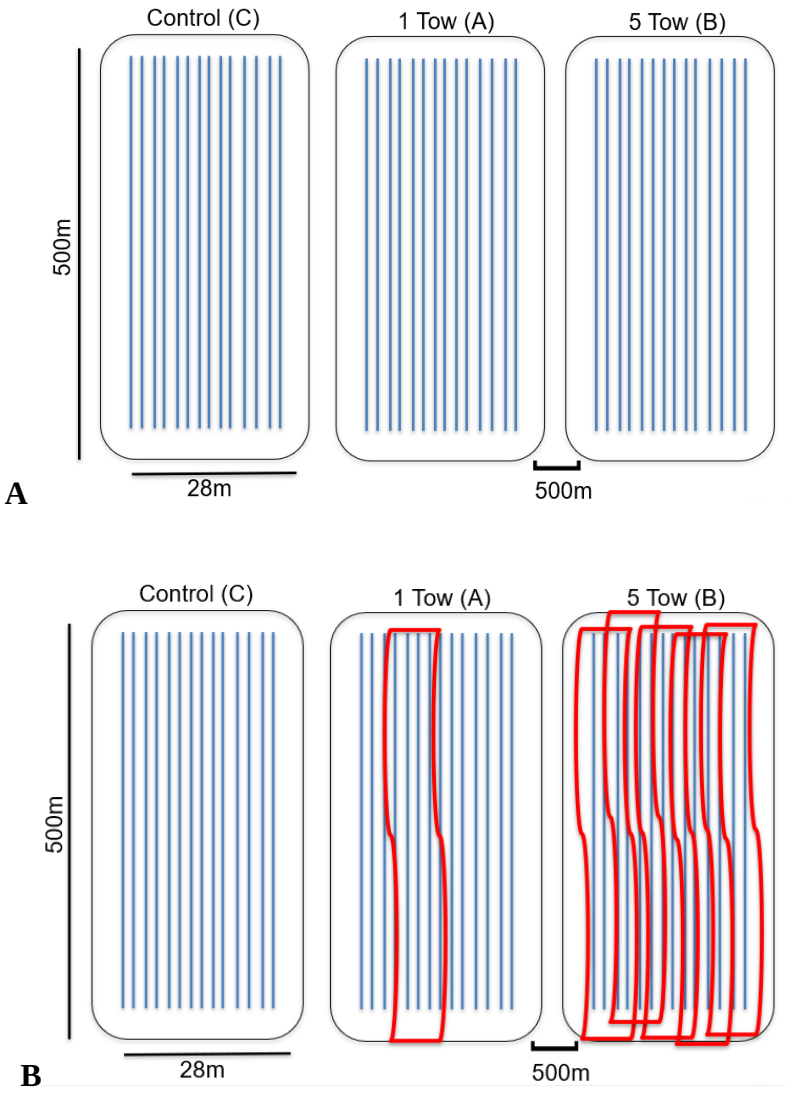


Figure 1.4: A) The AUV mission design before dredge treatment. B) The AUV mission design after treatment with dredge scars illustrated in red.

Survey lines were 500m in length and separated by 2m. The vehicle maintained a steady altitude of 2.5m and a constant propulsion rate of 600rpm. Photos were shot at 3.75fps and down sampled 50% resulting in an actual frame rate of 1.875fps, however survey coverage still exceeded 100% across sites. For this study 18 AUV missions were conducted resulting in the acquisition of 263,529 photos, 252,000m2 of digital image coverage, and 432,000m2 of side scan sonar coverage. The one tow sites were later omitted from the analysis of this study due to the lack of effect observed in previous studies (Ferraro, 2016).

Commercial Dredge

The fishing vessel Christian and Alexa provided a commercial style scallop dredge for the study that was outfitted according to National Marine Fisheries Service (NMFS) criteria (Figure 1.5). The dredge was a 4.6m (15ft) wide turtle deflector dredge equipped with 4" rings and rock chains ranging in size from ½" to 5/8" chain. The dredge was towed from a block on the starboard bow at a vessel speed between 4 and 5kts (2-3.5m/s). All catch from the dredge treatments was enumerated into bushel basket counts on deck and then moved into the starboard cutting box of the boat for processing. Catch processing consisted of length measurement and counts for an array of biological demographics including sea scallops, bycatch finfish species, skates, and other macroinvertebrates. All samples were processed and retained on board until the vessel left the study area to ensure that discarded catch did not reenter the study sites. The catch from each tow was then discarded off site.

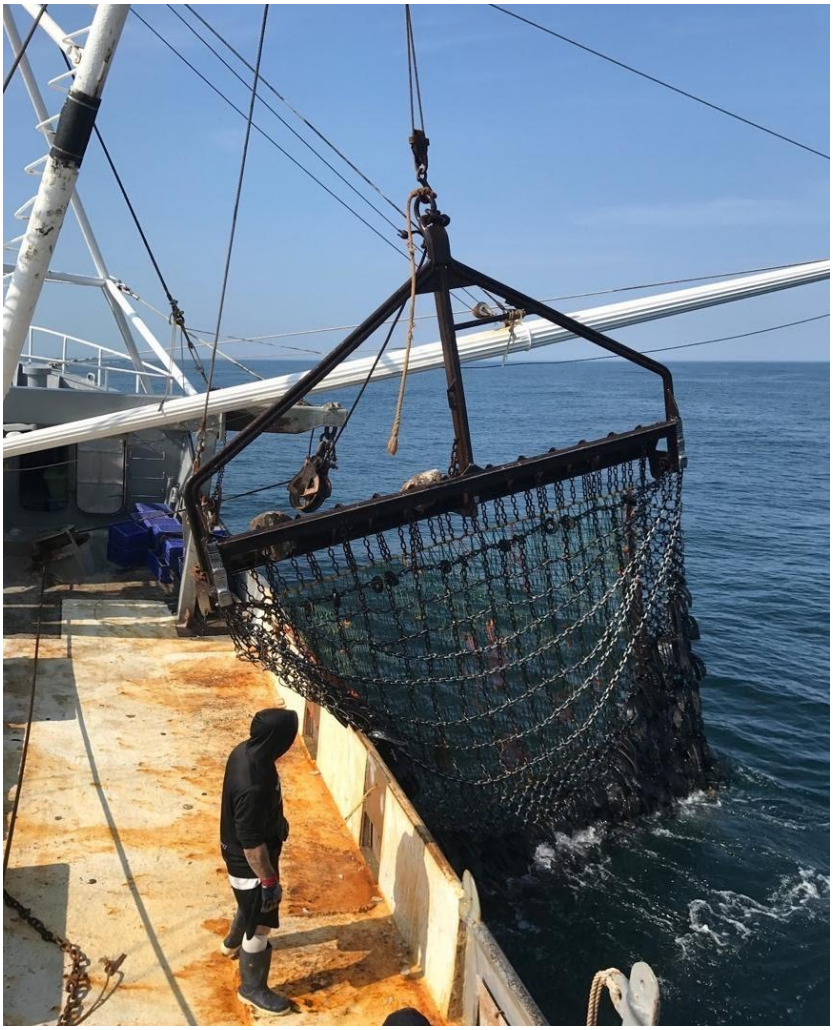


Figure 1.5**:** The starboard turtle deflector dredge of F/V Christian and Alexa used for the dredge treatments of this study. Photo courtesy of Danielle Ferraro.

Data were collected via an electronic length board connected to the VIMS FEED data collection interface. Data collected during catch processing were backed up to a solid-state drive and transferred to the University of Delaware after returning from sea.

Camera Parameters and Image Quality

The camera used for this study was essentially identical to the camera of the previous incidental mortality studies (Ferraro et al. 2017), but the camera did receive a sapphire lens and flash strobe upgrade for this study. These upgrades did not change any of the camera's image parameters. As a result of using the same camera as previous studies horizontal field of view (41.19^{o}) and resolution are the same as in past studies (Ferraro 2016). Due to the variable weather conditions, the water clarity and image quality of each AUV mission is different. AUV missions were conducted from sunrise until after sunset, and as such the amount of light available at depth and its interaction with the light provided by the strobe caused a lack of uniformity in the illumination of images. As such raw images were retained and post processed for image quality normalization using the Retinex image filter after the completion of each mission while at sea as performed in past studies (Ferraro et. al., 2017). After enhancement the images of this study possessed a halo like ring of clarity surrounded by pixelated margins, much like a fisheye lens (Figure 1.6). This is attributed to the change in the strobe light of the vehicle and resulted in decreased image quality when compared to that of past studies, but this effect was unavoidable, and images were annotated regardless of the distorted margins.

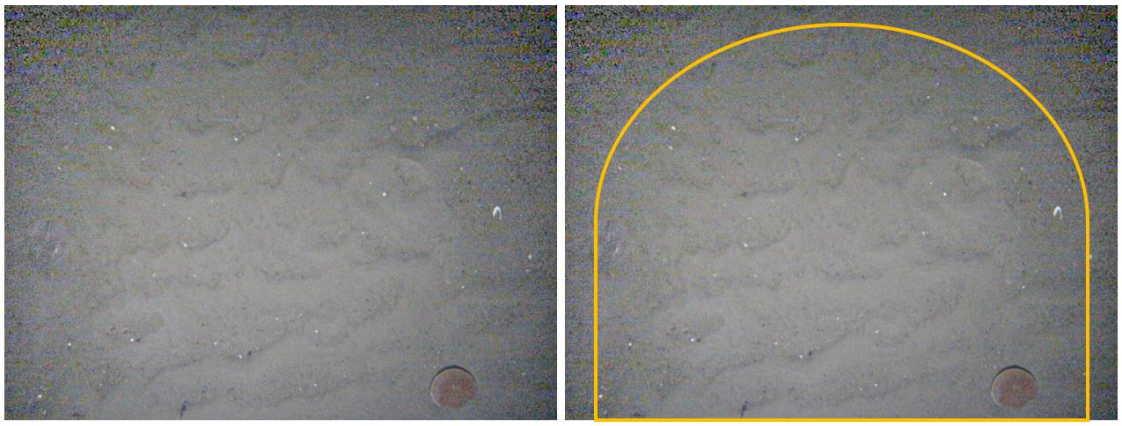


Figure 1.6: An enhanced AUV image with the margins highlighted to show the pixilation and distortion.

Past Manual Annotation Efforts

Image processing in the form of image annotation is required for analysis scallop incidental mortality. Manual annotation is hosted within the University of Delaware developed Scallop Image Annotation System (SIAS) interface (Figure 1.7). SIAS is an interactive guided user interface (GUI) hosted on the University of Delaware server "Robots", allowing annotators access to the interface anywhere that internet is available (www.robots.udel.edu/Scallop).

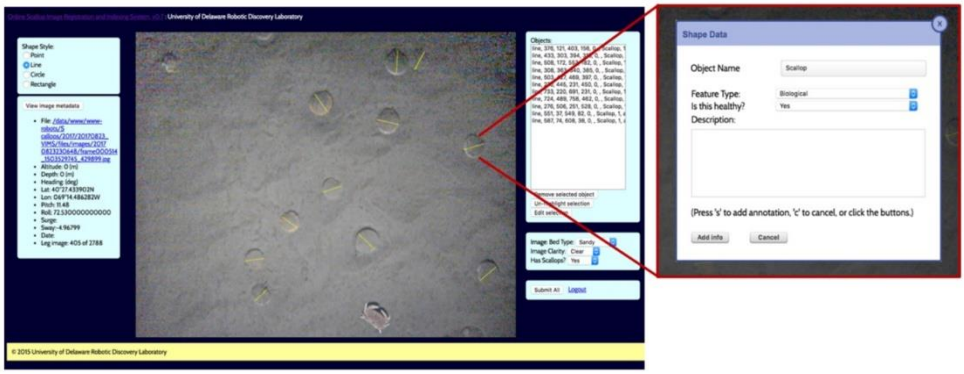


Figure 1.7**:** The University of Delaware Scallop Image Annotation System interface. Yellow lines indicate shell height measurements for individual scallops. Upon completion of a measurement the dialogue box outlined in red appears and allows for species level identification and health status determination for the measured individual. A separate dialogue box in the lower right-hand corner allows for the annotation of substrate classification for each image.

Manual annotation efforts thus far have included the annotation of AUV mission sites within the Mid Atlantic Bight, Nantucket Lightship, and Elephant Trunk Closed Areas. In 2012 an annotation team annotated 16,451 scallops from 249,564 MAB images in 120 hours (Walker et al. 2013). This effort annotated 100% of the database for the 2012 study and served as a proof of concept for manual annotation for image processing of AUV based optical survey imagery. This rate was accepted and used as a baseline for the time estimates of a later incidental mortality study (Ferraro et al. 2017). The team of annotators for the 2017 study consisted of 15 individuals who annotated 133,246 images from the ETCA and 38,614 images from the NLCA with point detections and health status. The images from these sites were down sampled to every fourth image to increase processing efficiency. However, the annotation team invested 1,150 person hours of work and only accounted for about half of the image set for the study. The 2014 IM images

were later reannotated to include shell height measurements. Images from all experimental treatments were reannotated for the NLCA image set, but only the images from the control and five tow treatments were reannotated for the ETCA. Sites one and two within the ETCA were annotated completely, while site 3 of the ETCA and all 3 sites in the NLCA were down sampled to every fourth image for reannotation. The most recent annotation effort was another line annotation effort for the combined VIMS/UD tow duration and dredge efficiency project within the NLCA (NA17NMF4540044). This effort utilized the same user interface as the 2017 IM study and aimed to size and count scallops from the dredge efficiency portion of the study. The team of 24 annotators completed annotation of 31,089 images within the timeline of the project, covering 20 of the 30 total dredge tow/AUV mission pairs.

Another annotation effort was completed for predator species from 25,036 images located within the Nantucket Lightship Closed Area of the past IM studies (Philips, 2019, Ferarro et al., 2017). Targeted species included rock crabs (*Cancer irroratus*), spiny dogfish (*Squalus acanthias*), summer flounder (*Paralichthys dentatus*), moon snails (*Euspira spp*.), ocean pout (*Zoarces americanus*), and sea stars (*Astropecten americanus* and *Asterias vulgaris*). In total 12,168 annotations were completed for the above species, but statistical analysis of the annotations did not yield viable results because of having only a small sample of detections relative to the spatial scale of the images annotated.

ScallopScan Development

The ScallopScan Convolutional Neural Network (CNN) was developed in collaboration with Dr. Christopher Rasmussen and graduate student Jiayi Zhao to automatically detect scallops from the AUV imagery of past studies. The detector was first based upon the YOLOv2 object classification architecture, and it was trained to detect a single object class deemed “Healthy Scallop”. The YOLO image classification system operates by scanning each image once and outlining possible object detections with a bounding box and pseudoprobability value of confidence in its detection of the desired target class on a 0 to 1 scale (Figure 1.8).

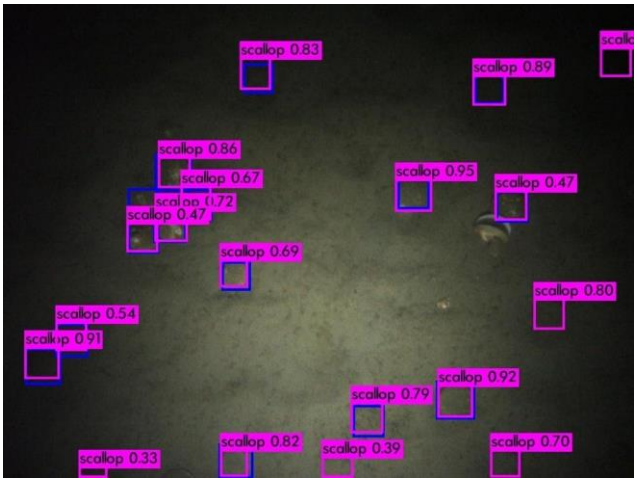


Figure 1.8: An example ScallopScan output image with bounding boxes, object class, and detection probabilities displayed over individual scallop detections. ScallopScan detections are signified by pink bounding boxes and instances of human annotation are outlined in blue. Photo source: 2017 Mission 94

These threshold values can be used as a filter to limit the amount of false positive detections by establishing thresholds for the omittance of low probability scallop detections. ScallopScan has recently been upgraded to the YOLOv3 architecture, an

upgrade that provided enhanced detection speed over previous versions without sacrificing precision and accuracy observed in version two (Redmon and Farhadi, 2018).

Training the ScallopScan network relies upon manual annotation inputs as proxies for true positive detections. During early studies manual annotations were made as points near the center of the scallop shell surface during image processing. This made implementing manual annotations a challenge as the data for scallop detections simply gave a point location for each scallop rather than a dimension or area. A separate image processing network deemed the "Heatmap" network was created to allow these point annotations to be used as a training set for the ScallopScan network. The Heatmap network is an image auto encoder that operates via assigning pixels a color value. These color values are normalized based upon the probability of each pixel comprising a scallop shell. The resulting output from the heatmap network portrays pixel clusters likely to be scallops in red and other objects such as substrate in cooler colors like green and blue. Based upon the heatmap outputs, a scallop image mask was created to allow bounding box detections to take place. These foundational developments in scallop detection laid the framework for the YOLOv3 based ScallopScan network.

ScallopScan was used for the first time during the 2017 collaborative dredge efficiency project after being upgraded to the YOLOv3 framework. In order to assess the performance of the latest ScallopScan iteration 19,469 images of the dredge efficiency dataset were reserved to serve as a test image set. One way of assessing the performance of deep learning neural networks is to calculate the precision and recall values for the test set of images. Precision is the ratio of true positives possible in the dataset and the sum of

the true positive and false positive annotations of the network. Recall is the ratio of the number of true positives in the dataset and the sum of the true positive and false negative annotations of the network (Figure 1.9).

$$\text{Precision} = \frac{\text{True Positive Detections}}{\text{True Positive Detections + False Positive Detections}}$$

$$\text{Recall} = \frac{\text{True Positive Detections}}{\text{True Positive Detections + False Negative Detections}}$$

Figure 1.9: Precision and Recall Formulas.

Precision and recall trade off, meaning you cannot have an increase in one without a decrease in another. Results from testing the network are plotted as a Precision-Recall (PR) Curve, with precision on the y-axis and recall on the x-axis. Average precision (AP) is another statistic for assessing network results. Average precision is defined as the area under the PR curve (Tan, 2019). The way that precision-recall curves are derived is by using cutoffs for the pseudoprobability values of predicted target classes to reach the highest average precision value possible for the image set. For instance, ScallopScan tested with an AP of 0.925 at the 0.30 confidence threshold cutoff. This means that when ScallopScan was bound to a detection confidence of 0.30 only bounding

an object as a scallop if it was more than 30% sure an object was a scallop, its precision on average across all images was 0.92 on a 0 to 1 scale. (Figure 1.10).

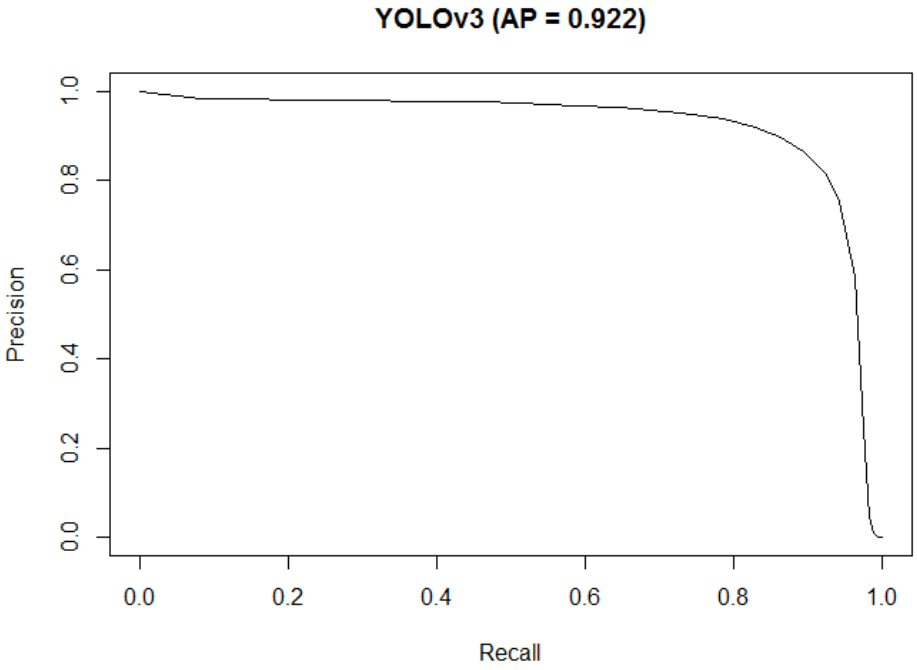


Figure 1.10: AP curve of ScallopScan showing 0.92 value for test images.

Our human annotation team must score at least 95% (AP = 0.95) for precision and recall on the test PowerPoint to be able to freely annotate raw images, so an AP of 0.92 (92%) for ScallopScan means that the algorithm is close to being as accurate at detecting scallops as annotators. Following these promising test results, the CNN annotated 294,768 images containing 6,333,478 scallops from the remaining images of the dredge efficiency image set with a mean detection confidence of 0.63. Additionally, this processing was completed at a rate of 12.4 frames per second with a precision value of 0.897 and recall of 0.863.

Side Scan Sonar Processing

Side scan sonar data were collected by the Gavia AUV during each mission of this study. Sonar was collected at 1800khz with a 10m range per side to ensure proper overlap for creating a sonar mosaic of each treatment site. All sonar data was uploaded into SonarWiz 6 (Chesapeake Technologies) to create side scan mosaics in accordance with methods developed in past studies (Ferraro, 2016). The primary application of sonar data is to detect the area of the dredge scar. Dredge scar areas were targeted by examining the side scan mosaic in SonarWiz and identifying an area of evenly swept seafloor approximately 15ft (4.6 m) wide, or roughly the same width as the mouth of the commercial dredge (Figure 1.11). Dredge scars were then digitized via the polyline feature tool in SonarWiz and exported as text files (.csv), shape files(.shp), and Google Earth layers (.kmz).

Post processing to digitize the dredge scar allows for image sorting based on each image centroid position relative to the dredge scar area spatially. This spatial sorting can then be modified to examine the dynamics of the dredge scar area more specifically, including examination of the dredge scar area itself, the margins of the dredge scar area, and the imaged area outside of the area of dredge-seafloor interaction (Ferraro 2016). In addition to the dredge path itself, side scan sonar also allowed for the broader characterization of sediment type for each site based upon the backscatter intensity of the sonar return. Ground truthing of sediment classification is also made possible by backreferencing the georeferenced images taken by the vehicle.

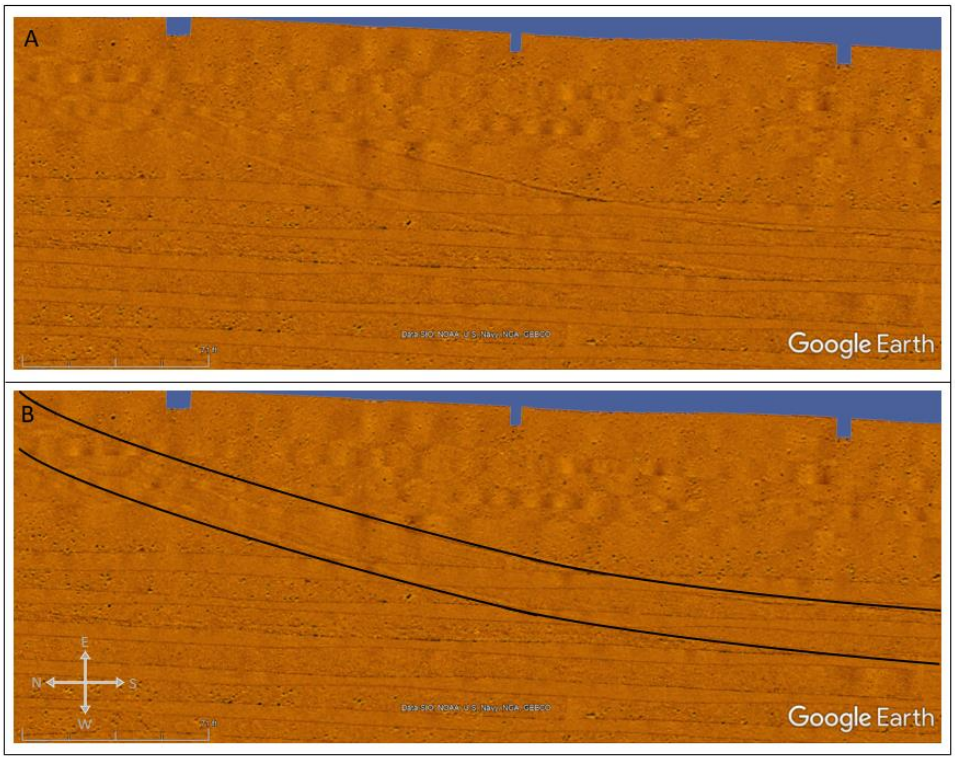


Figure **1.**11: A) Unhighlighted, raw side scan sonar mosaic containing a dredge scar from survey Site 1, B2. B) Highlighted dredge scar inside the same side scan sonar mosaic as illustrated in “A)”

Chapter 2

METHODS

Image Processing and Annotation

Post cruise, images were backed up to the Dogfish server at the University of Delaware. The raw images were then enhanced using the same retinex code of Ferarro 2016 (http://www.fmwconcepts.com/imagemagick/retinex/). Enhanced images were then uploaded onto the web-based server Robots. Although all the 263,539 images from the AUV missions were uploaded to SIAS, only certain missions were examined for this study. Mission selection was optimized based upon the results of past studies that found minimal effect from the one tow experimental treatment (Ferraro 2016, Ferraro et al. 2017). As such only the images of the control and five tow treatments of the three experimental sites of this study were examined. This left 13 missions and 165 mission legs for manual annotation. The images of these legs were down sampled to 25% for annotation and annotated in the SIAS interface for incidence of healthy and compromised scallops in accordance with the methods of past studies (Ferraro et. al. 2017).

Two annotators, Molly Struble and I, completed the annotation of the images of this data set. Both of us received training and experience through annotating images included in past IM analysis efforts (Ferraro 2016). Annotation efforts began in March 2020 and concluded in January 2021. Scallops were annotated by creating a line from the

center of the hinge to the shell margin. Scallops on the margins of images were only annotated if more than half of the scallop was visible in the image. If the shell height line was not visible due to shell damage or the position of the scallop in an image, then a line was made across the longest visible section of shell. After measurement, the health status of each individual was recorded. Data was also collected for the substrate composition, image clarity, and scallop prevalence of each image annotated regardless of scallop prevalence. After the completion of image processing, the raw data from each image and scallop annotated were retrieved from the online MySQL relational database and compiled into spreadsheet format.

Additional data were derived from the raw database by manipulating the pixel values of the shell height line and the centroid of each scallop as performed in past studies (Walker et. al. 2016). Shell height was calculated using the same protocol as Trembanis et al. 2019 (Figure 2.1).

$$SH = 1000\, W_{pixel} \sqrt{(x^2 - x^1) + (y^2 - y^1)}$$

Figure 2.1: Formula for deriving shell height from pixel values of annotated shell height lines.

Sea scallop locations were then derived by converting the latitude and longitude centroid of each image into UTM. Knowing the location of the image centroid in meters and the size of each pixel in meters allows for the derivation of the location of any pixel in the image by multiplying the pixel value of the scallop centroids by the size of a pixel in millimeters. The location of each scallop then becomes the result of adding the easting (x) and northing (y) offset of the center pixel coordinate of each scallop shell height line from the image centroid (Figure 2.2).

Scallop Centroid Easting = (Pixel Size in m * Pixel x coordinate value) + Image Centroid Easting

Scallop Centroid Northing = (Pixel Size in m * Pixel y coordinate value) + Image Centroid Northing

Figure 2.2: Formula for the location of a scallop in UTM.

Both derived and raw image data received data quality control in the form of outlier removal. Image quality control included removing images with unusual altitude values and clarity issues. Scallop quality control removed any individual exhibiting anomalous shell height parameters or missing crucial data such as health status.

Incidental Mortality Calculation

Once scallops were annotated, sized, and located, incidental mortality was determined for each experimental treatment. The same incidental mortality definition as

used in past studies was also employed for this one (Ferraro et al 2017, Trembanis et al. 2019), where *c* is incidental mortatlity, $compromised_t$ and $compromised_0$ are the number of compromised scallops before (time 0) and after the experimental treatment (time t), and $total_t$ and $total_0$ are the total number of scallops before and after the experimental treatment (Figure 2.3).

$$c = \frac{compromised_t}{total_t} - \frac{compromised_0}{total_0} \times 100$$

Figure 2.3: Formula for incidental mortality constant, *c*

In addition to IM calculations other comparative statistics were run in the R coding software (R Core Team, 2020)). Two proportion Z-tests were run to determine if the proportion of compromised scallops significantly differed before and after treatment across sites.

Side Scan Sonar

The processed side scan sonar mosaics and exported dredge scar polygons produced for this study previously were used to filter processed images by their location relative to the dredge scars at each of the 5 tow sites. Using the "sp" package of R (Pebesma, 2005) images were determined to either be inside of the scar polygons and a 3 meter buffer zone or outside of the scar based upon their centroids. After scar filtering

has been performed IM values can be calculated for the dredge scar area and a chosen buffer zone, while also isolating the area outside of the scar for further baseline analysis.

ScallopScan Implementation

The ScallopScan convolutional neural network was shown the images of the 2017IM cruise following its implementation in the 2017 dredge efficiency project. To quantify the relationship between the annotations of ScallopScan as they compare to those of human annotators for the images of this study, the platforms were compared on a scallop count per site and count per image basis. Additional sizing analysis was performed by utilizing the same methods as outlined in Trembanis et al. (2019) (Figure 2.4).

$$SH = W_{pixel} \frac{(X_2 - X_1) + (Y_2 - Y_1)}{2}$$

Figure 2.4: Formula for calculating shell height from ScallopScan bounding boxes.

Chapter 3
RESULTS

Incidental Mortality Study on Hard Substrate

The manual annotation of the images from Closed Area I resulted in the processing of 38,813 images containing 31,972 scallops. Of the scallops annotated 1,968 were marked as “compromised” and 30,004 were marked as “healthy” (Table 1). When calculated on the scale of an entire study site the highest incidental mortality value was at site 1 where it was found to be 3.43%. The highest percent compromised scallops at site 1 was 9.32% and observed in mission 63. At site 2 incidental mortality was calculated at -0.017% after treatment and the percent compromised scallops after treatment was 3.46%. At site 3 incidental mortality was calculated to be -0.056%. Negative IM values indicate a decrease in the number of compromised scallops observed after treatment. The percent compromised scallops after treatment at site 3 was 11.99%. The mean percent change in compromised scallops for the 5 tow treatments was calculated to be 0.54 ± 1.51 and the mean percent change in compromised scallops for the control treatments was found to be -3.17 ± 1.58 (Figure 3.1).

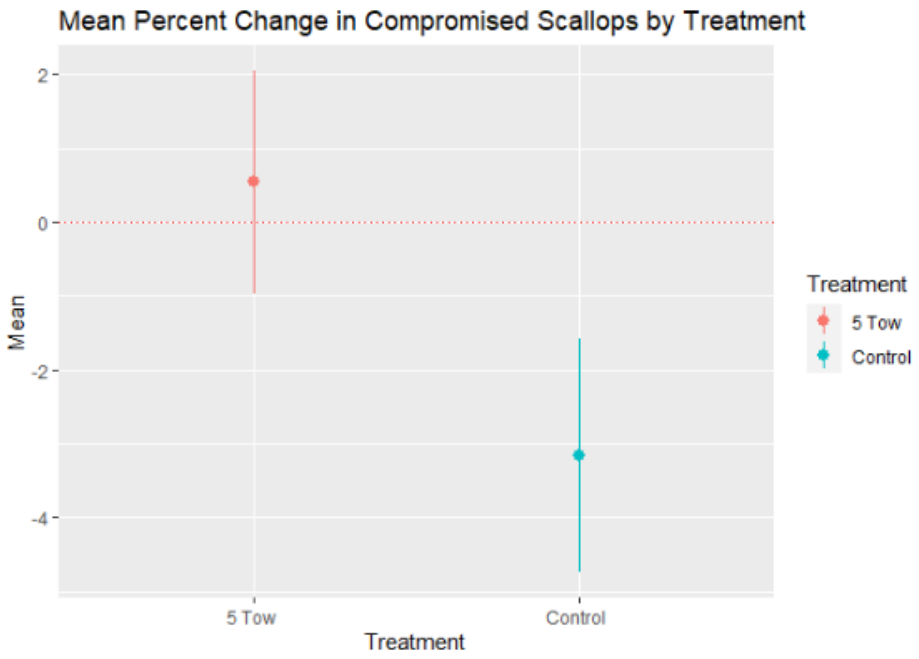


Figure 3.1: Mean change in percent compromised scallops by treatment

When tested using a two-proportion z test the increase in compromised scallops at the site 1 5 tow site was not found to be significant (p-value = 0.99). Two proportion z tests at the site 2 and three 5 tow sites revealed a significant decrease in and no significant change in the proportions of compromised scallops, respectively. Data from all 3 of the control sites consistently showed decreases in the percent compromised scallops detected after treatment. Due to turbidity during the mission at site 3 the second control mission could not be meaningfully annotated.

After observing the results from the five tow treatments when looking at the sites in total, scallops were filtered as to their location relative to the dredge scars at each site.

Site 1 had the largest prevalence of dredge scar features with 3 individual dredge scars inside the survey area (Figure 3.2).

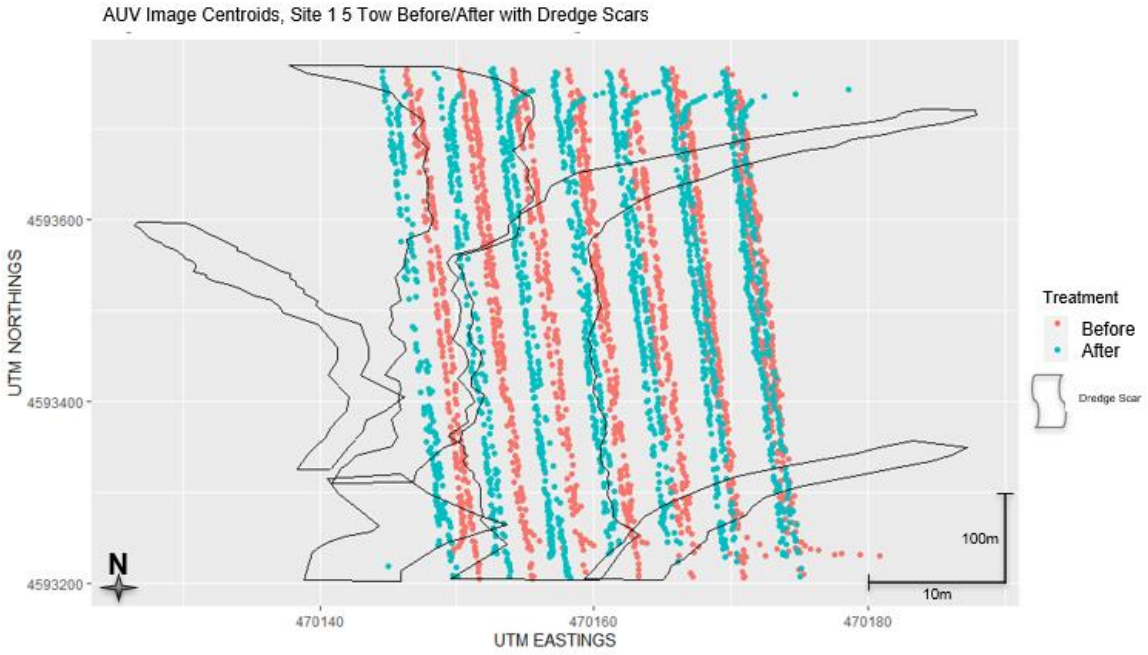


Figure 3.2: The sea scallop centroids of both the before (red) and after (blue) site 1 five-tow missions plotted with dredge scar bounds (black).

Site 3 had 3 dredge scars as well, but one was already on the seafloor at the time of the pre-treatment survey. Site 2 only had one scar within the survey area. At site one 52.32% of the scallops lay within the dredge scar area, and the percent compromised inside the scar increased from 6.99% before treatment to 13.69% after treatment (Table 2). The calculated IM value inside the dredge scar was 6.70% at site 1, a slightly higher result than when observing the site in its entirety. The percent compromised outside the

scar also slightly increased at site 1 from 4.66% to 5.21%. Site 2 showed decreases in the percent compromised after treatment both in and outside the scar. The number of scallops detected for site 2 increased after treatment inside and outside the scar, however. At site 3 the percent compromised scallops increased by less than a percent from 14.61% to 15.15% inside the scar and decreased from 9.92% to 8.71% outside the scar (Table 3).

Filtering images by their location relative to the dredge scar area significantly increased the proportion of compromised scallops after the five-tow treatment for site 1. The proportion of compromised scallops inside the scar rose from 6.99% to 13.69% following treatment at site 1, which overlaps with the lower end of the expected IM values for hard substrate. However, the dredging performed at this site does not completely represent a true five tow treatment as there are only 3 scars in the survey area and the overlap between them is minimal. This result does suggest that under true high fishing pressure scenarios IM may lie in the literature range of 10-20% in areas of hard substrate (Caddy, 1973, Murawkski and Serchuk 1989).

Despite the challenges set forth by high turbidity and low image quality at site 3, some anecdotal evidence from the site suggests that literature values for incidental mortality remain conservative yet appropriate. Annotation efforts on the first mission of the site 3 control site returned a value of 17.78% compromised scallops, but the follow up mission could not be annotated due to high turbidity and resulting poor image quality. At the 5-tow mission site the percent compromised scallops was found to be 12.04% before and 11.98% compromised after treatment for the site. The lack of change, but initially high proportion of compromised scallops at the study site coincides with the presence of

a preexisting dredge scar running in the east -west direction inside the 5-tow AUV survey area. This foreign scar suggests that fishing effort was taking place relatively recently in comparison to the timing of the experimental trials at this site given the presence of a detectable dredge scar in the side scan sonar mosaic from the pre-treatment survey. Only 2 of the 5 dredge scars inflicted for the purpose of the experiment at site 3 overlapped with the survey area and no significant effect was found from the experimental treatment in the 5-tow area. Filtering the scallops by their location relative to the scar increased the proportions of compromised scallops inside the scar to 14.61% before treatment and 15.15% after treatment, meeting the bounds of the literature estimates for harder substrates set forth by earlier studies (Caddy, 1973, Ferraro et. al. 2017). However, outside the scar compromised scallop proportions remained generally high as well with values of 9.92% before treatment and 8.71% after treatment, suggesting that mortality values at site 3 were ambiently high before further perturbation by our experimental trials.

Sediment Classification

Of the 38,813 images annotated across all 3 sites 54.46% received the substrate classification of gravel. The next highest sediment annotation type was sand at 30.38%. Rocky images only comprised 4.36% of the annotated image set. There were noticeable changes in the substrate composition of each study site within the span of an AUV mission (Figure 3.3). The scallops annotated for this study were largely annotated as lying on hard substrates with 70.2% of the scallops being from images marked as "Gravel" substrate(Table 4). In total, 20.5% of the images of this study were marked as

"Sandy" and 2.8% were marked "Rocky"(Table 5). Of the compromised scallops, 77.4% came from gravel substrate and 6.8% of them came from rocky images, meaning 84.2% of the compromised scallops came from hard substrates.

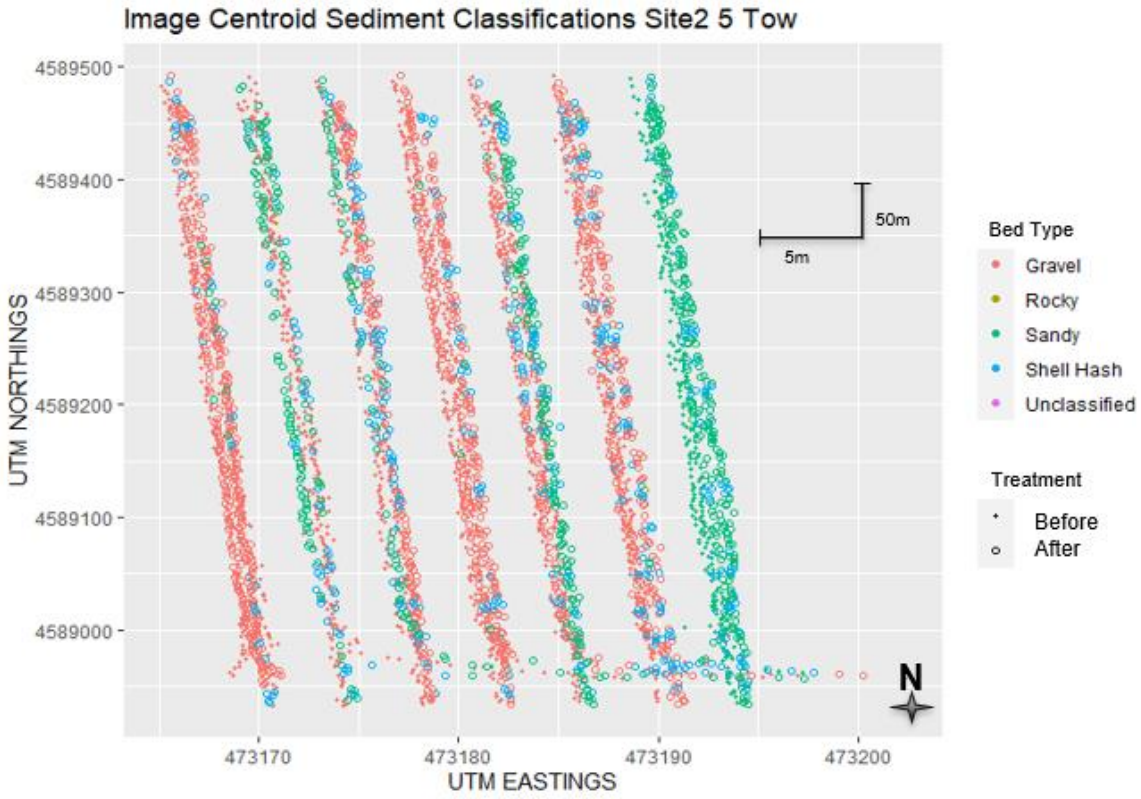


Figure 3.3: AUV image centroids color coded by sediment annotation. Centroid shapes determine whether points are from before (points) or after (circles) the experimental treatment.

ScallopScan Shell Height and Scallop Counts

At the AUV mission level ScallopScan scallop counts mirrored those of human annotators, although they were often conservative in comparison (Figure 3.4). Scallop count per image values were also close between the point and line annotations.

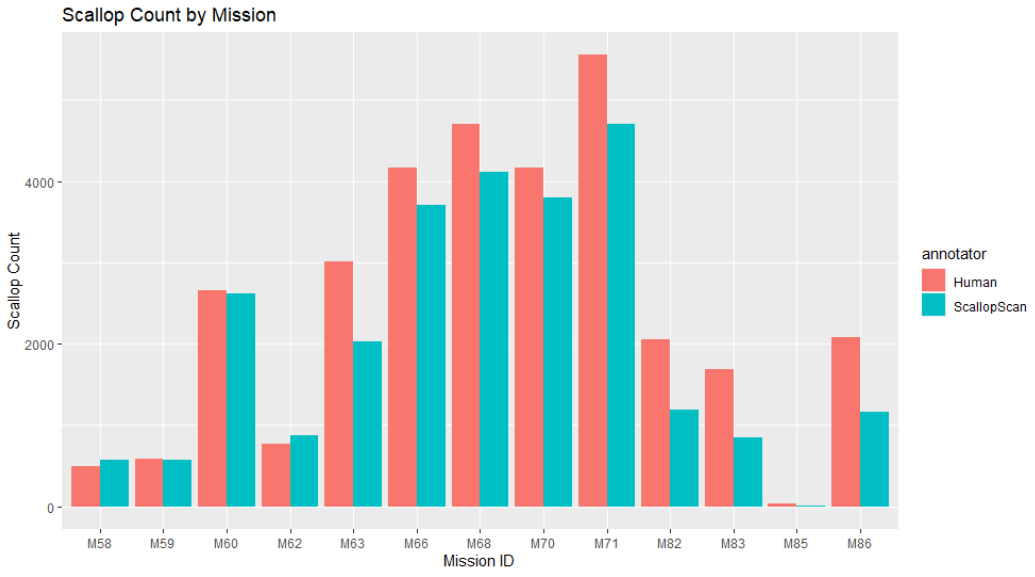


Figure 3.4: Scallop count by AUV mission for human annotators (red) and ScallopScan (blue).

Variation is present between the shell height distributions for ScallopScan and those of human annotators. At site 1 the mean shell height for ScallopScan was 128.9mm. This is larger than the largest scallop annotated manually across all sites, and the mean shell height for manually annotated scallops at site 1 five-tow area was 72.9mm. The shell

height distributions for ScallopScan do mirror those of the dredge, however (Figure 3.5). The mean shell height for the dredge at the site 1 five-tow area was 139.3mm, also larger than any scallop measured by human annotators.

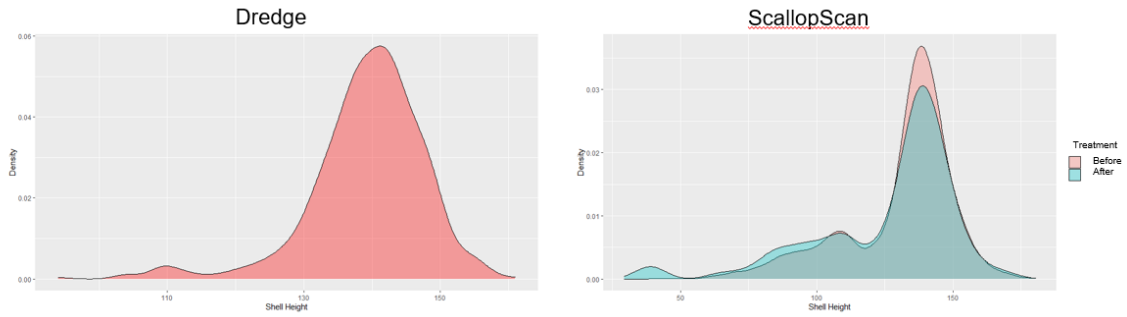


Figure 3.5: Left: Dredge scallop shell height distribution in millimeters. Right: ScallopScan shell height distribution in millimeters.

Sea scallop shell height also seems to play a role in detectability, as smaller individuals saw decreased numbers of compromised scallops after treatment than their larger counterparts, even under control conditions for scallops annotated by humans (Figure 3.6). The human detected sea scallops were placed into three size bins based upon their derived shell heights. Large 90-120mm compromised scallops were less likely to be detected after treatment when compared to control values for the same shell height bin, but the margin of error for these estimates is high and overlaps between the control and five tow sites.

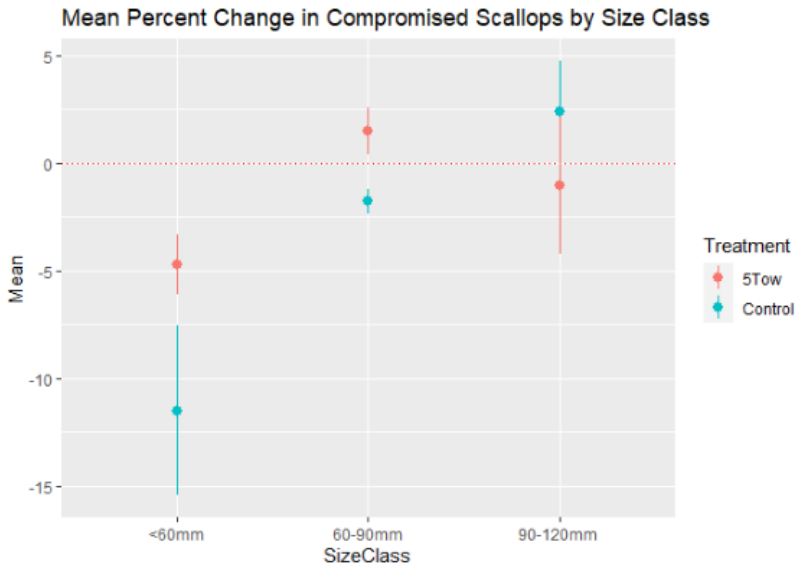


Figure 3.6: A dot plot illustrating the percent change in compromised scallops by size class for control and 5 tow treatments. Error bars represent the standard error of the mean.

The detection of medium sized compromised scallops in the 60-90mm range decreased between treatments for control sites on average but increased slightly for five tow treatments with a mean value of 1.5133 ± 1.077%. Small individuals less than 60mm were the least likely to be seen more than once, with a mean change in percent compromised scallops calculated to be -11.50 ± 3.94% at the control sites. These results show a steady increase in the percent compromised scallops detected after treatment for control sites as a function of shell height. At five tow sites this same increase is mirrored, except at the largest shell height bin.

Chapter 4
DISCUSSION

Incidental Mortality Study

Overall, we found significantly different IM values across our three study sites after what should have been uniform fishing pressure between treatments. Site 1 was the most consistent with results from past studies, although the increase in compromised scallops observed in the five-tow treatment and resulting incidental mortality calculation yielded a value (3%) more consistent with soft substrates than hard. The decrease in compromised scallops at Site 2 combined with a lack of effect at site 3 do little to support the findings of past studies or the hypotheses put forth by this study and potentially reflects a lack of uniformity in our execution of the 5-tow mission design.

Nuanced spatial analysis in the form of dredge scar filtering allowed for a greater understanding of the dynamics in and around the areas directly impacted by the dredge treatments of this study, but the dredge treatments themselves did not overlap with the survey areas as put forth in the experimental design, further exacerbating issues with observing an effect from the experimental treatments. This result suggests that an experimental design less reliant on precision in navigation and repetition would be desirable for future studies. Site 3 stands out as a model for assessing ambient mortality values for a piece of seafloor that has experienced fishing pressure without the need for

an additional experimental treatment. The mortality observed at site 3 did not change significantly after the experimental treatment, and the time lag between the foreign dredge scar and the time that we arrived on site are unknown. However, mortality was still observed in the dredge scar area and the mortality values for the site were consistent with literature values despite the lack of overlapping dredge tows at the site.

Site 1 creates the most complete illustration of the intended experimental design as it contains the most dredge scar activity and demonstrated predictable changes to the sea scallop shell height distribution remaining on the seafloor following the dredging treatment. Although the mean shell height of 72mm for site 1 is relatively small in the scale of commercially exploitable scallops, the removal of larger shell height individuals was observed after trial and coincided with an increase in the number of compromised individuals across shell heights (Figure 4.1). This could be a result of the removal of larger shell height individuals by the dredge during treatment in conjunction with a culling action and general displacement being forced upon smaller individuals at the same study site during the process of dredging.

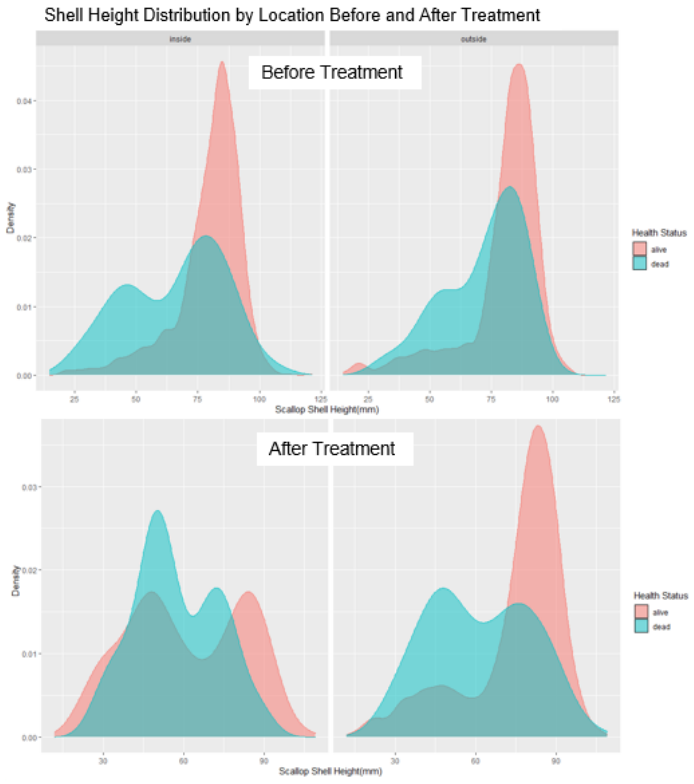


Figure 4.1: Shell height distributions before (top) and after (bottom) treatment broken down by location relative to the dredge scar. Healthy scallops comprise the red distribution and compromised scallops comprise the blue distributions.

These changes were observed inside and outside of the dredge scar but were much more noticeable within the dredge scar area. The decrease in the percent of compromised scallops observed across the control sites is likely attributed to navigational error and image down sampling. Despite having adequate overlap to down sample as in past studies the image footprints show a lack of overlap between control mission replicates. This issue

is more widespread than just the control missions though, with the 5 tow missions of sites 1 and 3 experiencing similar variability.

The site 2 five tow missions illustrate more consistent AUV navigation than the other sites and treatments. This observed overlap correlates more with literature documentation of the capabilities of the vehicle (Ferraro et al 2016, Walker et al 2016, Singh et. al. 2013). However, proportions of compromised individuals at this site decreased after the five-tow treatment for the site as a whole and when filtered by inside or outside of the dredge scar area, contrary to expectations. Shell height distributions also remained constant before and after treatment at the five-tow site. Upon examining the image counts of the legs of the 5 tow missions and noting the time lag of roughly 4.5 hours between the AUV missions some discrepancies become apparent. The AUV was operating on a fixed propellor speed and was thus affected by tidal current flows while navigating the mission. This effect can be observed in the oscillation of the image counts for each leg of an AUV mission (Table 6). The oscillation in image count combined with relatively small sample sizes within sites and down sampling the images annotated to 1 in 4 could have resulted in the variability observed in scallop counts before and after treatment. There was also only one dredge scar overlapping with the survey area at site 2, meaning that the experimental design was more like that of a one tow site of past studies. The signal of one tow treatments has been regarded as minimal in the literature (Ferraro et. al. 2017) and this was the fundamental reason why only the control and five tow sites of this study were considered for IM analysis. The results of this study site reinforce the difficulty in deciphering a signal in trials with only a single dredge tow and show that

even if the navigation of the vehicle is acceptable there are other factors that can still create challenges for data analysis.

Subsequently, the results from sites 1 and 3 suggest that IM can still be assessed despite navigational issues with the AUV platform. However, the time investment and navigational precision required for both a fishing vessel and an AUV to replicate dredge tows and photographic surveys respectively across weather and tidal cycles during the span of a study leave little room for error in data collection efforts.

Having only three replications of the control and 5 tow treatments also constrains data analysis, and in the case of this study resulted in one of each possible outcome: an increase in the proportion of compromised scallops after treatment (site 1), a decrease in the proportion of compromised scallops after treatment (site 2) and no effect on the proportion of compromised scallops after treatment (site 3). Small sample sizes are another hallmark of this dataset. Having small scallop sample sizes could amplify the "noise", or lack of signa observed in past studies. For compromised scallops small sample size could explain the decrease in the proportions observed after treatment as 4 of the 6 control trials had sample sizes less than 100. A greater number of replications with larger prevalence of dredge impact zones and higher scallop densities would be largely beneficial to the experimental design of this study.

Another aspect of the IM analysis effort that has been continually referenced is the time and personnel investment required to process image sets of this size. For this study only two annotators were utilized, which helped to limit time investment but caused

progress to move more slowly than it may have with a full annotation team. On the other hand, a small team of annotators provides for less inter-annotator variability and thus more internally consistent annotations.

Another inherent limitation of the collected dataset lies in the assumption of mortality associated with annotating a scallop as "compromised". Scallops were marked as compromised if a visible crack was present in the shell, if the scallop had been pushed into the sediment, or the scallop was flipped upside down. However, the physiology of scallops allows for repair to minor shell lacerations which could result in some proportion of the cracked compromised scallops to remain alive despite being marked for use in IM calculations. Additionally, scallops are capable of locomotion and some of the flipped individuals could have righted themselves after being marked as compromised as well. Some injuries following dredging are also internal, such as the hinge of the scallop being packed with sand because of interacting with the dredge. This could result in some individuals that appear as healthy individuals to have been compromised scallops. These annotation outcomes likely comprise a small proportion of the total dataset but could add another explanation for the lack of signal observed in the trials of this study.

Autonomous Underwater Vehicle Variability

The navigational error from the AUV suggests that the reported capabilities of the AUV platform as used in this study may only be realized under ideal current and turbidity conditions. These results also suggest that the platform may be more prone to

navigational discrepancies because of environmental influence than previously understood. Large geological features such as rocks and sand waves are likely culprits for the variability in altitude noted previously. Past studies completed with the Gavia AUV report sub meter navigational precision and highly repeatable mission design but were performed in the mid-Atlantic and on Georges Bank where currents are anecdotally known to be less intense than in Closed Area I and the area surrounding the great south channel. Additionally, an unprecedented level of turbidity was present in the images of this study in comparison to the image sets of past studies. Observed variability in image count of the legs within a mission is further evidence for strong currents during the trials of this study and these currents likely contribute to some of the issues with image quality observed throughout the image processing of this study.

Largely, the initial aim of this study was to rigorously examine the dynamics of incidental mortality on harder substrates than previous studies in and around Closed Area I. In setting such goals the importance of satisfying the assumptions of the experimental design become important. Throughout all vehicle activities logs of navigational and environmental data assure adherence to the experimental design and allow for examination of the performance of the Gavia AUV system for this study. After observing minimal effect from the experimental treatment when looking at the study sites in their entirety and experiencing noticeable variability in the proportions of healthy and unhealthy scallops at the control sites of the study the quality of the image set came under scrutiny. Upon plotting the image centroids of the site 1 and site 3 5 tow missions the drift of the vehicle off course follows a similar pattern – the initial survey line is on

course and has high overlap with the previous missions' images but as the vehicle proceeds to subsequent survey lines the amount of overlap between image sets decreases (Figure 4.2).

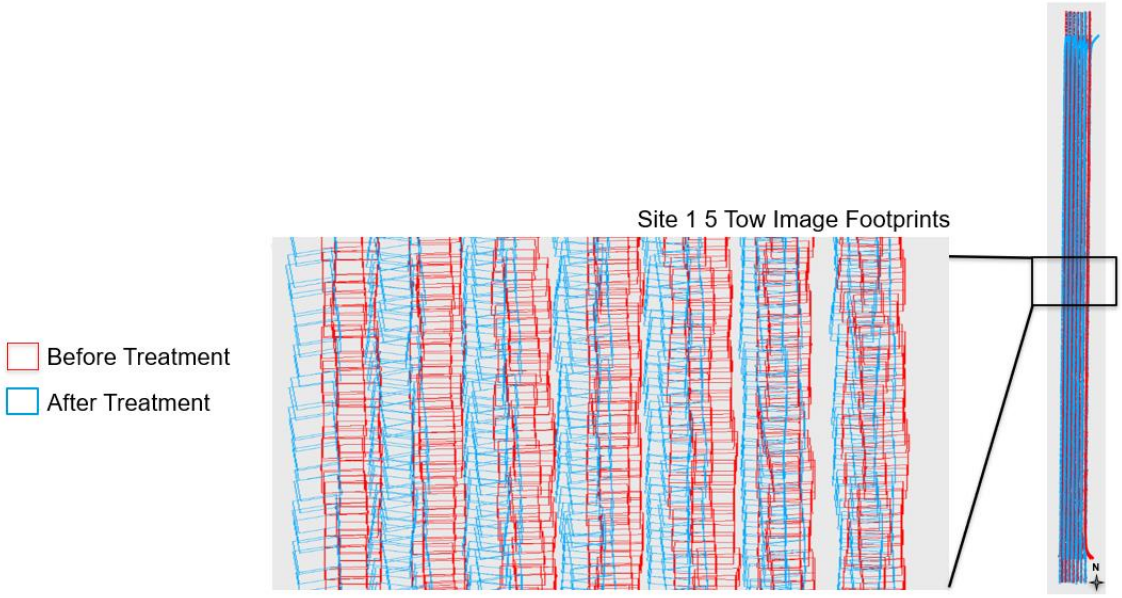


Figure 4.2: Google earth kml display of image footprints from the Site 1 5 Tow treatment.

This decrease in overlap worsens as the mission proceeds and is consistent with error associated with the internal navigation system (INS) of the AUV. Additionally, vehicle altitude is also assumed to stay within a meter of its fixed 2.5m altitude. The 95 confidence intervals for each study site show that altitude values were within a meter of the mean value at each site, however the mean value of altitude differed between each site. As a result, sizing derivations for each site were calculated separately based upon the

mean altitude of the missions within. When calculating scallop shell height from altitude derived pixel values for the upper and lower bounds of the confidence intervals at each site shell height was found to vary by less than a millimeter, thus negating the suspected effects of variable altitude for most of the scallops sized in this study.

Sediment Classification

Variablitiy and a lack of visible sorting of sediments about the site illustrates that sediments are highly variable on the spatial scale of the sites of this study and presents one of the possible explanations as to why mortality values are variable between the sites themselves and the treatments within. Site 2 provides an example of highly mobile sediment and bedform feature on the scale of days (Figure 4.3).

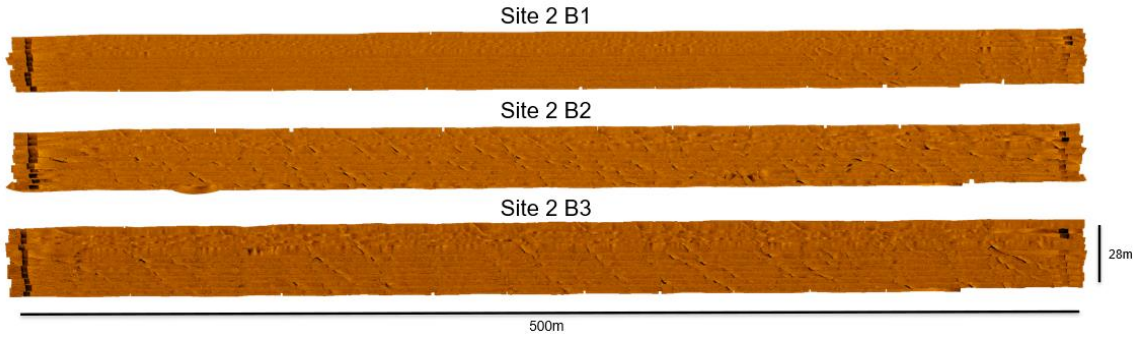


Figure 4.3: Side scan sonar mosaics from the site 2 five-tow treatment AUV missions.

The time lag between the B1 and B2 trials was more than a day and resulted in large textural changes in the seafloor about the study site in the side scan sonar data.

There was a shorter time lag on the scale of hours between the B2 and B3 trials but changes in the seafloor also appear between these two sonar surveys as well. This illustrates a highly dynamic seafloor that likely accounts for much of the variability in sediment annotation between the images of the AUV missions of site 2. The presence of a dredge scar at site 3 shows that although things were highly dynamic at site 2, site 3 has a more stable sediment distribution. This inter-site and intra-site variability further illustrates the spatial heterogeneity of the sea scallop grounds.

ScallopScan Shell Height and Scallop Counts

Possessing a largely functional ScallopScan neural network that had been trained and tested on scallop detections form the ETCA and NLCA at the start of this study ended up rewarding us by making the leap to a new study area and producing scallop count numbers that were consistent with those of our manual annotation effort. Albeit the ScallopScan network is only detecting “healthy” scallops, but compromised scallops such as those flipped over by the dredge appear similarly to a healthy albino individual to the untrained eye. As such some compromised individuals are likely to fall into the ScallopScan detections for this study. There is no denying the need for integration of automated detection of compromised scallops for ScallopScan to be viable in contributing to IM calculations. Although it is a simple maneuver for human annotators to make the jump to annotating another object class, the same cannot be said for convolutional neural networks such as the YOLO based ScallopScan system. Introducing another degree of freedom to a neural network system complicates the object detection workflow and increases the potential for error in detection. However, ScallopScan

remains a valuable tool for processing large batches of images in short periods of time. Its precision and recall levels have reached a point that its estimations can be taken for a conservative approximation of the true scallop abundance for an AUV mission.

Experimentation with the ScallopScan system is also still ongoing and contributions to its progress through this study have been largely limited by time and personnel constraints. However, the annotations of this study have been converted from pixel values of shell height lines to bounding boxes that could be ingested into future neural network experiments. The images of this study possess coarser substrates than previous studies and as a result the detection of scallop look-alikes such as larger stones and clams are more likely than in previous image sets that the network had been tested on. A lack of false positive detections illustrates the tradeoff between time investment into training and testing neural network systems for this application, as making detections for the single object class, “healthy scallop” was still possible for a new and more dynamic area of seafloor than previously experimented on without additional training (Table 9). Automated image detection serves to alleviate the largest bottleneck to image analysis studies, as the rate at which today’s neural networks can process through images trump even the fastest of human annotators. The results of this study support the need for further training and testing of the ScallopScan system. The use of the ScallopScan or a similar system in a more portable and modular format for implementation on the AUV itself in the field or off the vehicle on field computers would be largely beneficial to future experiments of this kind. Rapid acquisition of ballpark scallop populations and

distributions could help with site selection and ecological studies seeking to examine scallop populations in larger spatial domains with modest time investment.

Recommendations for Future Incidental Mortality Studies

Given the technological advancements of the sea scallop fishery and the requirement for vessel monitoring systems (VMS) to be recording during the process of harvest, perhaps a more novel approach to investigating IM *in-situ* could be to target areas of seafloor within managed fishery areas that have already had a high prevalence of recent scallop vessel tracks. Performing side scan sonar surveys in suspected areas of high traffic would allow for the creation of study sites based upon the presence or absence of dredge scars, and the opposite could be true for controls. Streamlining data collection in this way could allow for more study sites to be selected in the timeline of a cruise and would rely less upon navigational precision in an at sea environment for assessment of incidental mortality.

The results of this study provide evidence of highly dynamic conditions within the scallop grounds of Closed Area I and the surrounding open bottom area north of the Great South Channel. High mobility was observed in the sediment distributions of the study sites, which shifted spatially on the scale of meters and temporally on the scale of hours during our trials. Variability in the execution of the experimental design and the constraints of accurately annotating incidences of mortality in the population of scallops of this study made the detection incidental mortality especially challenging despite improvements in data processing and lessons learned from studies of the past. Our IM estimates are also conservative when compared to the literature, and further support the

theory that mortality is a highly variable and challenging parameter to assess rigorously with a MBACI experimental design.

TABLES

Table 1: Scallop Sample Sizes by Mission for Human Annotators

Mission ID	Treatment	Number of Scallops	Number of Healthy Scallops	Number of Compromised Scallops	Percent Healthy	Percent Compromised
58	Site 1: Control, Before	497	459	38	92.35%	7.65%
59	Site 1: Control, Before	586	539	47	91.98%	8.02%
60	Site 1: 5 Tow, Before	2653	2497	156	94.12%	5.88%
62	Site 1: Control, After	772	748	24	96.89%	3.11%
63	Site 1: 5 Tow, After	3016	2735	281	90.68%	9.32%
66	Site 2: 5 Tow, Before	4162	3946	216	94.81%	5.19%
68	Site 2: Control, Before	4712	4561	151	96.80%	3.20%
70	Site 2: Control, After	4162	4095	67	98.39%	1.61%
71	Site 2: 5 Tow, After	5551	5359	192	96.54%	3.46%
82	Site 3: 5 Tow, Before	2060	1812	248	87.96%	12.04%
83	Site 3: Control, Before	1682	1383	299	82.22%	17.78%

85	Site 3: Control, After	41	41	0	100.00%	0.00%
86	Site 3: 5 Tow, After	2078	1829	249	88.01%	11.99%
Totals		31972	30004	1968		

Table 2: Five Tow Before Missions Dredge Scar Summary

	Number of Scallops Inside Scar	**Number of Scallops Outside Scar**	**Percent of Scallops Inside Scar**	**Percent Compromised Inside Scar**	**Percent Compromised Outside Scar**
Site 1	1388	1265	52.32%	6.99%	4.66%
Site 2	1879	2283	45.15%	5.69%	4.77%
Site 3	931	1129	45.19%	14.61%	9.92%

Table 3: Five Tow After Missions Dredge Scar Summary

	Number of Scallops Inside Scar	**Number of Scallops Outside Scar**	**Percent of Scallops Inside Scar**	**Percent Compromised Inside Scar**	**Percent Compromised Outside Scar**
Site 1	1461	1555	48.44%	13.69%	5.21%
Site 2	1917	3634	34.53%	4.59%	2.86%
Site 3	1056	1022	50.82%	15.15%	8.71%

Table 4: Image Counts by Sediment Classification and Scallop Prevalence

	Sandy	**Rocky**	**Gravel**	**Shell Hash**	**Unclassified**	**Totals**
Images with Scallops	3807	613	11755	720	39	16934
Images Without Scallops	7985	1078	9383	443	2990	21879
Totals	11792	1691	21138	1163	3029	38813

Table 5: Image Count Percentages by Sediment Classification and Scallop Prevalence

	Sandy	**Rocky**	**Gravel**	**Shell Hash**	**Unclassified**	**Totals**
Images with Scallops	9.81%	1.58%	30.29%	1.86%	0.10%	43.64%
Images Without Scallops	20.57%	2.78%	24.17%	1.14%	7.70%	56.36%
Totals	30.38%	4.36%	54.46%	3.00%	7.80%	100%

Table 6: AUV mission image counts and scar filtering values.

MissionID	Treatment	MissionLeg	Totalimages	ImagesOutsideScar	ImagesInsideScar	PercentInScar
58	Site1_C1	20170817135046	779	767	12	2
58	Site1_C1	20170817142559	194	193	1	1
60	Site1_B1	20170817181849	787	0	787	100
60	Site1_B1	20170817182441	1129	0	1129	100
60	Site1_B1	20170817183225	778	17	761	98
60	Site1_B1	20170817183814	1139	29	1110	97
60	Site1_B1	20170817181112	1121	48	1073	96
60	Site1_B1	20170817180521	799	89	710	89
60	Site1_B1	20170817184602	767	272	495	65
60	Site1_B1	20170817185147	1165	652	513	44
60	Site1_B1	20170817185940	775	580	195	25
60	Site1_B1	20170817190527	1187	910	277	23
60	Site1_B1	20170817191327	785	629	156	20
60	Site1_B1	20170817191914	1166	969	197	17
60	Site1_B1	20170817192705	773	675	98	13
60	Site1_B1	20170817193255	1166	1028	138	12
63	Site1_B2	20170818125049	829	0	829	100
63	Site1_B2	20170818130409	818	1	817	100
63	Site1_B2	20170818124338	1042	4	1038	100
63	Site1_B2	20170818125648	1073	11	1062	99
63	Site1_B2	20170818131006	1110	231	879	79
63	Site1_B2	20170818131739	807	175	632	78
63	Site1_B2	20170818123733	836	308	528	63
63	Site1_B2	20170818122932	1165	440	725	62
63	Site1_B2	20170818132331	1116	819	297	27
63	Site1_B2	20170818133112	792	592	200	25
63	Site1_B2	20170818133705	1150	918	232	20
63	Site1_B2	20170818134457	771	631	140	18
63	Site1_B2	20170818135845	773	668	105	14
63	Site1_B2	20170818135042	1183	1036	147	12
65	Site2_A1	20170818183823	856	75	781	91
65	Site2_A1	20170818184435	1042	101	941	90
65	Site2_A1	20170818185143	849	207	642	76
65	Site2_A1	20170818183127	994	245	749	75
65	Site2_A1	20170818185750	1055	299	756	72
65	Site2_A1	20170818182511	866	441	425	49
66	Site2_B1	20170818201117	1183	0	1183	100
66	Site2_B1	20170818200526	780	25	755	97
66	Site2_B1	20170818195733	1165	50	1115	96
66	Site2_B1	20170818195147	759	69	690	91
66	Site2_B1	20170818201918	780	263	517	66
66	Site2_B1	20170818194346	1178	478	700	59
66	Site2_B1	20170818193759	768	400	368	48
66	Site2_B1	20170818202507	1168	615	553	47
66	Site2_B1	20170818203303	762	709	53	7
66	Site2_B1	20170818203850	1189	1159	30	3

MissionID	Treatment	MissionLeg	Totalimages	ImagesOutsideScar	ImagesInsideScar	PercentInScar
67	Site2_A2	20170819122414	918	79	839	91
67	Site2_A2	20170819121023	903	80	823	91
67	Site2_A2	20170819123047	1051	107	944	90
67	Site2_A2	20170819121651	1067	110	957	90
70	Site2_C2	20170819185125	1014	1014	0	0
71	Site2_B2	20170819205318	1171	0	1171	100
71	Site2_B2	20170819204727	770	28	742	96
71	Site2_B2	20170819203932	1157	44	1113	96
71	Site2_B2	20170819203339	784	61	723	92
71	Site2_B2	20170819202558	1123	427	696	62
71	Site2_B2	20170819210116	782	370	412	53
71	Site2_B2	20170819202007	793	405	388	49
71	Site2_B2	20170819210707	1200	761	439	37
71	Site2_B2	20170819214847	1192	1192	0	0
82	Site3_B1	20170821111029	1464	159	1305	89
82	Site3_B1	20170821112011	694	148	546	79
82	Site3_B1	20170821110454	696	177	519	75
82	Site3_B1	20170821115447	641	229	412	64
82	Site3_B1	20170821112550	1673	599	1074	64
82	Site3_B1	20170821114229	1854	665	1189	64
82	Site3_B1	20170821105529	1397	566	831	59
82	Site3_B1	20170821113653	656	417	239	36
82	Site3_B1	20170821104959	704	467	237	34
82	Site3_B1	20170821102033	801	712	89	11
82	Site3_B1	20170821103519	714	647	67	9
82	Site3_B1	20170821102630	1314	1229	85	6
82	Site3_B1	20170821101922	145	145	0	0
83	Site3_C1	20170821143342	1436	1436	0	0
84	Site3_A1	20170821170537	1101	237	864	78
84	Site3_A1	20170821171305	795	215	580	73
84	Site3_A1	20170821165949	790	265	525	66
84	Site3_A1	20170821171851	1120	574	546	49
84	Site3_A1	20170821165235	1053	637	416	40
84	Site3_A1	20170821172625	788	529	259	33
84	Site3_A1	20170821173210	1166	1015	151	13
85	Site3_C2	20170821194540	720	720	0	0
86	Site3_B2	20170821220721	1135	92	1043	92
86	Site3_B2	20170821220136	788	86	702	89
86	Site3_B2	20170821215418	1074	160	914	85
86	Site3_B2	20170821223416	1214	387	827	68
86	Site3_B2	20170821221501	769	260	509	66
86	Site3_B2	20170821222837	742	271	471	63
86	Site3_B2	20170821214827	812	297	515	63
86	Site3_B2	20170821214129	1004	589	415	41
86	Site3_B2	20170821222040	1185	736	449	38
86	Site3_B2	20170821211105	682	631	51	7
86	Site3_B2	20170821211554	937	868	69	7
86	Site3_B2	20170821212837	971	906	65	7
86	Site3_B2	20170821212229	870	827	43	5
86	Site3_B2	20170821210951	152	152	0	0

MissionID	Treatment	MissionLeg	Totalimages	ImagesOutsideScar	ImagesInsideScar	PercentInScar
87	Site3_A2	20170821234215	642	146	496	77
87	Site3_A2	20170821234730	1708	462	1246	73
87	Site3_A2	20170821233156	1563	472	1091	70
87	Site3_A2	20170821232630	696	390	306	44
87	Site3_A2	20170821235852	327	188	139	43
87	Site3_A2	20170822000432	287	233	54	19
87	Site3_A2	20170822001603	135	126	9	7
87	Site3_A2	20170822003846	284	284	0	0
87	Site3_A2	20170822005538	371	371	0	0

Table 7: ScallopScan scallop count by mission

Mission ID	**Treatment**	**Raw Number of Scallops**	**Downsampled Number of Scallops (25%)**
58	Site 1: Control, Before	2317	579
59	Site 1: Control, Before	2304	576
60	Site 1: 5 Tow, Before	10487	2622
62	Site 1: Control, After	3483	871
63	Site 1: 5 Tow, After	8136	2034
66	Site 2: 5 Tow, Before	14829	3707
68	Site 2: Control, Before	16434	4109
70	Site 2: Control, After	15196	3799
71	Site 2: 5 Tow, After	18786	4697
82	Site 3: 5 Tow, Before	4770	1193
83	Site 3: Control, Before	3378	845
85	Site 3: Control, After	47	12
86	Site 3: 5 Tow, After	4634	1159
Totals		104801	26203

REFERENCES

Caddy, J. F. 1973. Underwater observations on tracks of dredges and trawls and some effects of dredging on a scallop ground. Can. J. Fish. Aquat. Sci. 30:173–180.

Ferraro, D.M., Trembanis, A.C., Miller, D.C., and D.M. Rudders. 2017. Estimates of Sea Scallop (Placopecten magellanicus) Incidental Mortality from Photographic Multiple Before—After-Control—Impact Surveys. J. Shellfish Res. 36(3): 615-626

Jenkins, S. R. & A. R. Brand. 2001. The effect of dredge capture on the escape response of the great scallop, Pecten maximus (L.): implications for the survival of undersized discards. J. Exp. Mar. Biol. Ecol. 266:33–50.

Murawski, S., Serchuk, F. 1989. Environmental effects of offshore dredge fisheries for bivalves. ICES Statutory Meeting 1989.

Northeast Fisheries Science Center (NEFSC). 2018. 65th Northeast Regional Stock Assessment Workshop (65th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 18-11; 659 p.

Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rasmussen, C., Zhao, J., Ferraro, D., and Trembanis, A. 2017. Deep Census: AUV-Based Scallop Population Monitoring. Workshop on Visual Wildlife Monitoring at the International Conference on Computer Vision, October 22-29, 2017, Venice, Italy, 2017: 2865-2873.

Redmon, J., and A. Farhadi. 2018. YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Redmon, J., and A. Farhadi. 2017. YOLOv39000: Better, faster, stronger. In: Proceedings IEEE Computer Vision and Pattern Recognition. Pp. 7263-7271.

Rudders, D., & Roman, S. 2018. An Assessment of Sea Scallop Abundance and Distribution in the Nantucket Lightship Closed Area and Surrounds: Final Report. Virginia Institute of Marine Science, College of William and Mary. http://dx.doi.org/doi:10.21220/m2-pa13-za07

Rudders, D. B., Sally A. Roman, Art Trembanis and Danielle Ferraro. 2019. A Study to Assess the Effect of Tow Duration and Estimate Dredge Efficiency for the VIMS Sea Scallop Dredge Survey. VIMS Marine Resource Report No. 2019-4. 60 pp.

Stokesbury, K.D.E., Carey, J.D., Harris, B.P., and O'Keefe, C.E. 2011. Incidental fishing mortality may be responsible for the death of ten billion juvenile sea scallops in the mid-Atlantic. Mar. Ecol. Prog. Ser. 425, 167-173.

Tan, R. 2019. Breaking down mean average precision (mAP). URL https://towardsdatascience.com/breaking-down-mean-average-precision-map-ae462f623a52

Trembanis, A. Miller, D., Rudders, D. and D. Ferraro. 2017. Final report: Incidental mortality estimates of sea scallops from AUV based BACI surveys. URL https://www.nefsc.noaa.gov/coopresearch/pdfs/FR-14-0073.pdf

Trembanis, A., Rudders, D., Ferraro, D., Tipton, H. 2019. A Study of Incidental Mortality in Sea Scallops Investigating Predator Responseand Size Selective Rates of Mortality from BACI Image Surveys. Submitted to the NMFS NEFSC Cooperative Research Program

Walker, J.H., Trembanis, A.C., and Miller, D.C., 2016. Assessing the use of a camera system within an autonomous underwater vehicle for monitoring the distribution and density of sea scallops (Placopecten magellanicus) in the Mid-Atlantic Bight. Fish Bull. 114, 261-273.

Walker, J.H., 2013. Abundance and size of the sea scallop population in the Mid-Atlantic Bight. Master's thesis, 272 p. Univ. Delaware, Newark, DE.

Yochum, N. & W. D. Dupaul. 2008. Size-selectivity of the northwest Atlantic sea scallop (Placopecten magellanicus) dredge. J. Shellfish Res. 27:265–271.

Appendix A

SIDE SCAN SONAR MOSAICS OF STUDY SITES

Site 1, C1

Site 1, C2

Site 1, B1

Site 1, B2

Site 2, C1

Site 2, C2

Sie 2, B1

Site 2, B2

Site 2, B3

Site 3, C1

Site 3, C2

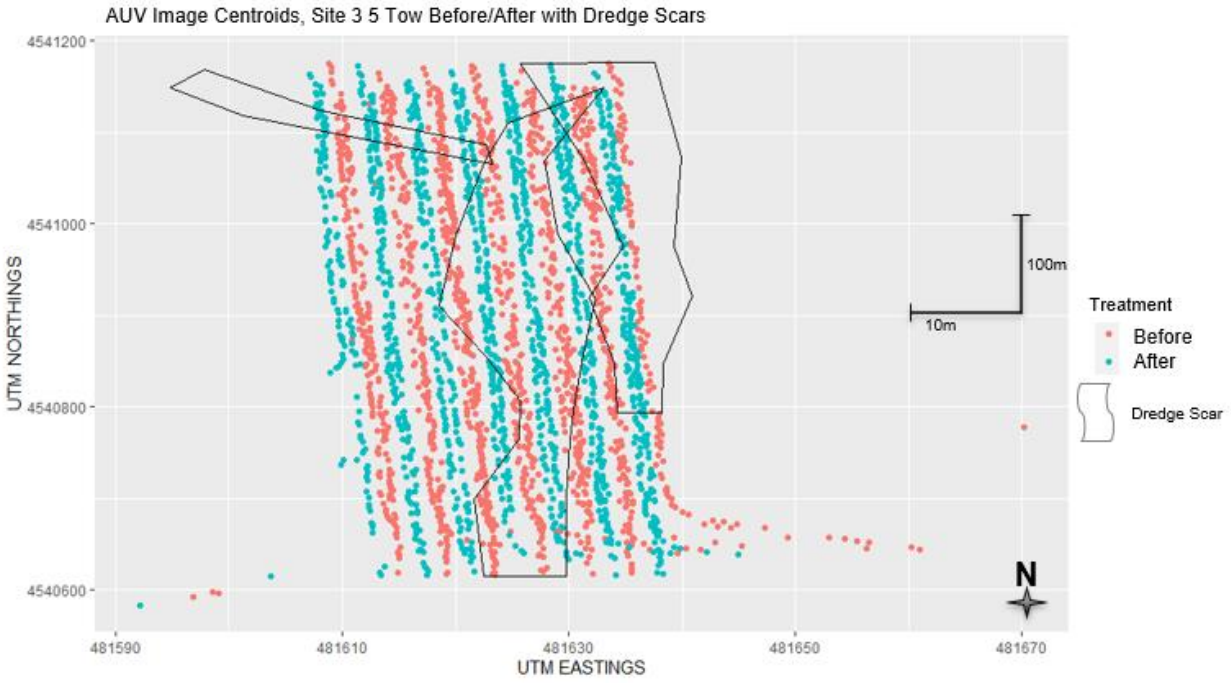
Site 3, B1

Site 3, B2

Appendix B

DREDGE SCARS

Dredge Scars of Site 2



Dredge Scars of Site 3