Spring 2020

# An Analysis of Dredge Efficiency for Surfclam and Ocean Quahog Commercial Dredges

Leanne Poussard

# AN ANALYSIS OF DREDGE EFFICIENCY FOR SURFCLAM AND OCEAN QUAHOG COMMERCIAL DREDGES

by

Leanne M. Poussard

A Thesis
Submitted to the Graduate School,
the College of Arts and Sciences
and the School of Ocean Science and Engineering
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Science

Approved by:
Dr. Eric N. Powell, Committee Chair
Dr. Daniel R. Hennen
Dr. Chet Rakocinski

| | | |
|---|---|---|
| Dr. Eric Powell<br>Committee Chair | Dr. Joe Griffitt<br>Director of School | Dr. Karen S. Coats<br>Dean of the Graduate School |

May 2020

# ABSTRACT

Between 1997 and 2011, The National Marine Fisheries Service conducted 50 depletion experiments to estimate survey gear efficiency and stock density for Atlantic surfclam (*Spisula solidissima*) and ocean quahog (*Arctica islandica*) populations using commercial hydraulic dredges. The Patch Model was formulated to estimate gear efficiency and organism density from the data. The range of efficiencies estimated is substantial, leading to uncertainty in the application of these estimates in stock assessment. Analysis of depletion experiment simulations showed that uncertainty in the estimates of gear efficiency from depletion experiments was reduced by higher numbers of dredge tows per experiment, more tow overlap in the experimental area, a homogeneous as opposed to patchy distribution of clams in the experimental area, and the use of gear of inherently high efficiency. Simulations suggest that adapting the experimental protocol during the depletion experiment by adjusting tow number and degree and dispersion of tow overlap may substantively reduce uncertainty in the final efficiency estimates.

Known values of four metrics for each field experiment were compared to metrics from the 9,000 simulations in the simulation dataset to determine which experiments diverge from those in the simulation dataset, and which experiments were likely to have high error in the efficiency estimate. The error metrics used implicate a subset of experiments that are outliers, biasing the efficiency estimates for the entire dataset.

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Dr. Eric Powell, for his continuous support of my masters study and research and for opening the doors to many diverse and exciting research opportunities. His guidance helped me in all the time of research and writing of this thesis. I am grateful for my time being part of his lab and for the opportunity to conduct research all over the country. I would also like to thank my committee members, Dr. Dan Hennen and Dr. Chet Rakocinski for their encouragement and insightful comments throughout the writing of this thesis. I would especially like to thank Dan for his mentorship when I was working at the NEFSC, and I appreciate the support and friendship from the staff there.

I am grateful to the community of students, faculty, and staff at GCRL for their support and friendship. I am especially grateful to the members of my lab, Sara Pace, Kathleen Hemeon, and Laura Solinger, for all the fun times and for making Ocean Springs a home. I'd also like to thank my friends and colleagues at VIMS for their advice, encouragement and for starting me off on this grad school adventure.

# DEDICATION

This thesis is dedicated to my Aunt Rita DiLisio, who passed away at the time of this writing. Thank you for always supporting me, telling me how proud you were of my achievements in and out of school, and showing me boundless love. Until we meet again!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I – Efficiency Estimates from Depletion Experiments for Sedentary Invertebrates: Evaluation of Sources of Uncertainty in Experimental Design

## 1.1 Introduction

The ocean quahog (*Arctica islandica*) and the Atlantic surfclam (*Spisula solidissima*) support substantial fisheries on the northeast U.S. continental shelf. Ocean quahogs are reputed to be the longest-lived non-colonial marine species (Butler et al. 2013) - they are certainly the most abundant of the very-long-lived species - and, on the U.S. East coast continental shelf, are typically found offshore in deep water, between 30 m and 220 m (NEFSC 2017b) with life spans exceeding 250 years (Pace et al. 2017). The Atlantic surfclam has a lifespan of about 30 years and is found in more inshore waters, typically between 8 and 66 m depth (NEFSC 2017a). They coexist with ocean quahogs along their offshore range boundary that approximately follows the 15°C summer bottom water temperature isotherm (NEFSC 2017b, Powell et al. 2020). Surveys conducted in the 2000s show evidence that a range shift is occurring as the western North Atlantic warms, with surfclams invading deeper water, presently often occupied by ocean quahogs, throughout the mid-Atlantic region (Hofmann et al. 2018, Powell et al. 2019).

The ability to accurately estimate abundance from the National Marine Fisheries Service (NMFS) ocean quahog (*Arctica islandica*) and Atlantic surfclam (*Spisula solidissima*) stock assessment surveys benefits from empirical estimates of the efficiency of the survey dredge (NEFSC 2003, Powell et al. 2007). Gear efficiency is defined as the probability that an organism in an area intersected by the dredge will be caught (Hennen et al. 2012). However, efficiency is a key source of uncertainty in stock assessments. Efficiency estimates are notoriously variable (Vølstad et al. 2000, Powell et al. 2007, Hennen et al. 2012, Morson et al. 2018) at least in part because little is known about how experimental variables can influence the efficiency of a dredge. Efficiency estimates have been obtained for a range of dredge types, including oyster dredges (Powell et al. 2007, Morson et al. 2019), crab dredges (Vølstad et al. 2000, Bohrmann & Christman 2012, Wilberg et al. 2013), and scallop dredges (Beukers-Stewart & Beukers-Stewart 2009, Lasta & Iribarne 1997). These are all dry dredges designed to harvest epibenthic animals. In contrast, commercial surfcalm and ocean quahog fisheries, as well as the stock assessment surveys for both stocks use hydraulic dredges. Hydraulic dredges are designed to harvest infaunal clams by using water pressure to liquefy the sediment, thereby penetrating deeply into the sediment and exhuming the clams (Da Ros et al. 2003, Hauton et al. 2007, Meseck et al. 2014). Hydraulic dredges are efficient in comparison to dry dredges (Thórarinsdóttir et al. 2010).

A typical hydraulic dredge is a large rectangular box constructed of steel bars evenly spaced apart mounted on skids and towed along a seabed (Lambert & Goudreau 1996, Meyer et al. 1981). A cutting blade in front of the dredge digs into the sediment as high-pressure water is pumped through a series of jets from a manifold, serving to liquefy the sediment, thus permitting the dredge to be towed with little resistance through the surficial sediment and thereby increasing the catchability of the target bivalve species (Gilkinson et al. 2003). Parker (1971) provides a historical account of the development of hydraulic dredges in the surfclam fishery.

Despite the increased focus on quantitative stock assessments in recent years and the industrial success of the hydraulic dredge, gear efficiency is still an uncertain parameter that is affected by many variables, including the size frequency of clams in the population, current force and direction, sediment density, and bottom roughness (fine scale bathymetry). Little is known about exactly how these factors might cause variation in efficiency estimates for hydraulic dredges.

Depletion experiments are commonly used to estimate gear efficiency and density of the target organism in the benthos (Leslie & Davis 1939, Skalski et al. 1983, Lasta & Iribarne 1997, Gedamke et al. 2005, Wilberg et al, 2013). Depletion experiments consist of deploying the gear multiple times in a target area, allowing the catch per tow to decline as a result of decreasing organism density. This rate of decline is used to estimate gear efficiency and the initial abundance of the organism.

A series of depletion experiments was conducted between 1997 and 2013 by academic and industry collaborators on commercial and survey vessels to estimate the efficiency of the commercial clam dredges and infer the efficiency of the National Marine Fisheries Service survey dredge (NEFSC 2010c, 2013). The depletion experiments were carried out at locations specified in Appendix 3 of NEFSC (2017a) (Figures 2.7 and 2.8).

The Patch Model (Rago et al. 2006) was developed to analyze depletion experiments and estimate gear efficiency, stock abundance, and dispersion of organisms in a target area. The Patch Model has been important in informing stock assessments of commercially exploited populations of Atlantic surfclam, ocean quahog, monkfish (*Lophius americanus*; (NEFSC, 2010a)) and Atlantic sea scallop (*Placopecten magellanicus;* (NEFSC 2010b, NMFS 2009). Hennen et al. (2012) examined the performance of the Patch Model under a range of conditions and found that uncertainty in dredge position and distribution of dredge tow overlap in the experimental area were important contributors to the uncertainty in estimates of dredge efficiency.

The correction for dredge efficiency continues to be a primary source of uncertainty in the estimation of stock abundance for both clam species. In this study, we extend the analytical approach of Hennen et al. (2012) to develop metrics that can be used to guide retrospective evaluation of the effectiveness of experimental design of previous depletion experiments and to proffer an improved experimental design for future dredge efficiency estimates. To do so, a simulation protocol (Hennen et al. 2012) is implemented to test Patch Model efficiency estimates under a variety of conditions involving experiment methodology and dispersion and density of the target species to ascertain the characteristics of depletion experiments that contribute to the accuracy and precision of efficiency estimates.

## 1.2 Methods

### 1.2.1 ***The Patch Model***

The Patch Model estimates capture efficiency (the probability of capture of an organism intersected by the dredge), and density of organisms in the target area (numbers per $m^2$) by tracking the relative depletion (reduction in catch) over the tow series. Capture efficiency is theoretically a measurable characteristic of the gear. Here, we examine the influence of the number of tows in an experiment, the spatial distribution of organisms in the benthos, the density of organisms in the benthos, the degree of overlap of the multiple dredge tows conducted during the experiment, and the influence of the inherent ('true') efficiency of the gear on the uncertainty in the estimate of dredge efficiency obtained from the experiment.

Simulated depletion experiments follow a typical in-field experimental design used for actual depletion experiments conducted using hydraulic dredges designed to capture surfclams and ocean quahogs. A long rectangular area is chosen, on average about 10 dredge widths wide (25-38 m), and 400-800 m (1200-2400 ft) long. For the simulations described herein, the box dimension was taken as 960 m x 45 m, the narrow dimension being about 12 times the width of the present survey dredge. A series of overlapping dredge tows are taken across the selected area, with the dredge hitting bottom at one of the short edges of the rectangle and being retrieved at the other short edge. Ideally, the dredge is towed over the same ground multiple times while covering the majority of the demarcated area. The experiment requires the assumption that all catches are random samples and that no transport of organisms into or out of the study site occurs during the experiment (Leslie & Davis 1939). The tow paths, catch, and fishing effort are recorded for each tow. Over a series of tows, the catch per tow should decrease; this rate of decline is proportional to the efficiency of the dredge (Hennen et al. 2012). For example, if the rate of decline is rapid, the dredge is highly efficient.

### 1.2.2 ***Patch Model Estimates of Efficiency***

The Patch Model inputs are the tow series of area swept, the observed catch, and the hit matrix. The spatial domain in the model is defined as the smallest rectangular area that contains every tow in the experiment. Typically, this is marginally smaller in the short dimension and longer in the long dimension that the original specified rectangular area as vessels shy away from the lateral boundaries as they tow and inaccuracy in dredge deployment and retrieval routinely extends tows across the narrow ends of the box. Any particular point in this domain can be touched by the dredge 0 to *n* times after *n* tows. The rectangular area is subdivided into a grid of points that is used to calculate the hit matrix which records the number of times any point was contacted by the dredge.

The backbone of the Patch Model is the ability to calculate the catch per tow, the density of organisms in the area after each tow, and the cumulative catch for any tow from the initial conditions of the experiment. The equation

$$E(c) = e\left(\frac{a}{A}\right)N \quad (1)$$

describes the expected catch in a sample from a closed population given *e,* the probability of capture of an individual given an encounter with the dredge, $\frac{a}{A}$, the area swept by the tow divided by the total area, and *N*, the number of individuals in the population in the defined area. Substituting $q$ for $e\left(\frac{a}{A}\right)$ and adding the elements of time and space allows for the calculation of expected catch in any tow *i* given initial density and the cumulative catch from previous samples, $T_{i-1}$:

$$E(C_i) = q(N_0 - T_{i-1}). \quad (2)$$

Rago et al. (2006) incorporated the fraction of cells in the defined area that were hit by the tow, the hit matrix, into the equation, giving the expected CPUE for depletion tow *i* as

$$E(C_i) = (EAS)D_o \quad (3)$$

where $D_0$ is the initial density of the target organism in the area, and EAS is the effective area swept, the total area swept ($m^2$) by the dredge in tow *i* taking into account points hit by the dredge in previous tows. EAS is calculated as:

$$EAS = ea_i \sum_{j=1}^{i} f_{i,j}(1 - e\gamma)^{j-1} \quad (4)$$

where $e$ is capture efficiency, $a_i$ is the area swept by tow *i,* $f_{i,j}$ is the fraction of cells hit by the dredge *j* times, and $\gamma$ is the ratio of dredge width and cell size, or in other words, the fraction of the cell the dredge swept. In Rago et al. (2006) and NEFSC (2010c), the study site is subdivided into small square cells about twice the width of the dredge. Hennen et al. (2012) set $\gamma$ to 1 by reducing the cells to finely spaced points, which results in improved accuracy and precision of efficiency estimates.

A negative binomial distribution is used to describe the distribution of catch, as it accounts for extra variation in observed catches and can take into account catch from previous tows when estimating catch in tow $i$. This method uses the cumulative spatial pattern of animal removals to define the probability of capture and expected catch per tow. The negative binomial distribution of catch can be expressed as a function of $D_o$ (initial density of organisms), $k$ (the dispersion parameter), and $EAS$ (the effective area swept in tow $i$) (Rago et al. 2006): thus,

$$Pr(Ci|D_0, k, EAS) = \left(\frac{k}{D_0(EAS)+k}\right)^k \left(\frac{D_0(EAS)}{D_0(EAS)+k}\right)^{C_i} x \prod_{j=1}^{C_i} \frac{k+j-1}{j} \quad . \tag{5}$$

The log likelihood function gives the likelihood of the dispersion parameter, initial density, capture efficiency, and fraction of the cell hit, given the data for catch and area swept.

$$LL(D_0, k, e, \gamma \mid Ci, EAS) = k\sum_{i=1}^{I}(\log(k) - \log(D_0(EAS) + k)) + \sum_{i=1}^{I}(\log(D_0(EAS)) - \log(D_0(EAS) + k)) + \sum_{i=1}^{I}\sum_{j=1}^{Ci}\log(k + j - 1) - \sum_{i=1}^{I} Ci! \tag{6}$$

Rago et al. (2006) utilized the hit-matrix approach to simulate the number of clams caught in the dredge in each tow. The fractions $f_{i,j}$ are part of a square $n$ x $n$ hit matrix consisting of one row vector for each depletion tow and one column vector for each cell hit at least once. Each row represents an entire tow, thus the cells in a single row always sum to 1. Organisms remaining in a cell that was hit by the dredge are assumed to be mixed randomly within the cell after each tow. Hennen et al. (2012) changed the definition of the hit matrix, using points 10 cm apart rather than a grid of cells. The redefined $f_{i,j}$ is the ratio of the number of points hit *j* times by the end of each tow divided by the total number of points hit during the tow.

### 1.2.3 ***Patch Simulation Protocol***

Five thousand four hundred synthetic depletion experiments were generated using a defined distribution of clams with a given density and a set of dredge tow paths across the defined experimental area. The Simulation Model allows the dredge with an assigned efficiency to catch clams in its path. The Patch Model uses the catch per tow data generated from the simulated depletion experiment to estimate gear efficiency, clam distribution, and clam density in the area. The Simulation Model allows for various inputs to be treated as predictors of model performance by examining the effects of initial stock abundance, distribution, and fishing behavior on stock removal and Patch Model efficiency estimates.

The Simulation Model allows for the adjustment of factors some of which are normally unalterable or difficult to evaluate in designing an actual field experiment. Normally, the distribution and density of organisms on the bottom is unknown, although test tows in the area of interest might provide some information as to the uniformity of clam distribution. Dredge tow paths can be partly, but not completely controlled as tide and wind conditions affect vessel performance. The number of tows required is not known *a priori*.

In this study, the Simulation Model was specifically used to evaluate how the number of tows and amount of tow overlap in an experiment, the distribution and abundance of organisms, and the inherent gear efficiency affect the efficiency estimates using the Patch Model. A complete block design was implemented to support statistical analysis. Blocks included 4 levels of clam distribution (a relatively uniform distribution (no patches: NP), clams distributed in vertical bands (P), clams distributed primarily in the lower half of the area (HP), and clams distributed in a triangular wedge (T)) (see depictions in Figure 1.1), 3 numbers of tows per simulation (10, 20, 40) (Figure 1.2), 3 levels of clam density (0.75, 1.5, and 3 clams $m^{-2}$) (Figure 1.3), and 3 levels of inherent gear efficiency (0.2, 0.6, and 0.9) (Figure 1.4). Fifty simulations were run for each set of the 4 blocked variables, 108 tetrads of block variables in total, each followed by a Patch Model estimation of efficiency and density. Henceforth, each simulated depletion experiment will be called a simulation, and a set of 50 simulations for 1 tetrad of block variables will be called an experiment.

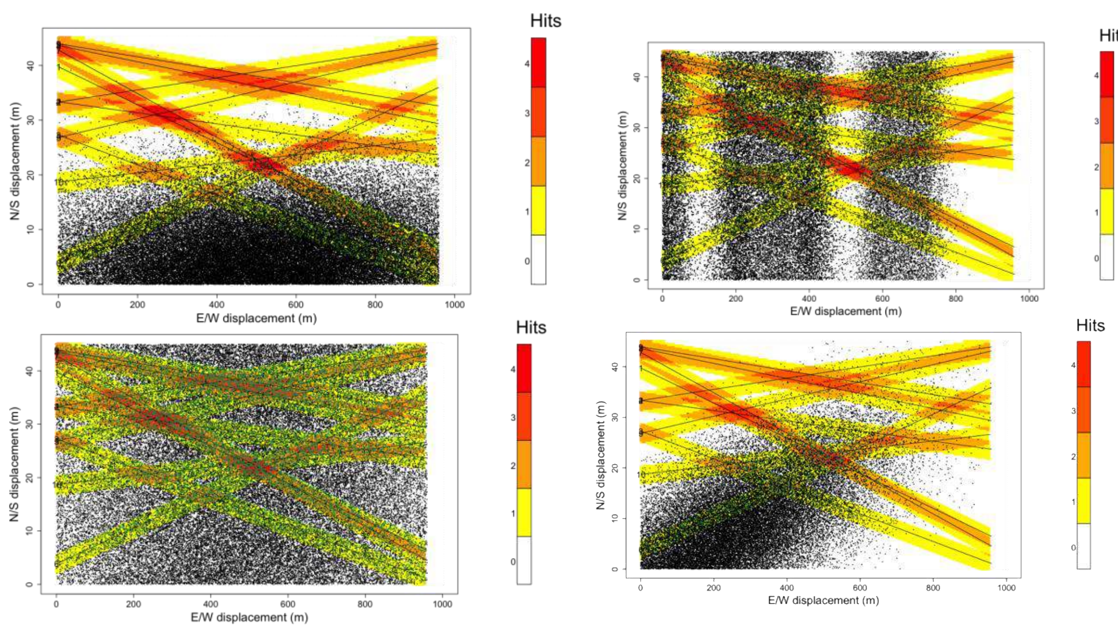
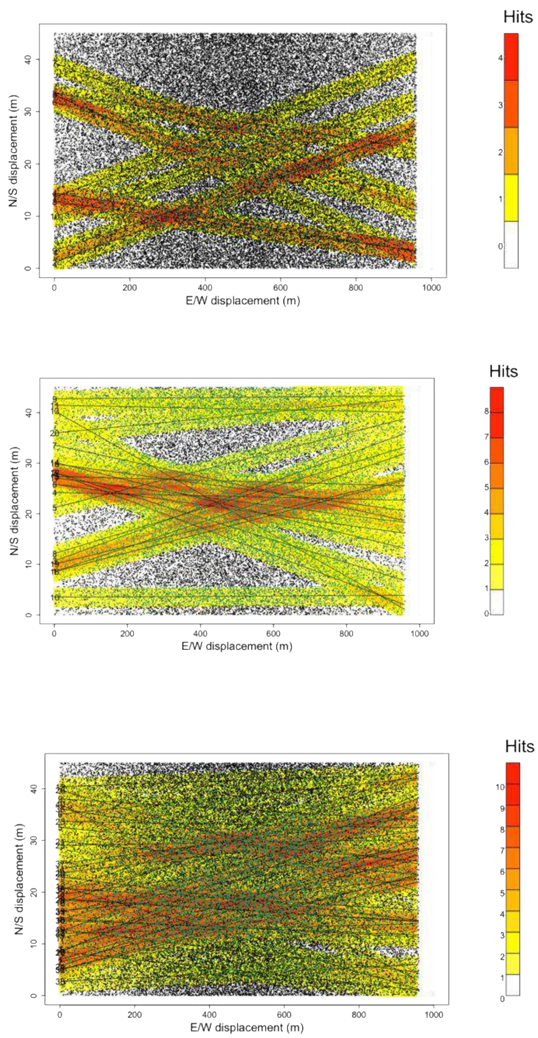



Figure 1.1 The different clam distributions with dredge tow paths, the straight colored lines, passing through the area. Colors denote the amount of overlap (number of hits) in the dredge paths. Dots are clams. Top left: a biased clam distribution with highest densities in half the area (HP). Top right: clams distributed in even vertical bands (P). Bottom left: not-patchy, clams distributed relatively uniformly (NP). Bottom right: clams distributed in a triangle wedge from south west to north east across the area (T).



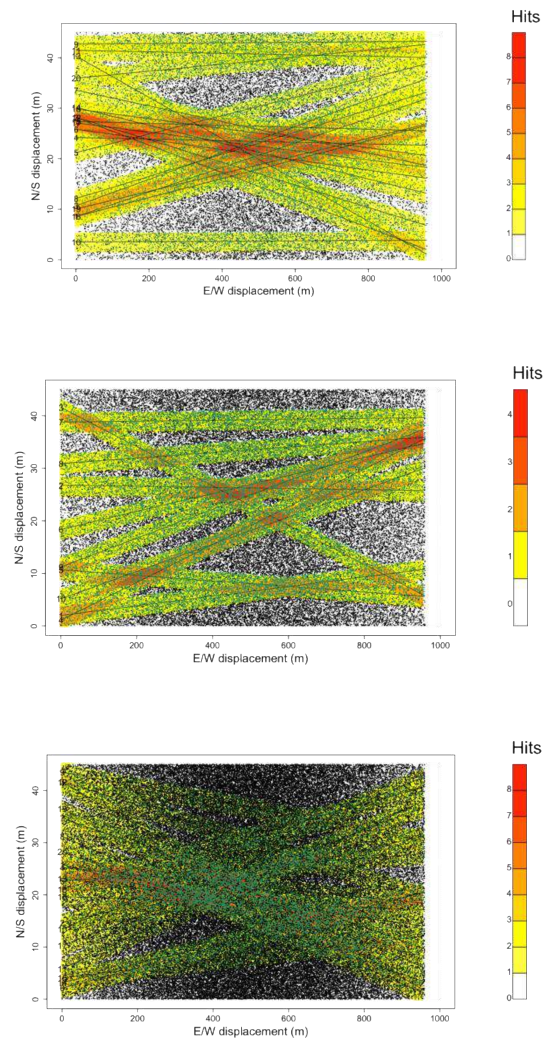




Figure 1.2 (Left) From top to bottom; an example using the NP distribution of the number of tows per simulation: 10, 20, and 40. The straight colored lines are dredge tow paths; colors denote the amount of overlap in the dredge paths. The dots are the clams.

Figure 1.3 (Right) From top to bottom; an example using the NP distribution of the different levels of clam density: 0.75, 1.5, and 3 clams $m^{-2}$. The straight colored lines are the dredge tow paths; colors denote the amount of overlap in the dredge paths. The dots are the clams.

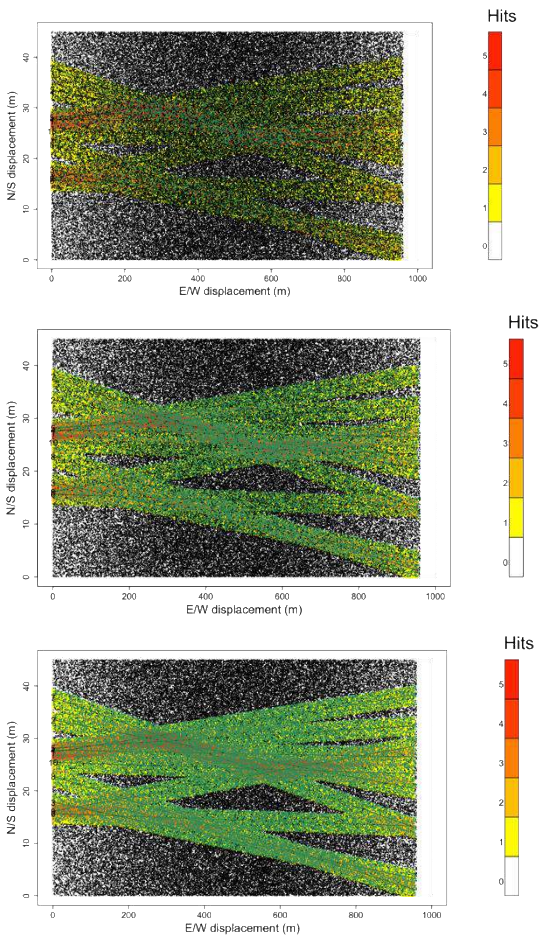



Figure 1.4 From top to bottom; an example using the NP distribution with a density of 3 clams $m^{-2}$ of the levels of true gear efficiency, 0.2, 0.6, and 0.9. Green dots are caught clams, black dots are uncaught clams.

For each experiment, 50 sets of tow paths were randomly generated for each tow number (50 tow paths were generated with 10 tows, 50 tow paths with 20 tows, and 50 tow paths with 40 tows), using the *runif* function in R (R Core Team, version 3.6.0) The same set of tow paths was used for every experiment with the same number of tows. The tow paths were generated by randomly generating a start point ($x_o$,$y_o$) and an end point ($x_1$,$y_1$) at the short ends of the rectangle (e.g., Figure 1.1) and linearly interpolating the tow path between the 2 points. The rectangular experimental area was kept constant at 960 x 45 m.

The depletion rectangle was populated with clams by placing clams randomly at position coordinates within the target area as specified from a multivariate normal distribution using the *mvrnorm* function in R. The number of organisms in the area is determined by the density level specified for the experiment. Catch was simulated by randomly sampling organisms within the tow paths with a probability of capture equal to the assigned efficiency of the dredge. When an organism is encountered by the dredge, a uniform (0,1) random number is drawn. If that number is less than the assigned efficiency value, the organism is considered captured and removed from the area. If the number is greater than the true efficiency, the organism is not captured, and it remains in the area.

#### 1.2.4 ***Statistics***

Type III SS ANOVAs were used to analyze main effects and interaction effects of tow number, clam distribution, clam density, and degree of tow overlap on the Patch Model error in the efficiency estimates, and the coefficient of variation of the efficiency estimate (CV) for simulations. An important evaluation of the simulation is the difference between the estimated efficiency obtained from the simulation and the true efficiency declared for the simulation. In the ideal case, the Patch Model would return the same efficiency it was given. The error in the efficiency estimate it designed to identify how closely the model was able to do that. The percent error in efficiency was calculated from the Patch model estimate of efficiency, *EstEff*, and the inherent efficiency specified, *TrueEff*, as:

$$Error = \frac{EstEff - TrueEff}{TrueEff} X\ 100. \qquad (7)$$

The CV was calculated as the standard deviation of the efficiency estimate divided by the mean of the efficiency estimates from the Patch Model (Equation 6):

$$CV = \frac{\sigma}{\acute{m}} X\ 100. \quad (8)$$

To evaluate the overlap in tows for each simulated experiment, the effective area swept (EAS) was calculated for each tow. The EAS measures the amount of previously towed ground over which the dredge was towed for each tow (Equation 4). A lower EAS indicates that more of the dredge tow path passed over ground that had been swept by the dredge in previous tows.

## 1.3 Results

### 1.3.1 ***Effects of tow number per simulation, clam density, and clam distribution on efficiency***

Simulations with higher numbers of tows and more even distributions of clams produce more reliable efficiency estimates. Clam density does not influence the accuracy of the efficiency estimate, but it can combine with an irregular clam distribution to decrease the precision of the efficiency estimate. At an inherent efficiency of 0.6, clam distribution, tow number, clam density, and their pairwise interaction terms significantly affected the error in efficiency estimates (Table 1.1). At inherent efficiencies of 0.9 and 0.2, significant effects of clam distribution, tow number and their interaction on the error in efficiency were retained, but clam density no longer exerted a significant main effect nor did any of its pairwise interactions.

Table 1.1

| P Values: Response = Error in Efficiency Estimate | Efficiency | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | <0.0001 | <0.0001 | 0.006 |
| Density | - | <.0001 | - |
| Number of Tows | <.0001 | <.0001 | <.0001 |
| Distribution x Density | - | <.0001 | - |
| Density x # Tows | - | - | - |
| Distribution x # Tows | <0.0001 | <0.001 | <0.0001 |
| Density x Distribution x # Tows | - | - | 0.017 |

P values from Type III SS ANOVA conducted on error in Patch Model efficiency estimates for 5393 simulations. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation. Only significant ($a$<0.5) results are shown, nonsignificance is denoted by a dash (-).

At inherent efficiencies of 0.9 and 0.6, clam distribution, tow number, and their interaction terms significantly affected the CV of efficiency estimates (Table 1.2). At an inherent efficiency of 0.2, the main effects of clam distribution and tow number still strongly influenced the CV of efficiency estimates, whereas the interaction term was barely significant.

Table 1.2

| P Values: Response = CV for Efficiency Estimate | Efficiency | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | <0.001 | 0.002 | 0.006 |
| Density | - | - | - |
| Number of Tows | <.001 | <.0001 | <.0001 |
| Distribution x Density | - | - | - |
| Density x # Tows | - | - | - |
| Distribution x # Tows | 0.046 | <0.001 | <0.0001 |
| Density x Distribution x # Tows | - | - | 0.0165 |

P values from Type III SS ANOVA conducted on CV in Patch Model efficiency estimates for 5393 simulations. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation. Only significant ($a$<0.5) results are shown, nonsignificance is denoted by a dash (-).

EMmeans analysis showed that significantly different error in efficiency estimates are primarily associated with low tow numbers, lower inherent efficiency levels, and more irregular clam distributions (Figure 1.5).

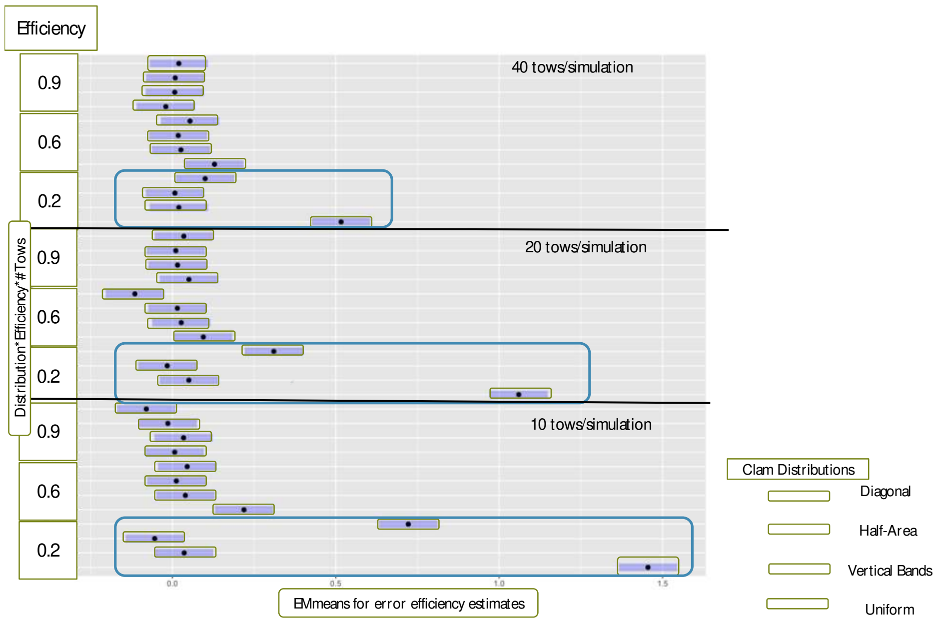



Figure 1.5 Plot of estimated marginal means (EMmeans) for error in the estimates of efficiency for groups of simulations conducted under a variety of efficiency levels, tows per simulation, and clam distributions. Each bar is 150 simulations, 50 for each of the 3 density levels, with the corresponding set of parameters. Colors correspond to the distribution of clams, Numbers on the y-axis are the true efficiency levels, horizontal black bars separate experiments by number of tows per simulation. Blue outlines emphasize the increase in error as the true efficiency decreases and the number of tows per simulation decreases

The efficiency estimates from the triangle (T) and half area (HP) clam distributions are typically significantly different from the efficiency estimates from the non-patchy (NP) and vertical banded (P) clam distributions. In the same vein, experiments run with 10 tows per simulation had more variability in the efficiency estimates than experiments run with 20 and 40 tows per simulation. The error in efficiency estimates at 10 tows per simulation was much higher than the error in efficiency estimates at 20 and 40 tows per simulation, and significant differences existed in the T and HP distributions for experiments with 10 tows, 20 tows, and 40 tows per simulation at all true efficiency levels. In contrast to the T and HP distributions, the EMmeans for the error in efficiency estimates for the NP and P distributions were not significantly different across differing tow numbers and efficiency levels.

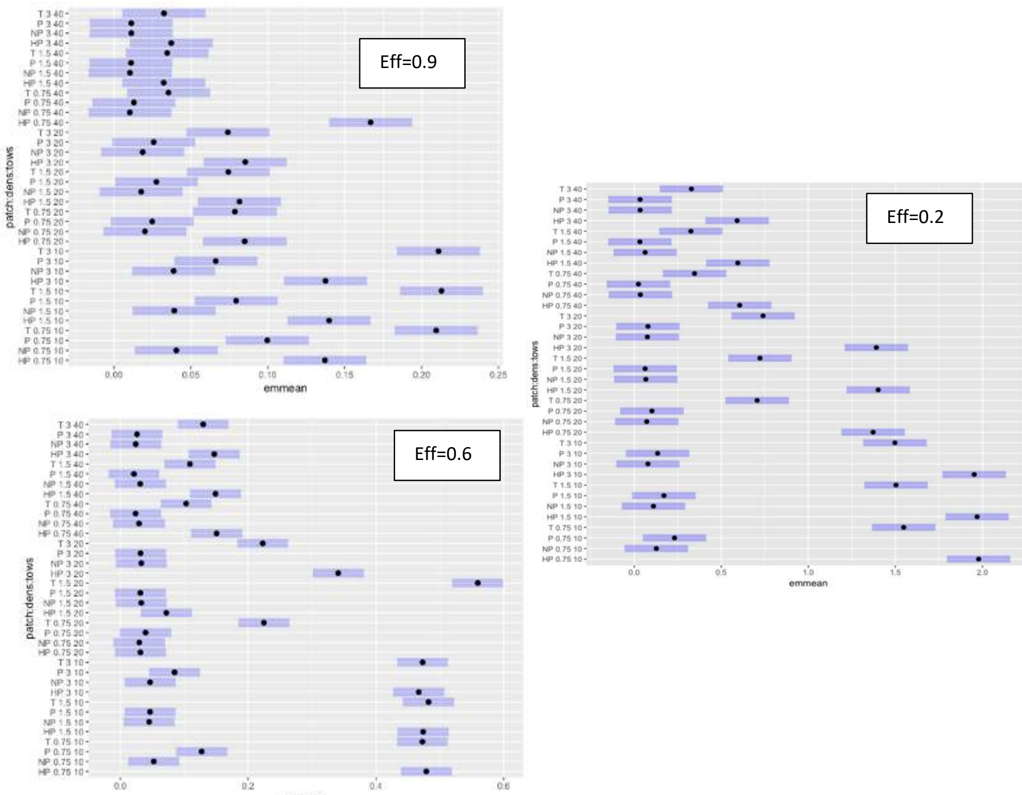



Figure 1.6 Estimated marginal means (EMmeans) of the absolute value of the error in efficiency estimates with clam distribution, clam density, and number of tows as variables. Each bar is 50 simulations for each parameter. Top is an efficiency of 0.9, middle is an efficiency of 0.6, and bottom is an efficiency of 0.2. Note the different scales on the x-axis for each efficiency level. Ordinate labels correspond to the distribution-density-number of tows.

The error in efficiency estimates varied significantly among density levels at an efficiency of 0.6, as density exerted a significant main effect at this efficiency level (Table 1.2, Figure 1.6). Only one experiment at an efficiency of 0.6 revealed a significant difference in error in efficiency estimates due to density, however; that one being the set of simulations with the T (triangle) clam distribution, 20 tows per simulation, and 1.5 clams $m^{-2}$.

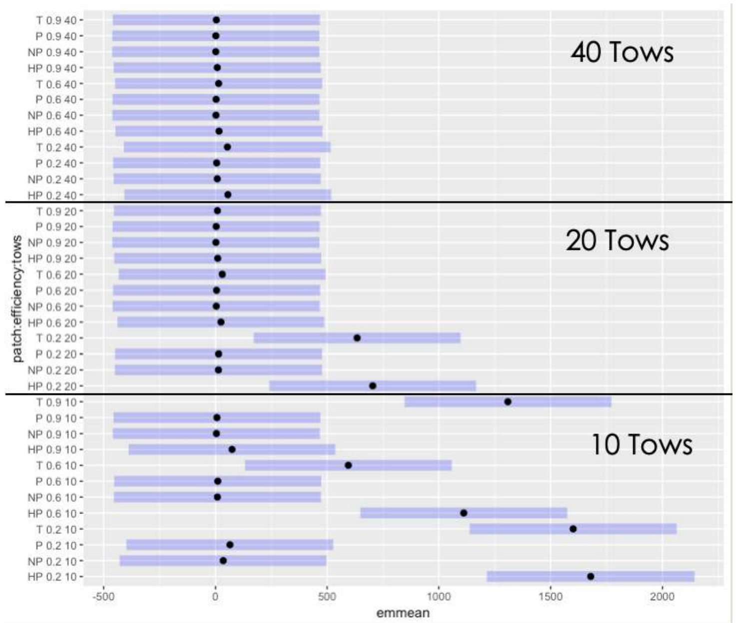



Figure 1.7 Estimated marginal means (EMmeans) of the CVs of the efficiency estimates of simulated depletion experiments with clam distribution, clam density, and number of tows as variables. Each bar is 150 simulations, 50 for each clam density level. The figure is divided vertically by the number of tows in each simulation (from top to bottom: 40, 20, 10). Ordinate labels correspond to the distribution of clams – true efficiency – number of tows per simulation.

Comparisons of CVs among the different tow numbers and clam distributions for each declared true efficiency using EMmeans are shown in Figure 1.7. At an inherent efficiency of 0.9, the EMmeans for CVs for the experiment with the T (triangular wedge) clam distribution and 10 tows was significantly different from all other CV EMmeans. At an inherent efficiency of 0.6, experiments with 10 tows, the T and HP (half of the area) clam distributions did not differ significantly from each other but each differed significantly from all over CV EMmeans. At an efficiency of 0.2, the T and HP experiments with 10 tows have CVs that are no longer significantly different from the HP experiment with 20 tows. None of the experiments conducted with 40 tows per simulation had significantly different CVs.

### 1.3.2 *Effect of dredge tow overlap on Patch Model efficiency estimates*

The mean effective area swept (EAS) per tow was calculated for each simulation as an estimate of the degree of overlap amongst dredge tows (Figure 1.8). At lower tow numbers, dredge overlap and clam distribution significantly affected the error in efficiency estimates. At all inherent efficiency levels for experiments with 10 tows, the clam distribution and the interaction between clam distribution and EAS exerted significant effects on the error in efficiency estimates (Table 1.3). EAS as a main effect was also significant at true efficiencies of 0.2 and 0.9. In contrast, EAS did not exert a significant main effect on the error in efficiency estimates at an inherent efficiency of 0.6, indicating that the amount of tow overlap did not significantly affect the error in the efficiency estimates for these simulations. No significant main or interaction effects were observed with simulations with 20 tows, whereas, at an inherent efficiency of 0.9, experiments with 40 tows produced a significant main effect on the error in efficiency for clam density, and all interactions exerted significant effects (Table 1.4). As in the case with 20 tows, in contrast, no significant main effects or interaction terms were observed at inherent efficiencies of 0.6 and 0.2 with 40 tows.

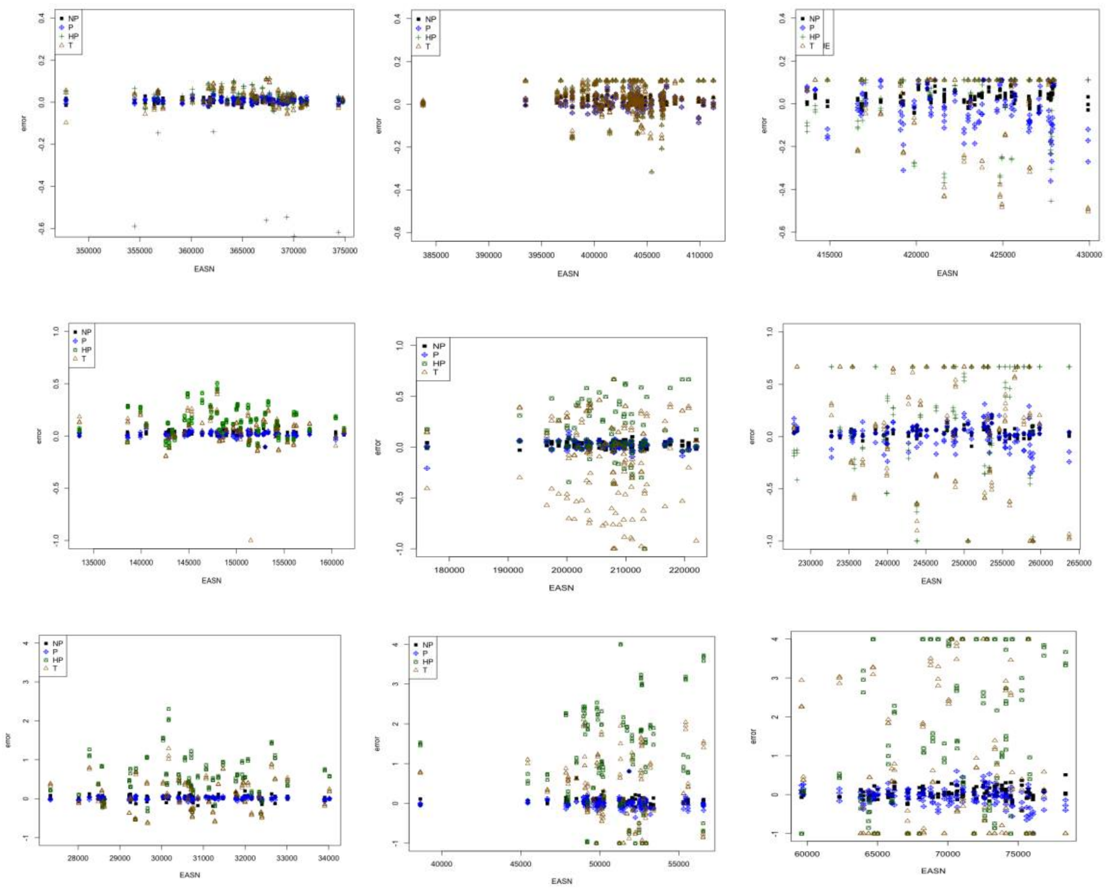



Figure 1.8 Error in efficiency estimates as a function of EAS for simulations. Plots are orientated into columns based on the number of tows per simulation: 40 tows (left), 20 tows (middle) and 10 tows (right). Plots are oriented into rows by efficiency levels: 0.9 (top), 0.6 (middle), 0.2 (bottom). Thus, the upper left plot provides results for 40 tows per simulation and an inherent efficiency of 0.9. Note that the ordinate range varies substantially by row.

The CVs of the efficiency estimates were evaluated to determine if the CVs were significantly affected by the same factors that significantly affected the error in efficiency

Table 1.3

| P Values: Response = Error in Efficiency Estimate | Efficiency, 10 tows/simulation | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | <0.0001 | <0.0001 | <0.0001 |
| Density | - | - | - |
| EAS | 0.025 | - | <.001 |
| Distribution x Density | - | - | - |
| Density x EAS | - | - | - |
| Distribution x EAS | <0.0001 | <0.0001 | <0.0001 |
| Density x Distribution x EAS | - | - | - |

P values from Type III SS ANOVA conducted on error in Patch Model efficiency estimates for 1799 simulations with 10 tows. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation. Only significant (*a*<0.5) results are shown, nonsignificance is denoted by a dash (-).

Table 1.4

| P Values: Response = Error in Efficiency Estimate | Efficiency, 40 tows/simulation | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | - | - | - |
| Density | - | - | 0.011 |
| EAS | - | - | - |
| Distribution x Density | - | - | 0.002 |
| Density x EAS | - | - | 0.013 |
| Distribution x EAS | - | - | 0.06 |
| Density x Distribution x EAS | - | - | 0.003 |

P values from Type III SS ANOVA conducted on error in Patch Model efficiency estimates for 1798 simulations with 40 tows. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation.Only significant (*a*<0.5) results are shown, nonsignificance is denoted by a dash (-).

estimates. Higher CVs were observed with 10 tows than with 20 and 40 tows for all efficiency levels (Figure 1.7). At an inherent efficiency of 0.9, simulations with 10 tows produced a significant main effect for clam distribution on the CV of the efficiency estimates. The EAS exerted a significant main effect at an inherent efficiency of 0.6. No main effects or interaction terms proved to be significant with an inherent efficiency of 0.2

(Table 1.5). For simulations with 20 tows, no variable produced a significant effect on CV nor were any interaction terms significant. Experiments with 40 tows exhibited significant main effects and interactions only when the inherent efficiency was set at 0.9. In this case, clam density and most pairwise interactions were significant (Table 1.6).

Table 1.5

| P Values: Response = CV for Efficiency Estimate | Efficiency, 10 tows/simulation | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | - | - | 0.024 |
| Density | - | - | - |
| EAS | - | 0.006 | - |
| Distribution x Density | - | - | - |
| Density x EAS | - | - | - |
| Distribution x EAS | - | - | - |
| Density x Distribution x EAS | - | - | - |

P values from Type III SS ANOVA conducted on CV in Patch Model efficiency estimates for 1799 simulations with 10 tows. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation. Only significant (*a*<0.5) results are shown, nonsignificance is denoted by a dash (-).

Table 1.6

| P Values: Response = CV for Efficiency Estimate | Efficiency, 40 tows/simulation | | |
|---|---|---|---|
| Variable | 0.2 | 0.6 | 0.9 |
| Distribution | - | - | - |
| Density | - | - | 0.033 |
| EAS | - | - | - |
| Distribution x Density | - | - | 0.006 |
| Density x EAS | - | - | 0.038 |
| Distribution x EAS | - | - | - |
| Density x Distribution x EAS | - | - | 0.008 |

P values from Type III SS ANOVA conducted on CV in Patch Model efficiency estimates for 1799 simulations with 40 tows. Columns are the true efficiency levels. Rows are the parameters tested: clam distribution, clam density, and effective area swept (EAS) in each simulation. Only significant (*a*<0.5) results are shown, nonsignificance is denoted by a dash (-).

## 1.4 Discussion

The simulations show that low tow number, certain patchy distributions, and low effective area swept (EAS) generate the largest deviations in estimated efficiency from the true efficiency. High tow numbers, which also ordinarily generate low EAS (indicating more dredge overlap), and uniform clam distributions routinely conduce highly accurate efficiency estimates. Save for rare occurrences, clam density has no significant influence on the efficiency estimate.

The error in efficiency estimates, the CVs, and the differences in efficiency estimates across all four defining variables - tow number, inherent efficiency, clam distribution, and clam density - indicate that the depletion experiments yielding the most accurate efficiency estimates are those characterized by a high number of tows. No significant differences in CV were observed for simulations with 20 tows and 40 tows, but experiments with 10 tows per simulation were more likely to be distinguished by higher uncertainty (Figure 1.7). The error in efficiency estimates, a direct evaluation of the ability of the Patch Model to return a known efficiency, shows the same pattern, with a clear trend towards improved performance as tow number increases. Even at patchy distributions and low densities of clams, the Patch Model is more likely to produce an accurate efficiency estimate with a high number of tows, thus tow number is a controlling variable capable of mitigating the influence of conditions inducing uncertainty (Figures 1.6 and 1.8). Simulation results show little improvement in the accuracy of efficiency estimates between experiments with 20 and 40 tows, especially with relatively uniform distributions of clams in an area, but the differential becomes apparent when certain highly biased clam distributions are present.

A metric describing the error in efficiency estimates is more informative for understanding the causes of variation in efficiency estimates than is the CV. Unfortunately, while the convenience of assigning an inherent efficiency for a dredge in a simulated depletion experiment is appreciated, that luxury is not afforded to fisheries scientists during a field depletion experiment. Thus, the error estimate used here, while valuable in assessing the results of simulations, is not available as a metric for distinguishing reliability in efficiency estimates for field experiments. The Patch Model outputs include an efficiency estimate obtained from the maximum likelihood estimate and a corresponding standard deviation for the estimate. For a field experiment, the CV is the only option for quantifying the accuracy of the Patch Model efficiency estimate.

Clam density rarely had a significant effect on the CV or error in efficiency estimates. The few significant results are associated with clam distributions in which clam density varies strongly across the narrow dimension of the depletion domain and how these distributions interact with tow path overlap in the areas with clams, rather than an inherent effect of clam density overall. Generally, a hydraulic dredge should be equally efficient over a wide range of clam densities. The simulations support this expectation.

Effective area swept (EAS) provided useful information in describing the likely error in efficiency estimates at 10 tows per simulation. Tow overlap significantly affected outcomes at low tow number. At 20 or 40 tows per simulation, the error in the efficiency estimate is not significantly affected by the degree of tow overlap. Thus, EAS and tow number provide equivalent expectations about the quality of a depletion experiment assuming that the vessel operator is limited in their ability to carefully determine tow location and direction. In field experiments, tow location and tow linearity are commonly controlled as the experiment proceeds (NEFSC 2007); thus, we did not simulate patterned distributions of dredge tow paths in this study. Error in the assumed position of the dredge has previously been shown to reduce accuracy and precision in the estimation of efficiency (Hennen et al. 2012, Wilberg et al. 2013).

This analysis confirms findings from Hennen et al. (2012) that under low dredge-efficiency conditions, the number of tows per simulation and the amount of overlap in the dredge tow paths significantly affect the error in efficiency estimates to a greater degree than they do for gear with inherently high efficiency. Table 1.3 shows that at 10 tows per simulation the distribution of clams and the degree of dredge overlap quantified by the EAS significantly affect the error in the efficiency estimates, but this is not the case at 20 and 40 tows. Across all efficiency levels, the most error in the efficiency estimates is observed at low tow numbers.

Higher error values are seen when clam distributions are less uniform across the domain of the experiment. This is true across all efficiency conditions and numbers of tows per experiment, with the highest error values observed under combinations of low tow numbers per experiment, low inherent efficiency, and strongly non-uniform clam distributions. When clams are irregularly distributed across the depletion rectangle and the numbers of tows is low, the dredge tows may overlap multiple times in an area with low clam density, biasing the efficiency estimate. Increasing tow number counteracts this problem by encouraging a more even dredge tow overlap throughout the area. Given this, the fact that EAS did not significantly affect error or CV in efficiency estimates at higher numbers of tows per simulation is not surprising.

However, even experiments with 40 tows can occasionally produce poor results. For example, one set of 50 simulations with 40 tows per simulation, a clam distribution biased towards the lower half of the depletion area with a clam density of 0.75 $m^{-2}$ produced higher error and CV values than other 40-tow experiments with the same clam density and distribution. (Figure 1.9). Here, by chance, a high amount of dredge overlap in the portion of the domain with low clam density produced inaccurate efficiency estimates. The tendency for outliers to occur is seen across simulations with all tow numbers, with increasing frequency as tow number declines, but is most pronounced with the HP clam distribution. In this case, the upper half of the experimental area has very few clams, generating a higher likelihood for tow overlap to occur in the low-density portion of the domain. Simulations show that the effect of patchiness is only pronounced when the patchiness is orthogonal to the short dimension of the depletion domain, thereby reducing the probability that a tow will fairly sample the range of patchiness in the domain.

Simulations suggest that a good depletion experiment can be characterized as having a high number of tows, between 20 and 40, with dredge paths that overlap multiple times, but distributed evenly throughout the studied area. Multiple tows in an experiment reduce the likelihood that the dredge paths will overlap multiple times exclusively over ground with low densities of clams when the domain is characterized by aggregated clam distributions. The influence of clam distribution is sufficiently pervasive that it could be beneficial to have potential sites for depletion experiments evaluated remotely by divers or video to get an understanding of how clams are distributed in space. Of course, such a capability may reduce the need for the depletion experiment. A recent analysis used Habcam camera system tows along with dredge tows to estimate abundance of sea scallops, for example (Miller et al. 2019), and Thórarinsdóttir et al. (2010) employed divers to ascertain the efficiency of hydraulic clam dredges in shallow water.

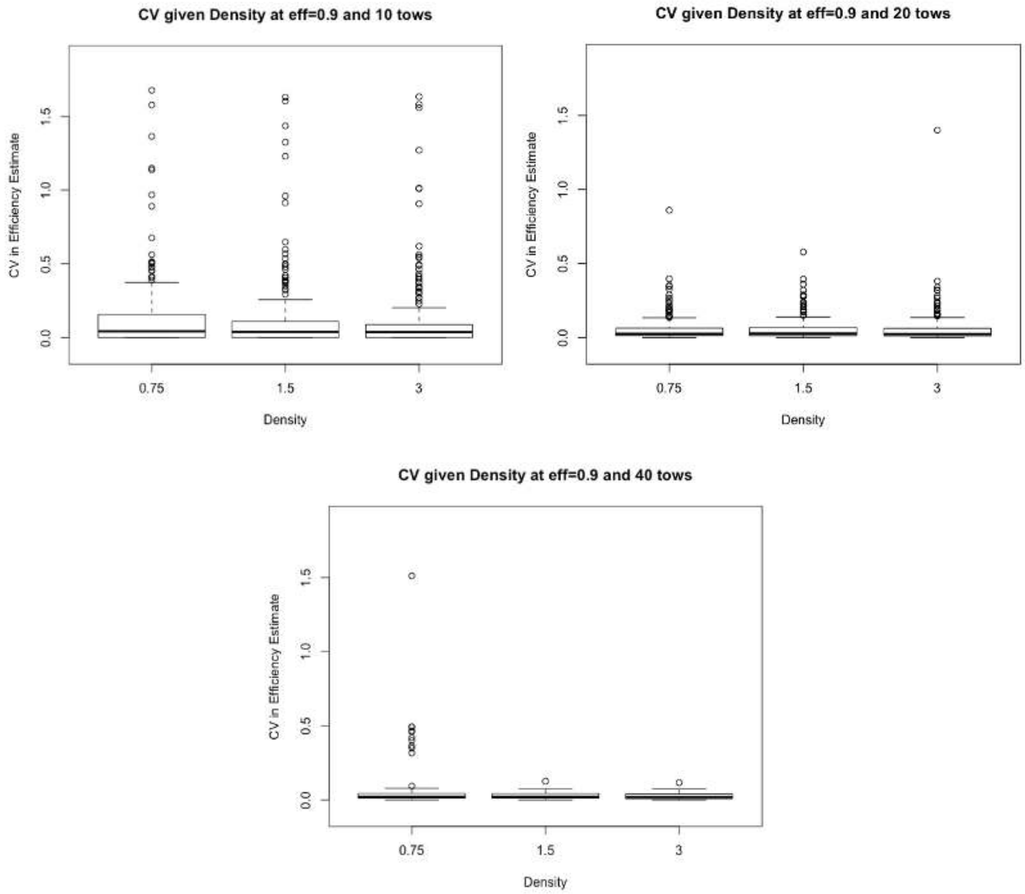



Figure 1.9 Boxplots for CV in efficiency estimates as a function of clam density for 1798 simulations with inherent efficiencies of 0.9. The plots are orientated by number of tows per simulation, from top to bottom: 10 tows, 20 tows, 40 tows. The bottom and top borders of the box represent the first and third quartile. The whiskers represent the 10th and 90th percentiles.

Comparison of simulated depletion experiments to in-field depletion experiments is difficult because little empirical data will be available in the field to guide experimental design. The range of variability in the simulated clam distributions and densities includes the range observed in the field surfclam and ocean quahog experiments summarized in NEFSC (2017a, 2017b), supporting the validity of the simulated experiments. Depletion experiments are costly and time consuming at sea. In-field experiments of 20 or more tows typically take 10-20 hours of continuous sampling and place an inordinate burden on the captain and the scientific crew. Accordingly, experimental design aimed at limiting the number of tows to the degree possible is critical. Unfortunately, tow number dominates the determinants of error in efficiency estimates and the secondary but also critical effector, clam dispersion, is inherently difficult to observe. Clam density is of little consequence, but inherent gear efficiency is consequential. Hydraulic dredges, being inherently high efficiency gears, limit that degree of consequence, thus placing emphasis on tow number and the degree of tow overlap in experimental design.

Gear efficiencies are influenced by a wide number of relatively uncontrollable factors such as sediment type, bottom current force, sea state, etc. These add uncertainty that cannot be easily constrained. Yet, a correction for gear efficiency is frequently the largest correction factor in determining true abundance from the survey index. A variety of gear calibration methods have been used, including diver quadrat sampling (Powell et al. 2007, Thórarinsdóttir et al. 2010, Morson et al. 2018), hydraulic patent tongs (Chai et al. 1992, Mann et al. 2004), and video (Giguère & Brulotte 1994), all of which are highly efficient sampling methods in shallow water or for epibenthos. Options are limited for infauna on the continental shelf, however, with the depletion experiment being a method of choice. Thus, attention to reducing uncertainty in depletion experiments is important.

Given attention to a reasonable dispersion of tows in the depletion rectangle, how many tows are enough for a depletion experiment to produce an accurate efficiency estimate? Certain metrics may inform this decision during the depletion experiment, assuming that data collected are evaluated using the Patch Model continuously during the experiment. Of greatest use may be the trends in CV and EAS to measure the uncertainty of the dredge efficiency estimate and the amount of dredge overlap as the depletion experiment continues. Adjustments in tow number are readily made on the fly and EAS can be modified at least to some degree by controlling tow location relative to the existing hit matrix that can be updated with each succeeding tow. Preliminary simulation work suggests that adapting the experimental protocol during the depletion experiment may substantively reduce uncertainty in the final efficiency estimate.

CHAPTER II – An Analysis of Existing Depletion Experiments

## 2.1 Introduction

The implementation of a definitive measure of dredge efficiency for shellfish survey data substantially improves the estimation of abundance. Commonly, depletion experiments are used to estimate gear efficiency and population density in a target area (Gedamke et al. 2005, Lasta & Iribarne 1997, Leslie & Davis 1939, Skalski et al. 1983, Wilberg et al. 2013), although other quantification methods have also been used (Chai et al. 1992, Ragnarsson & Thórarinsdóttir 2002, Morson et al. 2018) . Efficiency estimates exist for a range of dredge types, including oyster dredges (Morson et al. 2018, Powell et al. 2007), clam dredges (Pezzuto et al. 2010), crab dredges (Vølstad et al. 2000; Wilberg et al. 2013), and scallop dredges (Beukers-stewart & Beukers-stewart, 2009; Lasta & Iribarne, 1997). These are dry dredges however, that are designed to harvest epibenthic animals. By comparison, highly efficient hydraulic dredges are the primary gear type used to harvest infaunal clams (Da Ros et al. 2003, Hauton et al. 2003, Moschino et al. 2003, Gilkinson et al. 2005, Meseck et al. 2014). Hydraulic dredges increase the catchability of the target organism by using water pressure to liquefy the sediment (Da Ros et al. 2003; Gilkinson et al. 2003; Hauton et al. 2007; Meseck et al. 2014).

A series of depletion experiments was conducted between 1997 and 2013 on commercial clam vessels targeting the Atlantic surfclam (*Spisula solidissima*) and the ocean quahog (*Arctica islandica*) to estimate the efficiency of both commercial hydraulic dredges and the National Marine Fisheries Service (NMFS) survey hydraulic dredge. The locations of these depletion experiments are specified in Appendix 3 of NEFSC (2017a) (Figures 2.7 and 2.8). As is often the case in measures of dredge efficiency (Vølstad et al.

2000, Powell et al 2007, Hennen et al. 2012, Wilberg et al. 2013, Morson et al. 2018), individual experiments varied widely in their estimates of efficiency. Very little is known about how environmental and sampling variables influence the efficiency of a hydraulic dredge, and these are probably the source of between-experiment variation. Estimates of gear efficiency based on depletion experiments are a key source of uncertainty in the stock assessments for the Atlantic surfclam and ocean quahog stocks as a consequence.

The ocean quahog and the Atlantic surfclam support substantial fisheries on the northeast U.S. continental shelf and are harvested exclusively by hydraulic dredges. A typical hydraulic dredge is a large rectangular box between 8 and 13 ft wide constructed of evenly spaced steel bars that is towed over a seabed (Lambert & Goudreau 1996, Meyer et al. 1981). A manifold at the head of the dredge distributes high-pressure water provided by an onboard water pump through a connecting hose. The water is focused through a series of jets onto the seabed, liquefying the sediment and unearthing the clams for easy capture by the dredge. Hydraulic dredges, widely used in clam fisheries, are necessary for the commercial harvest of ocean quahogs and surfclams due to the depth and distance from the shore at which these species are found and the fact that the fishery is based on a high volume-low unit cost product. Thus rapid and efficient capture methods are economically essential.

Ocean quahogs are typically found offshore in deep water, between 30 m and 200 m on the U.S. east-coast continental shelf (NEFSC 2017b) and are the longest-lived non-colonial marine species (Butler et al. 2013). The Atlantic surfclam is found closer to shore between 8 and 66 m deep and has a lifespan of about 30 years. Surfclams coexist with ocean quahogs along their offshore range boundary that approximately follows the 15°C summer bottom water temperature isotherm (NEFSC 2017a, Powell et al. 2020).

The Patch Model was developed to analyze the results of depletion experiments to estimate the efficiency of capture of sedentary species such as surfclam and ocean quahogs (Rago et al. 2006). The Patch Model has been important in informing stock assessments of commercially exploited populations of Atlantic surfclam, ocean quahog, monkfish (*Lophius americanus*: NEFSC, 2010a) and Atlantic sea scallop (*Placopecten magellanicus*: NEFSC, 2010b; NMFS, 2009). Over the span of 14 years, 50 depletion experiments were carried out along the U.S. Mid-Atlantic continental shelf off the coast of Massachusetts, Long Island, New Jersey, and the Delmarva Peninsula to determine the efficiency of hydraulic dredges used in the surfclam and ocean quahog fishery and by the NMFS survey vessel. The Patch Model estimates capture efficiency (the probability of capture for an organism in the tow path), and average density of organisms in the target area (numbers per $m^2$) by tracking the relative depletion (reduction in catch) over the tow series. Theoretically, capture efficiency is a measurable characteristic of the gear (Hennen et al. 2012).

A field depletion experiment is designed to have the dredge towed over the same ground repeatedly while covering the majority of a predetermined area; typically, in the experiments considered herein, a long rectangular area, on average about 10 dredge widths wide (23-24 m) and 400-1000 m long. A series of overlapping dredge tows are taken across the selected area, with the dredge hitting bottom at one of the short edges of the rectangle and being retrieved at the other short edge. The experiment requires the assumption that all catches are random samples and that no transport of organisms into or out of the study site occurs during the experiment (Leslie & Davis 1939). The catch per tow, the dredge positions along each tow, and the fishing effort are recorded for each tow. Over a series of tows, the catch per tow will decrease; this rate of decline is proportional to the efficiency of the dredge (Hennen et al. 2012). If the rate of decline is steep, the dredge is highly efficient.

Field depletion experiments can take hours to complete and require much effort on the part of scientists and crew on the ship; thus it is important to know if the experiments that have been conducted produced reliable efficiency estimates for the gear used and to evaluate characteristics leading to poor performance that might be avoided in future endeavors. The NMFS depletion dataset is unique; at the time of this writing, no other depletion dataset is this expansive. Analyzing the NMFS hydraulic dredge depletion experiments may provide an improved understanding of the sources of uncertainty in efficiency estimates derived in this way, including characteristics of the experimental protocol and the environmental factors affecting gear efficiency.

## 2.2 Methods

### 2.2.1 *The Patch Model*

To estimate the catchability coefficient, the depletion experiments permit correction of survey catch using the equations $N = SA/q$ and $q = \frac{\alpha e}{A}$, where $N$ is stratum stock abundance or biomass and $SA$ is the swept area average of all tows in a stratum. The catchability coefficient $q$ is obtained from α, the area swept by the sampling gear, $e$, the dredge efficiency, and $A$, the spatial domain of the estimates (Paloheimo & Dickie 1964). The area swept by the dredge is calculated as the distance the dredge is towed times the width of the dredge.

The expected catch of organisms in any tow $i$, $E(C_i)$, given initial density of the target organisms ($D_0$) and the cumulative catch from previous tows, $T_{i-1}$, can be calculated as:

$$E(C_i) = q(D_0 - T_{i-1})\ . \tag{1}$$

Incorporating the portion of the area that has already been hit by the dredge prior to tow *i*, also known as the hit matrix, gives the expected catch for tow *i* as:

$$E(C_i) = (EAS)D_o \tag{2}$$

where $D_0$ is the initial density of the target organism in the area and EAS is the effective area swept defined as the total area swept ($m^2$) by the dredge in tow *i* taking into account the portion of the experimental area hit by the dredge in previous tows. EAS is a measure of tow overlap in a depletion experiment: higher EAS values indicate more untouched area being covered by the dredge in each tow, and lower EAS values indicate more dredge overlap in each tow. EAS is calculated as:

$$EAS = ea_i \sum_{j=1}^{i} f_{i,j}(1 - e\gamma\ )^{j-1} \tag{3}$$

where *e* is the capture efficiency as estimated by the Patch Model, $a_i$ is the area swept by tow *i*, $f_{i,j}$ is the fraction of the area $a_i$ that was hit by the dredge *j* times in previous tows, and γ is the ratio of the cell size and the dredge width. Rago et al. (2006) divided the experimental area into cells twice the width of the dredge. Hennen et al. (2012) removed γ by reducing the cells to points, eliminating the need to calculate cell size, which results in improved accuracy and precision of efficiency estimates.

Accounting for extra variation in observed catches and taking into account catch from previous tows when estimating catch in tow *i* requires the use of a negative binomial distribution to describe the catch distribution. This method uses the cumulative spatial pattern of animal removals to define capture probability for each organism. The negative binomial distribution of catch can be expressed as a function of $D_o$ (initial density of organisms), $k$ (the dispersion parameter), and EAS (the effective area swept in tow $i$) (Rago et al, 2006): thus,

$$Pr(Ci|D_0, k,\ EAS) = \left(\frac{k}{D_0(EAS)+K}\right)^{k} \left(\frac{D_0(EAS)}{D_0(EAS)+k}\right)^{C_i} x \prod_{j=1}^{C_i} \frac{k+j-1}{j} \quad . \tag{4}$$

The log likelihood function gives the likelihood of the dispersion parameter, initial density, capture efficiency, and fraction of the cell hit, given the data for catch and area swept.

$$LL(D_0, k, e, \gamma \mid Ci, EAS) = k \sum_{i=1}^{I}(log(k) - log\ (D_0(EAS) + k)) + \sum_{i=1}^{I}(log\big(D_0(EAS)\big) - log\ (D_0(EAS) + k)) + \sum_{i=1}^{I} \sum_{j=1}^{Ci} log(k + j - 1) - \sum_{i=1}^{I} Ci! \tag{5}$$

### 2.2.2 *Simulated Datasets*

Poussard et al. (in prep.) report the results of 9,000 simulated depletion experiments conducted in a block design in which animal density, true dredge efficiency, the number of tows per experiment, and the dispersion of animals on the bottom were varied (Table 2.1).

Table 2.1

| | | | | | |
|---|---|---|---|---|---|
| **True Efficiency** | 0.9 | 0.6 | 0.2 | | |
| **Number of Tows** | 40 | 25 | 20 | 15 | 10 |
| **Clam Density (# m$^{-2}$)** | 0.75 | 1.5 | 3.0 | | |
| **Clam Distribution** | NP | HP | P | T | |

Metrics used in the simulation analysis block design of Poussard et al. (in prep.). All combinations of the four parameters were simulated: 50 simulations for each tetradic combination were conducted. Clam distribution is denoted as NP: uniform across the area, P: patches oriented across the narrow dimension, HP: patches oriented longitudinally, and T: patches of a triangular nature emanating from one side of the rectangle (see Poussard et al. (in prep) Figure 2.1).

The Patch Model provides 4 useful metrics of error measure for comparison besides the estimates of efficiency and density. The four metrics are the average effective area swept (EAS), the overlap score describing tow overlap, the coefficient of variation (CV) for the efficiency estimate, and the CV of the *k* parameter, the negative binomial dispersion parameter. The CVs were calculated as the delta method standard deviation of the Patch Model estimates divided by the means of the estimates obtained from the log likelihood equation (Equation 5):

$$CV=\frac{\sigma}{\bar{m}}x100. \quad (6)$$

The overlap score (OS) is a metric describing tow overlap that does not depend on estimated efficiency, the number of tows in an experiment, or the spatial dimensions of the site. OS was used to compare simulated depletion experiments and field experiments in this analysis. OS is derived directly from the hit matrix (Hennen et al. 2012) where the *n* rows equal the number of tows in the experiment and the *m* columns are the number of points touched *m* times previously. The most possible overlap for any depletion site would be the exact duplication of the longest tow in each sequence (the row with the most total points touched) *n* times ($OS_{max}$). For tow *i*:

$$OS_i = \sum_{h=i}^{m}(p_{i,h}h) \quad (7)$$

where $p_{i,h}$ are the number of points in the hit matrix row *i* and column *h*. The OS for each tow sequence is then

$$OS = \frac{\sum_i^n OS_i}{OS_{max}} \text{ X } 100 \quad (8)$$

Where *n* is the total number of tows in the sequence. A higher value of OS equates to more overlap in the tow structure of an experiment. OS is correlated with the number of tows ($r^2 = 0.75$; $P<0.0001$), but retains more information about the tow structure than simply the number of tows.

For the simulations, where the true efficiency was already known, Poussard et al. (in prep.) calculated the percent error in efficiency from the Patch Model estimate of efficiency, EstEff, and the inherent efficiency specified in the simulation, TrueEff, as:

$$\text{Error} = \frac{\text{EstEff} - \text{TrueEff}}{\text{TrueEff}} \text{ X } 100. \quad (9)$$

Analysis of simulated depletion experiments by Poussard et al. (in prep.) showed that the uncertainty in the estimate of gear efficiency from depletion experiments was reduced by higher numbers of dredge tows per experiment, increased tow overlap in the experimental area, a homogeneous as opposed to a patchy distribution of clams in the experimental area, and the use of gear of inherently high efficiency. The results of these simulations were compared to the field depletion experiments using the set of 4 metrics to match the field experiments to simulated experiments with similar characteristics. The known errors in the set of comparable simulated experiments were then used as a proxy.

2.2.3 ***Application of Simulations: Error Estimates***

Experiments varied in the length of the depletion site and the width of the dredge used. For statistical analysis, EAS was standardized to a dredge width of 12.5 ft (3.8 m) and a site length of 960 m consistent with the simulation dataset of Poussard et al. (in prep.). All EAS values used were the average values per tow, rather than the total values, to take into account the large range in tow numbers among experiments. A Principal Components Analysis was conducted on the simulation dataset to determine if the 4 metrics describing depletion performance were correlated and, if so, to derive new orthogonal metrics. The data were standardized to a mean of 0 and a standard deviation of 1 and factors were designated using Varimax rotation. Factor loadings showed that each of the four metrics loaded on separate axes with loads exceeding 0.95; thus, the 4 metrics in their original form are orthogonal and provide independent information for evaluating experimental performance.

For each field experiment, the values of the 4 metrics were compared to the 9,000 simulations. Experiments were extracted from the simulation dataset by determining whether the values of each of the four metrics for a given in-field depletion experiment fell above or below the mean value for the metric from the simulation dataset. This generated a 4-digit integer sequence (e.g., 1011) for each field experiment with a 1 assigned if the field experiment metric fell above the mean of the simulated experiments metrics and a 0 if below. The same set of integer sequences were calculated for each simulation and compared to the mean of the metrics for all simulated experiments. Then the subset of simulations having the same sequence as the in-field experiment was extracted from the dataset. This process was repeated sequentially on each extracted subset, with the mean values for the simulated experiments being updated using only the extracted subset, until none of the final subset of simulations had the same 4-digit value as the field experiment in question. The subset of simulated experiments considered to be the most comparable to each field experiment was the subset immediately preceding this final null subset of simulated sites. This "most comparable" subset typically numbered 2-20 of the 9,000 simulations and was used to describe the average simulated four metrics and the average error in efficiency most appropriate for comparison to the known (Tables 2.2 and 2.3).

Table 2.2

| Depletion Experiment OQ00-02: estimated efficiency: 0.68234 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | CV | CV *k* parameter | OS | EAS ($ft^2$) | |
| Average Values from Simulations | | | 4.2199 | 36.6461 | 16 | 18926 0.98 | |
| Values from OQ00-02 | | | 30.8497 | 36.322 | 14 | 18481 9.19 | |
| Mean Absolute Error Estimate | | | 0.0247 | Range | 0.0002-0.0787 | | |
| Error in Efficiency Estimate | Density (#/$m^2$) | Clam Distribution | True Efficiency | CV Efficiency Estimate | CV *k* Parameter | OS | EAS ($ft^2$) |
| 0.0487 | 0.75 | NP | 0.6 | 12.1844 | 66.5056 | 14 | 154414.4 1 |
| 0.019 | 0.75 | NP | 0.6 | 11.5787 | 65.1144 | 15 | 147257.5 9 |
| 0.0515 | 0.75 | NP | 0.6 | 11.2896 | 66.3462 | 14 | 154300.5 9 |
| 0.0133 | 1.5 | NP | 0.6 | 1.4095 | 28.6311 | 14 | 150000.8 |
| 0.0133 | 0.75 | P | 0.6 | 1.7763 | 28.8672 | 14 | 152296.0 9 |
| 0.1033 | 3 | P | 0.6 | 15.1487 | 64.1773 | 14 | 153005 |
| 0.0702 | 0.75 | P | 0.6 | 4.6283 | 38.7079 | 16 | 198290.7 |
| 0.0551 | 0.75 | P | 0.6 | 3.5147 | 41.8403 | 17 | 192149.3 |
| 0.0552 | 0.75 | P | 0.6 | 4.0844 | 38.156 | 17 | 191426.3 |
| 0.0044 | 0.75 | P | 0.6 | 4.1349 | 36.6887 | 16 | 194741.5 9 |
| 0.0858 | 1.5 | P | 0.6 | 4.2764 | 34.0108 | 16 | 193992.0 9 |
| 0.1201 | 1.5 | P | 0.6 | 4.4265 | 31.439 | 16 | 197707.8 |
| 0.0179 | 1.5 | P | 0.6 | 3.6316 | 33.9349 | 17 | 192075.4 1 |
| 0.0059 | 1.5 | P | 0.6 | 4.0279 | 32.3748 | 17 | 191459 |
| 0.0014 | 1.5 | P | 0.6 | 3.4498 | 33.381 | 15 | 199825.4 1 |
| 0.0214 | 1.5 | P | 0.6 | 4.0042 | 36.1805 | 16 | 194861.9 1 |
| 0.0396 | 1.5 | P | 0.6 | 3.1085 | 34.008809 | 15 | 201317.0 0 |

Parameters for 18 simulations that best compared to the depletion experiment OQ00-02. Clam distribution is denoted as NP (uniform), P (vertical bands). OS is multiplied by 100 and truncated into integers. See Figure 1.1 in Poussard et al. (in prep).

Table 2.3

| SC04-01: estimated efficiency: 0.53334 | | | | |
|---|---|---|---|---|
| | CV | CV *k* parameter | OS | EAS ($ft^2$) |
| Average Values from Simulations | 11.3299 | 25.0282 | 16 | 93721.8 |
| Values from SC08-03 | 19.8354 | 28.0845 | 16 | 138041.61 |
| Mean Absolute Error Estimate | 0.8768 | Range | | 0.0017-7.440 |

| Error in Efficiency Estimate | Density (#/$m^2$) | Clam Distribution | True Efficiency | CV Efficiency Estimate | CV *k* Parameter | OS | EAS ($ft^2$) |
|---|---|---|---|---|---|---|---|
| 3.59 | 0.75 | HP | 0.2 | 19.6078 | 27.9661 | 17 | 56568.98 |
| 3.69 | 1.5 | HP | 0.2 | 18.7633 | 27.6673 | 17 | 56568.98 |
| 3.72 | 3 | HP | 0.2 | 18.2203 | 26.6898 | 17 | 56568.98 |
| 2.025 | 3 | HP | 0.2 | 1.3091 | 24.6763 | 15 | 30166.25 |
| 0.06 | 1.5 | P | 0.2 | 4.684 | 27.2273 | 15 | 30034.26 |
| 0.07 | 3 | P | 0.2 | 4.5514 | 24.6013 | 15 | 30034.26 |
| 0.065 | 3 | P | 0.2 | 4.3427 | 24.6724 | 16 | 28483.06 |
| 0.0133 | 1.5 | NP | 0.6 | 1.5049 | 26.3674 | 16 | 139917.59 |
| 0.06 | 1.5 | NP | 0.6 | 1.6352 | 26.7274 | 15 | 142804.09 |
| 0.035 | 1.5 | NP | 0.6 | 1.7552 | 25.419 | 16 | 138588.7 |
| 0.0533 | 1.5 | NP | 0.6 | 1.6139 | 27.3153 | 16 | 140525.7 |
| 0.045 | 1.5 | NP | 0.6 | 1.8341 | 25.9707 | 16 | 142581.3 |
| 1.725 | 0.75 | T | 0.2 | 26.422 | 25.7031 | 17 | 44666.07 |
| 1.079 | 1.5 | T | 0.2 | 28.3093 | 26.4591 | 17 | 44843.84 |
| 1.4523 | 1.5 | T | 0.2 | 28.1613 | 25.9795 | 16 | 46510.5 |
| 1.1149 | 3 | T | 0.2 | 28.8642 | 26.2663 | 17 | 44843.84 |
| 1.4665 | 3 | T | 0.2 | 27.6409 | 25.9012 | 16 | 46510.5 |

Parameters for 17 simulations that best compared to the depletion experiment SC04-01. Clam distribution is denoted as P (vertical bands) and T (diagonal across the area). OS is multiplied by 100 and truncated into integers. See Figure 1.1 in Poussard et al. (in prep).

Each simulation in the subset of simulations extracted was run using a specified dispersion of clams (see Figure 1.1 in Poussard et al. in prep). These were a relatively uniform distribution across the depletion rectangle (denoted as NP), patches oriented across the narrow dimension (P), patches oriented longitudinally (HP), and patches of a triangular nature emanating from one side of the rectangle (T). The fraction of chosen simulations assigned to each in-field experiment falling into each of these categories was obtained to describe possible clam dispersion characteristics in the area occupied by the in-field depletion experiment.

Comparisons to field experiments were made using 4 error estimates chosen to determine which of the in-field depletion experiments diverge the most from the identified “most comparable” simulations using the 4 integer test. Two error estimates describe how closely the 4 in-field experiment metrics derived from the field experiments agreed with the same metrics obtained from the extracted subset of the simulations, henceforth referred to as Err1 and Err2:

$$Err1 = \sum_{j=1}^{4} \frac{abs(observed - expected)}{expected} \tag{10}$$

$$Err2 = \sum_{j=1}^{4} \frac{(observed - expected)^2}{expected}\ . \tag{11}$$

where the observed metric is obtained from the field experiment and the expected metric is the average value of the extracted simulations.

Err3 is the average percent error obtained from the simulation subset obtained by comparing the simulated estimate of efficiency with the known efficiency used in the simulation (Equation 9). Err3 was modified as a simple difference between the averages obtained from the simulation subset as Err4 (Kleisner et al. 2017):

$$Err4 = abs(obseff - trueeff\ ) \tag{12}$$

Caveat lector; no metric exists that can definitively estimate the accuracy of an in-field depletion experiment, as the true efficiency perforce is unknown. The four error estimates relate attributes of a large set of simulated experiments which use combinations of 4 different depletion experiment characteristics to describe how precisely the Patch Model estimate of efficiency returned the known efficiency specified in the simulation. In this study, we use these four error estimates to identify in-field experiments which have characteristics that resemble the 4 performance characteristics in the simulations of Poussard et al. (in prep.): the efficiency CV, the k parameter CV, the OS, and the EAS. As the simulation study did not evaluate all possible experimental conditions (e.g., all possible tow numbers, or all possible true efficiencies), we cannot affirm that the comparisons provided by the 4 error estimates validly identify subpar in-field experiments; we can only suggest that the forensic evidence casts increased suspicion on certain experiments as ones of dubious accuracy, as the metrics from the in-field experiments resemble metrics from simulation experiments that performed poorly or deviated substantially from the suite of metrics provided by the most similar of the simulations.

### 2.2.2 *Statistics*

Unless otherwise indicated, statistics used SAS Version 9. Field experiments that fell at or above the 80$^{th}$ percentile for one or more of the 4 error estimates were compared to the remaining experiments falling below the 80$^{th}$ percentile using a Wilcoxon rank sum test (Sokal & Rohlf 1998) to determine if the suspicious subset of in-field experiments were a random subset of all in-field experiments, as determined by the 4 error estimates and other metrics as earlier described.

The relationship between descriptors of Patch Model performance, including efficiency and density estimates, and descriptors of the experiment such as location, depth, and target species were resolved using correspondence analysis (Clausen 1998). For this purpose, continuous variables were classified into quartiles (1-4) or halves (1-2) (Table 2.4). Table 2.4 identifies the variables used to specify the coordinate system for the correspondence analysis and a series of supplementary variables assigned coordinate positions (Clausen 1998). Of note, the error terms were all designated supplementary variables.

Pearson correlations (R Core Team, version 3.6.0) were conducted on variables describing the in-field experiments to determine how factors such as dredge width, experiment area width, number of tows, year, and latitude correlated with Patch Model efficiency, density, and k parameter estimates.

Table 2.4

| Correspondence Analysis Legend | | | |
|---|---|---|---|
| Patch Model Outputs | | Species<br>O: Ocean quahog, S: Surfclam | |
| | | Region<br>LI: Long Island, NJ: New Jersey, DMV: Delmarva | |
| E | Efficiency | | |
| D | Density | Error Terms (Supplementary Variables) | |
| K | *k* parameter | R12 | Err1 |
| C | CV Efficiency | R22 | Err2 |
| N | CV Density | R32 | Err3 |
| P | CV *k* parameter | R42 | Err4 |
| Experiment Descriptors | | Clam Distributions (Supplementary Variables) | |
| S | EAS | NP2 | Non-Patchy Clam Distribution |
| T | Tows | PP2 | Patchy Distribution |
| L. | Latitude | HP2 | Half-Patchy Distribution |
| | | NT2 | Triangular Distribution |
| Z | Depth | | |

Variables used in correspondence analysis. Error estimates and clam distributions were entered as supplementary variables (Clausen 1998). E,D,K,C,N,P,S,T,L, and Z were entered as quartiles; only quartiles 1 and 4 are shown in the graphs. Clam distributions were entered as halves; only the upper halves are shown in the graphs. Error estimates were entered as 1(below the $80^{th}$ percentile) or 2 (at or above the $80^{th}$ percentile

## 2.3 Results

### 2.3.1 ***Field Depletion Experiment Characteristics***

Table 2.5 provides the mean and median efficiency estimates, density estimates, and *k*-parameter estimates for the 50 depletion experiments. The mean value of the efficiency estimates for the 31 depletion experiments targeting surfclams is 0.635 (Figure

2.1) and the mean value of the efficiency estimates for the 19 depletion experiments targeting ocean quahogs is 0.586 (Figure 2.2).

Table 2.5

| Ocean Quahog (N=19) | | | | | |
|---|---|---|---|---|---|
| | Efficiency | Density (#/m$^2$) | *k* Parameter | EAS (ft$^2$) | OS |
| Mean | 0.586 | 1.184 | 7.724 | 116701.2 | 17.433 |
| Median | 0.629 | 0.094 | 6.165 | 92941.7 | 17.270 |
| Average Standard Deviation | 0.113 | 0.646 | 3.045 | | |
| Average Coefficient of Variance | 0.357 | 16.907 | 0.613 | | |
| **Surfclam (N=31)** | | | | | |
| | Efficiency | Density (#/m$^2$) | *k* Parameter | EAS (ft$^2$) | OS |
| Mean | 0.635 | 1.496 | 12.097 | 146077.2 | 22.330 |
| Median | 0.59 | 0.738 | 5.689 | 78852.3 | 19.143 |
| Average Standard Deviation | 0.131 | 1.786 | 3.011 | | |
| Average Coefficient of Variance | 0.206 | 12.855 | 0.351 | | |

Mean, median, mean standard deviation as estimated by the Patch Model, the effective area swept (EAS), the number of tows, and the mean CV for depletion parameters efficiency, density, and *k* for the 50 field depletion experiments.

The mean density estimate for surfclam depletion experiments is 1.496 clams m$^{-2}$ (Figure 2.3) and the mean density estimate for ocean quahog depletion experiments is 1.184 clams m$^{-2}$ (Figure 2.4). These densities are well above the average stock density for both species as the depletion experiments were purposely sited in high-density areas. The mean k-parameter estimate for the surfclam experiments is 12.097 (Figure 2.5) and the mean for the ocean quahog experiments is 7.724 (Figure 2.6).

Most depletion experiments targeting ocean quahogs were conducted at higher latitudes and at deeper depths than depletion experiments targeting surfclams (Table 2.6). For ocean quahog depletion experiments, higher efficiency estimates were produced further north (Figure 2.7). Surfclam depletion experiments produced higher efficiency estimates off the coast of New Jersey (Figure 2.8).

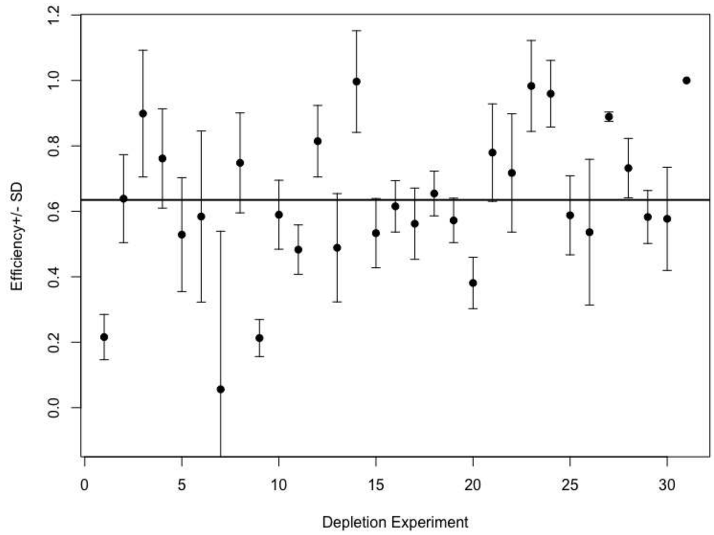



Figure 2.1 Efficiency estimates with standard deviations for the 31 depletion experiments targeting surfclams. Black horizontal line indicates the mean efficiency for all 31 experiments.

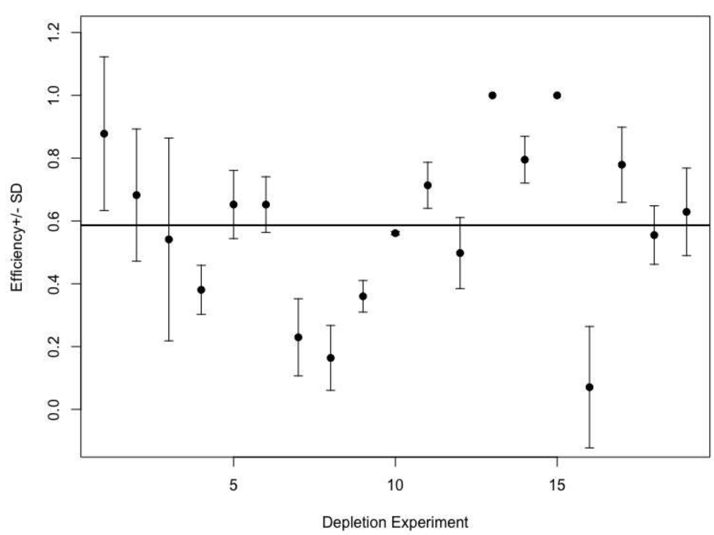



Figure 2.2 Efficiency estimates with standard deviations for the 19 depletion experiments targeting ocean quahogs. Black horizontal line indicates the mean efficiency for all 19 experiments.

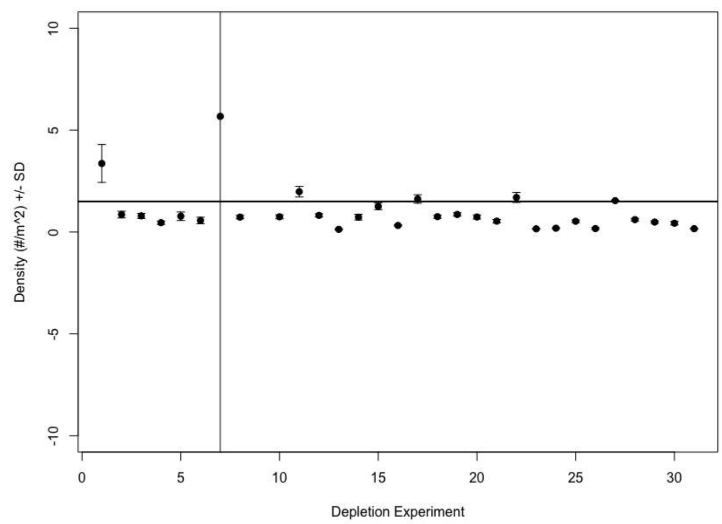



Figure 2.3 Density estimates with standard deviations for the 31 depletion experiments targeting surfclams. Black horizontal line indicates the mean density estimate for all 31 experiments.

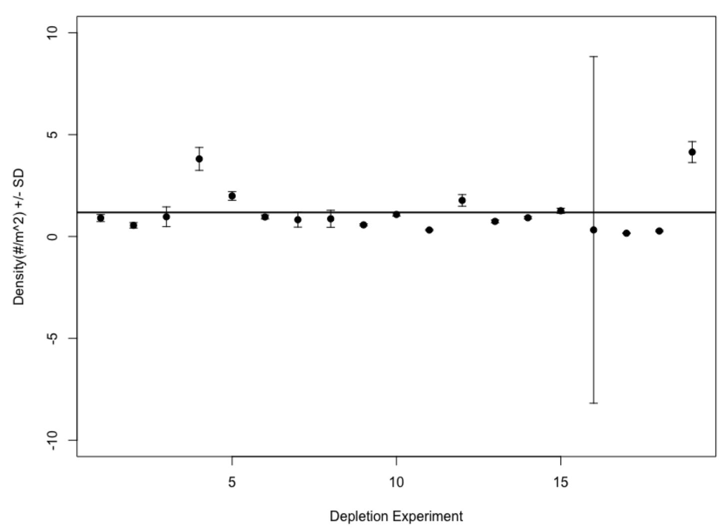



Figure 2.4 Density estimates with standard deviations for the 19 depletion experiments targeting ocean quahogs. Black horizontal line indicates the mean density estimate for all 19 experiments.

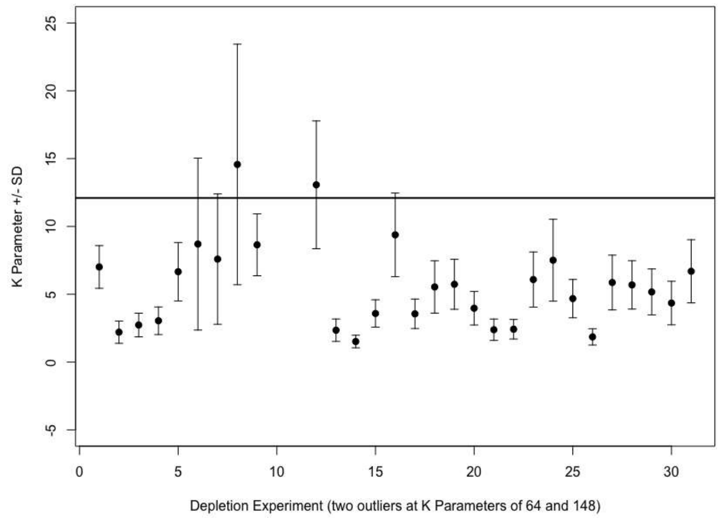



Figure 2.5 *k*-parameter estimates with standard deviations for 31 depletion experiments targeting surfclams. Two outliers at k-parameter values of 64 and 148 are not shown on this graph. Black horizontal line indicates the mean k-parameter estimate for all 31 experiments, including the two outliers not graphed.

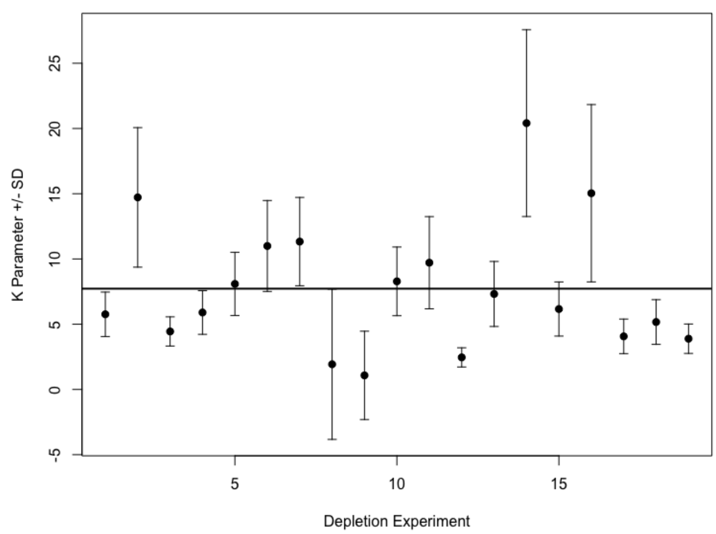



Figure 2.6 *k*-parameter estimates with standard deviations for 19 depletion experiments targeting ocean quahogs. Black horizontal line indicates the mean k-parameter estimate for all 19 experiments.

Table 2.6

| Experiment ID | Region | Dredge Width(ft) | Tows | OS | Year | Latitude | Longitude |
|---|---|---|---|---|---|---|---|
| SC1997-2(*2) | NJ | 8.33 | 39 | 7 | 1997 | 40.05317 | -73.83917 |
| SC1997-3 | NJ | 10.83 | 13 | 29 | 1997 | 39.39317 | -73.91033 |
| SC1997-4 | NJ | 10.83 | 18 | 18 | 1997 | 39.39317 | -73.91033 |
| SC1997-5(*1) | NJ | 8.33 | 17 | 15 | 1997 | 39.365 | -73.89833 |
| SC1997-6 | NJ | 8.33 | 19 | 14 | 1997 | 39.365 | -73.89833 |
| SC1999-2 (*1) | NJ | 10.83 | 4 | 53 | 1999 | 39.68133 | -73.74667 |
| SC1999-3 (*2*3*4) | NJ | 10.83 | 5 | 39 | 1999 | 39.68133 | -73.74667 |
| SC1999-4 | NJ | 10.83 | 6 | 54 | 1999 | 39.52133 | -73.77867 |
| SC1999-5 (*2) | DMV | 10.83 | 28 | 12 | 1999 | 36.902 | -74.97583 |
| SC1999-6 (*2) | NJ | 10.83 | 4 | 43 | 1999 | 39.56333 | -73.91167 |
| SC1999-7 | NJ | 10.83 | 10 | 20 | 1999 | 39.768 | -73.91633 |
| OQ00-01 (*4) | LI | 12.5 | 22 | 11 | 2000 | 40.60217 | -71.9875 |
| OQ00-02 (*1) | LI | 12.5 | 16 | 14 | 2000 | 40.3945 | -72.543 |
| OQ00-03 (*2) | LI | 10 | 27 | 7 | 2000 | 40.583 | -72.79683 |
| OQ02-01 | LI | 10 | 24 | 14 | 2002 | 40.72762 | -71.7373 |
| OQ02-02 (*1) | LI | 10 | 22 | 12 | 2002 | 40.10312 | -73.19108 |
| OQ02-03 | NJ | 10 | 20 | 14 | 2002 | 38.81491 | -73.81335 |
| OQ02-04 (*3*4) | DMV | 10 | 24 | 13 | 2002 | 37.88755 | -74.64486 |
| SC02-02 | NJ | 10.83 | 16 | 16 | 2002 | 40.10908 | -73.84423 |
| SC02-03 (*3*4) | NJ | 10.83 | 20 | 19 | 2002 | 39.26923 | -73.78116 |
| SC02-04 | DMV | 10.83 | 18 | 16 | 2002 | 38.85791 | -74.02778 |
| SC04-01 | NJ | 10 | 24 | 16 | 2004 | 39.28611 | -73.87778 |
| SC04-02 | NJ | 10 | 20 | 16 | 2004 | 39.58278 | -74.02778 |
| SC04-03 (*2*3*4) | DMV | 10 | 20 | 19 | 2004 | 38.27075 | -74.3792 |
| OQ05-01 (*1*2*3*4) | LI | 10 | 20 | 25 | 2005 | 40.51903 | -72.07617 |
| OQ05-02 (*1*3*4) | LI | 10 | 21 | 25 | 2005 | 40.38957 | -72.3895 |
| OQ05-03 | LI | 10 | 20 | 22 | 2005 | 40.6422 | -72.6517 |
| OQ05-04 | LI | 10 | 17 | 23 | 2005 | 40.6817 | -72.18147 |
| OQ05-06 (*1*2*3*4) | LI | 10 | 20 | 20 | 2005 | 40.0555 | -72.41673 |
| SC05-01 | NJ | 10 | 20 | 22 | 2005 | 39.2653 | -74.37947 |
| SC05-02 | NJ | 10 | 17 | 22 | 2005 | 39.56383 | -73.90364 |
| SC05-03(*1*2) | NJ | 10 | 20 | 19 | 2005 | 39.89733 | -73.90591 |
| SC05-04 (*3*4) | DMV | 10 | 20 | 23 | 2005 | 39.56972 | -73.54946 |
| SC05-05 (*4) | NJ | 10 | 17 | 7 | 2005 | 39.43615 | -73.3732 |
| OQ08-01 | LI | 12.5 | 17 | 29 | 2008 | 40.93762 | -72.04765 |
| OQ08-02 | LI | 12.5 | 17 | 18 | 2008 | 40.27445 | -72.84397 |
| OQ08-03 | SNE | 12.5 | 17 | 15 | 2008 | 41.02307 | -70.85472 |
| SC08-01 | NJ | 12.5 | 13 | 14 | 2008 | 39.30475 | -74.05158 |
| SC08-02 | NJ | 12.5 | 18 | 53 | 2008 | 39.18136 | -74.07645 |
| SC08-03 (*1) | NJ | 12.5 | 21 | 39 | 2008 | 39.60343 | -73.42194 |
| SC08-04 | NJ | 12.5 | 17 | 54 | 2008 | 39.81033 | -73.9149 |
| SC08-09 | NJ | 12.5 | 17 | 12 | 2008 | 39.31328 | -74.05285 |
| OQ11-01 (*2) | NJ | 12.5 | 10 | 43 | 2011 | 39.03003 | -74.05774 |
| OQ11-02 | NJ | 12.5 | 20 | 20 | 2011 | 39.89356 | -73.48104 |
| OQ11-02S | NJ | 12.5 | 18 | 11 | 2011 | 39.8925 | -73.475 |
| OQ11-05 (*2) | LI | 12.5 | 22 | 14 | 2011 | 40.13542 | -72.1201 |
| SC11-02 | NJ | 12.5 | 20 | 7 | 2011 | 39.89356 | -73.48104 |
| SC11-02S | NJ | 12.5 | 18 | 14 | 2011 | 39.8925 | -73.475 |

Table 2.6 Continued

| Experiment ID | Region | Dredge Width(ft) | Tows | OS | Year | Latitude | Longitude |
|---|---|---|---|---|---|---|---|
| SC11-03 (*1) | LI | 12.5 | 14 | 12 | 2011 | 40.567 | -73.081 |
| SC11-04 | LI | 12.5 | 17 | 14 | 2011 | 40.641 | -73.036 |

Metrics for in-field depletion experiments targeting ocean quahogs and surfclams between 1997 and 2011. Region is identified as LI - Long Island, NJ - New Jersey, SNE - Southern New England, DMV - Delmarva. Experiments found falling at or above the 80th percentile for each error estimate are denoted with an asterisk (*) followed by the number of the error estimate (1,2,3,4).

Over the 14 years that depletion experiments were conducted, methodology and gear changed. Dredge width, for example, gradually increased from 8.33 ft to 12.5 ft. The number of dredge tows used in each experiment varied through the years as well. The majority of experiments, especially in later years, used between 15 and 20 tows, but some experiments between 1997 and 2000 used as few as 4 dredge tows and as many as 39 tows (Figure 2.9).

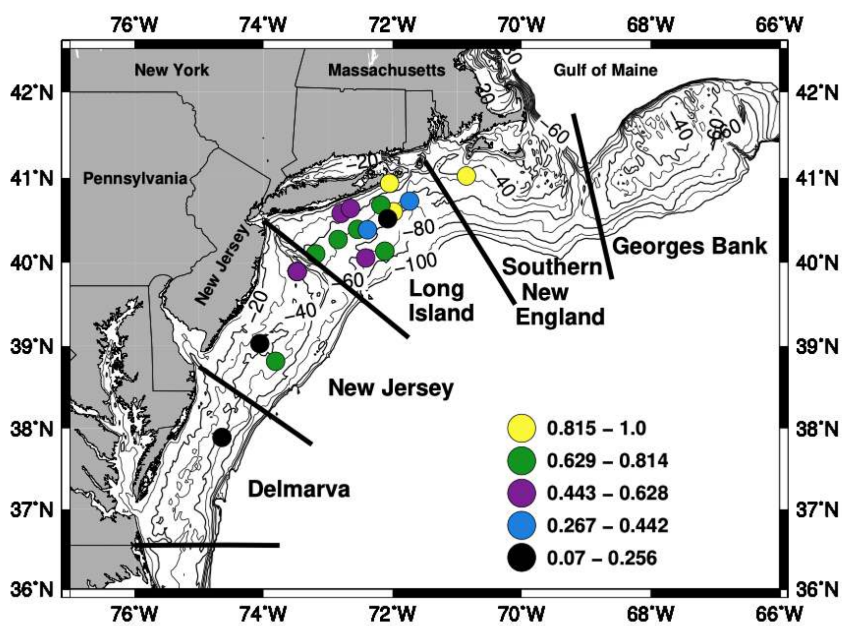



Figure 2.7 Locations of the 19 depletion experiments targeting ocean quahogs off the east coast of the US. Colors indicate Patch Model efficiency estimates for each depletion experiment. Boundaries on the continental shelf shown as fine lines represent the survey strata used prior to 2018.

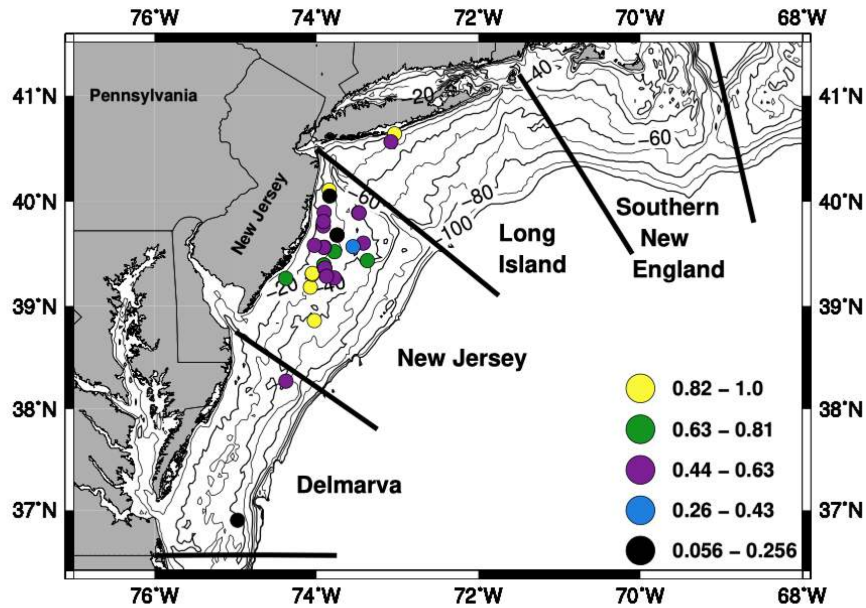



Figure 2.8 Locations of the 31 depletion experiments targeting surfclams off the east coast of the US. Colors indicate Patch Model efficiency estimates for each depletion experiment. Boundaries on the continental shelf shown as fine lines represent the survey strata used prior to 2018.

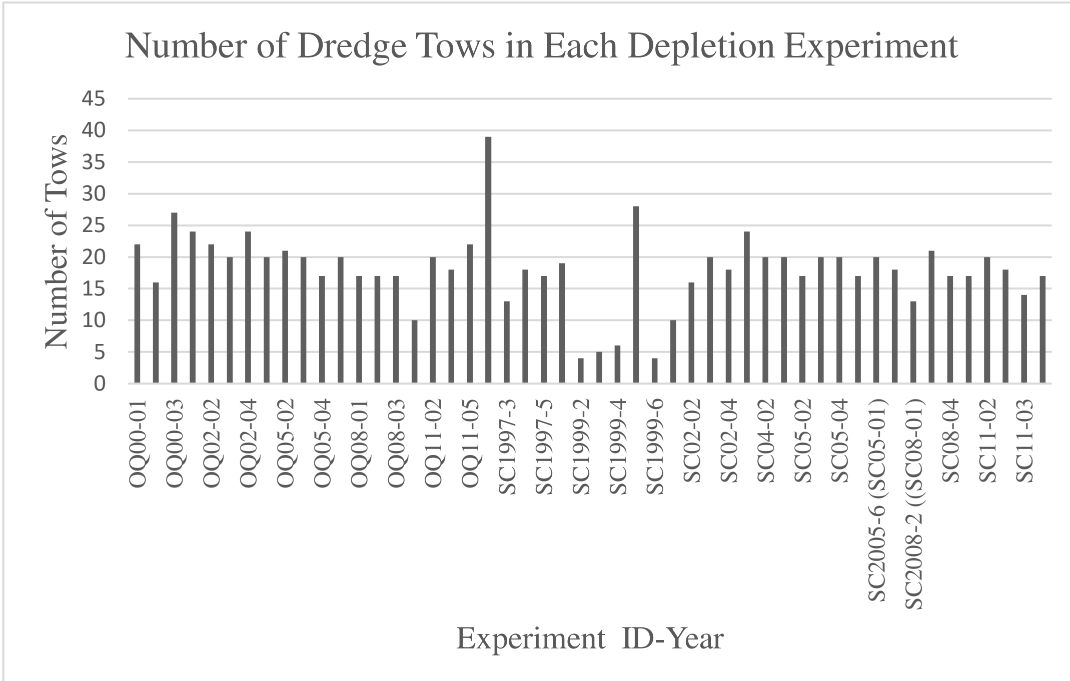



Figure 2.9 Graph of number of dredge tows in each depletion experiment. X axis is each depletion experiment ID, organized as the stock the depletion experiment was conducted on (OQ or SC), the last 2 numbers in the year the experiment was conducted in, and the order in which that experiment was conducted in that year.

### 2.3.2 ***Correlation Analysis***

Efficiency estimates for ocean quahog depletion experiments are significantly positively correlated with latitude (see Figure 2.7) and the width of the dredge (Figure 2.10). Efficiency is incorporated into the equation to calculate EAS, therefore the correlation between efficiency and EAS is expected and correlations between efficiency and other variables will be reflected by correlations between EAS and those same variables. Year is incorporated into the correlation analysis to see how parameters changed over time. As noted, dredge width increases with year, and tow number and depth decrease over time. The CV of the efficiency estimate (Equation 6) is negatively correlated with the number of tows and strongly positively correlated with the CV of the density estimate (Figures 2.10-

2.11). In surfclam depletion experiments, as opposed to ocean quahog experiments, the CV of the *k* parameter is significantly positively correlated with the CV of the density estimate. In the case of surfclams, no correlation exists between latitude and the efficiency estimates, but density estimates are negatively correlated with the latitude and efficiency estimates.

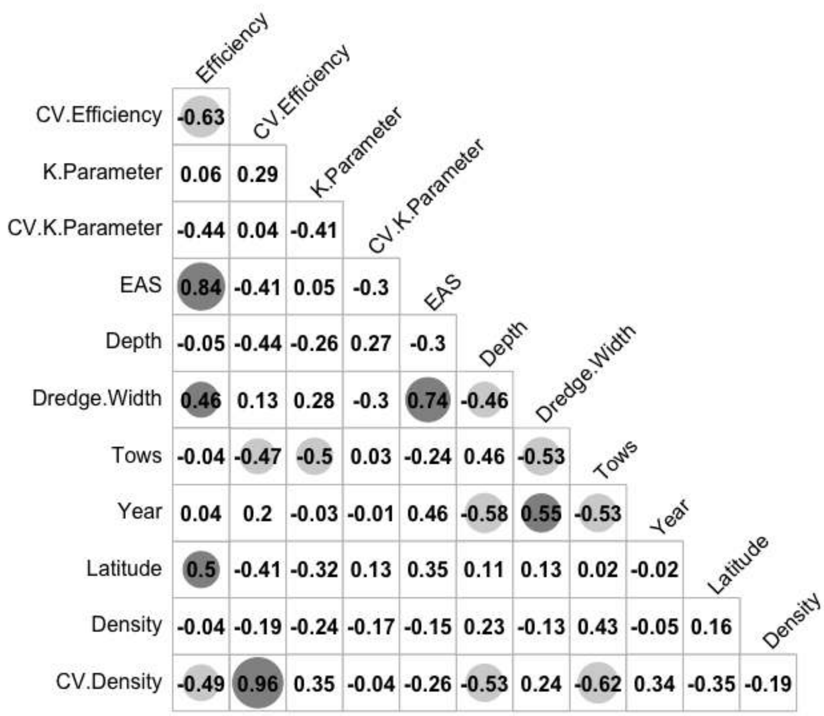



Figure 2.10 Correlogram for experiments targeting ocean quahogs. Numbers in the squares are Pearson's correlations. Significant correlations ($\alpha \leq 0.05$) are denoted by gray circles. Positive correlations are dark gray; negative correlations are light gray.

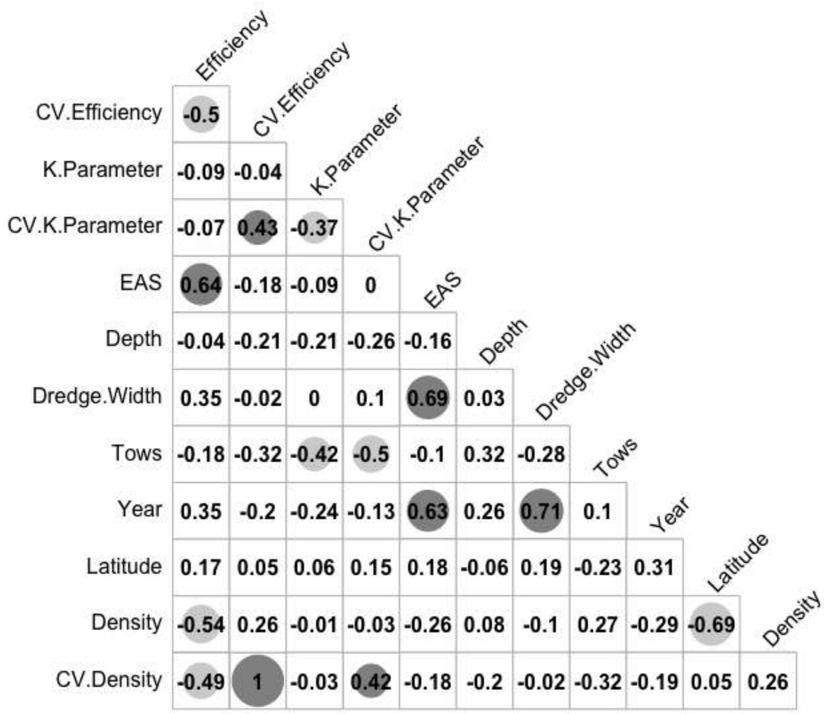



Figure 2.11 Correlogram for experiments targeting surfclams. Numbers in the squares are Pearson's correlations. Significant correlations (α≤0.05) are denoted by gray circles. Positive correlations are dark gray; negative correlations are light gray.

### 2.3.3 ***Error Estimates and Wilcoxon Tests***

In-field depletion experiments that fell at or above the 80th percentile of their respective "most comparable" simulated experiments, for one or more of the four error estimates are denoted by asterisks in Table 2.6. Of the 50 depletion experiments, 25 fell at or above the 80th percentile for one or more of the error estimates. Experiments falling at or above the 80th percentile for error terms Err1 (Equation 10) and Err2 (Equation 11) are experiments that differed substantively from the chosen subset of simulations for one or more of the four metrics describing the depletion experiments, the efficiency CV, the *k* parameter CV, the number of tows, and the EAS. These in-field experiments were not well described by the most similar subset of simulations. The possibility that the range of values for EAS might influence the differential in the results for Err1 and Err2 was tested by recomputing Err2 using $\log_e$(EAS). The set of experiments flagged by Err2 did not change. Field experiments falling at or above the 80th percentile for error estimates Err3 (Equation 9) and Err4 (Equation 12) were characterized by simulation subsets for which the error between the simulated efficiency estimate derived from the Patch Model and the true efficiency used for the simulation was large; that is, by simulation cases where the Patch Model poorly estimated the known efficiency used in the simulation.

In-field experiments flagged by an error metric produce parameter estimates that fall at or above the 80th percentile for the specified error metric (Table 2.7). Experiments flagged by Err2, Err3, and Err4 have lower average and median efficiency estimates than experiments identified by Err1. The *k*-parameter estimate is much higher for experiments flagged by Err1 than for the experiments identified by the other error estimates in terms of the mean, but not the median. The average standard deviation estimate, but not the median standard deviation estimate, for the *k*-parameter for Err2 is higher than found for the other error terms. The mean CVs in density estimates for Err2, Err3, and Err4 are much higher than for Err1, but the median CVs in density are not substantively different across error terms.

The relationships between the in-field experiments flagged by one or more error estimates with the rest of the dataset were evaluated using Wilcoxon rank sums tests (Table 2.8). Experiments flagged by Err1 did not differ significantly from the remaining experiments for any of the measured depletion parameters. In each case, the identified

Table 2.7

| | N | Efficiency | Efficiency SD | Efficiency CV | Density (# $m^{-2}$) | Density SD |
|---|---|---|---|---|---|---|
| **Err1 Average** | 10 | 0.544 | 0.134 | 27.256 | 0.859 | 0.153 |
| Median | | 0.581 | 0.117 | 21.624 | 0.568 | 0.111 |
| **Err2 Average** | 10 | 0.348 | 0.162 | 138.639 | 3.300 | 6.255 |
| Median | | 0.357 | 0.105 | 29.327 | 1.294 | 0.451 |
| **Err3 Average** | 8 | 0.342 | 0.153 | 135.848 | 1.525 | 6.181 |
| Median | | 0.371 | 0.111 | 28.280 | 0.846 | 0.246 |
| **Err4 Average** | 10 | 0.440 | 0.162 | 113.378 | 1.364 | 4.971 |
| Median | | 0.435 | 0.118 | 25.299 | 0.846 | 0.191 |
| | **Density CV** | ***k* Parameter** | ***k* Parameter SD** | ***k* Parameter CV** | **EAS ($ft^2$)** | **OS** |
| **Err1 Average** | 210.624 | 4.515 | 2.133 | 93.098 | 84059.6 | 22 |
| Median | 182.321 | 5.747 | 1.999 | 34.873 | 57200.2 | 20 |
| **Err2 Average** | 3961.220 | 0.205 | 2.612 | 57.375 | 36936.5 | 22 |
| Median | 262.005 | 6.373 | 1.709 | 30.241 | 34259.8 | 20 |
| **Err3 Average** | 1359.642 | 4.283 | 2.652 | 103.321 | 35585.5 | 23 |
| Median | 205.656 | 3.008 | 3.311 | 32.811 | 33467.9 | 22 |
| **Err4 Average** | 1125.329 | 4.241 | 2.371 | 88.917 | 57730.2 | 22 |
| Median | 189.541 | 3.008 | 1.471 | 31.758 | 43281.3 | 21 |

Average and median values for depletion experiment parameters for the experiments falling at or above the 80th percentile for each error estimate. Efficiency SD, Density SD, and *k* parameter SD are Patch Model metrics from the maximum likelihood equation (Eq 5). CV metrics are Eq 6.

Table 2.8

| | **Err1** | **Err2** | **Err3** | **Err4** |
|---|---|---|---|---|
| **Variable** | Pr > \|Z\| | Pr > \|Z\| | Pr > \|Z\| | Pr > \|Z\| |
| **Efficiency** | - | 0.0002 | 0.0003 | 0.0105 |
| **Efficiency CV** | - | 0.0079 | 0.0179 | 0.0138 |
| **Density** | - | 0.0098 | - | - |
| **Density CV** | - | 0.006 | - | 0.0481 |
| ***k* Parameter** | - | - | - | - |
| ***k* Parameter CV** | - | - | - | - |
| **EAS** | - | 0.0025 | 0.0002 | 0.0085 |
| **OS** | - | - | - | - |

Wilcoxon Rank Sums test results for depletion experiment variables classified by error estimates. Nonsignificance (α≤0.05) is denoted by a dash (-).

experiments were distributed randomly within the full set of in-field experiments with respect to the different parameters tested. In dramatic contrast, for experiments flagged by Err2, Err3, and Err4, the Patch Model efficiency estimate, the CV for the efficiency estimate, and the EAS values differed significantly from the remaining in-field depletion experiments. Experiments flagged by Err2 and Err4 differ significantly from the rest of the dataset with respect to the CV of the density estimate, and the 10 experiments flagged by Err2 differed significantly from the rest of the dataset in the estimated density. Interestingly, all 8 experiments flagged by Err3 were among the 10 flagged by Err4, yet Err4 experiments display a significantly different CV in the density estimates whereas the 8 flagged by Err3 do not.

2.3.4 ***Correspondence Analysis***

Correspondence analysis shows that variance in descriptor metrics is primarily explained by the first 2 axes (Figure 2.12). Table 2.4 describes the abbreviations in the chart. Dimension 1 (Figures 2.12 and 2.14) is determined primarily by Patch Model metrics including the estimate of efficiency, the CV of the efficiency estimate (Equation 2.6), the CV of the density estimate, the width of the dredge, and the EAS (Table 2.9).Table 2.10 describes how each variable falls out on the axes. The experiments flagged by the error terms (R1-4) and clam distributions (NP4, PP4, HP4, NT4) are included as supplementary variables. Low EAS (indicating more dredge overlap), low efficiency estimates, high CV values for efficiency and density estimates, and smaller dredge sizes, along with experiments falling at or above the $80^{th}$ percentile for error estimates Err2, Err3, and Err4, fall on the positive (right) side of Dimension 1. High efficiency estimates, high EAS, larger dredge sizes, and low CVs for the efficiency and density estimates fall on the negative (left) side of Dimension 1.

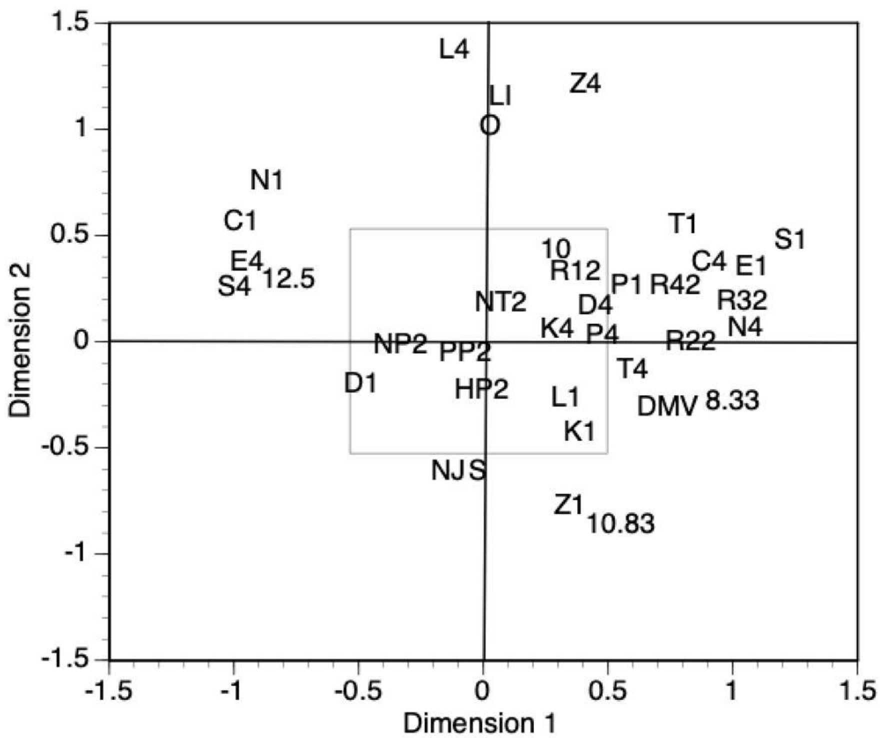



Figure 2.12 Correspondence analysis for the depletion dataset. The variables are denoted as a letter (defined in Table 2.4) and a number, 1 or 4, describing if that value is in the highest (>75% percentile) or lowest (<25th percentile) quartile of the data. A 2 denotes values in the upper 50th percentile of the data. Gray box demarcates the area from -0.5 to 0.5 on the x and y axes.

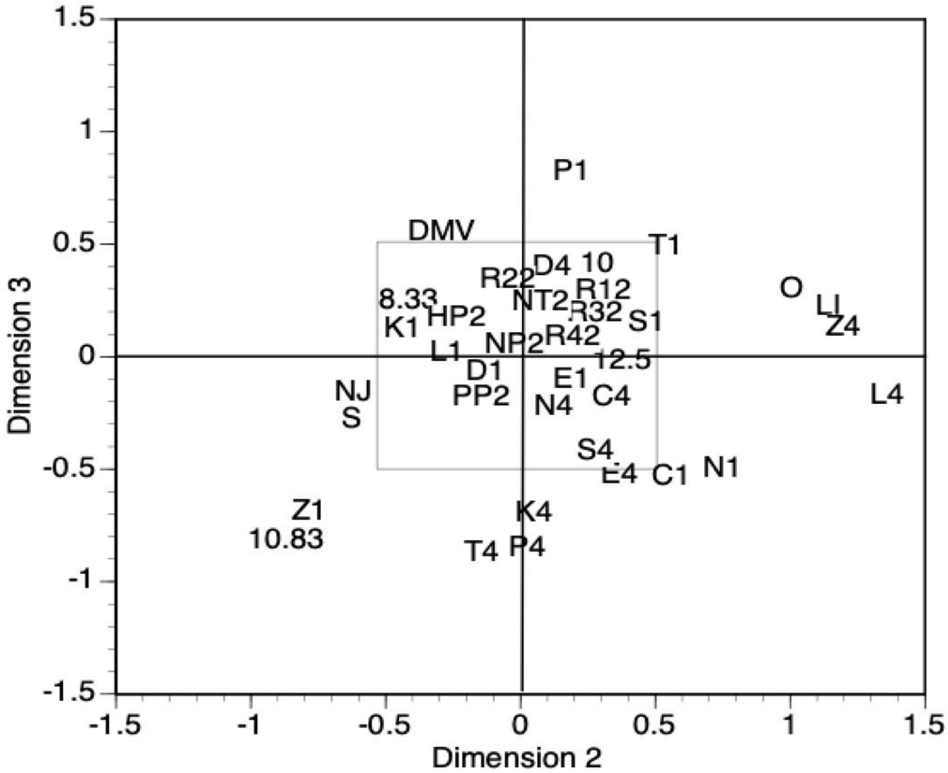



Figure 2.13 Correspondence analysis for the depletion dataset. The variables are denoted as a letter (defined in Table 2.4) and a number: 4 if that value is in the highest (>75% percentile) or 1 if it is in the lowest (<25th percentile) quartile of the data. A 2 denotes values in the upper 50th percentile of the data. Gray box demarcates the area from -0.5 to 0.5 on the x and y axes.

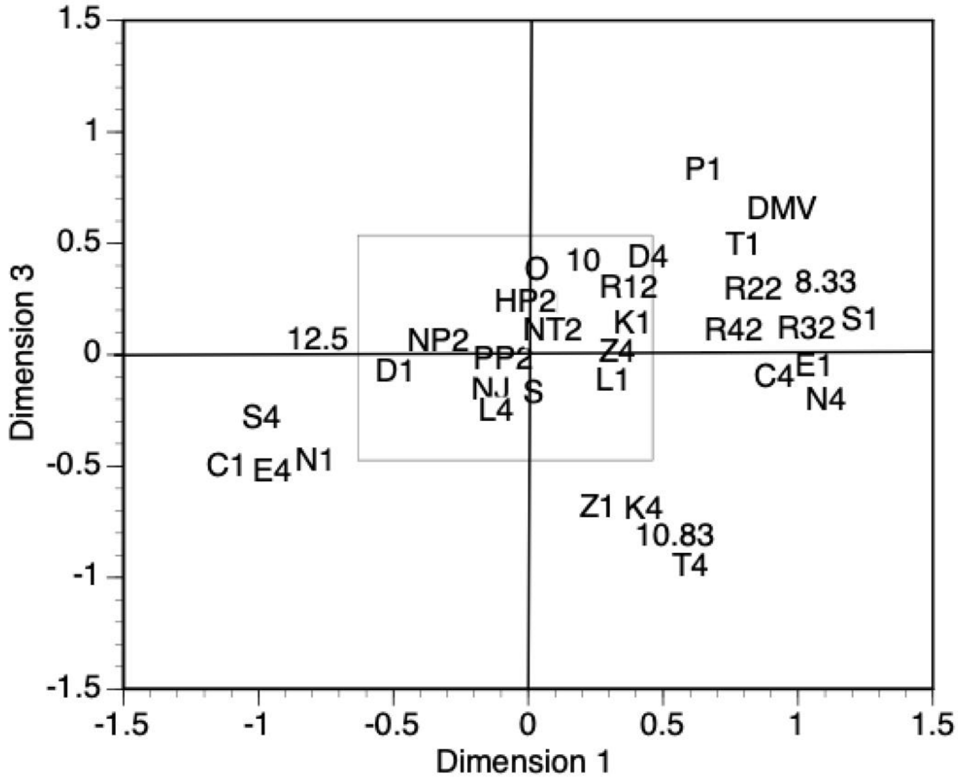



Figure 2.14 Correspondence analysis for the depletion dataset. The variables are denoted as a letter (defined in Table 2.4) and a number: 4 if that value is in the highest (>75% percentile) or 1 if it is in the lowest (<25th percentile) quartile of the data. A 2 denotes values in the upper 50th percentile of the data. Gray box demarcates the area from -0.5 to 0.5 on the x and y axes.

Table 2.9

| | Err 2 | | Err3 | | Err4 | |
|---|---|---|---|---|---|---|
| | < 80th Percentile | ≥ 80th Percentile | < 80th Percentile | ≥ 80th Percentile | < 80th Percentile | ≥ 80th Percentile |
| **Efficiency** | | | | | | |
| Mean | 0.683 | 0.348 | 0.669 | 0.342 | 0.661 | 0.440 |
| Median | 0.653 | 0.357 | 0.652 | 0.371 | 0.645 | 0.435 |
| **CV Efficiency** | | | | | | |
| Mean | 18.522 | 138.639 | 24.774 | 135.848 | 24.837 | 113.378 |
| Median | 16.720 | 29.327 | 17.325 | 28.280 | 16.720 | 25.299 |
| **CV *k* Parameter** | | | | | | |
| Mean | 42.245 | 57.375 | 34.214 | 103.321 | 34.360 | 88.917 |
| Median | 33.066 | 30.241 | 32.817 | 32.811 | 32.817 | 31.758 |
| **Density (# $m^{-2}$)** Mean | 0.898 | 3.299 | 1.350 | 1.525 | 1.382 | 1.364 |
| Median | 0.735 | 1.294 | 0.750 | 0.846 | 0.750 | 0.846 |
| **CV Density** | | | | | | |
| Mean | 155.565 | 3961.220 | 832.706 | 1359.642 | 864.938 | 1125.329 |
| Median | 130.905 | 262.005 | 132.933 | 205.656 | 130.905 | 189.541 |
| **OS** | | | | | | |
| Mean | 20 | 22 | 20 | 23 | 20 | 22 |
| Median | 18 | 20 | 18 | 22 | 22 | 21 |
| **EAS** | | | | | | |
| Mean | 204060.8 | 1029185.0 | 427195.8 | 64007.3 | 435046.6 | 105241.8 |
| Median | 172768.5 | 68921.2 | 172768.5 | 58405.1 | 170872.0 | 64757.9 |

A comparison of mean and median estimates of depletion experiment parameters between experiments at and above and those below the $80^{th}$ percentiles for error estimates Err2, Err3, and Err4.

Table 2.10

| **Dimension 1** | | | **Dimension 2** | | | **Dimension 3** | | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | **Negative (<-0.5)** | **Positive (>0.5)** | **Variable** | **Negative (<-0.5)** | **Positive (>0.5)** | **Variable** | **Negative (< -0.5)** | **Positive (>0.5)** |
| Dredge Width | 12.5 | 8.33 | Dredge Width | 10.83 | N/A | Dredge Width | 10.83 | N/A |
| OS | N/A | Low + High | OS | N/A | Low | OS | High | Low |
| Efficiency | High | Low | Species | Surfclam | Ocean Quahog | *k* Parameter | High | N/A |
| CV Efficiency | Low | High | CV Efficiency | N/A | Low | Region | N/A | DMV |
| EAS | High | Low | Region | NJ | LI | CV *k* Parameter | High | Low |
| CV *k* Parameter | N/A | Low | Depth | Low | High | Depth | Low | N/A |
| Err2, Err3, Err4 | N/A | High | Latitude | N/A | High | | | |
| CV Density | Low | High | CV Density | N/A | Low | | | |
| Region | N/A | DMV | | | | | | |

Variables that fall on each of the 3 dimensions with loading factors ≤ -0.5 or ≥ 0.5 according to the correspondence analysis. Error terms Err2, Err3, and Err4 are supplementary variables in this analysis. Err2, Err3, and Err4 are included as supplementary variables in the analysis.

Dimension 2 (Figures 2.12 and 2.13) is categorized by the species (ocean quahog and surfclam) and other variables relating to the location of the depletion experiments for the two species, such as depth, latitude, and region. The positive values are variables relating to ocean quahog depletion experiments, such as higher latitudes and deeper depths. Negative values are variables relating to surfclam depletion experiments, lower latitudes (Figure 2.8) and shallower depths. Ocean quahog experiments were typically conducted further north (Figure 2.7) than surfclam experiments and the species is generally found at deeper depths than are surfclams. Dimension 3 (Figures 2.13 and 2.14) is characterized by variation in the *k* parameter (the negative binomial dispersion parameter) and the CV of the *k* parameter estimate. OS (denoted as T on the correspondence analysis plots) is interesting because low and high OS fall on the positive portion of Dimension 1 but are clearly separated by Dimension 3 indicating that tow number exerts a complex influence on outcomes.

Correspondence analysis clearly reveals the relationships earlier identified by the Wilcoxon tests and by the Pearson correlations. The three errors, Err2, Err3 and Err4, which were shown to be highly significant in the Wilcoxon analyses fall on the positive side of Dimension 1 along with the metrics significantly influenced by them. Err1, which did not demonstrate significant differences in the Wilcoxon tests, falls near the origin in all three dimensions, indicating that the experiments identified by this error estimate are more or less randomly distributed throughout the in-field depletion dataset. A tendency for larger dredges to be associated with improved experimental performance is obvious from Figure 2.12; however, the influence of dredge size is complex as the various dredge sizes do not fall in order of size on Dimensions 1 or 2. Very likely, dredge size to some extent is conflated with other variables such as species, year, and depth, being determined more by boat availability and increased familiarity of the crew and scientific staff with depletion experiment methodology over time than experiment performance, with the clear exception of the largest dredge size. The fact that species falls near the origin on Dimensions 1 and 3 shows the similarity in efficiency estimates for the two species, which are separated essentially solely by depth.

The parameters describing clam distribution (NP,NP,HP, NT in the Figures 2.12, 2.13, and 2.14) do not fall on any axis and are grouped in the middle of the correspondence analysis graphs on all dimensions. Although clam distribution clearly affects the outcome of individual experiments as observed through simulation analysis (Poussard et al. in prep.), this effect is distributed across the experimental spectrum, distributing uncertainty in a relatively random way with respect to the in-field experimental dataset.

## 2.4 Discussion

### 2.4.1 ***Forensics on Efficiency Estimates***

The four error estimates identify in-field depletion experiments that have attributes that engender misgivings as to their quality. Since the 4 metrics, the efficiency CV, the *k*-parameter CV, the overlap score (OS), and the average EAS used to generate two of the error estimates (Err1 and Err2) are orthogonal to each other, identification of a subset of experiments based on Err1 and Err2 suggests that these experiments are characterized by an unusual distribution of these 4 descriptive metrics. A close fit to the values of these 4 metrics was not found amongst the 9,000 simulations of Poussard et al (in prep.) which covered a wide range of experimental protocols and field conditions of clam dispersion (Table 2.1). The absence of a close fit generates reason to suspect that these experiments may be uninformative or at least have produced inaccurate efficiency estimates. Error estimates 3 and 4 relate to an inferred error in the efficiency estimates, also gleaned from comparison to the simulation dataset of Poussard et al. (in prep.). All 8 experiments flagged by Err3 were also flagged by Err4, as these two metrics are very similar. These offer independent, but still suppositional evidence of poor performance. These experiments may be uninformative or at least have produced inaccurate efficiency estimates. Ultimately, due to the forensic nature of the error estimates, the inference that these experiments produced uninformative or inaccurate efficiency estimates cannot be affirmed. In aggregate, however, the evidentiary weight points to a subset of in-field experiments of lower quality than the remainder.

Interestingly, the experiments identified by error estimates Err1 which might identify suspect experiments, exerts less influence on the final objective of determining the efficiency of hydraulic dredges. The distribution of these experiments is unbiased relative to the remaining experiments, regardless of the metric used for comparison (Table 2.6). The same cannot be said for error estimates Err2, Err3, and Err4. The series of 16 depletion experiments that fall at or above the $80^{th}$ percentile for error estimates Err2, and Err3, and Err4 are shown to be clearly biased relative to the remaining experiments based on Wilcoxon rank sums tests (Table 2.8) and this bias is re-enforced by correspondence analysis (Figures 2.12 and 2.14). In addition, the direction of bias is noteworthy. Experiments identified by error estimates Err2, Err3, and Err4 are characterized by lower efficiency estimates on average, and their inclusion may bias the overall efficiency estimates used to inform stock assessments.

In correspondence analysis, Err2, Err3, and Err4 also fall on the same dimensional axis as a lower EAS value. Low EAS and low efficiency generally occur together, as the efficiency value is a variable in the equation determining EAS (Equation 3). The relationship is well-documented by Poussard et al. (in prep.). This expectation is confirmed in the in-field depletion experiment dataset by Pearson correlation and demonstrated clearly in correspondence analysis (Figure 2.12). EAS is also positively correlated with year for ocean quahog experiments, and with dredge width for both ocean quahog and surfclam experiments (Figures 2.10 and 2.11). The relationship is driven by the largest dredge (12.5 ft); experiments with this dredge size clearly demonstrated superior performance.

Low OS falls on Dimension 1 along with low efficiency estimates, high uncertainty in the efficiency and density estimates, and the Err2, Err3, and Err4 error estimates. This confirms analysis from the simulation study that low tow numbers can produce an increase in uncertainty in Patch Model estimates. However, high OS also falls out on the positive side of Dimension 1, identifying OS as a complex metric in determining experiment performance. Higher OS, and by extension higher tow number, in a depletion experiment does not always reduce uncertainty in Patch Model estimates. An explanation for this discrepancy may come from the pragmatic efforts of a field experiment. Depletion experiments are costly in vessel time and crew effort, often requiring more than 8 hours of nearly continual dredging. Cost at sea was sufficient that adaptive time management during the experiment was directed at limiting tow number, albeit with limited empirical guidance to determine the stopping point for the depletion experiment. One consequence of adaptive time management during the depletion experiment was a decision to add tows if the experiment appeared not to be generating a clear and consistent reduction in catch per tow. Thus, higher tow numbers, and by extension higher OS, potentially were accorded to experiments of lower quality and this bias is borne out, as a consequence, by the positioning of T4 on the right side of Dimension 1 with high uncertainties in the efficiency and density estimates (Figure 2.12), precisely the opposite of expectation based on the clear improvement afforded by higher tow numbers in the simulation study of Poussard et al. (in prep.). Correspondence analysis demonstrates the danger of the use of adaptive decisions during depletion experiments without rigorous empirical determining criteria designed to optimize the cost and benefit of increased tow number. The danger of terminating a depletion experiment early based on a potentially misleading depletion curve is present as well.

In the correspondence analysis, Err2, Err3, and Err4 are associated with experiments characterized by smaller dredges, higher CVs for the efficiency estimates, and higher CVs for the density estimates. These characteristics co-occurring instill suspicion as to the quality of the results obtained from a subset of the depletion experiments. Essentially, experiments falling at or above the 80$^{th}$ percentile for Err2, Err3, and Err4 are associated with experiments that have low efficiency estimates and, for Err3 and Err4, high uncertainty in the efficiency estimates, strongly suggesting deletion of these experiments from further evaluation of the inherent efficiency of hydraulic clam dredges.

### 2.4.2 ***Estimation of Density***

Interestingly, experiments with high CV for density estimates are grouped with the low efficiency experiments identified by Err2, Err3, and Err4 in the correspondence analysis. Poussard et al. (in prep.) clearly show that efficiency and density are not correlated in simulated depletion experiments, a logical outcome based on an expectation that hydraulic dredges should be equally efficient whether used in low density or high density regions, even though, the Patch Model estimates of efficiency mathematically are negatively correlated with the density estimates. The apposition of high CV for the density estimate and low efficiency is likely a product of high uncertainty in the density estimate co-occurring with high uncertainty in the efficiency estimate. This could be indicative of an experiment design failing or environmental parameters not being conducive to estimating accurate and precise efficiency and density estimates.

The accuracy of the Patch Model density estimate was evaluated thoroughly in Hennen et al. (2012). The *k*-parameter, the negative binomial dispersion parameter, was not evaluated for accuracy in that study, however, because the distribution of clams in space was not created using a negative binomial distribution. The *k*-parameter is indirectly related to the distribution of clams and tow distance (Hennen et al. 2012). The simulations of Poussard et al. (in prep) show that the *k*-parameter estimates are higher with a uniform distribution of clams and lower with a more irregular distribution of clams. This parameter is influenced by the same conditions of the experiment that influence efficiency, but correspondence analysis clearly separates this variable from other variables such as the efficiency estimate, density estimate, depth, region, dredge width, and the CVs of the density and efficiency estimates. (Figures 2.13 and 2.14). Correspondence analysis identifies a tendency for high *k*-parameter and uncertainty in the *k*-parameter (the CV) to be associated with low OS and shallow depths. The latter however is almost certainly a byproduct of the tendency for surfclam experiments to have lower OS. The effect of low OS and hence low tow number dominates this association.

Poussard et al. (in prep.) showed clearly that the dispersion of clams on the bottom can cause a decrease in performance in the depletion experiment. This outcome is exacerbated by low tow number and low tow overlap. Despite the distribution of clams exerting a strong influence on the error in efficiency estimates in the simulation study, correspondence analyses jointly show that the error in efficiency estimates inferred for the in-field experiments is not correlated with the inferred distribution of clams. Of course, Poussard et al. (in prep.) tested only a subset of a vast number of possible clam dispersion patterns, but those tested were extreme cases. In practice, every experiment, no matter how many dredge tows were used and the degree of overlap in the tow paths, would appear to be equally susceptible to producing an unreliable efficiency estimate if the distribution of clams in the benthos is irregular. The fact that clam dispersion is a random effect for the in-field experiments despite its documented importance in determining outcomes is consistent with the fact that the locations for the experiments were chosen without any *a priori* knowledge of the local dispersion characteristics at the site.

2.4.3 ***Factors Affecting Field Outcomes***

The size of the dredge is related to the efficiency estimated, with larger dredges being used with experiments with higher efficiency estimates. Smaller dredges were used in many experiments and these contributed disproportionately to the subset identified by error estimates Err2, Err3, and Err4 (Figure 12). Dredge size and OS increased with year as well, so the possibility exists that the cause of the increase in precision of efficiency estimates that has to do with the year in which the experiment was conducted is an increased reliance on larger dredges in the experimental protocol. The majority of suspect experiments identified by the four error estimates were conducted in 1997, 1999, and 2005, and among these experiments are those categorized as having lower efficiency estimates with more uncertainty in the estimate. Although speculative, two possibilities may be forwarded explaining this trend. A wider dredge may be inherently more efficient as loss in efficiency is likely associated with the encounter of clams near the lateral edges of the dredge knife blade, and these clams are a lower fraction of the potential catch with the larger dredge. In addition, the narrow dimension of the depletion rectangle was generally set at 10 dredge widths; thus, the larger dredge was used to deplete larger regions which may have reduced the influence of small-scale variations in clam dispersion within the depletion rectangle. It is noteworthy that experiments conducted with the largest dredge were in later years, when depletion experiment methodology was more consistent among experiments, and produced higher efficiency estimates, yielding higher OS measurements. Accordingly, the improved performance cannot unequivocally be assigned to the larger size of the dredge used.

Location of the depletion experiment might also affect the efficiency estimate. Ocean quahog depletion experiments conducted off Long Island have higher efficiency estimates than experiments conducted further south. The relationship is shown objectively (Figure 2.7) and in correlation (Figure 2.10). The correspondence analysis does not show a significant relationship between latitude and the efficiency estimate, but this result accrues from the inclusion of high-efficiency surfclam experiments that took place further south (Figure 2.8). The relationship is not associated with dredge width, although efficiency and dredge width are significantly correlated for ocean quahog experiments (Figure 2.7). These experiments took place in deeper water, on the average, but correlation and correspondence analysis agree that depth, per se, does not influence outcomes. Edaphic factors may be invoked for the influence of region, but little information is available to make a determination.

Depth might be considered to be an effective variable determining the success of a depletion experiment for hydraulic dredges as these dredges are operated using an onboard water pump attached to the dredge by means of a large hose. The vessel is less maneuverable in deeper water due to the increased amount of hose required to maintain an adequate scope while dredging. Surprisingly, neither correlation analysis nor correspondence analysis offers any evidence for a significant correlation between depth and experimental performance or the final efficiency estimate. Depth related variables, in fact, fall orthogonally to experiment performance metrics and error estimates Err2, Err3, and Err4 in correspondence analysis.

### 2.4.4 ***The Efficiency of Hydraulic Dredges***

When 16 experiments (8 surfclam and 8 ocean quahog experiments) that fell at or above the 80$^{th}$ percentile for error estimates Err2, Err3 and Err4 are removed from the in-field depletion dataset, the mean efficiency estimate increased from 0.635 to 0.719 for surfclam experiments (Table 2.11). The median likewise rose substantially from 0.590 to 0.686 and the interquartile range, though remaining relatively unchanged in dimension, shifted to higher efficiency values. The mean efficiency estimate for ocean quahog experiments increased from 0.586 to 0.700, the median also rose from 0.629 to 0.667. The interquartile range was substantially reduced in dimension and also shifted to higher efficiency values. The efficiency estimates for the dataset after all experiments flagged by an error term are removed are included to show that Err1 experiments do not have efficiency estimates that are biased in either direction and do not meaningfully negate the trends established by the other three error terms. Interestingly, the mean and median efficiency estimates for these hydraulic dredges targeting surfclams and ocean quahogs are nearly identical. Neither the species nor the presence of one generally in deeper water than the other significantly influences the overall efficiency which stands at approximately 70% regardless of mean or median determination.

Table 2.11

| | | Mean | Standard Deviation | 1st Quartile | Median | 3rd Quartile |
|---|---|---|---|---|---|---|
| Ocean Quahog | Efficiency Estimates (All Experiments) | 0.586 | 0.260 | 0.381 | 0.629 | 0.779 |
| | Efficiency Estimates (8 flagged by Err 2,3,4 removed) | 0.700 | 0.177 | 0.595 | 0.667 | 0.787 |
| | Efficiency Estimates (9 flagged by all Error terms removed) | 0.707 | 0.196 | 0.561 | 0.683 | 0.795 |
| Surfclam | Efficiency Estimates (All Experiments) | 0.635 | 0.229 | 0.533 | 0.590 | 0.779 |
| | Efficiency Estimates (8 flagged by Err 2,3,4 Removed) | 0.719 | 0.171 | 0.583 | 0.686 | 0.889 |
| | Efficiency Estimates (13 flagged by all Error terms removed) | 0.740 | 0.179 | 0.583 | 0.725 | 0.899 |

Comparing mean, SD, median, and quartiles for all 19 ocean quahog and 31 surfclam experiments with the dataset after 16 experiments in the 80$^{th}$ percentile for error terms Err2, Err3, and Err4 were removed.

The Wilcoxon rank sums tests conducted on efficiency estimates between experiments falling at or above and below the 80$^{th}$ percentile for Err2, Err3, and Err4 show that the three groups of experiments have significantly different efficiency estimates and CVs from the remainder. Though these error estimates can only be used to infer experimental quality, they identify experiments with a range of questionable attributes which strongly implicate them as outliers biasing the efficiency estimates for the entire dataset. Removing these questionable experiments provides the best estimates of efficiency for these commercial hydraulic dredges and emphasizes that these are the most efficient dredges in use today.

## 2.5 Conclusions

Between 1997 and 2011, the National Marine Fisheries Service conducted 50 depletion experiments to estimate survey gear efficiency and stock density for Atlantic surfclam (*Spisula solidissima*) and ocean quahog (*Arctica islandica*) populations using commercial hydraulic dredges. A model formulated for this purpose, the Patch Model, was used to estimate gear efficiency and organism density. The range of efficiencies estimated is substantial, leading to uncertainty in the application of these estimates in stock assessment. A simulation protocol was developed to examine sources of uncertainty in Patch Model estimates. Analysis of simulations showed that uncertainty in the estimates of gear efficiency from depletion experiments was reduced by higher numbers of dredge tows per experiment, more tow overlap in the experimental area, a homogeneous as opposed to a patchy distribution of clams in the experimental area, and the use of gear of inherently high efficiency. Stock density was of lesser importance, though still contributing to estimated uncertainty. Simulations suggest that adapting the experimental protocol during the depletion experiment by adjusting tow number and degree and dispersion of tow overlap may substantively reduce uncertainty in the final efficiency estimates. Simulations also suggest that the pattern of population dispersion in the experimental area is, and will likely remain, an important source of uncertainty, which may, however, be mitigated by updating experimental design during the course of the experiment.

Known values of four descriptive metrics for each in-field experiment: the average effective area swept (EAS), the overlap score (OS) describing tow overlap, the coefficient of variation (CV) for the efficiency estimate, and the CV of the *k* parameter (the negative binomial dispersion parameter) were compared to metrics from the 9,000 simulations in the simulation dataset to determine which experiments diverge from those in the simulation dataset, and which experiments were likely to have high error in the efficiency estimate. The error metrics used implicate a subset of experiments that are outliers, biasing the efficiency estimates for the entire dataset. Though these error estimates can only be used to infer experimental quality, they identify experiments with a range of questionable attributes which strongly implicate them as outliers biasing the efficiency estimates for the entire dataset. When these outlier experiments are removed from the in-field depletion dataset, the mean efficiency estimate increased from 0.635-0.719 for surfclam experiments. The mean efficiency estimate for ocean quahog experiments increased from 0.586 to 0.700. The median values rose accordingly, from 0.590 to 0.686 for surfclam experiments and from 0.629 to 0.667 for ocean quahog experiments. The mean and median hydraulic dredge efficiency estimates for the surfclam and ocean quahog depletion experiments are almost identical. Neither the species, nor the fact that ocean quahogs are generally found in deeper water than surfclams substantially influences the overall efficiency of the dredge, which is estimated to be approximately 70%. The dispersion of clams inferred from simulation experiments suggests that clam distribution affects all experiments as a random factor increasing uncertainty in the estimate of efficiency. Removing the questionable experiments identified as outliers through the error metrics provides the best estimates of efficiency for these commercial hydraulic dredges and emphasizes that these are the most efficient dredges in use today.

REFERENCES


Beukers-Stewart, B. D., & Beukers-Stewart, J. S. (2009). Principles for the Management of Inshore Scallop Fisheries around the United Kingdom. Report to Natural England, Scottish Natural Heritage and Countryside Council for Wales. In *Marine Ecosystem Management Report no. 1*. Retrieved from http://eprints.whiterose.ac.uk/105473/

Bohrmann, T.F., & Christman, M.C. (2012). Evaluating sampling efficiency in depletion surveys using hierarchical Bayes. *Canadian Journal of Fisheries and Aquatic Sciences* 69, 1080-1090.

Butler, P., Wanamaker, A., Scourse, J., Richardson, A., & Reynolds, D. (2013). Variability of marine climate on the North Icelandic Shelf in a 1357-year proxy archive based on growth increments in the bivalve *Arctica islandica*. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *373*, 141–151. doi: 10.1016/j.palaeo.2012.01.016

Chai, A. C., Homer, M., Tsai, C-F,.& Goulletque, P. (1992). Evaluation of oyster sampling efficiency of patent tongs and an oyster dredge. *North American Journal of Fisheries Management 12*, 825-832.

Clausen, S-E. (1998) Applied Correspondence Analysis An Introduction. Sage Publications, Sage University Papers, Series N.07-121, 69 pp.

Da Ros, L., Nesto, N., Nasci, C., Moschino, V., Pampanin, D., & Marin, M. G. (2003). Biochemical and behavioural effects of hydraulic dredging on the target species *Chamelea gallina*. *Fisheries Research*, *64*, 71–78. doi: https://doi.org/10.1016/S0165-7836(03)00201-7

Gedamke, T., DuPaul, W. D., & Hoenig, J. M. (2005). Index-removal estimates of dredge efficiency for sea scallops on Georges Bank. *North American Journal of Fisheries*

*Management*, *25*, 1122–1129. doi: 10.1577/M04-111.1

Gilkinson, K. D., Fader, G. B. J., Gordon, D. C., Charron, R., McKeown, D., Roddick, D., Liu, Q. (2003). Immediate and longer-term impacts of hydraulic clam dredging on an offshore sandy seabed: effects on physical habitat and processes of recovery. *Continental Shelf Research*, *23*, 1315–1336. doi: 10.1016/S0278-4343(03)00123-7

Hauton, C., Howell, T. R. W., Atkinson, R. J. A., & Moore, P. G. (2007). Measures of hydraulic dredge efficiency and razor clam production, two aspects governing sustainability within the Scottish commercial fishery. *Journal of the Marine Biological Association of the United Kingdom*, *87*, 869–877. doi: 10.1017/S0025315407055300

Hennen, D. R., Jacobson, L. D., & Tang, J. (2012). Accuracy of the Patch model used to estimate density and capture efficiency in depletion experiments for sessile invertebrates and fish. *ICES Journal of Marine Science*, *69*, 240–249. doi: 10.1093/icesjms/fsr212

Hofmann, E. E., Powell, E. N., Klinck, J. M., Munroe, D. M., Mann, R., Haidvogel, D. B., Narváez, D. A., Zhang, X., & Kuykendall, K. M. (2018). An overview of factors affecting distribution of the Atlantic surfclam (*Spisula solidissima*), a continental shelf biomass dominant, during a period of climate change. J*ournal of Shellfish Research 37*, 821-831.

Kleisner, K., Fogarty, M.J., Mcgee, S., Hare, J., Morét, S., Perretti, C., Saba, V.S. (2017). Marine species distribution shifts on the U.S. Northeast Continental Shelf under continued ocean warming. *Progress In Oceanography*, *153*, 24-36.

Lambert, J., & Goudreau, P. (1996). Performance of the New England hydraulic dredge for the harvest of Stimson's surf clams (*Mactomeris polynyma*). *Canadian Industry Report of Fisheries and Aquatic Science*, *235*, 28.

Lasta, M. L., & Iribarne, O. O. (1997). Southwestern Atlantic scallop (*Zygochlamys patagonica*) fishery: assessment of gear efficiency through a depletion experiment. *Journal of Shellfish Research*, *16*, 59–62.

Leslie, P. H., & Davis, D. H. S. (1939). An attempt to determine the absolute number of rats on a given area. *Source Journal of Animal Ecology*, *8*, 94–113. doi: 10.2307/1255

Mann, R., Southworth, M., Harding, J. M., & Wesson, J. (2004). A comparison of dredge and patent tongs for estimation of oyster populations. *Journal of Shellfish Research*, *23*, 387–390.

Meseck, S. L., Mercaldo-Allen, R., Rose, J. M., Clark, P., Kuropat, C., Pereira, J. J., & Goldberg, R. (2014). Effects of hydraulic dredging for *Mercenaria mercenaria*, northern quahog, on sediment biogeochemistry. *Journal of the World Aquaculture Society*, *45,* pp. 301–311. doi: 10.1111/jwas.12114

Meyer, T. L., Cooper, R. A., & Pecci, K. J. (1981). The performance and environmental effects of a hydraulic clam dredge. *Maine. Fisheries Review*, *43(9)*, 14-22.

Miller, T. J., Hart, D. R., Hopkins, K., Vine, N. H., Taylor, R., York, A. D., & Gallager, S. M. (2019). Estimation of the capture efficiency and abundance of Atlantic sea scallops (*Placopecten magellanicus*) from paired photographic dredge tows using hierarchical models. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*, 847–855. doi: 10.1139/cjfas-2018-0024

Morson, J. M., Munroe, D. M., Ashton-Alcox, K. A., Powell, E. N., Bushek, D., & Gius,

J. (2018). Density-dependent capture efficiency of a survey dredge and its influence on the stock assessment of eastern oysters (*Crassostrea virginica*) in Delaware Bay. *Fisheries Research*, *205*, 115–121. doi: 10.1016/j.fishres.2018.04.012

NEFSC (2003). *Assessment of Atlantic surfclams. In Report of the 37th Northeast Regional Stock Assessment Workshop (37th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fisheries Science Center* Reference Document 03-16. Northeast Fisheries Science Center (NEFSC), Woods Hole, MA.597 pp

NEFSC. (2007). *Assessment of ocean quahogs. In Report of the 44th Northeast Regional Stock Assessment Workshop (44th SAW). Northeast Fisheries Science Center Reference Document 07-10*. Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 661 pp

NEFSC. (2010a). *Assessment of Atlantic monkfish. In Report of the 50th Northeast Regional Stock Assessment Workshop (50th SAW). Northeast Fisheries Science Center Reference Document 10-17.* Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 378 pp

NEFSC. (2010b). *Assessment of Atlantic sea scallops. In Report of the 50th Northeast Regional Stock Assessment Workshop (50th SAW). Northeast Fisheries Science Center Reference Document 10-17*. Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 316 pp

NEFSC. (2010c). *Assessment of Atlantic surfclams. In Report of the 49th Northeast Regional Stock Assessment Workshop (49th SAW). Northeast Fisheries Science Center Reference Document 10-03*. Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 383 pp

NEFSC. (2013). *Assessment of Atlantics Surfclams. In Report of the 56th Northeast Regional Stock Assessment Workshop (56th SAW). Northeast Fisheries Science Center Reference Document 13-10.* Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 868 pp

NEFSC. (2017a). *Assessment of Atlantic surfclams*. *In Report of 61st Northeast Regional Stock Assessment Workshop (61st SAW). Northeast Fisheries Science Center Reference Document 17-05.* Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 466 pp

NEFSC. (2017b). *Assessment of ocean quahogs. In Report of 63rd Northeast Regional Stock Assessment Workshop (63rd SAW). Northeast Fisheries Science Center Reference Document 17-10*. Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. 409 pp.

NMFS. (2009). Fisheries Economics of the United States. *United States Department of. Commerce, National Oceanic and Atmospheric Administration Technical Memorandum NMFS-F/SPO-118*, 172. Retrieved from https://www.st.nmfs.noaa.gov/st5/publication/index.html.

Pace, S. M., Powell, E. N., Mann, R., & Long, M. C. (2017). Comparison of age-frequency distributions for ocean quahogs *Arctica islandica* on the western Atlantic US continental shelf. *Marine Ecology Progress Series 585*, 81-98.

Paloheimo, J. E., & Dickie, L. M. (1964). Abundance and fishing success. *Rapport et Proces-Verbaux Des Réunions Conseil International pour l'Exploration de la Mer*, *155*, 152–163.

Parker, R. S. (1971). History and development of surf clam harvesting gear. *National Oceanic and Atmospheric Association Technical Report* NMFS CIRC-364, 15 pp.

Pezzuto, P. R., Schio, C., & Almeida, T. C. M. (2010). Efficiency and selectivity of the *Anomalocardia brasiliana* (Mollusca: Veneridae) hand dredge used in southern Brazil. *Journal of the Marine Biological Association of the United Kingdom 90*, 1455-1464.

Powell, E. N., Ashton-Alcox, K. A., & Kraeuter, J. N. (2007). Reevaluation of eastern oyster dredge efficiency in survey mode: application in stock assessment. *North American Journal of Fisheries Management*, *27*, 492–511. doi: 10.1577/M04-205.1

Powell, E. N., Ewing, A. M., & Kuykendall, K. M. (2020). Ocean quahogs (*Arctica islandica*) and Atlantic surfclams (*Spisula solidissima*) on the Mid-Atlantic Bight continental shelf and Georges Bank: the death assemblage as a recorder of climate change and the reorganization of the continental shelf benthos. *Palaeogeography Palaeoclimatoogy Palaeoecology 537*, #109205, 16 pp.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. URL https://www.R-project.org/.

Ragnarsson, S. A., & Thórarinsdóttir, G. G. (2002). Abundance of ocean quahog, *Arctica islandica*, assessed by underwater photography and a hydraulic dredge. *Journal of Shellfish Research 21,* 673-676.

Rago, P. J., Weinberg, J. R., & Weidman, C. (2006). A spatial model to estimate gear efficiency and animal density from depletion experiments. *Canadian Journal of Fisheries and Aquatic Sciences*, *63*, 2377–2388. doi: 10.1139/f06-121

Skalski, J. R., Robson, D. S., & Matsuzaki, C. L. (1983). Competing probabilistic models for catch-effort relationships in wildlife censuses. *Ecological Modelling*, *19*, 299–307. doi: 10.1016/0304-3800(83)90045-5

Sokal, R. R., & Rohlf, F. J. (1998). Biometry. The Principals And Practice Of Statistics In Biological Sciences. 3rd edition. W.H. Freeman and Company, New York.

Thórarinsdóttir, G. G., Jacobson, L., Ragnarsson, S. A., Garcia, E. G., & Gunnarsson, K. (2010). Capture efficiency and size selectivity of hydraulic clam dredges used in fishing for ocean quahogs (*Arctica islandica*): simultaneous estimation in the SELECT model. *ICES Journal of Marine Science* 67, 345-354.

Vølstad, J., Davis, G., Sharov, A., & Davis, B. (2000). A method for estimating dredge catching efficiency for blue crabs, *Callinectes sapidus*, in Chesapeake Bay. *Fishery Bulletin 98,* 410–420. Retrieved from papers2://publication/uuid/563EEE06-EDBE-4969-BDC2-09546680C2A7

Wilberg, M. J., Robinson, J. M., Rains, S. A. M., Humphrey, J. L., & Lipcius, R. N. (2013). Effects of location errors on estimates of dredge catchability from depletion based methods. *Fisheries Research*, *148*, 1–8. doi: 10.1016/j.fishres.2013.07.016