

Appendix B9: Summary of HabCam survey results for sea scallops and yellowtail flounder in the Nantucket Lightship Closed Area during 2009

The HabCam Group

Scott M. Gallager^{1, 2}, Amber D. York^{1, 2}

Richard Taylor^{2, 3}, Norman Vine²

Jonathan Howland¹, Steve Lerner¹

Karen Bolles^{3, 4}

(1) Woods Hole Oceanographic Institution

(2) Advanced Habitat Imaging Consortium

(3) Arnie's Fisheries, New Bedford, MA

(4) Ocean Explorium, New Bedford, MA

Conclusions

HabCam is a cabled optical and acoustic imaging system that is “flown” from a ship traveling at 5 kn at an altitude of 1 to 3 meters off the bottom while collecting high resolution still images at a rate of six images per second. Imaging rate provides ~50% overlap to allow for construction of image mosaics of the seafloor. A track approximately 100 nautical miles in length and 259,200 m² in area is imaged each 24 hour day while at sea. When operating continuously, HabCam samples nearly 2.5 times the area covered by a survey dredge.

Manual classification of the images provides the following information: 1) counts and measurements on sea scallops and groundfish (i.e. cod, haddock, flounders), epibenthic megafauna and many benthic infaunal species; 2) characterization of substrate; 3) observations on animal behaviors, inter- and intra-species interactions, biodiversity and community structure; 4) the ability to assess and monitor invasive species; 5) the tools to characterize oceanic properties (salinity, temperature, nutrients); and 6) the means to “map” the location of lost fishing gear (e.g. trawl and gillnets, lobster pots, and other miscellaneous fishing gear and parts). Automated methods for target classification are currently under development and will provide tools to reprocess archived results of image surveys as new technologies are developed.

Here we report on use of the HabCam camera system to: (1) conduct sea scallop surveys in the Nantucket Lightship Closed Area (NLSCA) as part of an effort to compare sampling technologies, (2) conduct dredge calibration with NOAA/NMFS vessels, and (3) conduct an analysis of inherent errors in camera calibration, scallop abundance estimates, and shell height measurements.

The objectives of the 2009 NLSCA survey were to estimate scallop abundance, shell height frequency distribution and biomass, and to estimate the distribution and abundance of yellowtail flounder in relation to substrate. A survey track line was designed as a modified spiral with track spacing from 2.6 to 1.3 nm. Total track line length was 348 nm with 1,235,251 images collected. Every 10th image was processed for a total of 123,000 images resulting in a total area covered of 0.187 nm² or 0.57% of the NLSCA. The density of scallops along survey tracks ranged between 0 and 23 scallops/m² with dense aggregations occurring at patch scales of about 400 and 900m. The raw values for scallop abundance on a per image basis were interpolated across the closed area using ordinary kriging. Total number of scallops in the NLSCA was 197,545,580. The overall mean for the closed area was 0.187 scallops/m² with a CV of 0.04.

Variance between cells ranged between 0.5, where the sampling density was highest, to over 0.7 where the sampling density was lower. A simple, alternative method to kriging for calculating total scallop counts is to multiply the mean abundance by the area. Our results using this approach is $0.187 \text{ scallops/m}^2 \times 1,142,280,000 \text{ m}^2 = 213,606,360$ scallops with a CV of 0.034, which is similar to the value estimated by kriging. Mean biomass per scallop was 32.9 g. Scallop meat biomass estimated by kriging in the closed area was 6,782 MT.

The HabCam system may be useful for mobile demersal fish, such as yellowtail flounder, in addition to sessile organisms. In the NLSCA, 124 observations of yellowtail flounder were made with the densest concentration in the central region. This region was also characterized by being mostly sand with patches of gravel. The most abundant aggregations of yellowtail were observed in the Southeast Part of CLAI and on the Canadian side with densities exceeding 0.14 fish/m^2 .

Since 2007, joint tows between the NMFS annual dredge survey and HabCam have been designed to compare scallop abundances and size estimates from the standard federal dredge survey with those derived from HabCam. HabCam estimates of scallop abundance were consistently greater than those for the dredge. Mean shell height measurements were similar between dredge and HabCam but the tails of the frequency distribution for HabCam are higher than those for the dredge indicating some inherent measurement error.

In June 2009, in addition to shadowing the Sharp with the Kathy Marie during Legs 1 and 2 of the annual scallop survey, HabCam was towed from the A-frame of the Sharp as part of routine dredge operations on Leg 3. Regression slopes between dredge and HabCam scallop abundances were 0.34 for Georges Bank stations and 0.46 for Mid Atlantic Bight stations. When the data are broken out by substrate regardless of region, the regression slope for sand was 0.35. For sand plus other substrate types such as shell hash the slope was 0.40 and on gravel it was 0.35. These results are a simple measure of the sampling efficiency of the dredge relative to HabCam but are biased low: see Appendix X for an unbiased approach. Results (dredge sampling efficiencies for sea scallops ~ 0.3 to 0.45) are similar to results from other studies. Moreover, they illustrate the potential for use of HabCam in directly estimating the sampling efficiency of other types of survey and fishing gear.

Errors associated with camera calibration and manual measurement of scallop shells on the computer screen were assessed. Camera resolution depends on altitude off the bottom and ranges between $0.37 - 0.89 \text{ mm/pixel}$. Following camera calibration intrinsic pixel error was ± 1.59 pixels resulting in a real-world error of $0.58 - 1.41 \text{ mm}$. Extrinsic errors associated with geometric projection of the image plane on the seafloor, taking into account of vehicle roll, pitch, and changes in altitude, produces real-world errors of $1.11 - 1.78 \text{ mm}$ under optimal water quality conditions in a test tank.

To estimate the level of error associated with manual screen measurement of scallops both within and between a given technician, we assigned four identical 4.2 nm long image transects containing 4,432 images from Western Great South Channel to four technicians. In most cases, the mean shell height within technician measurements were either accurate to the same number of pixels or within one pixel suggesting that within technician variability was extremely low. However, between technician variability was greater with an overall error ranging ± 4 pixels, which represents a real-world error of 3.0 to 7.1 mm . Therefore, measurement errors of scallop shell height are dominated by human extraction of data from the images.

Introduction

There is a great need in fisheries science to develop and utilize new tools and technologies that could help improve the assessment and management of our national marine resources. Coupled with this is a major change in approach from single species to ecosystem based management. The HabCam system was developed to move toward these goals.

HabCam is a seafloor imaging camera system mounted in a ten foot steel frame, and towed at about five knots 1 to 3 m off the ocean floor. HabCam is normally towed behind the F/V Kathy Marie, a New Bedford sea scallop vessel and can operate over the range of the continental shelf, 20 to 250 m depth. The HabCam Group consists of independent researchers, Woods Hole Oceanographic Institution engineers and scientists, and fishermen. HabCam was initially designed and constructed with funding from the Northeast Consortium with major improvements made with funding from the Scallop Research Set Aside Program. The initial goal of the HabCam project was to help improve sea scallop stock assessments by increasing the accuracy of scallop biomass estimates. Additional funding from the NOAA Integrated Ocean Observing Systems (IOOS) Program to support the Northeast Benthic-pelagic Observatory (NEBO) has greatly expanded the range of uses of the HabCam instrument. For example, small study areas have been revisited seasonally providing the baseline of an ecological time series.

Attributes of the HabCam system include: 1) acquisition of optical and acoustic imagery which can be viewed in “real time”, 2) the ability to count and measure scallops and groundfish (i.e. cod, haddock, flounders), 3) measurement of biodiversity and community structure, 4) the means to “map” where there are lost fishing gears (e.g. trawl and gillnets, lobster pots, and other miscellaneous fishing gear and parts), 5) characterization of substrate 6) measurement of oceanic properties (salinity, temperature, nutrients) 7) availability of data and data products online, and 8) relatively inexpensive operating costs. The HabCam system also has the ability to observe animal behaviors including inter- and intraspecies interactions as well as assess and monitor invasive species such as *Didemnum vexillum* and other epibenthic megafauna, and benthic infaunal species.

A historic record of images will be beneficial to understanding patterns, particularly in the implementation of ecosystem management schema. Further, because of direct industry participation, it may help to raise the confidence of the industry in stock assessment methods, monitoring capabilities, and management of our fisheries resources. The Habcam Group has also developed education activities and participated in various outreach projects. The group is currently collaborating with The Ocean Explorium, an education center and aquarium located in New Bedford, Massachusetts and local science teachers and educators.

Methods and Results

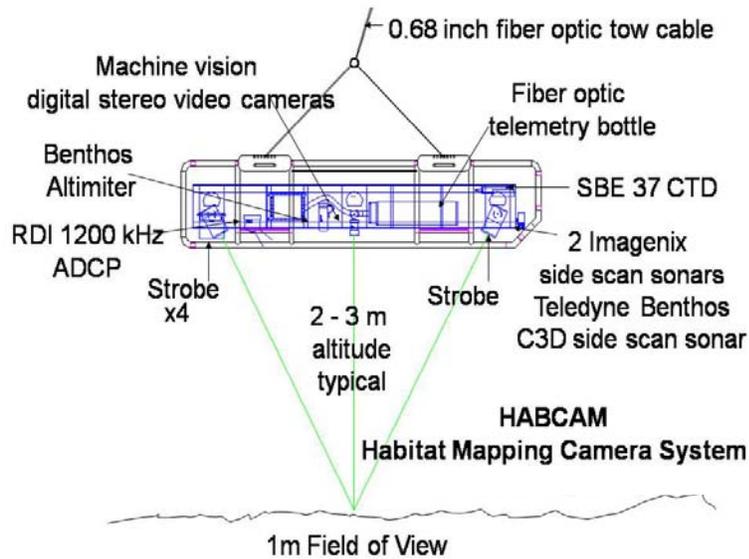
Onboard sensors include a high resolution machine vision GigE color camera, four xenon strobes, side scan sonar, CTD with temperature, salinity, chlorophyll, turbidity, and pH, and a variety of engineering sensors including vehicle roll, pitch, and heading (Fig. 1). All sensors are networked subsea and data transferred via a GigE network to the surface so that data are collected and sent to the ship in real-time where they are recorded, time stamped, and stored.

The HabCam imaging system is “flown” by an operator who controls the winch keeping the vehicle 1.5 to 3 meters off bottom while being towed at 4 to 5 knots (~2.5 m/sec). A track approximately 100 nautical miles is imaged each 24 hour day while at sea. Optical imagery is

collected at a width of approximately 1 to 1.25 meters (total ~200,000 m² /24 hr day). Images (1280x1024 pixels, 16 Bit) are acquired at 5-6 Hz providing a minimum of 50% overlap between images. Images are processed in real-time on the ship by color correcting raw 16 bit tiff images and converting them to 24 bit jpegs (Fig 2). Figure 2 represents a combination of existing data structures and what we envision as fully operational database.

The current NEFSC survey dredge is 8' wide dredge makes approximately 24, 15 min tows at 3.8 kt per day, covering about 4,500 m² per tow and 106,704 m²/day. Continuous operations with HabCam towing at 5 kt and producing 5 images per second with 50% overlap covers 259,200 m²/day. Thus, the spatial coverage of HabCam is nearly 2.5 times the area covered by the survey dredge.

We have implemented two simultaneous and complementary forms of image informatics

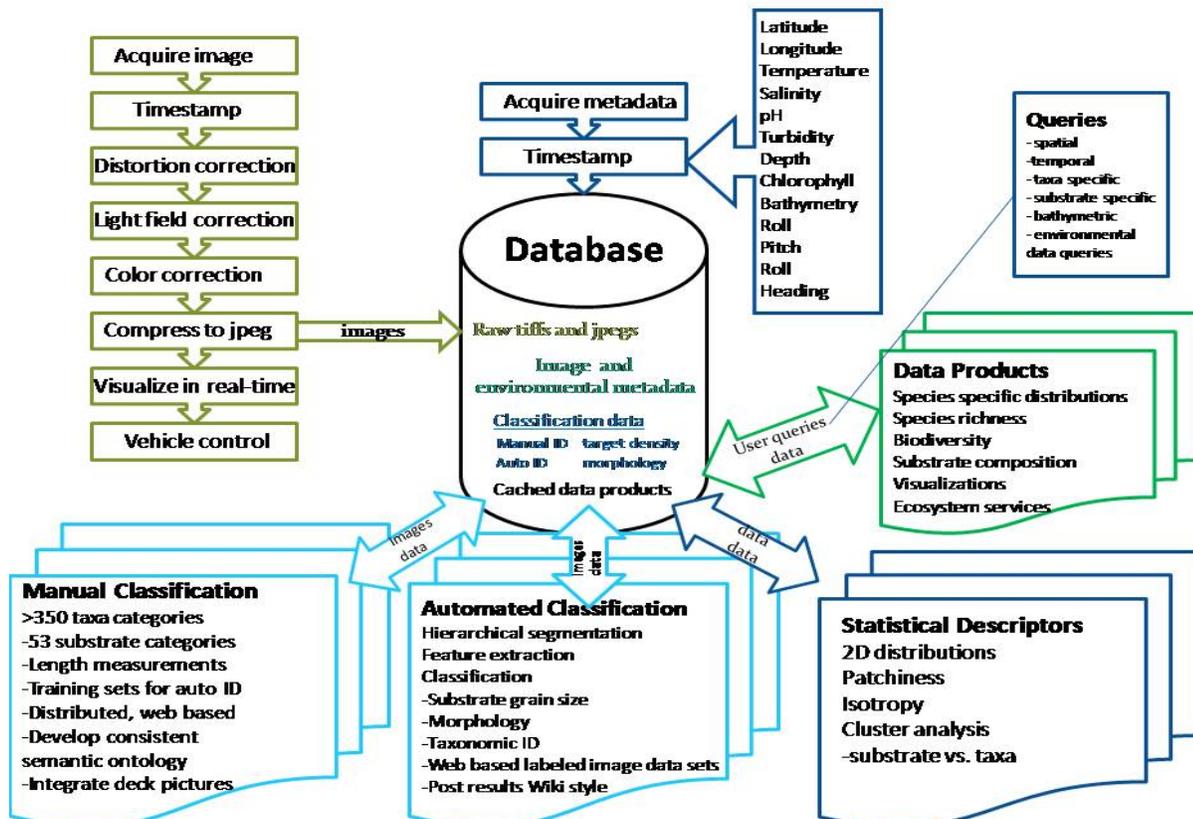


Appendix B9-Figure 1. The HabCam vehicle is towed on 0.68" fiber optic cable ~ 1.5-3 m from the bottom. The camera provides a field of view of 0.5-2 m². Four strobes flash synchronously with each image. Ancillary sensors include side scan acoustics, CTD, chlorophyll fluorometer, and CDOM fluorometer.

(i.e., extracting information from images): manual and automated classification. Manual

classification proceeds by having one or more operators review individual, or sets of, images to identify and measure target species using a GUI with point and click functionality. This allows about 60 to 200 images per hour per operator to be processed depending on image complexity and number of individual species being identified.

More than 460 taxa or taxonomic groups ranging in size between ~1 mm to 2 m have been observed and identified with HabCam. While taxonomic definitions used in image analysis are based on epibenthic organisms, a variety of infauna can typically be observed and quantified such as bivalve siphons, turbularian worms, burrowing shrimp, and some vertebrates (e.g., tilefish and their burrows). During manual operations, the operator also evaluates the substrate type in each image and categorizes it into one of 43 groups ranging from silt, sand, gravel, shell, cobble, boulder, and a variety of combinations. Development of approaches for automated



Appendix B9-Figure 2. Future iterations of the HabCam data workflow environment. Images and associated metadata enter the processing path from the left. Following preliminary image processing steps conducted in real-time (e.g., color correction), images are viewed and classified by scientists onboard. Results are entered into the database and used as training sets for automated classifiers. The database may be queried both spatially and temporally to build a set of data products shown on the right.

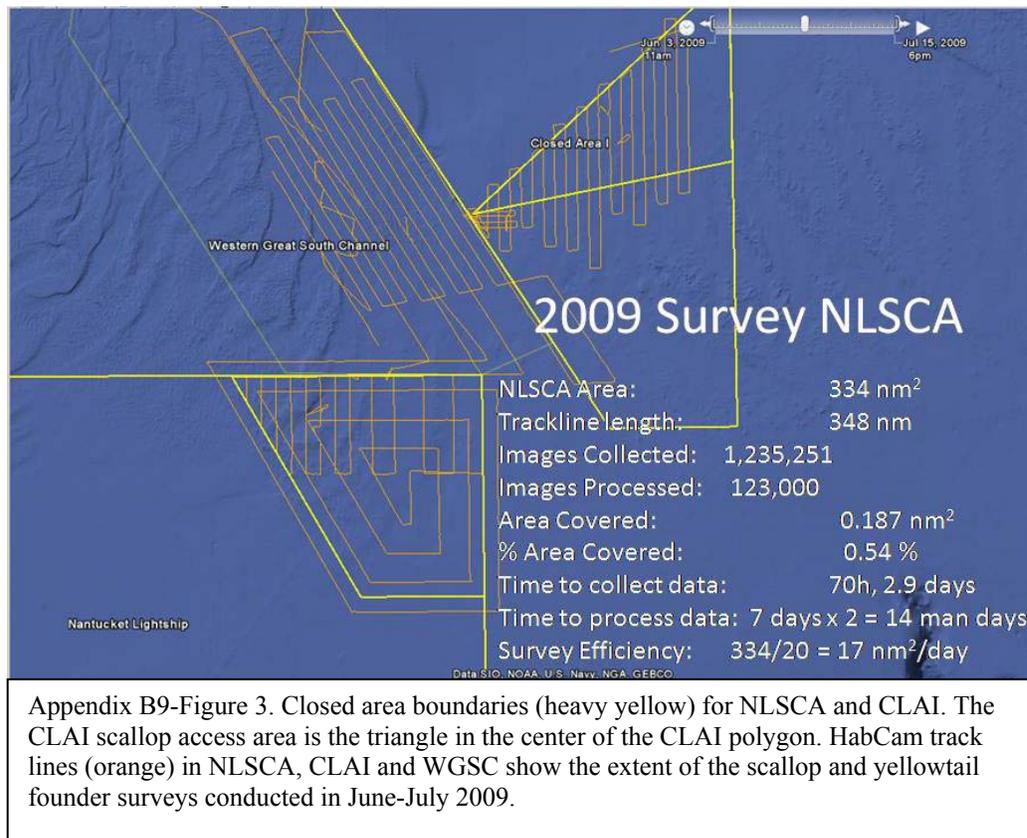
classification of targets and substrate is proceeding using the manually classified images as training sets. This is an area of ongoing research.

For the purpose of the scallop surveys reported here, images were classified manually by several technicians, who characterized substrate into the categories noted above and measured all scallops and groundfish, including yellowtail flounder. To speed the process, every 10th image was analyzed. Rationale for this strategy is discussed in the following sections.

2009 Nantucket Lightship Closed Area (NLSCA)

The objectives were to estimate scallop abundance, biomass, and shell height composition in the NLSCA using the data collected. A secondary objective was to estimate the distribution and abundance of yellowtail flounder in relation to substrate. The area of NLSCA is 334 nm². Total track line length was 348 nm with 1,235,251 images collected. A total of 123,000 images were processed for a total area covered of 0.187 nm². Total area sampled by HabCam was 0.57% of the NLSCA. The survey required 72 hours of continuous towing and approximately 27 person days to process the data, assuming 12 hour shifts.

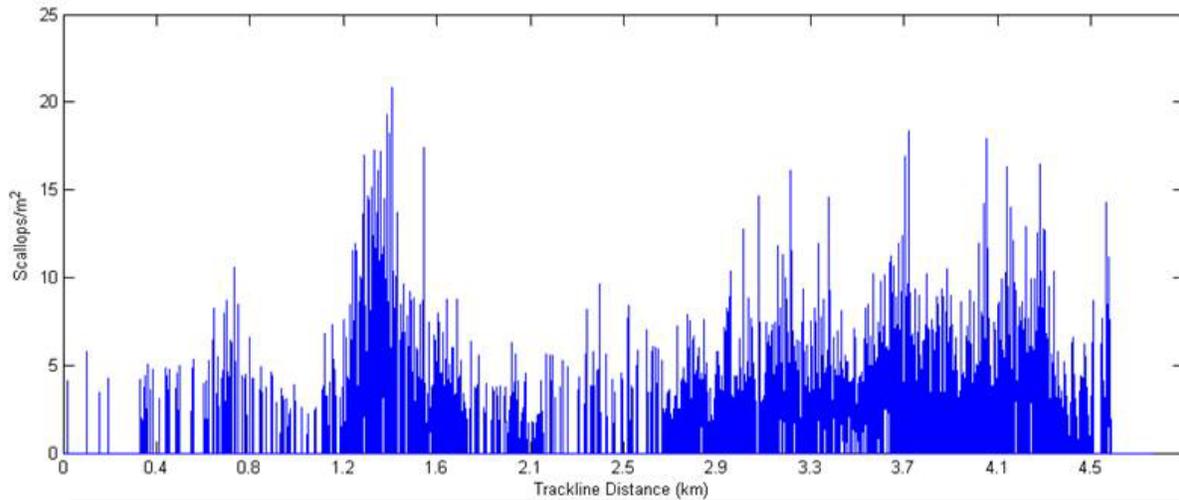
The survey track line was a modified spiral which started 1.3 nm outside the boundary of the closed area to allow interpolation of the final abundances without boundary influences (Fig. 3). The spiral was conducted around the border then continued at an interval of 2.6 nm.



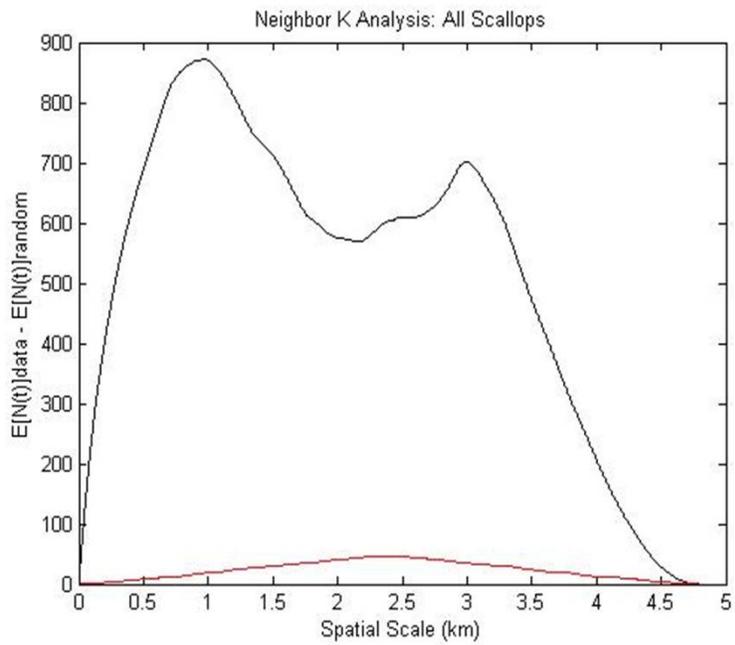
Following completion of the spiral, the vessel steamed to the northeastern corner and began a finer grid extending from the northern edge to the center of the closed area. This finer grid pattern began at an interval of 2.6 nm but compressed to 1.3 nm as the vessel approached the north-central section of the closed area. This was to provide higher resolution where prior knowledge indicated dense scallop abundance.

To assess multi-scale patchiness and to determine how many images should be processed, a preliminary transect from east to west 2 nm in length was processed by manually counting and measuring every other image. This provided information for calculating the appropriate image subsampling rate for processing the remainder of the spiral and for calculating a patchiness index for use in setting appropriate interpolation scales. Images were subsampled for the sake of efficient and fast data processing.

The density of scallops along survey tracks ranged between 0 and 23 scallops/m² (Fig. 4). It appeared that aggregations of scallops at densities between 5 and 20 /m² occurred in clumps at a spatial scale of 400m. Therefore, a Nearest Neighbor-k analysis was performed to establish the dominant spatial scales of patchiness. The Neighbor-k showed strong patchiness ranging from 400 to 900m and again at 3000m (Fig. 5).



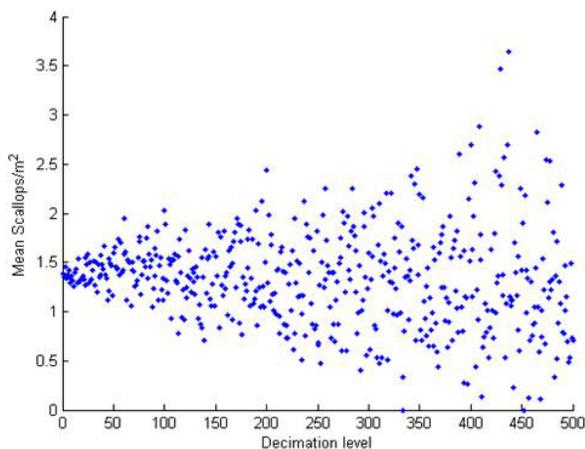
Appendix B9-Figure 4. Scallop abundance in the northern section of the NLSCA survey along a 4.5 km (2 nm) track. Every other image was classified manually to establish a baseline for subsampling and to estimate patchiness. Note the very patchy distribution ranging from 0 to >20 scallops per m².



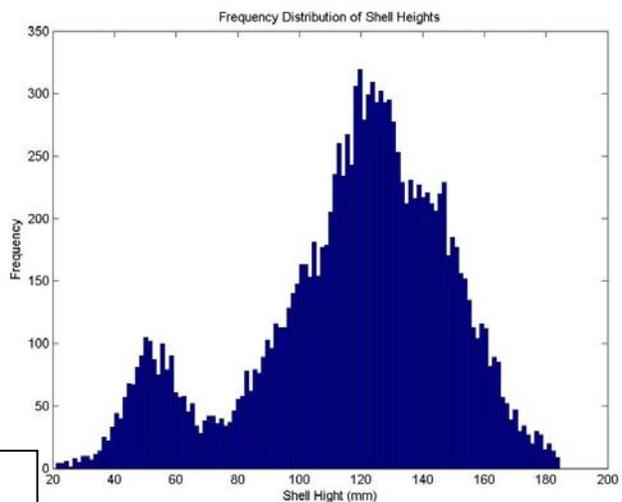
Appendix B9-Figure 5. Ripley's Neighbor-k analysis for one-dimensional patchiness was applied to scallop densities measured along the track (every other image was processed for this analysis). Spatial scale is on the x axis while the residual between observed and predicted nearest neighbor distance under randomness (1000 Monte Carlo simulations) is on the y axis. The first mode indicating a characteristic patch size is located at 700-900 m, and a second at 3.2 km. Both modes are well above the red line under randomness indicating these patch dimensions to be statistically significant at the 95% confidence level.

To determine an optimal image subsampling rate, data from every other image along the track was processed by extracting abundances and calculating the CV for density at sampling intervals ranging from every 4th, 8th, 10th, 12th etc. out to every 500th image. The mean and CV remained stable up to a subsample level of every 10th image. Therefore, the remaining spiral was processed at a rate of every 10th image (Fig. 6).

Over 12,900 scallop shell heights were counted and measured using MIP (Manual Identification Program developed by A.D. York), which allows users to quickly point and click on scallops to extract measurements and select substrate type from a menu. The shell height distribution was strongly bimodal with modes at 53 and 125 mm (Fig. 7). A third, less prominent mode was located at 142 mm.



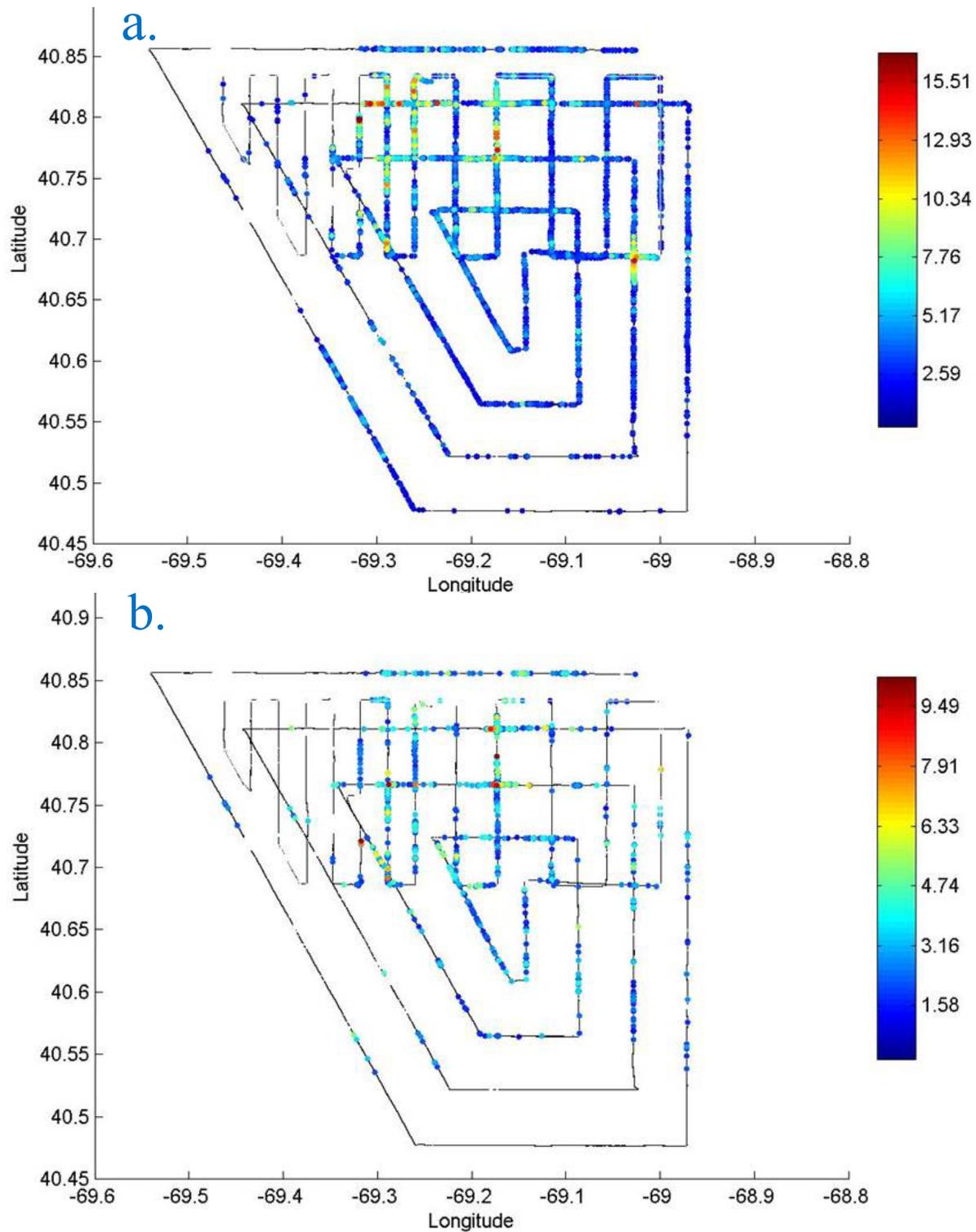
Appendix B9-Figure 6. The effect of subsampling (decimation) the continuous record of scallop abundance along the track line in Figure 3. Images were analyzed at subsampling rates of every 4th, 8th, 10th, 12th... out to 500 and the scallop abundance recalculated. Mean abundance is stable out to a subsampling rate of greater than 20 so a conservative level of processing every 10 image was chosen for the remainder of the analysis.



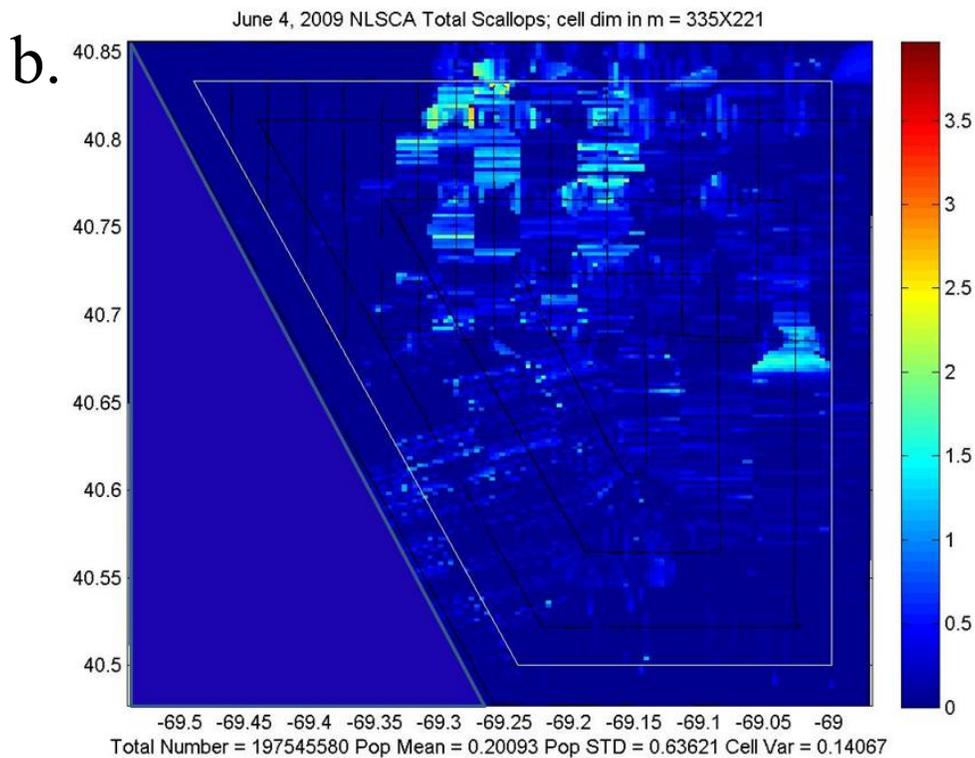
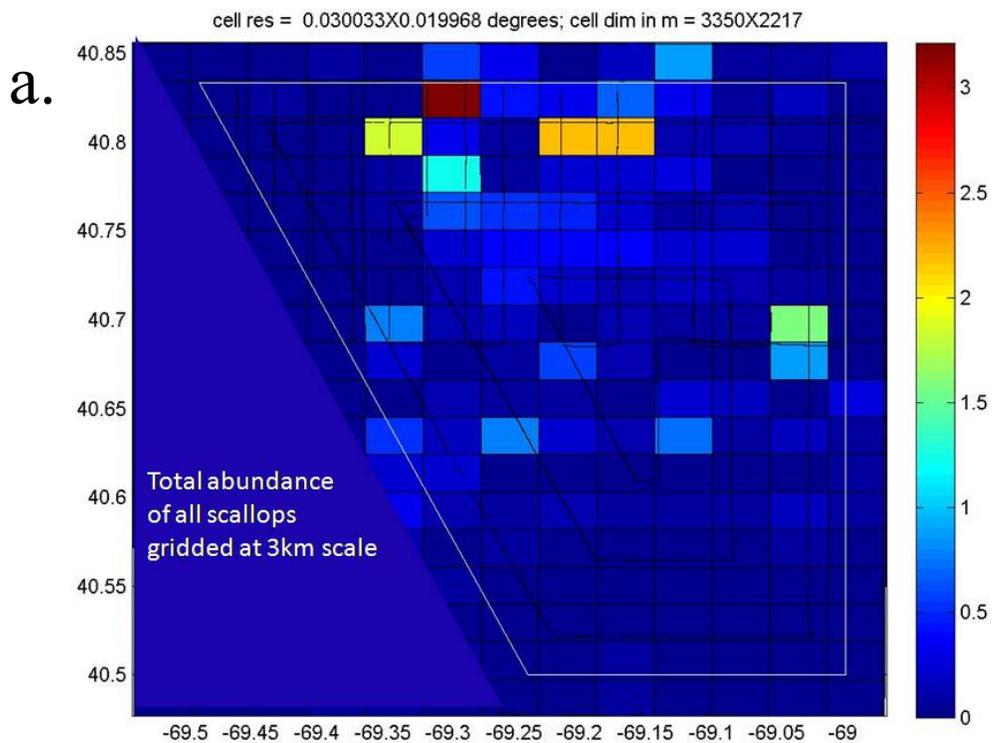
Appendix B9-Figure 7. Shell height frequency distribution for all scallops measured from the NLSCA survey. N=129,237. The shell height distribution was strongly bimodal with modes at 53 and 125 mm with a third, less prominent mode located at 142 mm.

The scallop density from every 10th image plotted as color coded dots for all scallops showed highest aggregations in the central upper third and in the central eastern region of the closed area (Fig. 8a). Scallops were sparse in the northwestern corner and southern regions. A similar plot for just those scallops with shell height between 20 and 65mm showed that small scallops were most abundant in the east central region of the closed area (Fig. 8b).

The raw scallop density per image was interpolated into rectangular grids at two scales using ordinary kriging based on cells of 3350x2217m and 335x221m. First, a semi-variogram was constructed to evaluate autocorrelation of the data. For the coarse and fine scales, the mean for each grid cell was color coded (Fig. 9a, b). Total abundance, mean, and variance for each grid cell were calculated. An overall CV was calculated by bootstrapping the standard error divided by the mean for each grid cell. Data were collected and kriged beyond the location of the closed area boundary, but results presented are only for the area within the NLSCA boundaries. The highest mean value was 4 and 6 scallops/m² for the coarse and fine scale grids, respectively.



Appendix B9-Figure 8. Raw abundance ($\#/m^2$) estimates on a per image basis for all scallops regardless of shell height. Each dot represents a single image with the abundance indicated by color. Where no dots exist, no scallops were observed. a) All Scallops ($\#/m^2$). b) Scallops with shell height less than 60mm.



Appendix B9-Figure 9. a) Kriged scallop densities at a scale of 3350x2217 m . Mean densities represented by color referenced to the color bar on the right. b) Kriged scallop densities at a scale of 335x221 m showing considerable patchiness at this fine

Total abundance within the closed area boundary was calculated by summing each of the grid cells that fell within the boundary. Those cells that were partially within the boundary were evaluated by including only the proportion of the cell falling within the boundary. Total scallops in the coarse and fine scale grids were 174,966,666 and 197,545,580 scallops, respectively. The discrepancy in total abundance between grid scales probably lies in the fact that the scallops were patchy at scales of 400-900 m as shown from the Nearest Neighbor-k analysis. The courser grid scale smoothes high density patch values over a larger area than the higher resolution grid. The finer grid, therefore, is providing a more representative view of the scallop distribution and also the most accurate estimate. The overall mean for the fine scale grid was 0.187 scallops/m² with a CV of 0.04. Variance between cells ranged from 0.5 where sampling density was greatest to over 0.7 where sample density was low (Fig. 10).



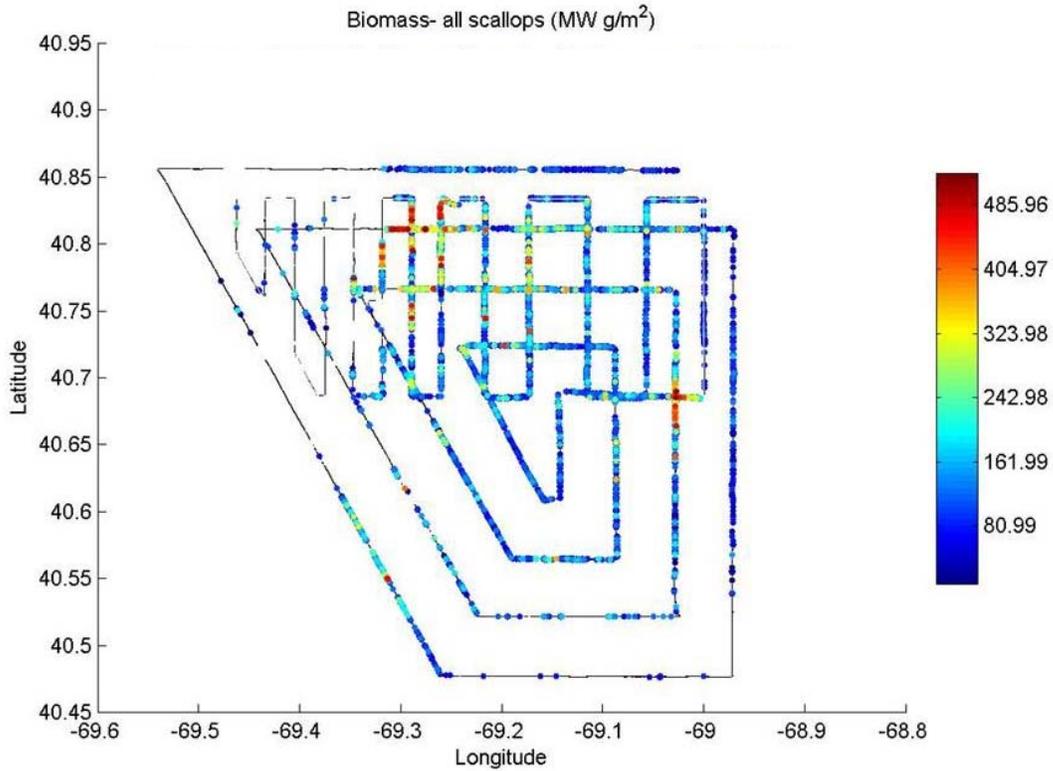
Appendix B9-Figure 10. Variance per cell for kriged scallop densities at the fine scale.

The weight of individual scallops used to estimate biomass was calculated using a shell height-meat weight relationship that included depth:

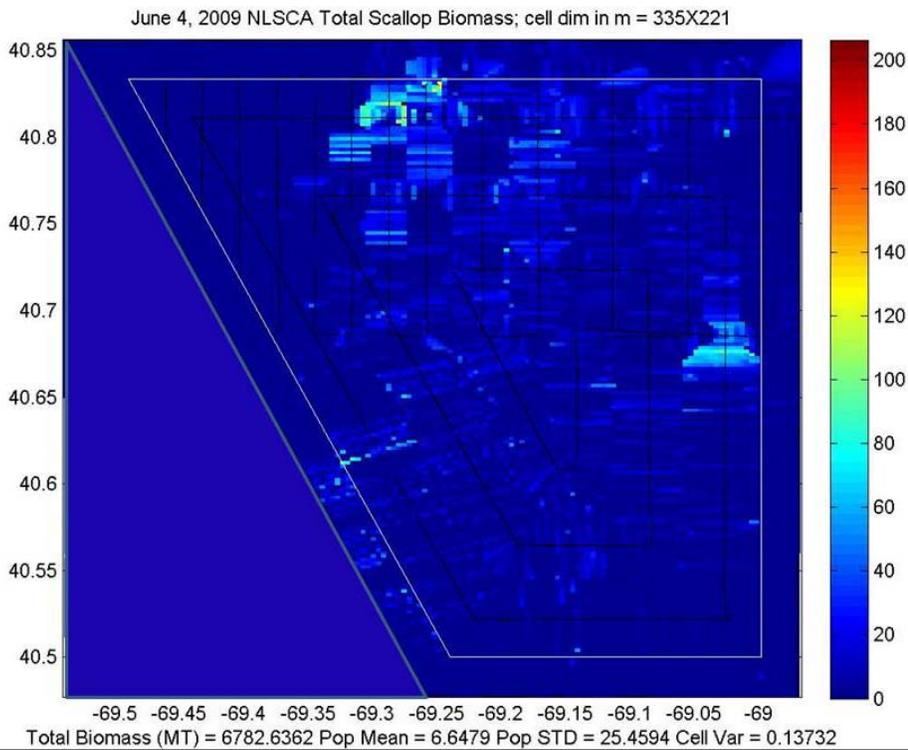
$$W = \exp(a+b \log(\text{SH}) + c \log(\text{depth}))$$

where W is weight (g), SH is shell height (mm), depth is in meters and the parameters a = -8.62, b = 2.95, and c = -0.51 (D. Hart, NEFSC, pers. comm.).

Mean biomass per scallop was 32.9 g. The basic pattern of distribution followed that of scallop abundance with dense scallop areas in the central northern region and in the central eastern region (Fig. 11). Biomass estimated by kriging at the fine scale 6,782 MT meats (Fig. 12).

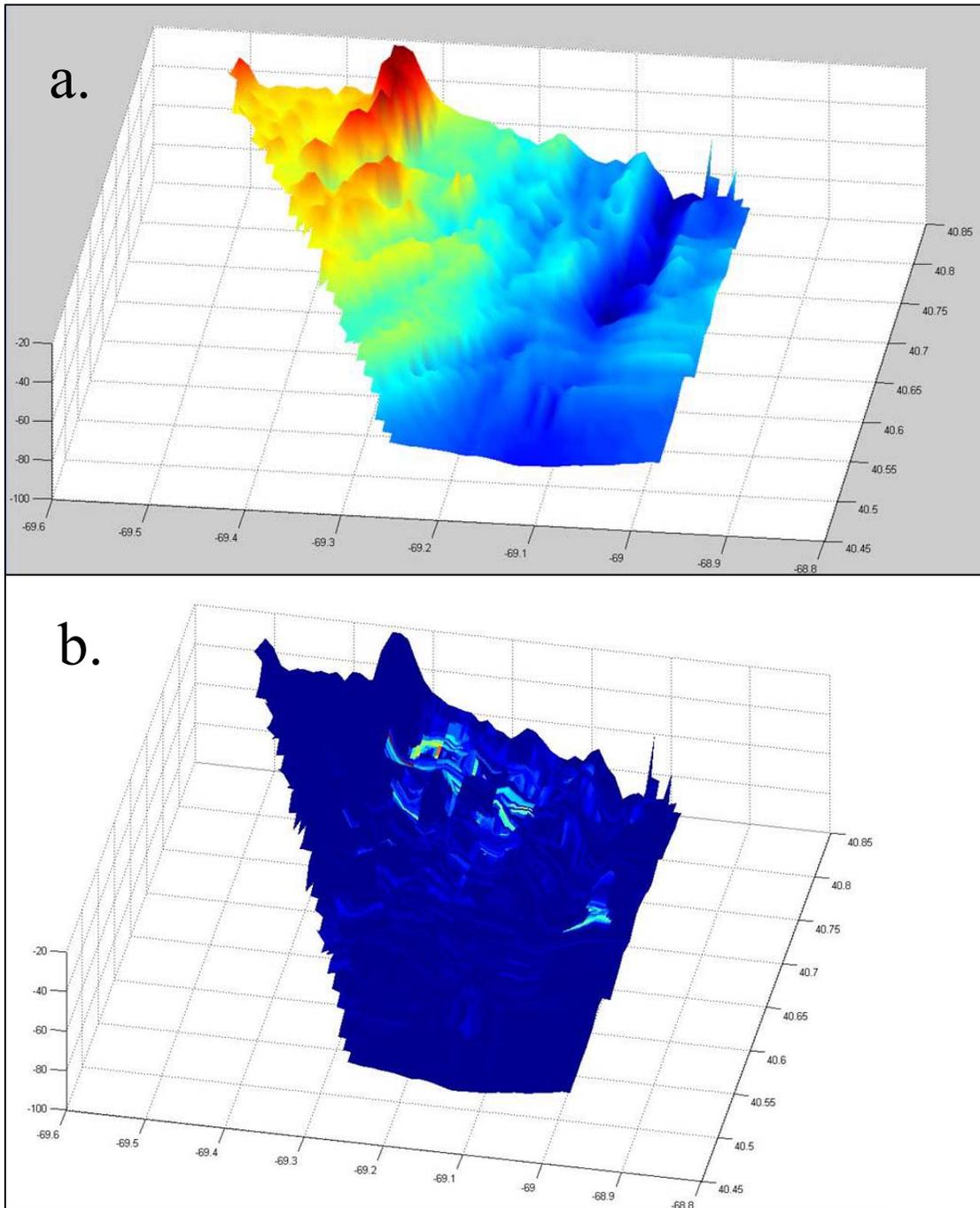


Appendix B9-Figure 11. Scallop biomass densities along the track line for all scallops of all sizes.

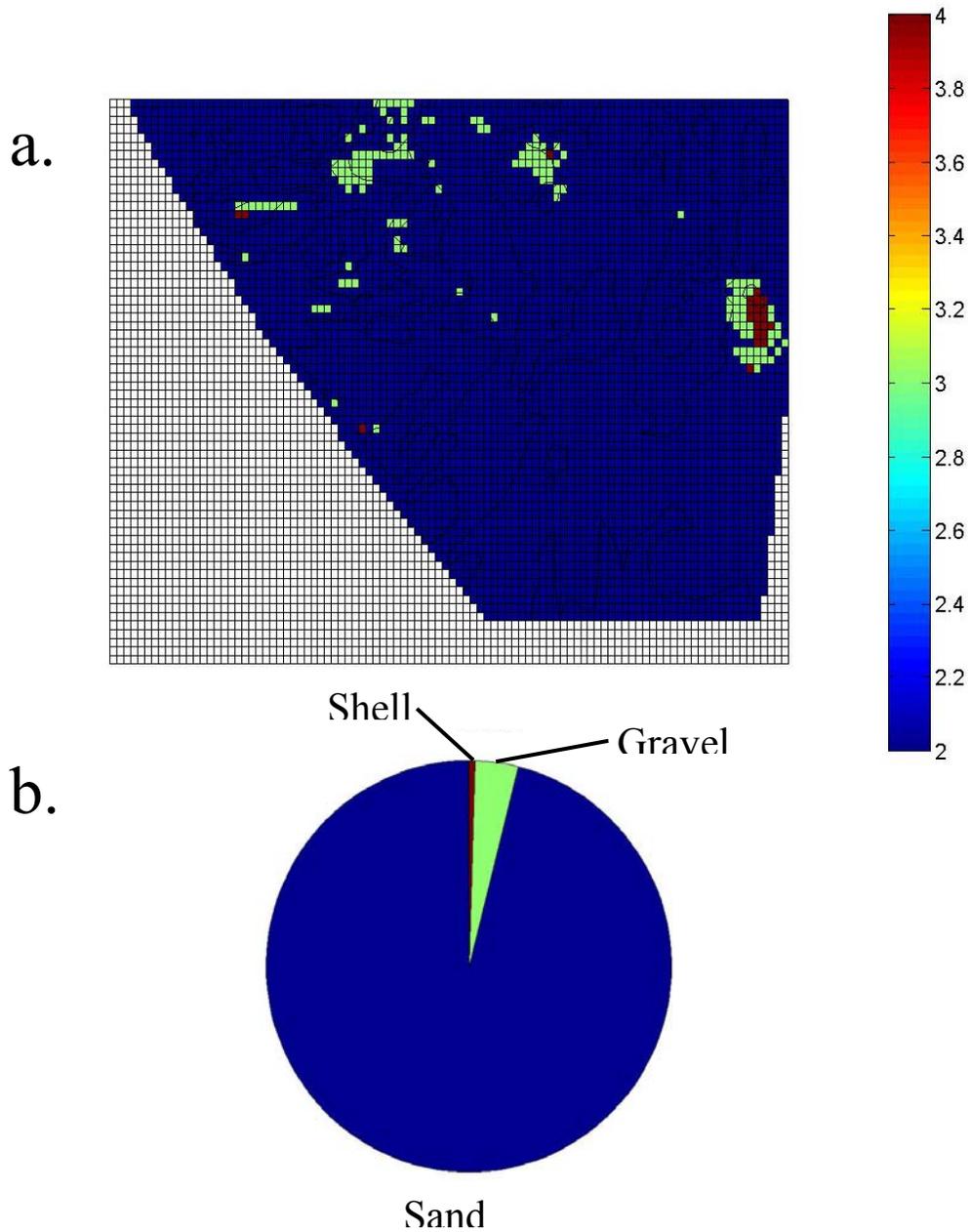


Appendix B9-Figure 12. Kriged biomass estimates generated at the fine scale for scallops of all sizes.

To examine relationships between density of scallops, depth, and substrate type, the depth from the ship's echosounder was linearly interpolated onto a uniform grid and colored as a function of depth (Fig. 13a). Sand waves in the northern central region and a trough in the central eastern region are notable. When scallop density was plotted as a color map over the interpolated depth data, it was clear that greatest densities were at the eastern base of the sand waves and just eastward of the trough, but not in the trough (Fig. 13b).

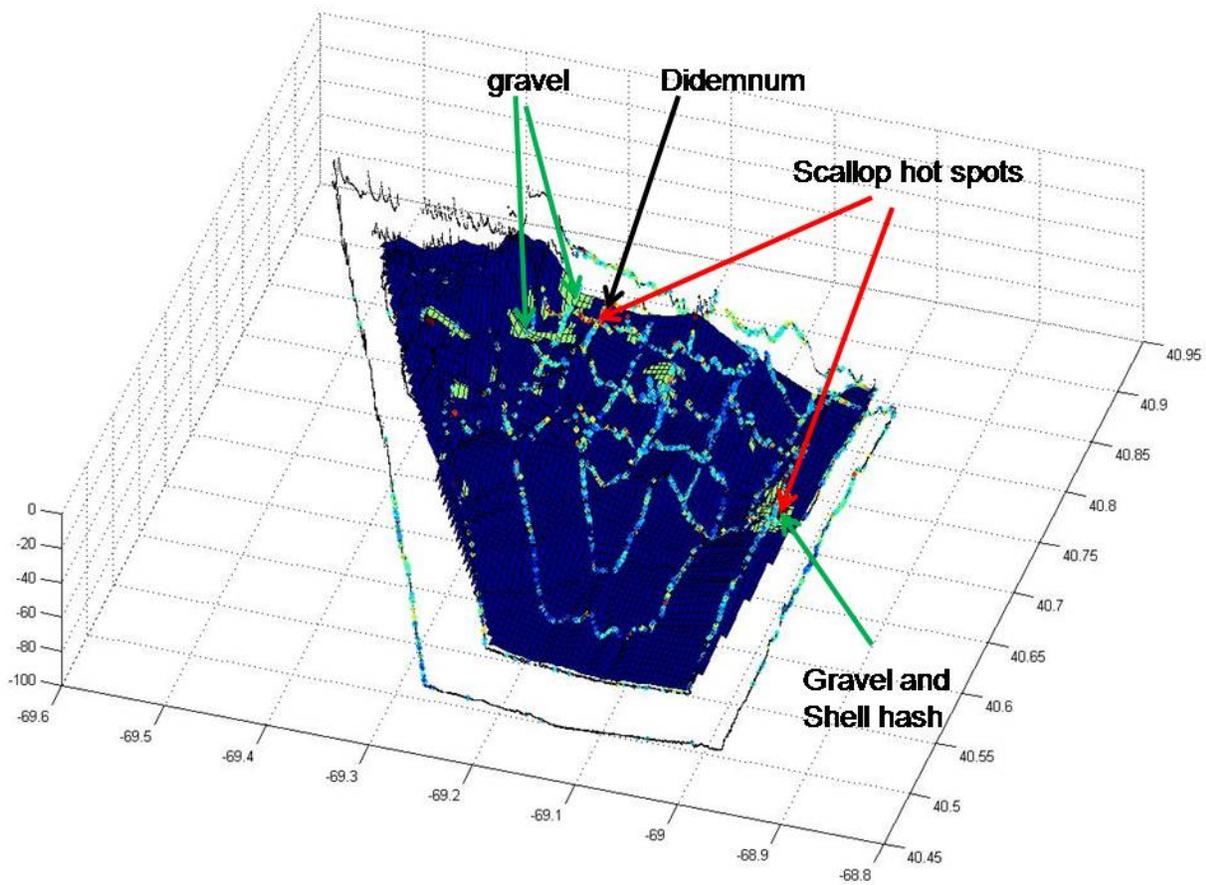


Appendix B9-Figure 13. a) Depth from the ship's sonar interpolated to a uniform grid.
 b) Depth with overlaid scallop abundance on the same color scale as in Fig. 12.
 Z axis is exaggerated for visualization purposes.

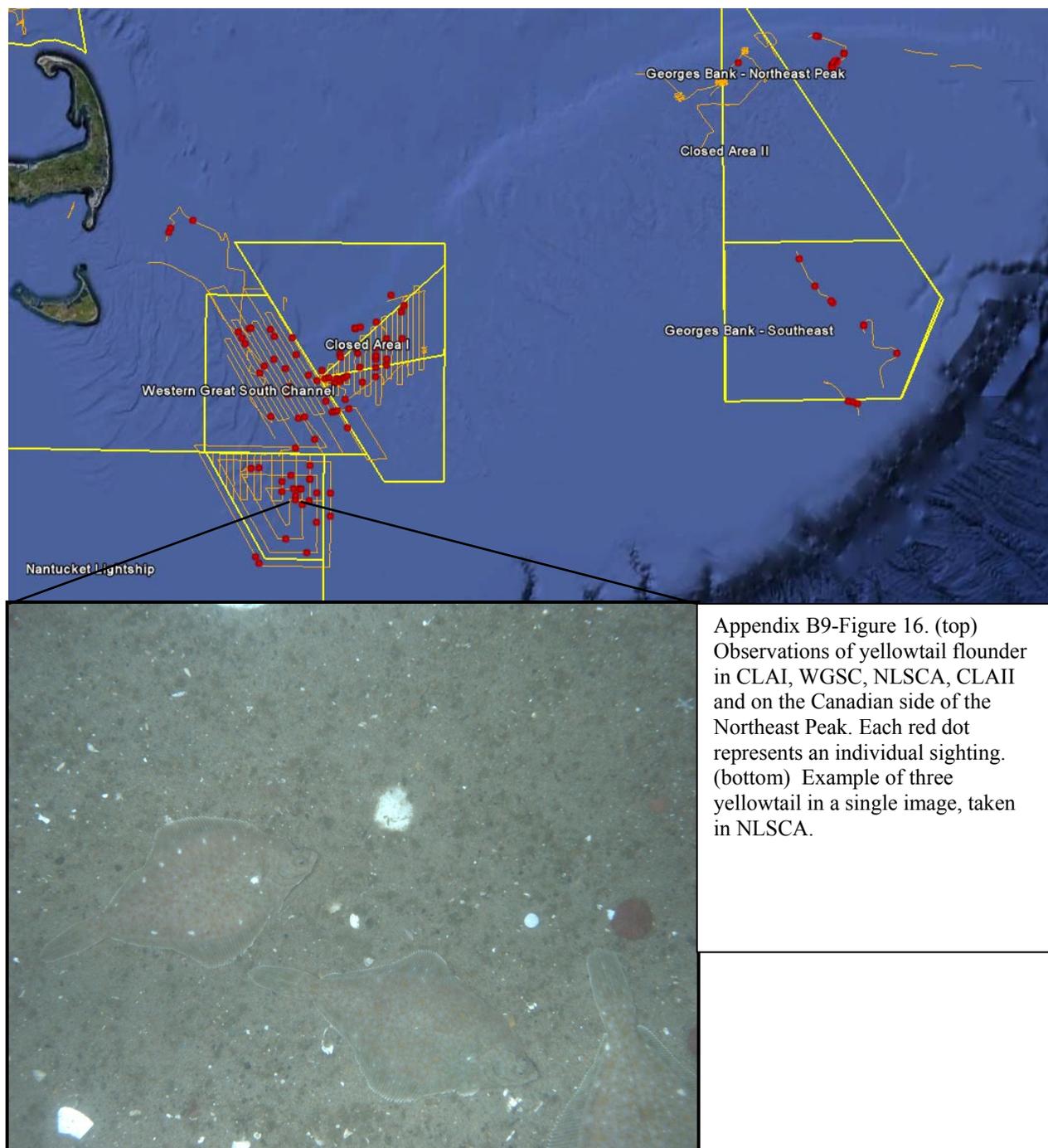


Appendix B9-Figure 14. a) Dominant substrate binned numerically into three categories (2) sand, (3) gravel, (4) shell. b) Sand dominated pie chart of substrate in NLSCA.

Substrate classifications were re-categorized into three numeric bins of dominant substrate: (2) sand, (3) gravel, and (4) shell. Dominant substrate categories include mixed substrate types, for example, “gravel” contains mixed substrate images such as gravel/sand and gravel/shell. Interpolation of these substrate categories across the NLSCA grid showed the entire area to be mostly sand (Fig. 14a,b). The greatest accumulation of sand/shell hash corresponded to areas of high scallop densities. The region to the central eastern side of the trough had notable sections of gravel, which is also where scallops were most abundant, particularly scallops less than 60mm in height. The combination of all three variables, scallop density, depth, and substrate (Fig. 15) provides a visualization of how scallop distribution is affected by these variables. HabCam data for the invasive tunicate *Didemnum vexillum* collected simultaneously with sea scallop data during the survey, illustrate spatial relationships of two species and substrate type and demonstrate the potential for use of data in ecological studies (Fig. 15).



Appendix B9-Figure 15. Depth overlaid with scallop abundance and substrate. Note location of invasive tunicate *Didemnum vexillum* in relation to high scallop densities.



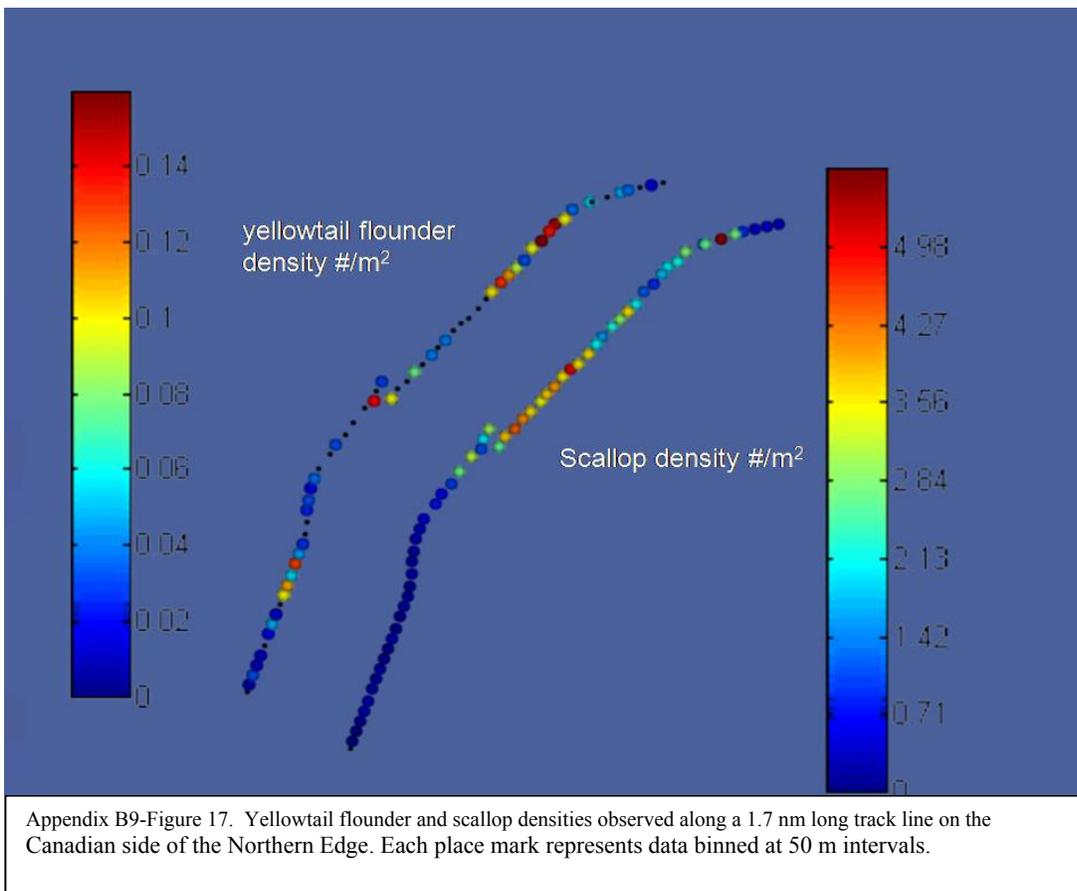
Appendix B9-Figure 16. (top) Observations of yellowtail flounder in CLAI, WGSC, NLSCA, CLAI and on the Canadian side of the Northeast Peak. Each red dot represents an individual sighting. (bottom) Example of three yellowtail in a single image, taken in NLSCA.

As an alternative method for calculating total population abundance of scallops without kriging or interpolating, one may simply use the overall mean observed in images multiplied by the total area. Our results using this approach is $0.187 \text{ scallops/m}^2 \times 1,142,280,000 \text{ m}^2 = 213,606,360$ scallops with a CV of 0.034.

In addition to sessile organisms, the HabCam system may be useful for imaging mobile demersal fishes. Yellowtail flounder were observed in NLSCA and other regions during our survey at relatively low densities (Fig. 16). In NLSCA, 124 observations were made with the densest concentration in the central region. This region was also characterized by being mostly

sand with patches of gravel. The most abundant aggregations of yellowtail were observed in the Southeast Part of CLAI and on the Canadian side with densities exceeding 0.14 fish/m². In some cases two or three fish were observed in a single image. Images from CLAI show yellowtail to be found on mostly sandy bottom with shell hash and occasionally on gravel.

An interesting relationship between scallop and yellowtail density was observed on the Canadian side of the Northern Edge. Data for yellowtail and scallops for the same track line are plotted alongside each other in Figure 17. Note that yellowtail appeared to be at highest densities where the abundance of scallops were low. This seemingly inverse relationship only holds for this track line at that point in time and is probably related more to substrate, food supply, reproduction, or environmental variables, than a true relationship between scallops and yellowtail. These results indicate the potential of HabCam data for use in fisheries management where, for example, the goal is to reduce bycatch of yellowtail during scallop fishing.



Joint ship operations

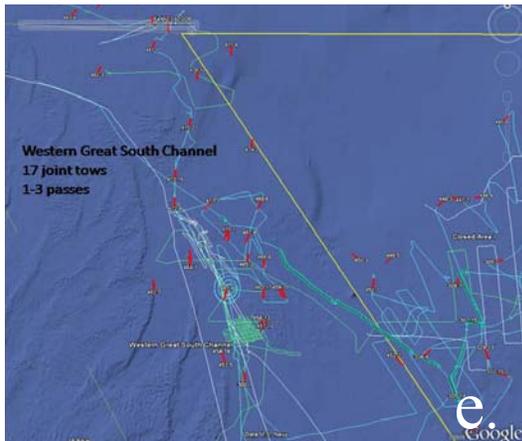
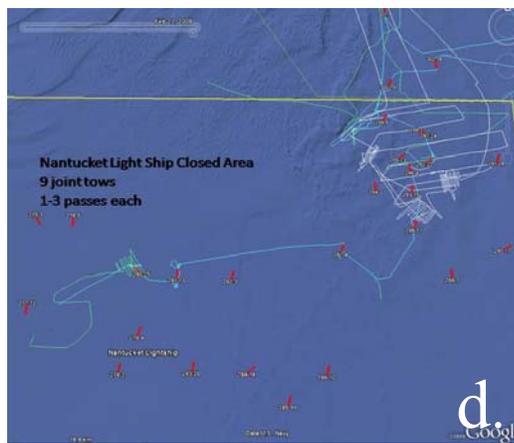
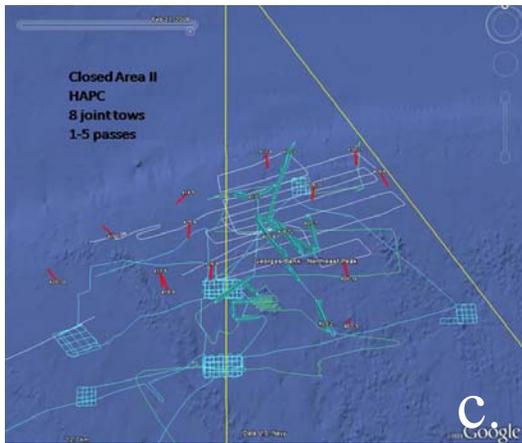
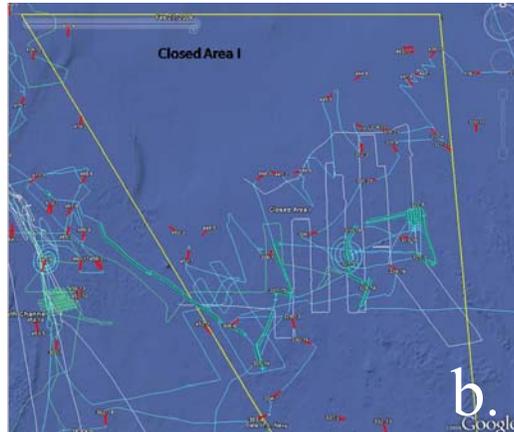
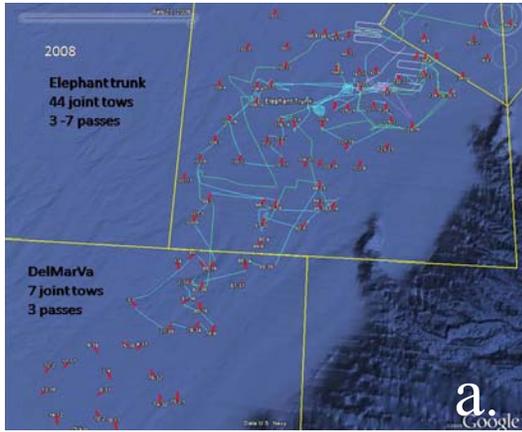
Since 2007, The HabCam Group has been collaborating with the NMFS in their annual scallop surveys by conducting paired tow experiments. These joint tows were designed to compare scallop abundances and size estimates from the standard federal dredge survey with those derived from HabCam imagery. Data will be presented here for 2008 and 2009.

In June and July 2008, The F/V Kathy Marie ‘shadowed’ the R/V Sharp on 113 total tows with 44 in the Elephant Trunk, 35 in CLAI, 8 in CLAI HAPC, 9 in NLSCA, and 17 in the proposed WGSC HAPC (Fig. 18). HabCam made at least three passes at over 50% of the NMFS stations, and in a few cases made up to seven. These multiple passes were designed to assess the variability of scallop density along each track and between multiple passes. Images from all passes were processed at a subsampled rate of every 10th image. This translates into processing about 1 m² for every 5 m of track line.

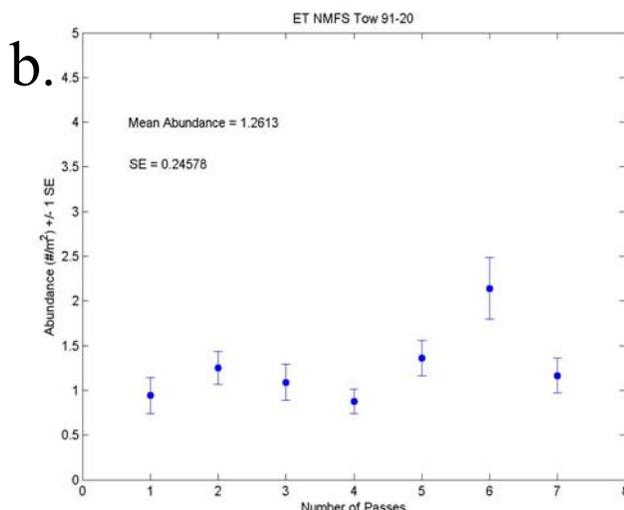
Within hours of conducting a dredge tow, the beginning and end points for the tow were communicated at sea via radio from the Sharp to the Kathy Marie. This allowed the captain of the Kathy Marie to line the vessel up on the dredge tow and follow a straight line from one end to the other. Because the absolute position of neither the dredge nor the HabCam vehicle was precisely known, our best efforts were to make multiple passes that coincided within about 50m of the dredge tow line and between each pass of HabCam. As an example, data for seven passes along the dredge tow for one station in Elephant Trunk (91) shows within and between variability of scallop densities observed by HabCam (Fig. 19). Although Pass 6 appears to be an outlier, results of a one-way ANOVA suggest that there is no significant difference between all 7 passes ($p < 0.001$).

HabCam estimates of scallop abundance were consistently greater than dredge counts, indicating that dredge efficiency is well less than 1. Mean shell height measurements were similar between dredge and HabCam (Fig. 20), but as will be discussed in the section under error analysis, the tails of the frequency distribution for HabCam are higher than those for the dredge indicating some inherent error in the measurement of shell heights. Count data tend to be accurate in optical surveys but some degree of body size measurement error is typical (Jacobson et al., 2010). This is an area of ongoing research.

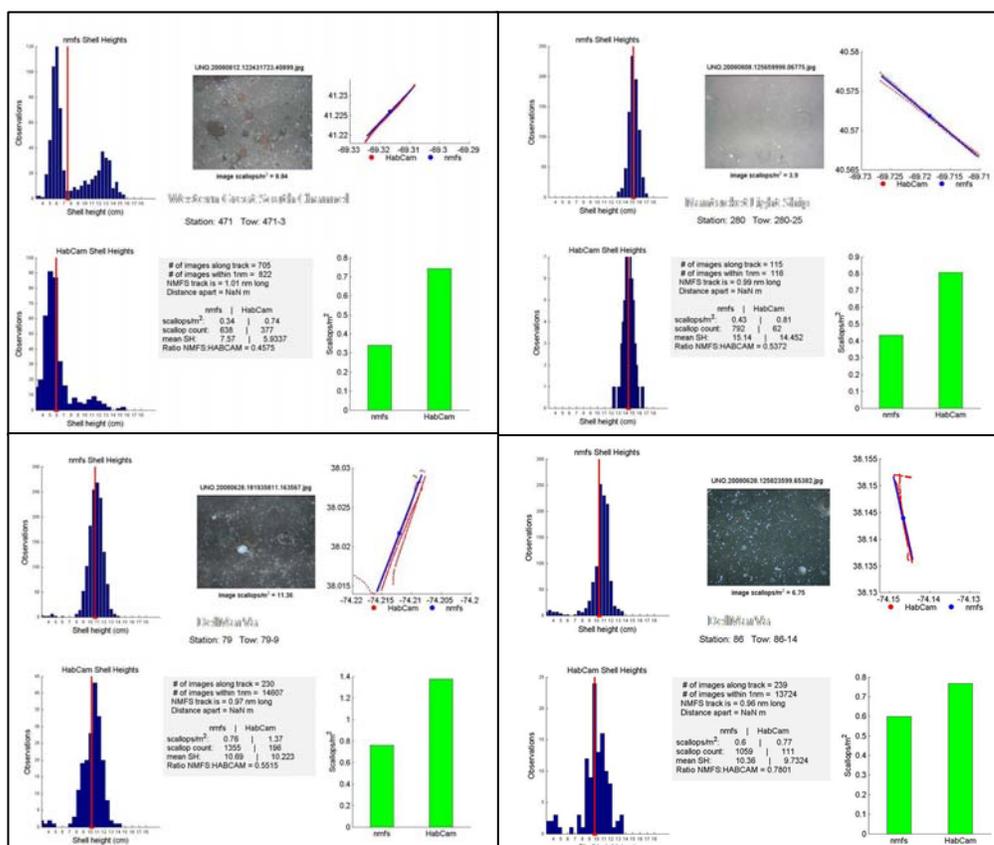
In June 2009, in addition to shadowing the Sharp with the Kathy Marie during Legs 1 and 2 of the annual scallop survey, HabCam was towed from the A-frame of the R/V Hugh R Sharp as part of routine dredge operations on Leg 3. This project was designed for comparison of HabCam data for sea scallops and yellowtail flounder with data from the standard dredge tows during Leg 3 of the 2009 NMFS Scallop Survey. Because of sea state and time considerations, HabCam was towed at and between 23 stations. HabCam collected a total of 787,832 images with a footprint of about 1 m² each. By area, 85,572 images were collected in CLAI, 216,809 images in CLAI, 183,070 images on the Canadian side of the Northern Edge of Georges Bank, and 302,381 images between stations. A final report has been filed with the NOAA CINAR office and Russell Brown at the NEFSC (HabCam Group, 2010).



Appendix B9-Figure 18. 44 Joint tows between R/V Sharp and HabCam on the F/V Kathy Marie in the (a) Elephant Trunk, (b) CLAI, (c) CLAI HAPC, (d) NLSCA, and (e) WGSC. Red lines are dredge tows, blue lines are HabCam track lines.



Appendix B9-Figure 19. (a) Federal dredge station 91 in ET with 1 nm tow shown in red. Seven passes by HabCam shown in blue. Multiple passes of HabCam were within 50m of each other. (b) Mean +/- SE of scallop abundance ($\#/m^2$) from each of the seven passes at station 91.



Appendix B9-Figure 20. Data from four joint stations illustrating the relationship between dredge and HabCam data. In each of the boxes shell height frequency distributions, mean abundance, and position along track for the dredge and HabCam are compared.

As each image represents about 1 m² and there is approximately 50% overlap, an area of about 242,000 m² was imaged. In an area on the Canadian side called the ‘seed box’, the density of small (50-60 mm) scallops was extremely high, upwards of 50 to 90 scallops per image (e.g. Fig. 21).

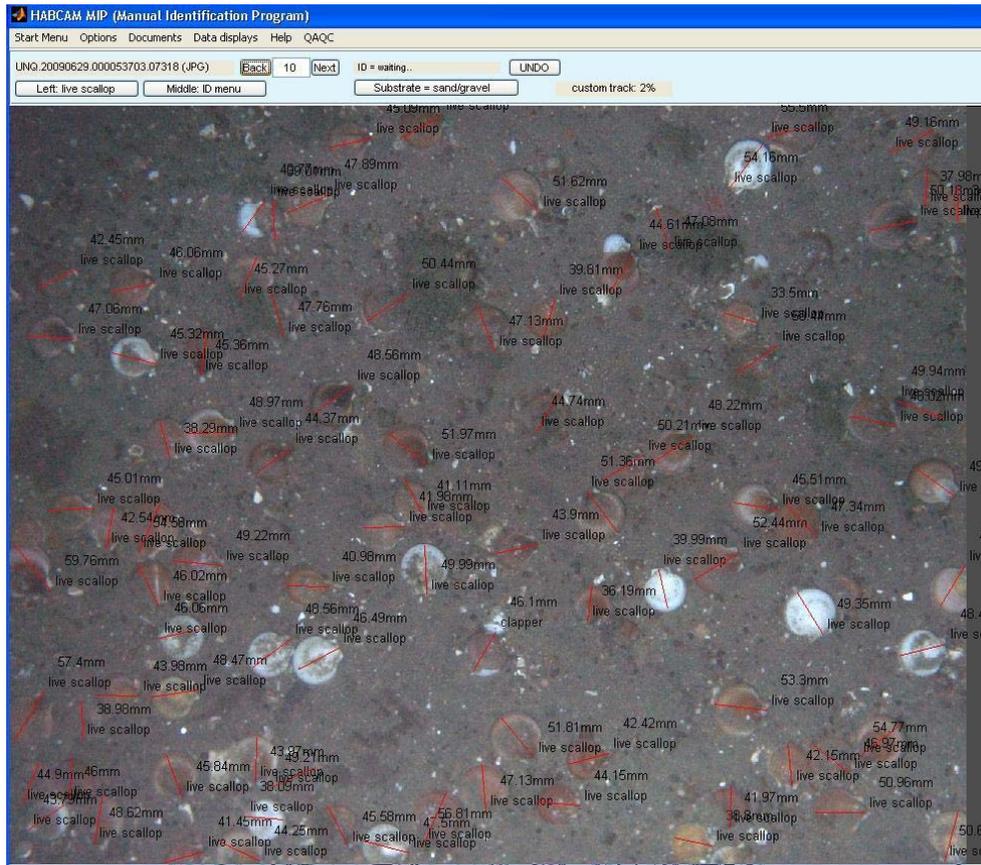
Shell height measurements from HabCam showed a strongly skewed distribution to the left with a mode of 55 mm and a mean of 79 mm (Fig. 22), indicating that this area was dominated by two year old scallops with relatively few older individuals.

Along track abundance of scallops at Station 404 ranged from 0 to well over 60/m² (Fig. 23). A Neighbor-k analysis of scallop distributions along the track in Fig. 4 showed that patchiness was significant at several spatial scales from 600 to 1000m (Fig. 24).

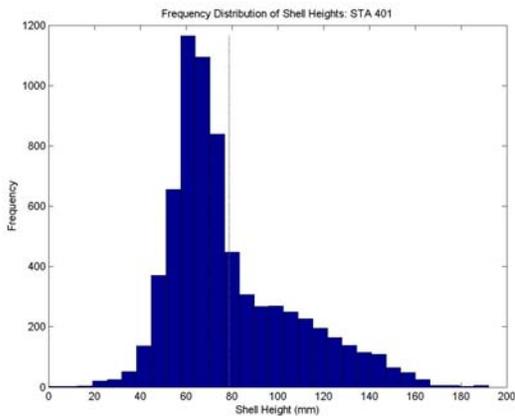
Yellowtail flounder were sparse but most abundant in the Southeast Part of CLAI and on the Canadian side of the Northern Edge (Fig. 25). Images show yellowtail to be found on mostly sandy bottom with shell hash and occasionally on gravel (Fig. 26).

Survey dredge capture efficiency is low relative to optical surveys and might be variable due to tow direction in relation to tidal currents, substrate composition, wire out, tow speed, and tow duration. To compare scallop abundances estimated by the NMFS dredge and HabCam, plots were generated by region and by substrate and include data for both 2008 and 2009 (Fig. 27). Georges Bank includes NLSCA, WGSC, CLAI and CLAI. Mid Atlantic Bight includes Elephant Trunk, Delmarva, and Hudson Canyon. Regression slopes were 0.34 for Georges and 0.46 for Mid Atlantic Bight. When the data are broken out by substrate regardless of region, the regression slope for sand was 0.35, for sand plus other substrate types such as shell hash it was 0.40, and on gravel it was 0.35. These slopes should modestly underestimate the sampling efficiency of the dredge relative to HabCam (due to errors in variables, i.e., that the x coordinates in the regression are uncertain since HabCam does not go over the exact same ground as the dredge). Results (dredge sampling efficiencies for sea scallops ~ 0.3 to 0.45) are similar to results from other studies. Moreover, they illustrate the potential for use of HabCam in directly estimating the sampling efficiency of other types of survey and fishing gear. Estimates from simple regressions are biased low because of errors in variables: see Appendix X for unbiased methodology.

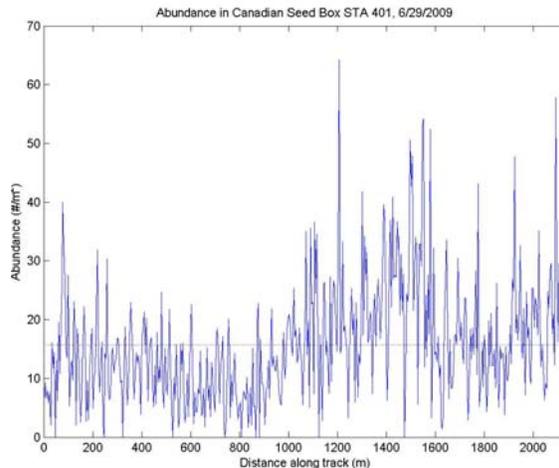
Bland-Altman plots are used to assess the correspondence between two forms of measurement for the same data and are constructed by plotting the differences between paired observations from two data sets against their mean. It was necessary to normalize the residuals for the sum of both dredge and HabCam samples. The mean residual for all data for 2008 and 2009 was 0.37 (Fig. 28), which is consistent with the regression analyses presented in Figure 27. The residuals are normally distributed between the limits of agreement suggesting that while there is a strong systematic bias, neither measurement approach is affected by abundance of scallops being measured.



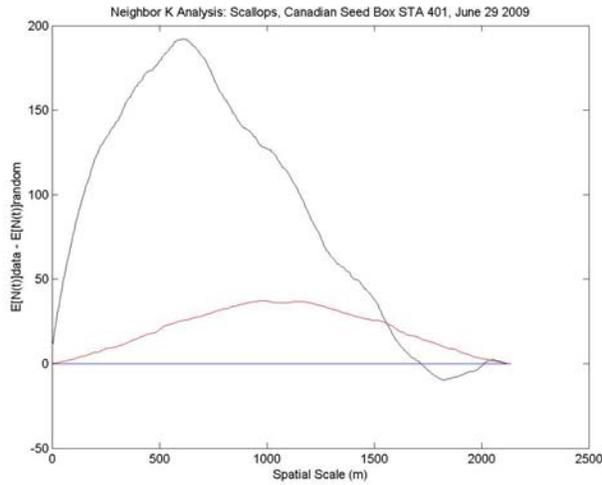
Appendix B9-Figure 21. The HabCam Manual Identification Program (MIP) while processing an image collected on the Canadian side near station 404 where 90 scallops were counted and measured in a single



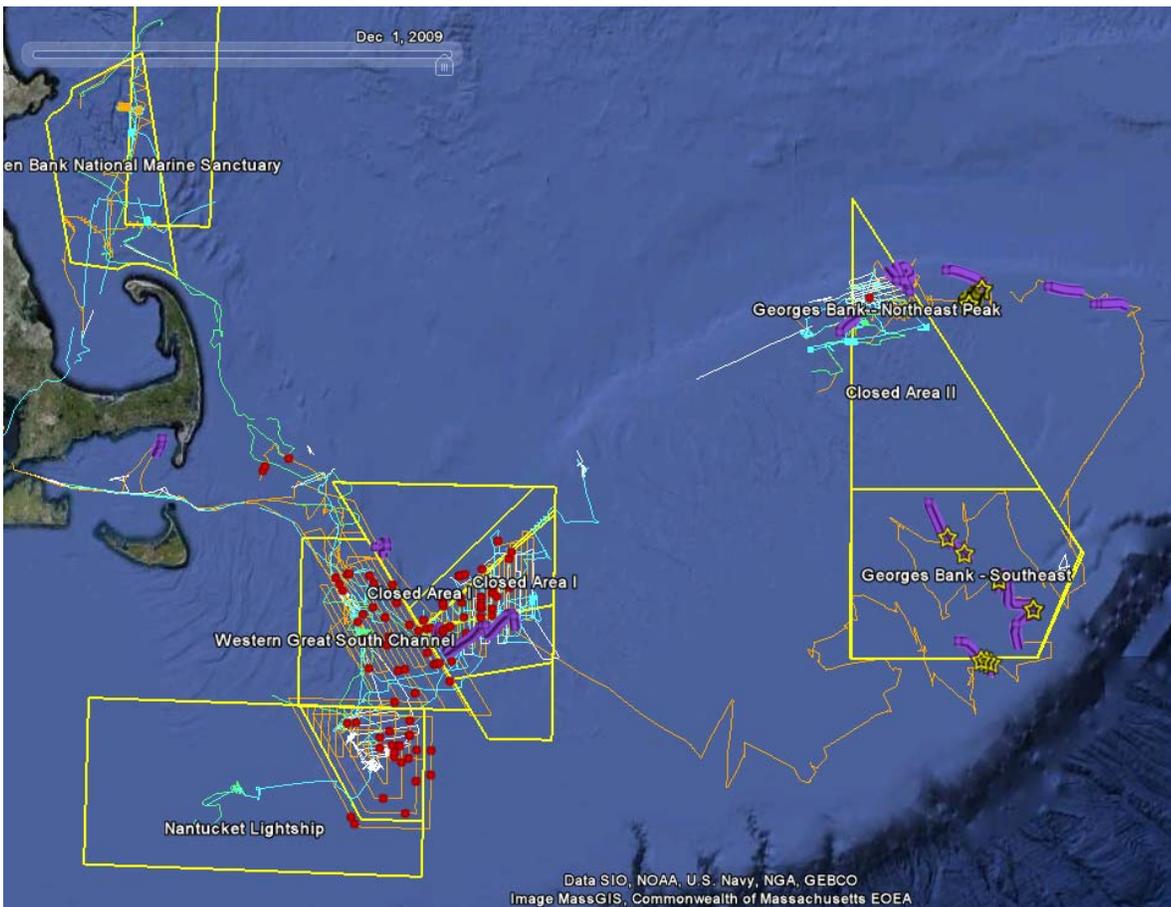
Appendix B9-Figure 22. Frequency distribution of scallop shell heights from HabCam images at Station 401 in the seed box.



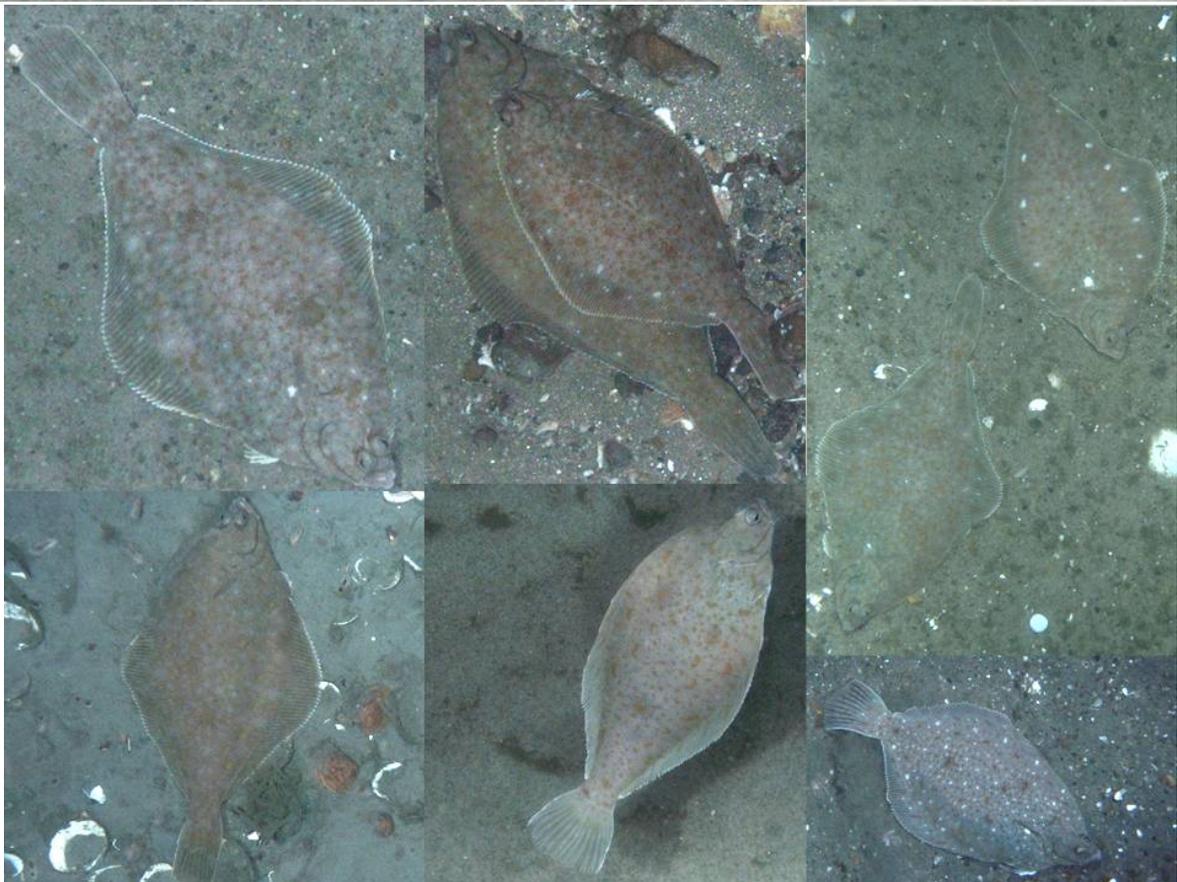
Appendix B9-Figure 23. Along track abundance of scallops at Station 401 in the 'seed box' on the Canadian side. Mean abundance was 16 scallops/m².



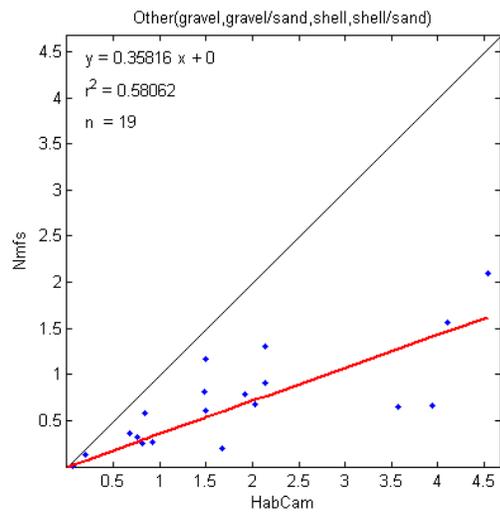
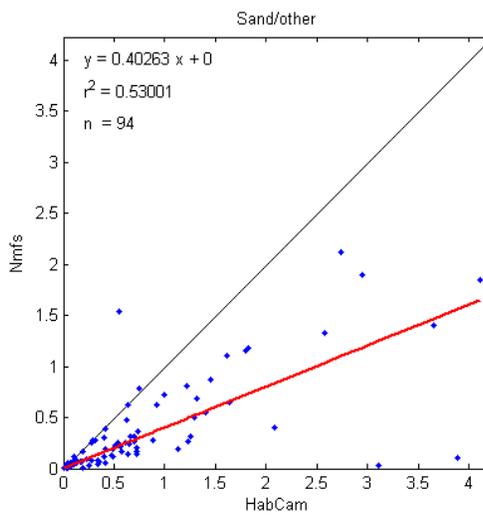
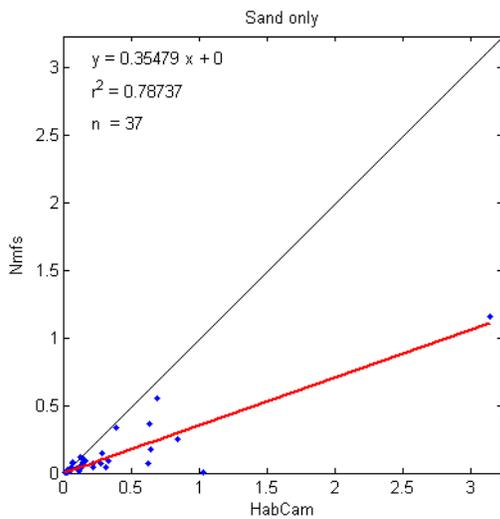
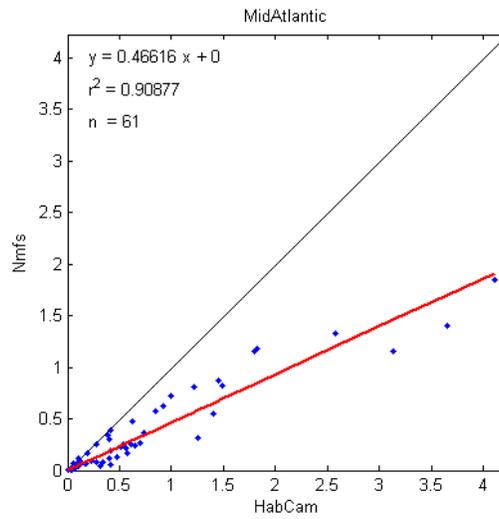
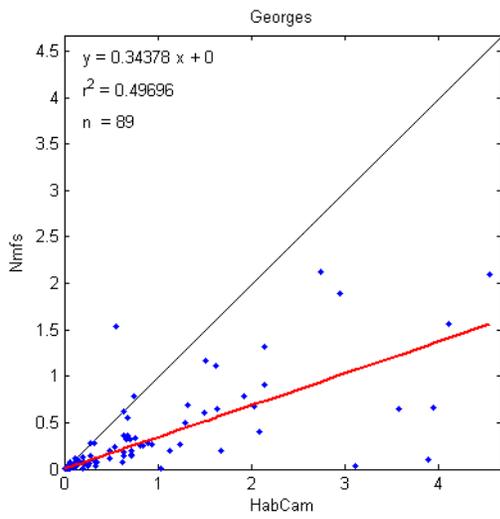
Appendix B9-Figure 24. Neighbor-k analysis of the distribution of scallops at Station 404. Note significant



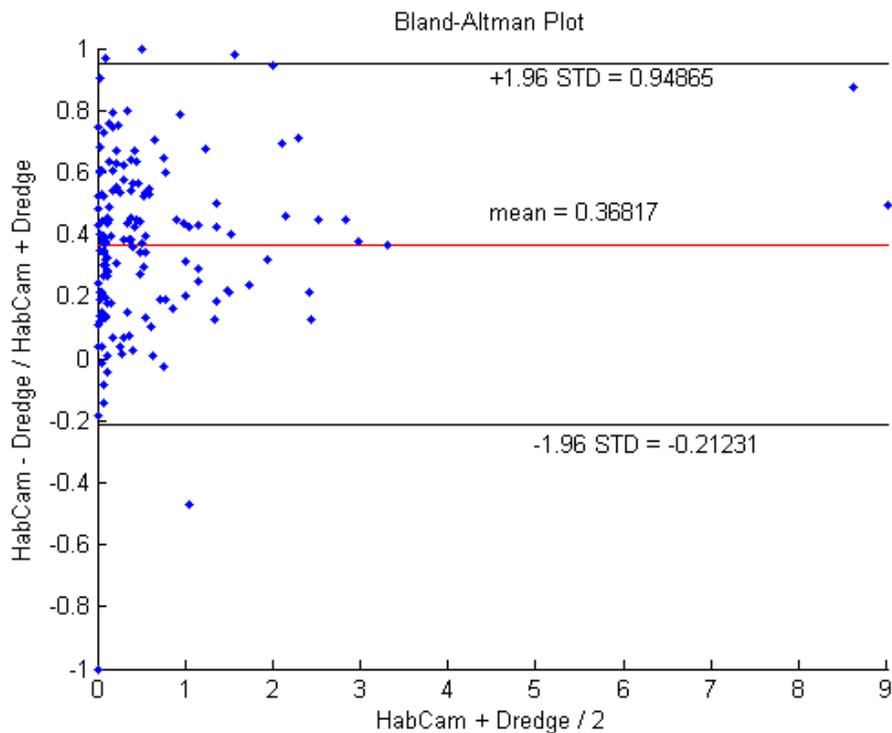
Appendix B9-Figure 25. Georges Bank and track lines of the R/V Hugh Sharp Leg 3 of the NMFS scallop survey (orange) and the regions where HabCam was deployed and collecting images (purple). Red dots and yellow stars are yellowtail flounder sightings



Appendix B9-Figure 26. Composite of example images of yellowtail flounder from Georges Bank.



Appendix B9-Figure 27. Regressions of dredge survey estimates against HabCam estimates of scallop densities for both Years 2008 and 2009. Each point represents a single 1nm tow in the Georges Bank (a) or Mid Atlantic Bight (b) areas. Data broken out by substrate type in sand (c), sand plus shell hash (d), and gravel (e). The one to one correspondence line is plotted in black.



Appendix B9-Figure 28. Bland-Altman plot of the residuals for all joint tows in 2008 and 2009. The y axis represents $(\text{Dredge}-\text{HabCam})/(\text{Dredge}+\text{HabCam})$ and the x axis is the mean of the two observations. The mean difference and the limits of agreement are also plotted.

Assessment of Real and Potential Errors Associated with HabCam Image Data

The sources of error to be assessed in this section are:

- a) Border rules for measuring and counting scallops
- b) Engineering Error
 - Calibration of Field of View (FOV), complete camera model for intrinsic parameters, estimation of in-water, focal length, principle point, and pixel error
 - Incorporation of extrinsic parameters for each image into calculation of FOV (area swept)- roll, pitch, heading, altitude
- c) Human Error
 - Analysis of measurement error both between individuals and within individuals using Intra Class Correlation
- d) Imaging Error
 - Scallop shells not orthogonal to camera axis
- e) Total Measurement Error
 - Analysis of shell height measurement error relative to NMFS dredge survey in NLSCA

f) Errors in interpolation of 1D data into 2D

- Kriging correlograms and variograms
- Variance within and between gridded cells
- Non-model based assessment of biomass

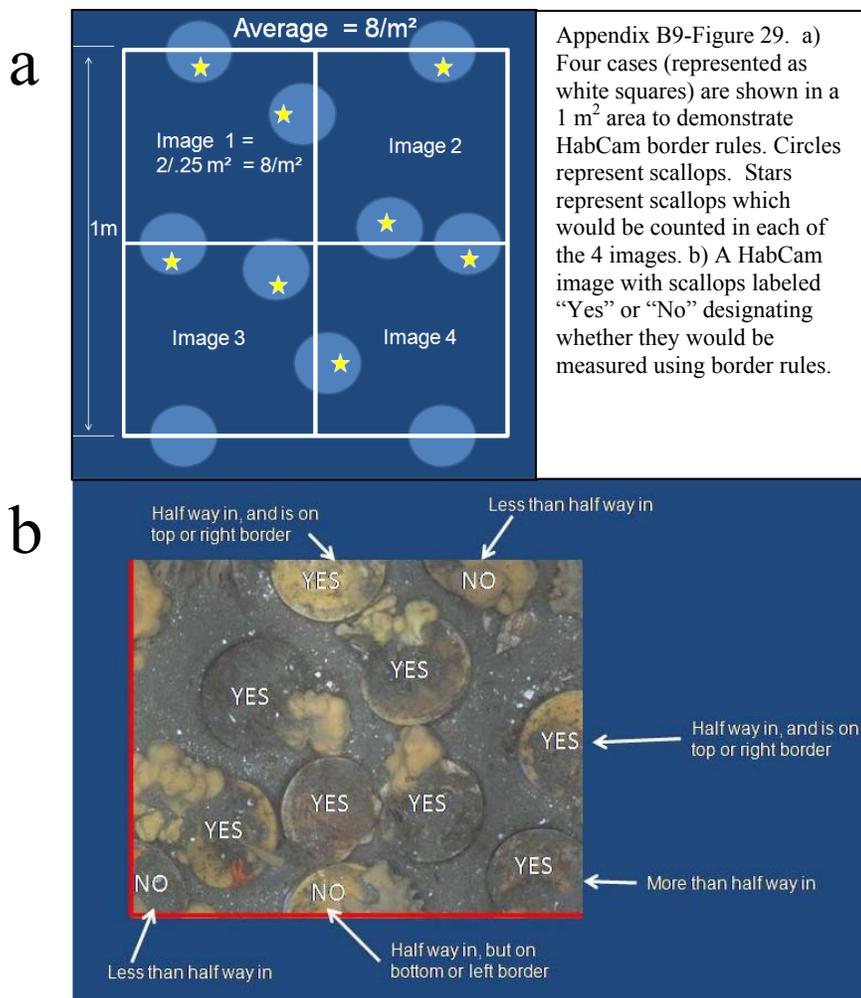
Each of these sources of potential error is discussed below.

HabCam border rules

The purpose of the HabCam rules for border effects is to reduce undercounting or over counting due to animals being on the edge of images. The desired outcome is to count scallops on the edge of images exactly half the time. To achieve this, the following rules, which result in counting only scallops that have their centroid in the image, are followed (Fig. 29):

Primary Rule: Count all organisms that are more than half way in the image.

Secondary Rule: If the organism is exactly half way in the image, count only the organisms that are half way in the top and right sides. This process is identical to that described for counting blood cells on a Spears-Levy Hemacytometer and eliminates the need for altering the field of view (FOV) on an image to account for image basis.



Engineering errors-Calibration of Field of View (FOV)

Calibration of an optical system must include a complete camera model for intrinsic parameters, estimation of in-water focal length, principle point, and pixel error, followed by image correction by employing extrinsic parameters collected for each image.

The intrinsic parameters for the HabCam camera were calculated using images of a 1m² target marked off at 10cm intervals in a 4 m deep seawater tank. The HabCam vehicle was positioned above the target at various altitudes (1-3m), roll, and pitch (0 and 20 degrees). Twenty eight images representing a range of positions were used for calibration of the camera with the Calibration Toolbox in Matlab.

Engineering errors-Intrinsic parameters

(based on 28 images of target at different altitudes and orientations)

Focal Length: $fc = [2773.25504 \ 2764.28859] \pm [7.18117 \ 7.13362]$

Principal point: $cc = [778.19667 \ 509.00401] \pm [4.13012 \ 3.80811]$

Skew: $\alpha_c = [0.00000] \pm [0.00000] \Rightarrow$ angle of pixel axes = 90.00000 ± 0.00000 degrees

Distortion: $kc = [-0.31591 \ 0.14388 \ 0.00070 \ 0.00138 \ 0.00000] \pm [0.00702 \ 0.02649 \ 0.00038 \ 0.00056 \ 0.00000]$

Pixel error: $err = [0.53035 \ 0.50489]$

The numerical errors are approximately three times the standard deviations

Intrinsic pixel error = +/- 1.59 pixels

Resolution range f (FOV): 0.37 – 0.89 mm/pixel

Intrinsic real-world error : 0.58 – 1.41 mm

These values provide error bounds on the resolution and accuracy of the camera system in water. Plots of the relative errors show that the camera CCD chip, lens and housing window are slightly out of alignment in both radial and tangential attitudes (Fig. 30). The pixel resolution is a function of FOV, which in turn is a function of altitude off the bottom. In calibrated screen measurement space, the overall measurement error is between 0.58 and 1.41 mm.

Distortion in each image is first corrected using the intrinsic parameters given above (Fig. 31).
 $KK = [fc(1) \alpha_c * fc(1) cc(1); 0 fc(2) cc(2); 0 0 1];$
where the KK matrix is the uncorrected image matrix.

$r2_extreme = (nx^2/(4*fc(1)^2) + ny^2/(4*fc(2)^2));$

$dist_amount = 1; \% (1+kc(1)*r2_extreme + kc(2)*r2_extreme^2);$

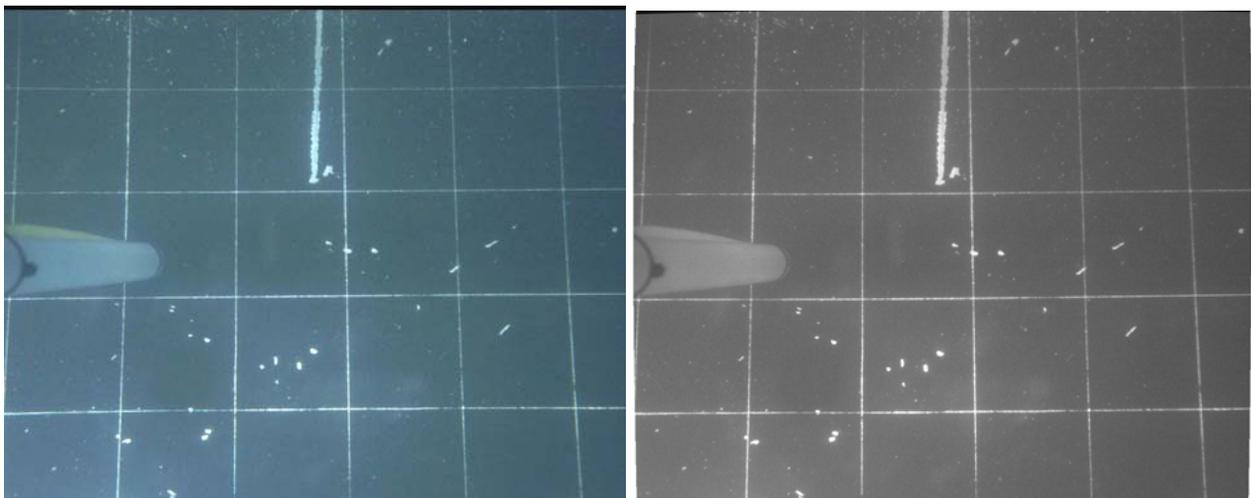
$fc_new = dist_amount * fc;$

$KK_new = [fc_new(1) \alpha_c * fc_new(1) cc(1); 0 fc_new(2) cc(2); 0 0 1];$

KK_new is the corrected image matrix.

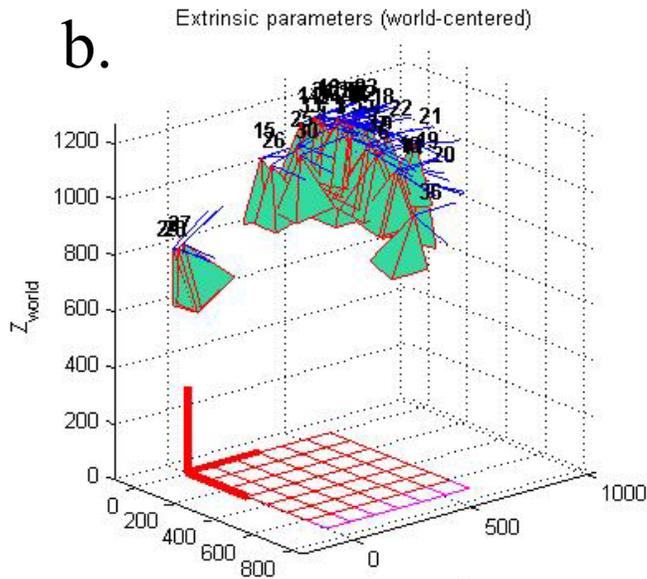
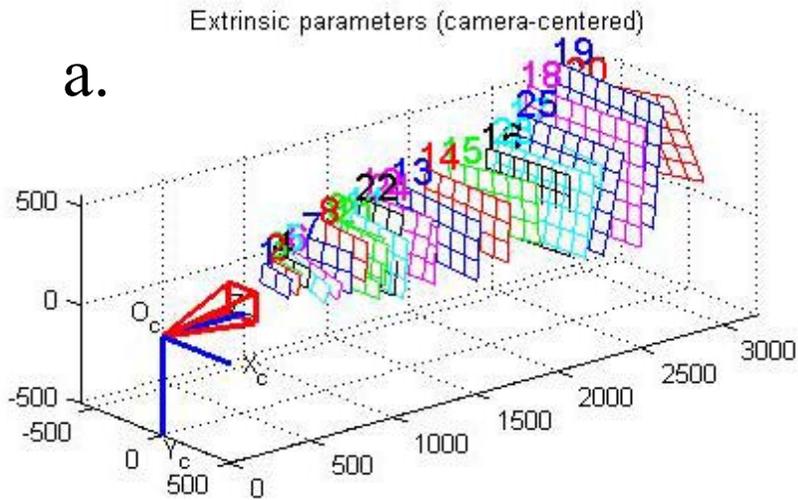
$[I2] = rect(I, eye(3), fc, cc, kc, KK_new);$

Where I is the distorted image and I2 is the undistorted image



Appendix B9-Figure 31. Correction of a distorted image of the calibration target in water (left) using intrinsic camera parameters. The corrected image (right) shows straight rather than curved lines particularly towards the corners of the image.

Extrinsic parameters relate to the combination of intrinsic parameters plus the orientation of the camera relative to the image plane. The calibration matrix is built up from the 28 views indicated in Figure 32.



Appendix B9-Figure 32.

- a) Camera centric views of 28 orientations and altitudes to build extrinsic parameter list.
- b) World centric views of 28 orientations and altitudes.

Engineering errors-Calculation of extrinsic parameters

Cross over points in the calibration chart are automatically detected and their locations in pixel space extracted before calculation of extrinsic parameters (Fig. 33).

The extrinsic parameters are encoded in the form of a rotation matrix (**Rc_ext**) and a translation vector (**Tc_ext**). The rotation vector **omc_ext** is related to the rotation matrix (**Rc_ext**) through the Rodrigues formula: **Rc_ext** = **rodrigues(omc_ext)**.

Let **P** be a point space of coordinate vector **XX** = [**X**;**Y**;**Z**] in the grid reference frame (**O,X,Y,Z**).

Let **XX_c** = [**X_c**;**Y_c**;**Z_c**] be the coordinate vector of **P** in the camera reference frame (**O_c,X_c,Y_c,Z_c**). Then **XX** and **XX_c** are related to each other through the following rigid motion equation:
XX_c = **Rc_ext** * **XX** + **Tc_ext**

In addition to the rigid motion transformation parameters, the coordinates of the grid points in the grid reference frame are also stored in the matrix **X_ext**.

Each image taken by HabCam has its own unique set of extrinsic parameters.

Extrinsic parameters for an example image:

Translation vector:

$$\text{Tc_ext} = [-225.840216 \quad -130.369514 \quad 608.628548]$$

Rotation vector:

$$\text{omc_ext} = [-2.148393 \quad -2.284790 \quad -0.123388]$$

Rotation matrix:

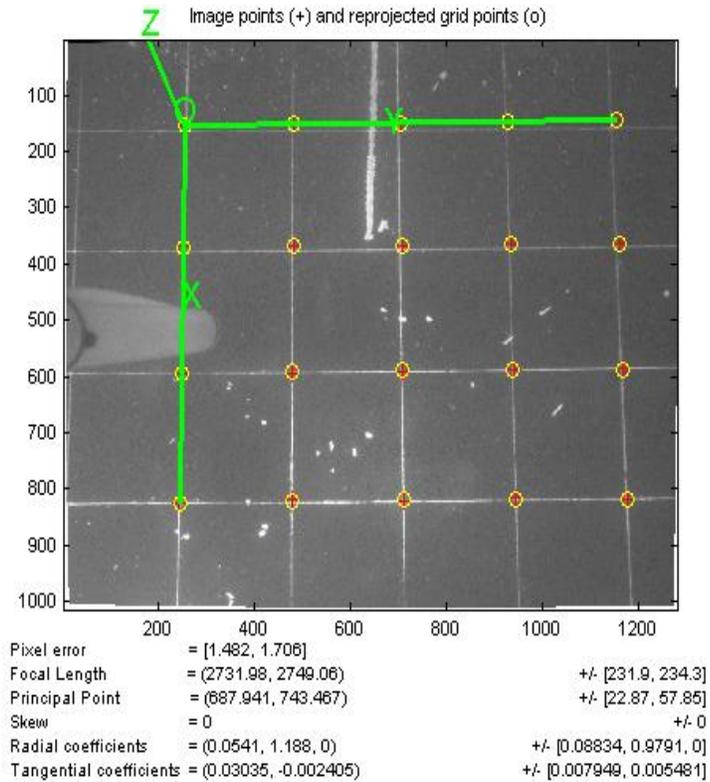
$$\text{Rc_ext} = [-0.062925 \quad 0.996680 \quad 0.051672 \\ 0.996448 \quad 0.059838 \quad 0.059254 \\ 0.055966 \quad 0.055217 \quad -0.996905]$$

$$\text{Reprojection Pixel Error:} \quad \text{err} = [2.00116 \quad 1.26492]$$

Extrinsic pixel error = +/- 2 pixels

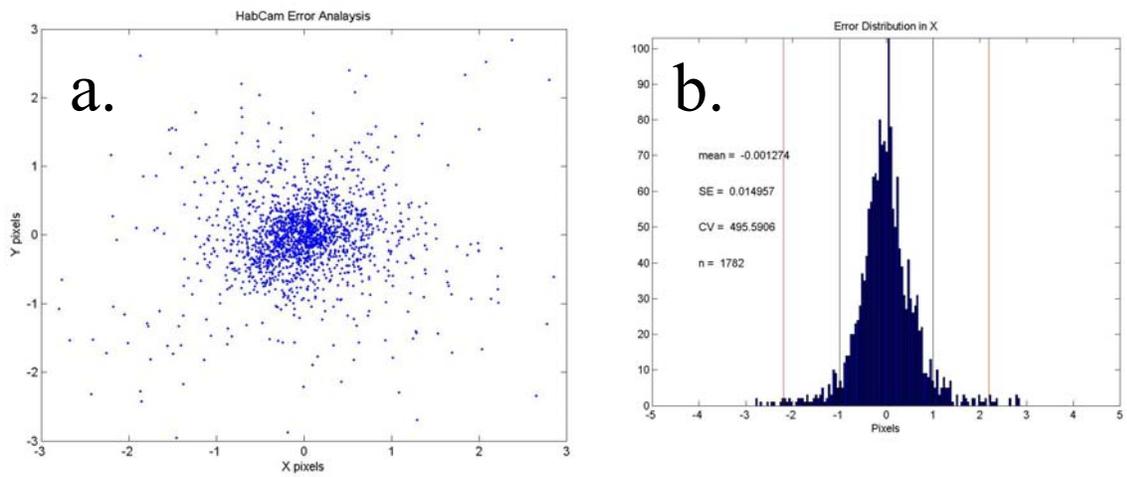
Resolution range (FOV): 0.37 – 0.89 mm/pixel

The extrinsic real-world error becomes: 1.11 – 1.78 mm (under best optical conditions)



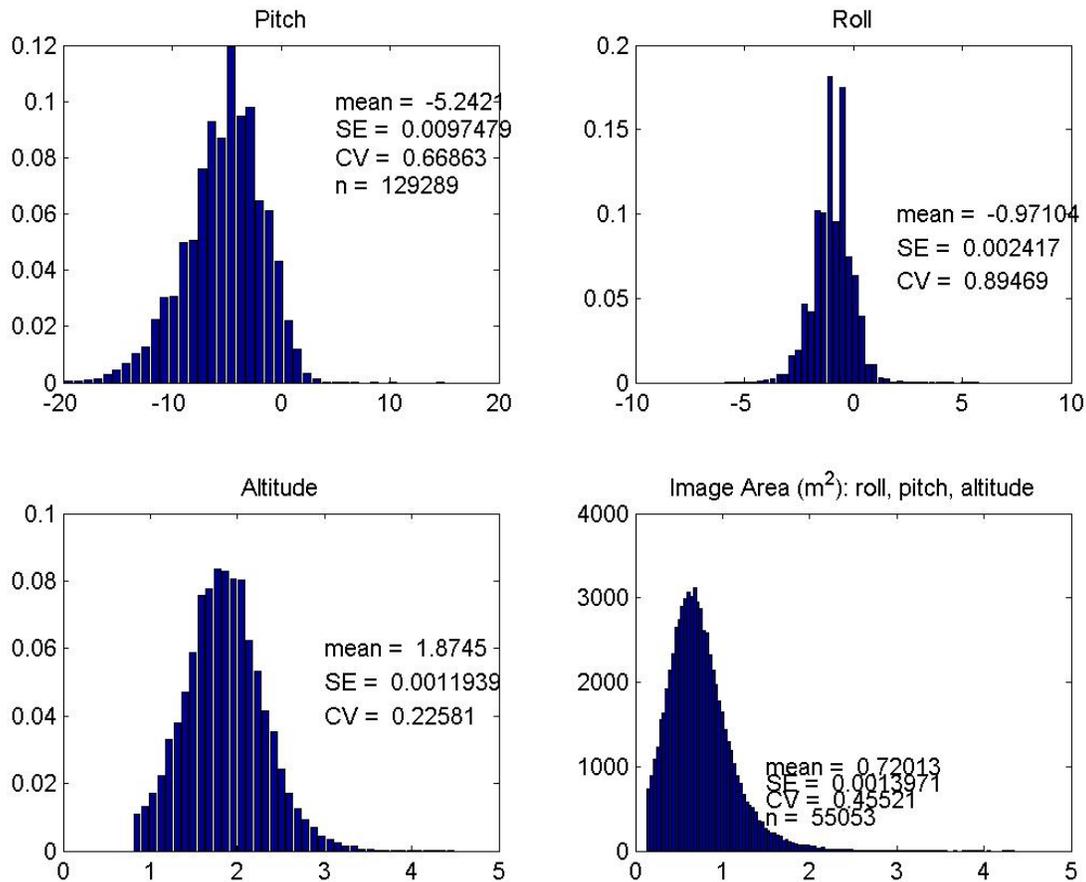
Appendix B9-Figure 33. Extrinsic parameters for image 24 above, as an example.

Pixel error in X and Y can be visualized as a scatter plot and frequency distribution (Fig. 34). Note that 99% of values are less than 2.2 pixels.



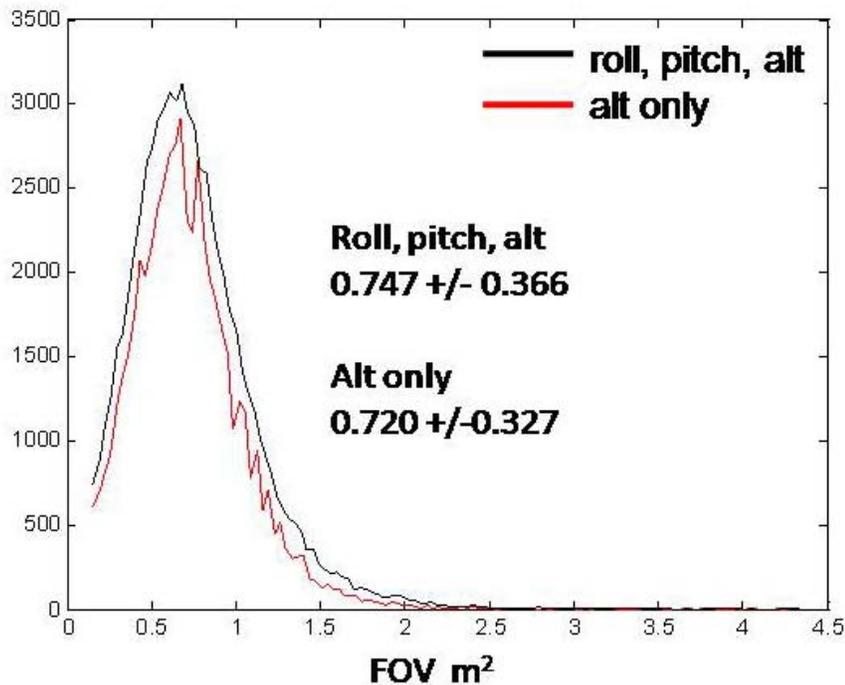
Appendix B9-Figure 34. (a) Scatter plot of pixel error around the origin. (b) Frequency distribution of pixel error along x axis.

Frequency distributions of pitch, roll, altitude, and image area (FOV) for the NLSCA 2009 survey based on 129,289 images are shown in Figure 35. The mean pitch was -5.24 degrees indicating that, on average, the nose of the vehicle pointed down slightly. Downward pitch is part of the system design and tends to stabilize the vehicle while underway. Mean roll was -0.97 with very little variation indicating the vehicle is quite stable, laterally. Altitude measurement varied from <1 to 4.5 m off the bottom with a mean of 1.87m. Images below 1 m were out of focus and removed from the image database. Images taken higher than 3 m were typically not sufficiently clear, due to turbidity, to be useful and were also not used. Taking roll and pitch into account using the extrinsic equations present above, the FOV ranged from 0.2 to >4m² with a mean of 0.72 m². 95% of the calculations for FOV fell between 0.4 and 1.5 m². Figure 36 shows a comparison between FOV calculated with and without the use of roll and pitch, i.e., directly from the altitude, only. Incorporation of roll and pitch into the geometric projection of the FOV has an effect of broadening and smoothing the frequency distribution of values without changing the mean.



Appendix B9-Figure 35. Frequency distributions of pitch, roll, altitude, and image area (FOV) for the NLSCA 2009 survey.

**Distribution of Image Area m² (FOV)
NLSCA 2009 Survey**

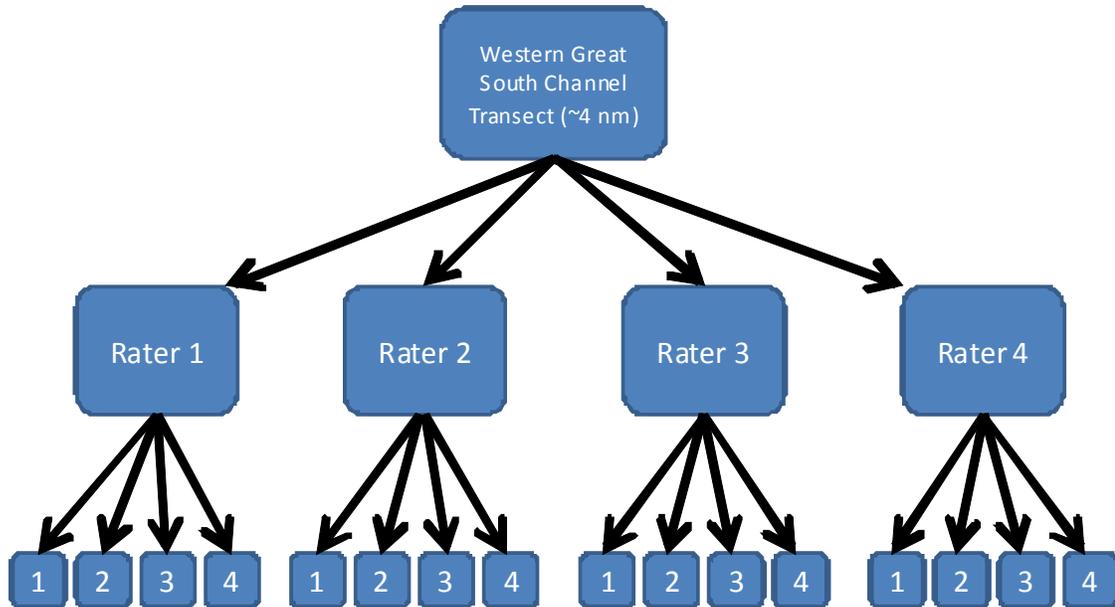


Appendix B9-Figure 36. Comparison between calculations of FOV with and without the effect of roll and pitch on geometric projection of FOV. Note broadening and smoothing of the distribution without a significant change in the mean, when projected geometry is used in conjunction with altitude.

Human errors: Analysis of measurement error both between individuals and within individuals
It was desired to estimate the level of error associated with the manual screen measurement of scallops both within a given technician and between technicians. The former would provide insight into measurement repeatability and the latter into systematic bias between individuals.

To accomplish this, we assigned four identical 4.2 nm long image transects containing 4,432 images from Western Great South Channel to six individuals (raters) (Fig. 37). Raters measured scallops using MIP under the same measurement rules as would be used under normal conditions (edge effects, height vs. width, etc).

ICC Intra Class Correlation Analysis



Two-way mixed effects model

$$X_{ij} = \mu + r_i + c_j = \mu + r_{ij} + e_{ij}$$

μ : population mean, r : row effects, c : column effects, e : residual effects

281 scallops x 4 raters x 4 passes = 4,496 measurements

Appendix B8-Figure 37. Inter Class Correlation analysis of scallop shell height measurements. Four individuals measured scallops from one transect four times.

Appendix B9-Table 1. Summary statistics in pixels. N = 277 for each run. A total of 4,432 scallops were measured. KLB, ADY, PK, and DPF are initials of the four raters.

rater	KLB				
	run1	run2	run3	run4	mean
mean	132.41	132.21	132.58	132.28	132
STD	31.61	31.61	31.44	31.69	
SE	1.89	1.91	1.88	1.9	
rater	ADY				
	run1	run2	run3	run4	mean
mean	128.48	128.46	128.41	128.75	128
STD	31.42	31.55	31.44	31.78	
SE	1.88	1.89	1.88	1.9	
rater	PK				
	run1	run2	run3	run4	mean
mean	135.57	135.88	134.95	134.28	135
STD	31.86	31.94	31.81	31.63	
SE	1.91	1.91	1.91	1.9	
rater	DPF				
	run1	run2	run3	run4	mean
mean	128.36	127.39	127.48	127.56	127
STD	31.74	31.63	31.45	31.57	
SE	1.9	1.9	1.89	1.89	

In most cases, mean within rater measurements were either accurate to the same number of pixels or within one pixel suggesting that within rater variability was extremely low (Table 1). Between rater variability was greater than within rater variability with mean values of 132, 128, 135, 127, providing a range of 135 to 127, or 8 pixels. Given the resolution range for varying FOV presented above (i.e., 0.37 – 0.89 mm/pixel), an error of 8 pixels represents a real-world error of 3.0 to 7.1 mm.

Inter and intra-Class correlations were analyzed using ICC, Intra Class Correlation analysis (McGraw and Wong, 1996). A two-way mixed effect model

$$X_{ij} = u + r_i + c_j = r_{cij} + e_{ij}$$

u: population mean, r: row effects, c: column effects, e: residual effects

was used to test the hypotheses that there is no difference between scallop measurements made by the same rater four times, and that there is no difference between individual raters.

ICC Type C-1: Tests the degree of consistency among measurements

$$r = (MSR - MSE) / (MSR + (k-1)*MSE);$$

$$F = (MSR/MSE) * (1-r^2)/(1+(k-1)*r^2);$$

$$df1 = n - 1;$$

$$df2 = (n-1)*(k-1);$$

$$p = 1-fcdf(F, df1, df2);$$

ANOVA Table					
Source	SS	df	MS	F	Prob>F
Time	143.5	3	47.8	6.28	0.0003
Group	40170.4	3	13390.1	3.36	0.0182
Ineratcion	480.6	9	53.4	7.01	0
Subjects (matching)	4398738.9	1104	3984.4	522.79	0
Error	25241.8	3312	7.6		
Total	4464775.2	4431			

$$r = 0.9842$$

$$LB = 0.9827$$

$$UB = 0.9857$$

$$p = 0$$

ICC Type A-1: Test the degree of absolute agreement among measurements.

$$r = (MSR - MSE) / (MSR + (k-1)*MSE + k*(MSC-MSE)/n);$$

$$a = (k*r0) / (n*(1-r0));$$

$$b = 1 + (k*r0*(n-1))/(n*(1-r0));$$

$$F = MSR / (a*MSC + b*MSE);$$

$$df1 = n - 1;$$

$$df2 = (a*MSC + b*MSE)^2 / ((a*MSC)^2 / (k-1) + (b*MSE)^2 / ((n-1)*(k-1)));$$

$$p = 1 - \text{fcdf}(F, df1, df2);$$

ANOVA Table					
Source	SS	df	MS	F	Prob>F
Time	143.5	3	47.8	6.28	0.0003
Group	40170.4	3	13390.1	3.36	0.0182
Ineratcion	480.6	9	53.4	7.01	0
Subjects (matching)	4398738.9	1104	3984.4	522.79	0
Error	25241.8	3312	7.6		
Total	4464775.2	4431			

$$r = 0.9796$$

$$LB = 0.9695$$

$$UB = 0.9856$$

$$p = 0$$

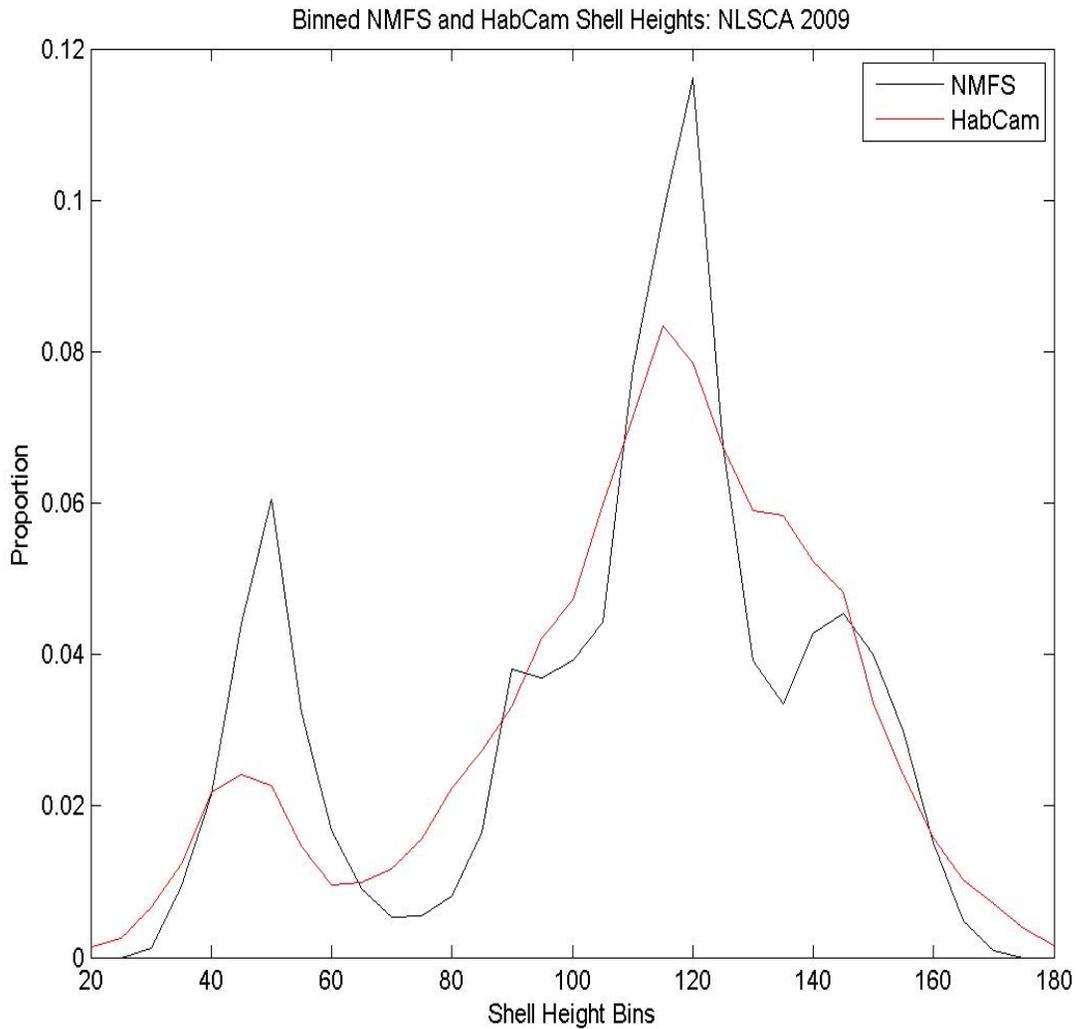
Summary

There is no difference in measurements made by the same individual raters or between individual raters.

Human errors-Analysis of shell height measurement error

Shell height measurements from HabCam images taken during the 2009 Nantucket Lightship survey were compared with shell height measurements made from 12 dredge tows on 2009 Leg 2 of the R/V Hugh Sharp during normal survey operations. The HabCam survey was conducted in early June 2009 while the NMFS survey was conducted in min July, 2009.

Frequency distribution for shell height measurements for HabCam and NMFS dredge survey show surprising similarity in overall pattern (Fig. 38). Since the NMFS survey was conducted about five weeks following the HabCam survey, the shift in mode of the NMFS data for small scallops can be accounted for by growth. The tails of the distribution for HabCam data are spread out more than for the NMFS data suggesting a source of measurement error. There is no indication of selectivity by either sampling approach.



Appendix B9-Figure 38. Shell height size frequency distributions for NMFS (black) and HabCam (red) measurements in the NLSCA. HabCam surveyed in early June while NMFS surveyed in mid July. Note a shift to the right of the mode for small scallops in NMFS data relative to that for HabCam probably due to growth.

Shell heights made from HabCam and NMFS dredge survey were analyzed using the approach described by Jacobson et al. 2010). Accuracy (RMSE, root mean square error) , bias (HabCam-Dredge), and precision (STD) were calculated.

HabCam measurements were positively biased relative to NMFS data by 3.7%. Percentage square root of the mean square error was 3.70%. Both NMFS and HabCam distributions were negatively skewed and more peaked relative to normal distributions.

stat		NMFS		HabCam
n		4,178		13,576
bias		NA		3.8
min		30		20
max		170		180
avg		106.4		110.3
%bias		NA		3.70%
STD		33.4		32.2
CV		31.4		27.90%
RMSE		NA		3.8
% RMSE		NA		3.70%
skewness (g1)		-0.59		-0.66
kurtosis (g2)		2.41		2.98

Summary of error analysis

Measurement error can come from a number of sources including intrinsic error in camera calibration, extrinsic error due to camera orientation and altitude relative to geometric projection on the image plane, and errors associated with human operators measuring scallops on the computer screen. Given the resolution range for varying FOV presented above (i.e., 0.37 – 0.89 mm/pixel), a real world for each source may identified.

Source	+/- mm error
Intrinsic	0.58 – 1.41
Extrinsic	1.11 – 1.78
Within operator	0.5
Between operator	3.0 – 7.1

Clearly, the magnitude of between operator errors dominates the overall potential for measurement error. However, once this source of variability is identified and characterized for each individual technician measuring scallops, a correction factor could be applied for each individual to normalize the results. In addition, automated counting and sizing of scallops and other targets is improving and will eventually be used in conjunction with manual measurements to produce more accurate results.