

# Final Report

## Sea Scallop Research

Grant: NOAA/NA06NMF4540261

Award Date: 11 Aug 2006

Start Date: 1 Aug 2006

End Date: 31 July 2007

Final Report: 29 Oct 2007

Project Title: **High-Resolution Video Survey of the Habitat and Sea Scallop Resource in the Elephant Trunk Closed Area.**

Principal Investigator: Kevin D. E. Stokesbury, Ph.D.

Address: School for Marine Science and Technology,  
University of Massachusetts Dartmouth,  
838 South Rodney French Blvd., New Bedford, MA, 02744-1221

Phone: (508) 910-6373

Fax: (508) 999-8197

Email: [kstokesbury@umassd.edu](mailto:kstokesbury@umassd.edu)

## 1. Project Summary

---

**Objectives:** We will examine the sea scallop abundance and evaluate their spatial distribution, size composition and density as well as habitat characteristics in the Elephant Trunk Closed Area using a high-resolution cooperative industry-based video survey.

---

**Methodology:** We will conduct two 7-day research video cruises in the Elephant Trunk Closed Area in the mid-Atlantic. The sampling procedure for these surveys will be a multistage systematic design with stations separated by approximately 2.2 km similar but a slightly larger scale than the 1999-2005 SMAST high-resolution (1.5 km) video surveys. These surveys will provide a highly precise and accurate estimate of size specific sea scallop density and a series of maps of the sea floor detailing the distribution of substrate, depth, live scallops, dead scallops, and macroinvertebrates (sponges, sea star, filamentous fauna).

---

**Conclusions:** Video surveys were conducted in the Elephant Trunk Closed Area to examine scallop abundance and evaluate their spatial distribution, size composition and density. The total sea scallop biomass in the Elephant Trunk Closed Area ranged from 97.1 to 126.4 million pounds and exploitable biomass ranged from 67.8 to 89.7 million pounds. The spatial distribution of scallops was highly aggregated and spatially segregated by size. Surficial substrates within the Elephant Trunk Closed area were dominated by sand and shell debris. These results were used to verify total exploitable biomass projected in Framework 18, and enabled the NEFSC and NMFS to adjust the trip estimates.

---

**Rationale:** In 2005 sea scallop fishing vessels were allocated 3 trips of 18,000 lbs each to the Hudson Canyon Closed Area designated in Framework 16 (6-42). Catch rates were much lower than anticipated based on projections from the 2002 NMFS dredge survey, and many vessels were unable to complete their 3 trips. New management regulations and alternatives had to be implemented. This process was extremely costly to fishermen, managers and the public. Projected densities in the Elephant Trunk Area are extremely high (as they were for the Hudson Canyon Area). The Framework 18 draft document estimates a total exploitable biomass of 55,300 mt (110.6 million lbs) for 2006. The preferred alternative is a harvest of 12,335 mt (27.2 million lbs) collected in 5 fishing trips per vessel in the 2007 fishing year. The SMAST video survey at the 5.56 km scale estimated a total biomass of 41,322 mt (91.1 million lbs) in May-June 2005. Thus considerable growth of the population must occur in this area to meet the projected exploitable biomass estimate in Framework 18. To ensure that a repeat of the 2005 Hudson Canyon fishing experience does not occur we propose this high-resolution video survey in May 2006 of the Elephant Trunk area. This survey will provide estimates of scallop density that are highly accurate and precise with CV's between 5 and 10% depending on scallop spatial distributions. We will provide preliminary estimates of density immediately to the NEFMC and the NMFS and quality controlled final estimates and size frequencies within 2 months of collection. These data will either verify the earlier projected biomass or enable the NEFMC and NMFS to adjust the trip estimates to the Elephant Trunk Area. Further these surveys will provide detailed maps of the habitat within this closed area prior to the fishing effort enabling a comparison Before-After-Control-Impact analysis similar to those conducted by SMAST for the Georges Bank closed areas.

---

## **2. Description of the issue/problem**

Projected total exploitable scallop biomass in the Elephant Trunk Closed Area in Framework 18 was extremely high at 55,300 mt (110.6 million lbs). The framework allocated 5 fishing trips per vessel in the 2007 fishing year to harvest 12,335 mt (27.2 million lbs). There was considerable uncertainty in the projections because a substantial majority of the scallops were young; true abundance of young scallops is difficult to estimate with a high degree of precision (letter from the Scallop PDT to the NEFMC and Patricia Kurkul, Regional Administrator entitled, "Updated exploitable biomass estimate for Elephant Trunk Area (ETA)" on 30 October 2006, attached document). In efforts to avoid hardships to fishermen, managers and the public resulting from overestimation of these projections under Framework 18, the Regional Administrator was given the authority to reduce the number of trips based on surveys conducted during 2006. We examined scallop abundance and evaluated their spatial distribution, size composition and density as well as habitat characteristics in the Elephant Trunk Closed Area using a high-resolution video survey. This work addresses the NEFMC and NMFS priorities of identifying the physical and biological variables associated with scallop and Essential Fish Habitat, improving information concerning scallop abundance estimates, and conducting high-resolution surveys examining distribution, recruitment, mortality and growth rate (such as size frequency) information.

### **Project goals and objectives:**

Using a high-resolution cooperative industry-based video survey, we examined the sea scallop abundance and evaluated their spatial distribution, size composition and density as well as habitat characteristics in the Elephant Trunk Closed Area. The sampling procedure for these surveys was a multistage systematic design with stations separated by approximately 2.2 km. This survey provided a highly precise and accurate estimate of size specific sea scallop density and produced a series of maps of the sea floor detailing the distribution of substrate, depth, live scallops, dead scallops, and macroinvertebrates (sponges, sea stars, filamentous fauna).

### **The problem addressed:**

We examined scallop abundance and evaluated their spatial distribution, size composition and density in the Elephant Trunk Closed Area using a high-resolution video survey to verify projected total exploitable biomass, enabling the NEFSC and NMFS to adjust the trip estimates.

## **3. Approach**

We conducted two video research cruises inside of a 4,676.2 km<sup>2</sup> sample area within the Elephant Trunk Closed Area (ETCA) from 23 August to 1 September 2006 (Fig. 1). We video surveyed 916 stations within the ETCA using a multistage systematic design with stations separated by approximately 2.2 km.

Commercial sea scallop fishing vessels were used as survey vessels. The underwater video surveys were conducted using the SMAST sampling Pyramid. The pyramid supports three live-feed S-VHS underwater video cameras (Stokesbury 2002, Stokesbury et al. 2004). The video cameras are configured such that two downward looking cameras view 3.235 m<sup>2</sup> and 0.8 m<sup>2</sup> (nested within the 3.235m<sup>2</sup> view) quadrats, and one side-looking camera provides a view across the quadrat parallel to the sea bed. Four quadrats were sampled at each station, increasing the sample area to 12.94 m<sup>2</sup>. The time, depth, number of live and dead scallops, latitude and

longitude were recorded at each station. After each survey the videotapes were reviewed in the laboratory and a still image of each quadrat was digitized. The field data were verified and the shell height (mm) of each scallop was measured in the still image using Image Pro Plus® software. Within each quadrat, macroinvertebrates and fish were counted and the substrate was identified (Stokesbury 2002). When possible fish and macroinvertebrates were identified to species, otherwise animals were grouped into categories based on taxonomic orders. Unidentified fish were grouped as “other fish.” Counts were standardized to individuals  $m^{-2}$ . Sponges, hydrozoa/bryozoa and sanddollars were recorded as present or absent within a quadrat.

### **Sea scallops (*Placopecten magellanicus*)**

#### Abundance and Density

Mean scallop density ( $m^{-2}$ ) and standard error were calculated using equations for a multi-stage sampling design (Cochran 1977, Krebs 1989, Stokesbury 2002).

The absolute number of scallops within a survey area was calculated by multiplying scallop density by the total area surveyed (Stokesbury 2002). Estimates of scallop meat weight in grams were derived from shell height (mm) frequencies collected during the survey and a shell height to meat weight regression fit to data from dissections of live scallops collected on 11/8/01 and 7/17/06, SARC 39, and VIMS survey. The mean meat weight was multiplied by the total number of scallops in the survey area to estimate the total biomass of scallop meats.

#### Shell Height Composition

Scallop shell heights were measured with ImagePro Plus software using still images digitized from the video survey footage (Stokesbury 2002, Stokesbury et al. 2004). Recent calibration experiments (Stokesbury et al., in review) show that the lens curvature corrections we used Stokesbury (2002), and Stokesbury et al. (2004) are unnecessary, therefore uncorrected shell height measurements are used in this analysis.

#### Spatial distribution

Distributions of scallops  $m^{-2}$  were plotted using Arcview. To facilitate visualization of the scallop, other macroinvertebrates, and substrate data the *Thessian tool* (ArcInfo®) was used to create a grid of polygons (each centered on a survey station). Each polygon was given the megabenthos and substrate attributes of the survey station it contained. This technique is simple and does not involve mathematical interpolation to create a surface for visualizing the data.

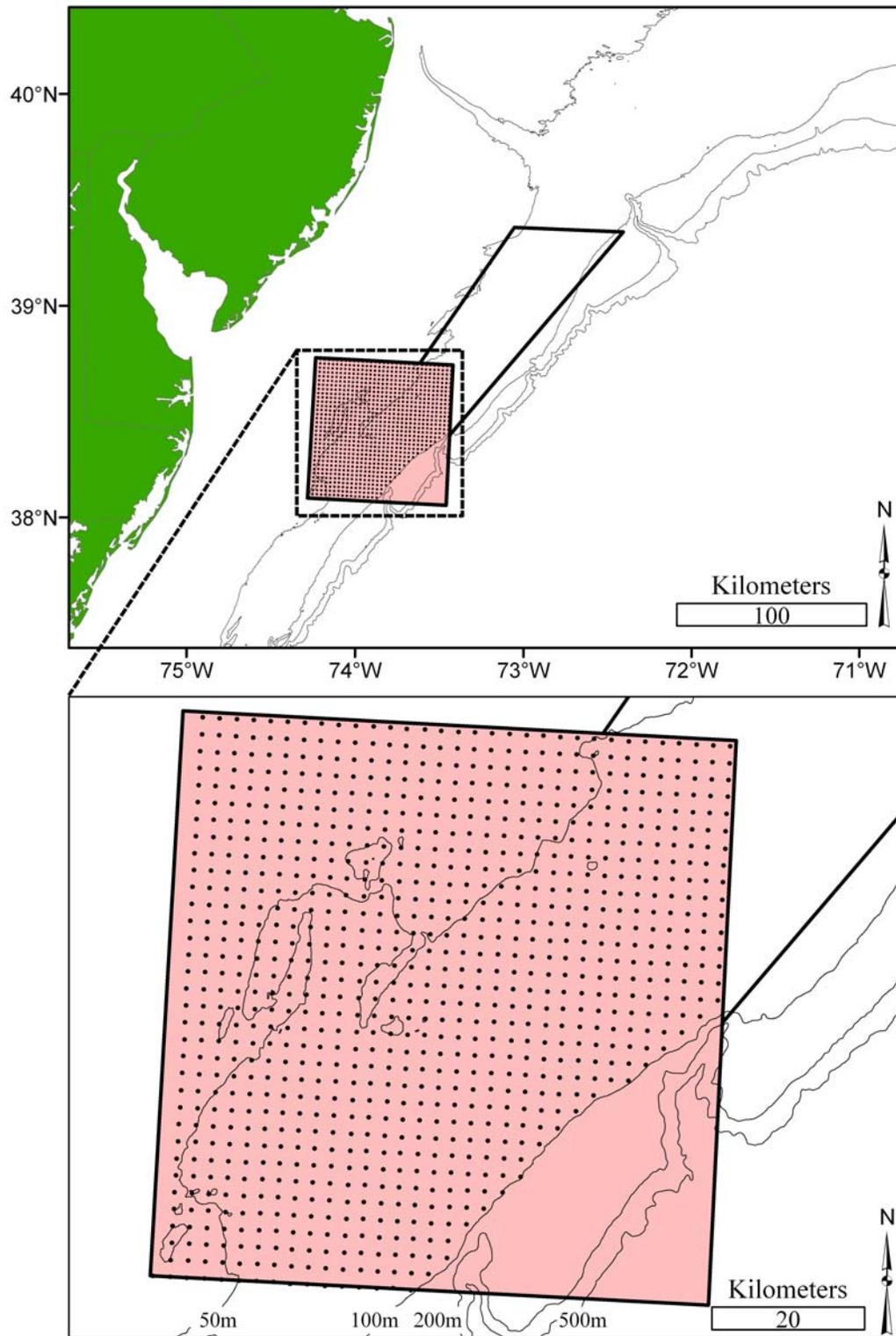


Figure 1. Map of the Elephant Trunk Area with the 2.2 km grid video surveys.

## Habitat

### Surficial Substrate

Substrates were visually identified and categorized based on the Wentworth particle grade scale (Wentworth 1922, Lincoln et al. 1992). We used texture, color, relief and structure in the video footage and still images, following the Wentworth scale (Fig. 2).

### Macrobenthos

Similar to scallops, sea star density ( $m^{-2}$ ) and standard error were calculated using equations for a multi-stage sampling design (Cochran 1977, Krebs 1989, Stokesbury 2002). The presence or absence of hydrozoans/bryozoans and sponges were examined in each survey using percent presence and area ( $km^{-2}$ ) occupied. The distinction between these groups and other megabenthic groups are that individual counts are not practical.

The sea stars species group includes: *Solaster endeca*, *Crossaster papposus*, *Leptasterias polaris*, *Asterias* spp., *Henricia* spp. The hydrozoans/ bryozoans species group includes: *Flustra foliacea*, *Callopora aurita*, *Electra monostachys*, *Cribrilina punctata*, *Eucratea loricata*, *Tricellaria ternata*, *Eudendrium capillare*, *Sertularia cupressina*, *Sertularia argentea*. The sponges species group includes: *Suberites ficus*, *Haliclona oculata*, *Halichondria panicea*, *Cliona celata*, *Polymastia robusta*, *Isodictya palmata*, *Microiona prolifera*.

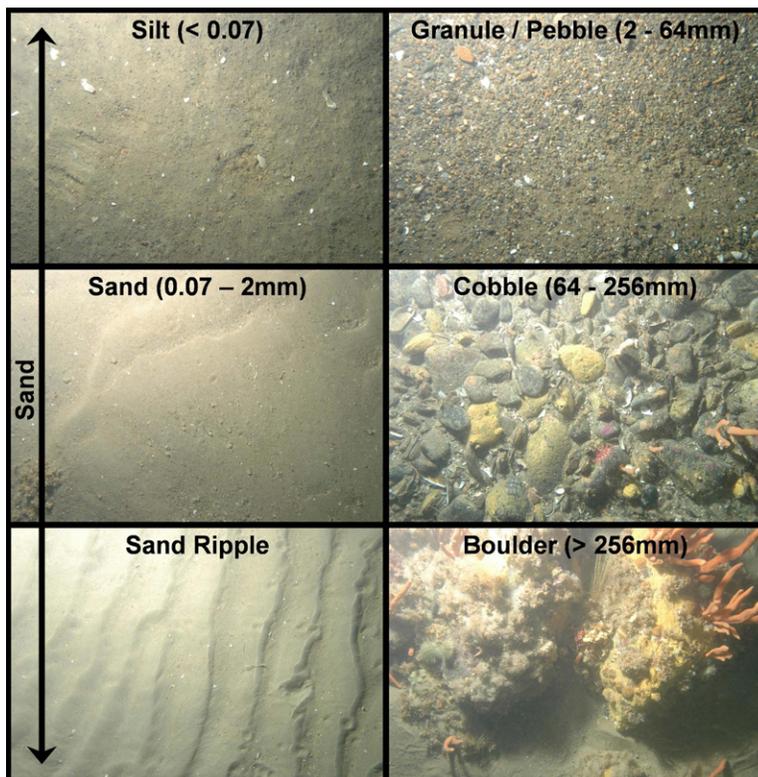


Figure 2. Digital still images of Sand, Granule/Pebble, Cobble and Boulder substrates including particle size ranges.

**4. Results** (the data provided to the scallop PDT and used in this report are attached to this report as an excel spread sheet).

We have completed the 14 harvest cruises proposed in grant NOAA/ NA06NMF4540261. Table 1 is a time-line listing the cruises completed.

The 14 harvest trips were completed before the end of the fishing year (2/28/07), the date below is the one on the checks received by the University.

<b>Date</b>	<b>F/V Conducting Harvest Trip (CAII)</b>
11/6/2006	Celtic
11/8/2006	Resolute
11/2/2006	Venture
<b>Date</b>	<b>F/V Conducting Harvest Trip (NLCA)</b>
11/3/2006	Endeavor
12/18/2006	Mary Anne
<b>Date</b>	<b>F/V Conducting Harvest Trip (Open Area)</b>
11/27/2006	Justice
2/26/2007	Venture
12/19/2006	Huntress
11/29/2006	Friendship
12/12/2006	Araho
12/6/2006	Chief
1/29/2007	Guidance
11/13/2006	Endeavor
1/19/2007	Silver Sea

Before each cruise I contacted Mr. Ryan Silva or Mr. Don Frie of the NMFS who provided a letter of authorization for the research and harvest cruises and notified the Coast Guard of our activities (weight-out sheets for these trips are provide in an attached PDF).

#### **Sea scallops (*Placopecten magellanicus*)**

##### Abundance and Density

The density of all scallops within the Elephant Trunk Closed Area was 0.62 scallop·m<sup>-2</sup> (SE = 0.034), whereas the density of scallops greater than the harvestable size (102 mm) was 0.34 scallop·m<sup>-2</sup> (Table 1).

Using four separate scallop shell height-meat weight relationships, the total sea scallop biomass surveyed within the Elephant Trunk Closed Area ranged from 97.1 to 126.4 million pounds (Table 1). The exploitable biomass, using a knife-edge selectivity for a 4" ring, ranged from 67.8 to 89.7 million pounds. The exploitable biomass, using the NMFS selectivity curve, ranged from 75.9 to 115.8 million pounds.

Table 1. The mean number of scallops·m<sup>-2</sup>, number of stations sampled on a 2.2 km grid with the SMAST video pyramid, standard error, CV%, estimated millions of lbs of scallop meats within the survey area. Four different scallop shell height – meat weight relationships were used to derive the biomass estimates, based on dissections from 11/8/01 and 7/17/06, SARC 39, and VIMS survey.

<b>MA ET 2006 (2.2 km grid)</b>						Estiamtion of biomass		number of	Biomass	
SMAST (cor)	per quad	per m	stations	SE	CV%	mwt lf (g)	Area sampled m <sup>2</sup>	scallops	mt	mill lbs
<b>Dis 11/8/01</b>	<b>2.00</b>	<b>0.62</b>	916	<b>0.0339</b>	5.49	<b>15.3</b>	4676243308	2886689338	44056	97.1
4" knife edge	<b>0.55</b>	<b>0.34</b>				<b>19.4</b>	4676243308	1587679136	30743	67.8
NMFS sel cur						<b>18.4</b>		1872080910	<b>34403</b>	75.9
<b>SARC 39</b>	<b>2.00</b>	<b>0.62</b>	916	<b>0.0339</b>	5.49	<b>19.9</b>	4676243308	2886689338	57325	126.4
4" knife edge	<b>0.55</b>	<b>0.34</b>				<b>25.6</b>	4676243308	1587679136	40688	89.7
NMFS sel cur						<b>24.3</b>		1872080910	<b>45406</b>	100.1
<b>VIMS</b>	<b>2.00</b>	<b>0.62</b>	916	<b>0.0339</b>	5.49	<b>18.2</b>	4676243308	2886689338	52510	115.8
4" knife edge	<b>0.55</b>	<b>0.34</b>				<b>23.7</b>	4676243308	1587679136	37550	82.8
NMFS sel cur						<b>22.4</b>		1872080910	<b>41863</b>	115.8
<b>Dis 7/17/06</b>	<b>2.00</b>	<b>0.62</b>	916	<b>0.0339</b>	5.49	<b>18.7</b>	4676243308	2886689338	54001	119.1
4" knife edge	<b>0.55</b>	<b>0.34</b>				<b>24.0</b>	4676243308	1587679136	38107	84.0
NMFS sel cur						<b>22.7</b>		1872080910	<b>42560</b>	93.8

### Spatial distribution

Scallops were aggregated in the Elephant Trunk Closed Area with the highest densities occurring between the 50 and 100 m isobaths (Fig. 3). Clapper (dead sea scallops with shells still attached) densities were low with the highest number of occurrences in the northeast corner of the Elephant Trunk Closed Area (Fig. 4).

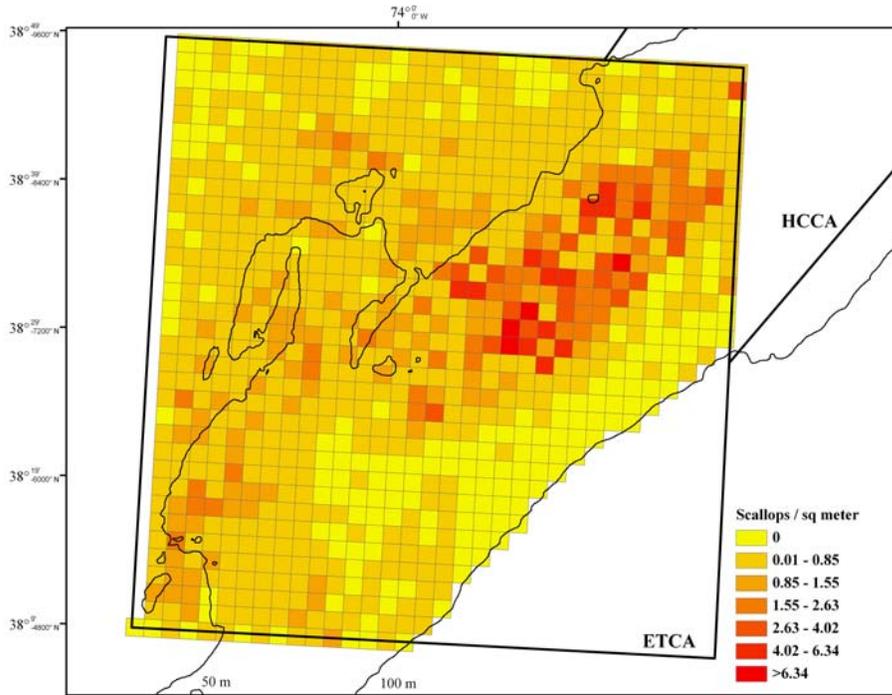


Figure 3. Sea scallop distributions (scallops  $m^{-2}$ ) in the Elephant Trunk Closed Area observed during the SMAST video survey from 23 August to 1 September 2006.

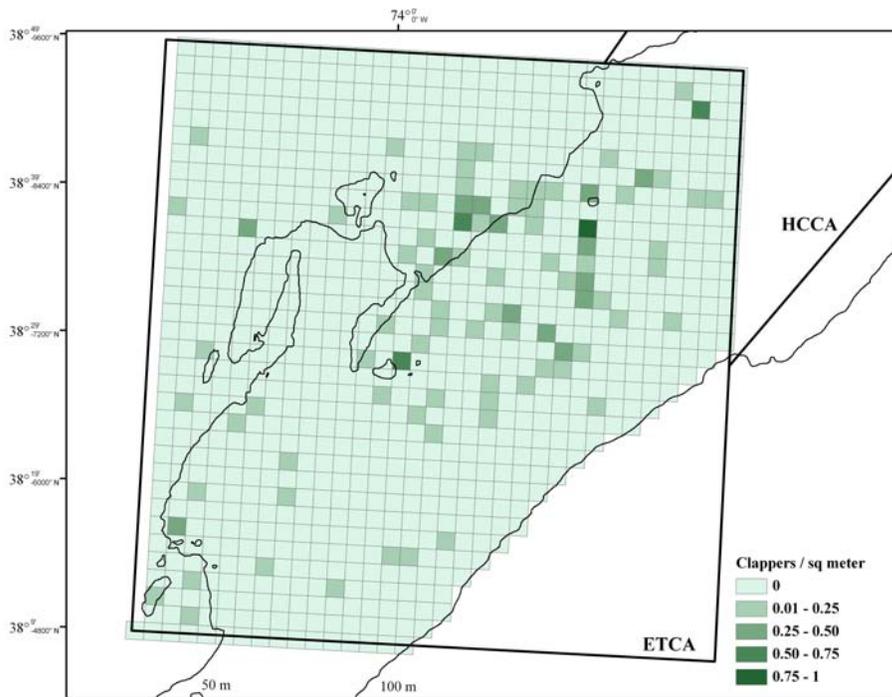


Figure 4. Clapper density (dead sea scallops with shells still attached; clappers  $m^{-2}$ ) in the Elephant Trunk Closed Area.

### Shell Height Composition

Shell heights were measured for a total of 4155 scallops (Fig. 5). The average size scallop was 102.6 mm (SD = 18.89 mm), with approximately 65% of the scallops greater than the harvestable size (102 mm).

There appears to be segregation of scallops by size (Fig. 6), which is examined in more detail in Supporting Doc. 1 (Rothschild et al. (in prep)). Back-of-the-envelope-calculations suggest prosecuting a relatively high fishing mortality on large scallops and a relatively low fishing mortality on small scallops in the Elephant Trunk Closed Area. There is a large concentration of small scallops (shell height less than 95 mm) located in the Northeastern portion of the Elephant Trunk Closed Area. Most of the larger scallops are in the western portion of the Elephant Trunk Area.

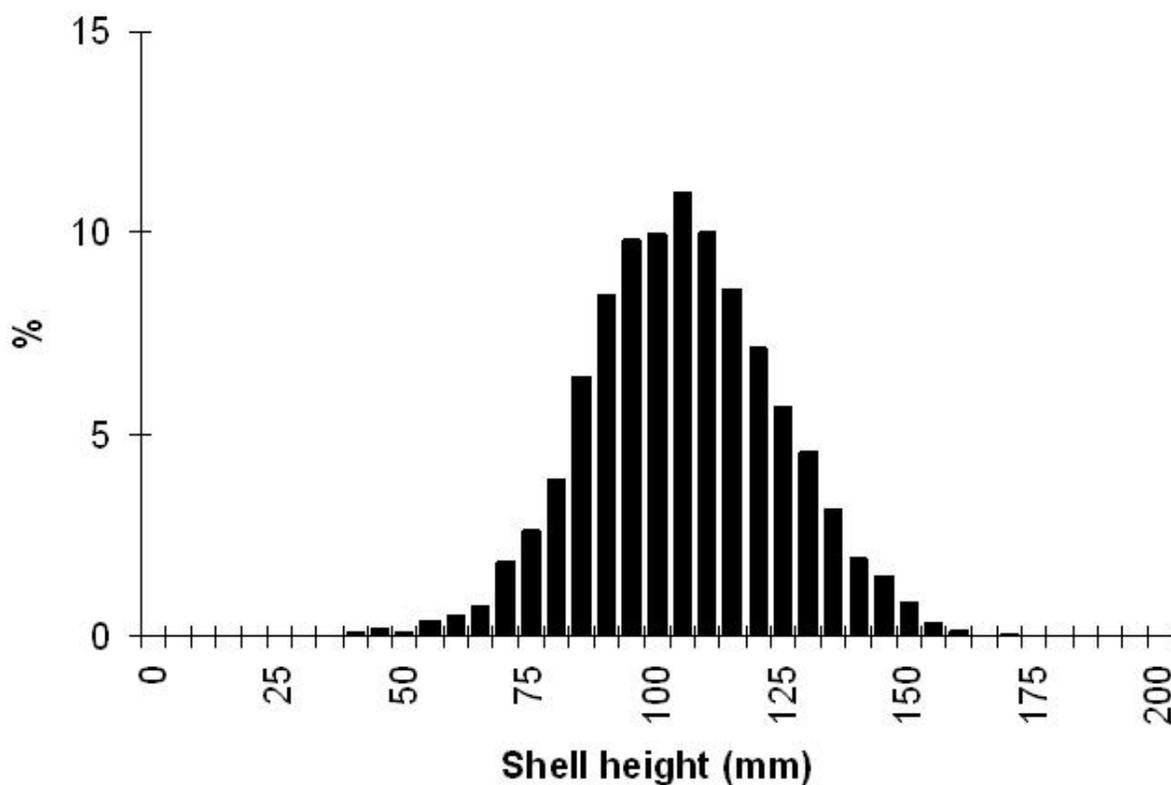


Figure 5. Shell height frequencies of sea scallops in the Elephant Trunk Closed Area.

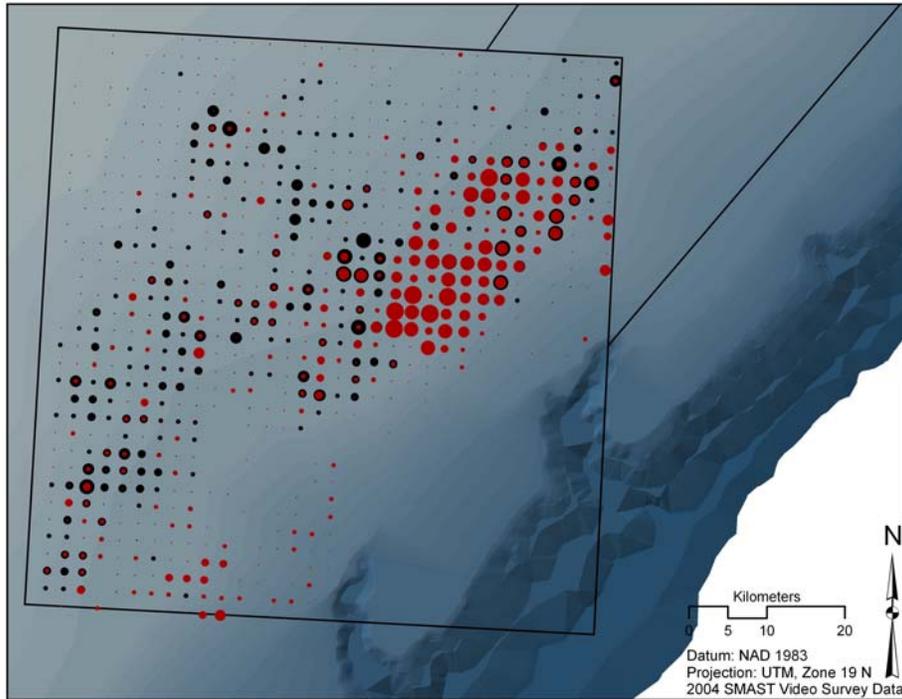


Figure 6. Size-specific sea scallop distributions (scallops·m<sup>-2</sup>) in the Elephant Trunk Closed Area. Red circles indicate scallops less than 95mm and black circles indicate scallops greater than 95mm.

## **Habitat**

### Surficial Substrate

Sand and shell debris dominated (> 99%) the Elephant Trunk Closed Area (Fig. 7). Out of the 3664 quadrats sampled, only 6 quadrats had gravel present and 1 quadrat had cobble present. Station depths ranged from 33 to 101 m and averaged 55.7 m (SD = 11.62 m) (Fig. 8).

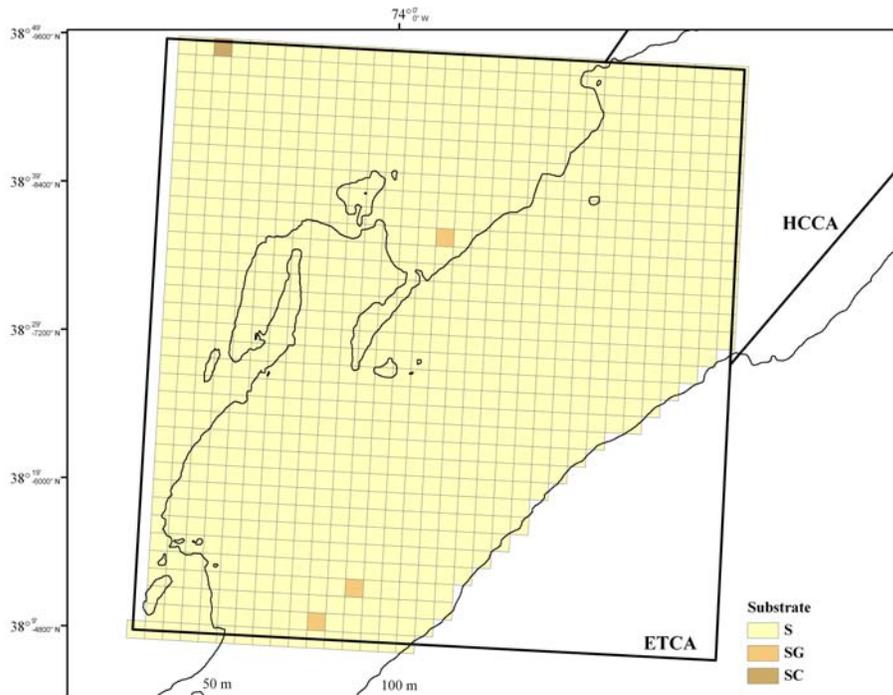


Figure 7. Substrate in the Elephant Trunk Closed Area.

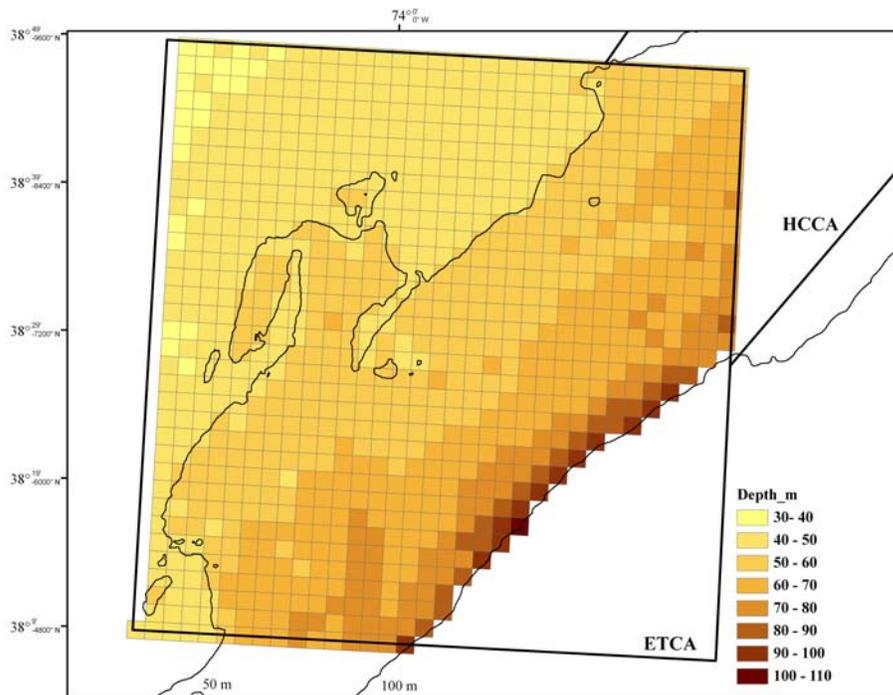


Figure 8. Depth (m) in the Elephant Trunk Closed Area.

### Megabenthos Densities, Presence/Absence

Sea stars were the most abundant macroinvertebrates observed in the Elephant Trunk Closed Area (1.16 sea stars·m<sup>-2</sup>; SE = 0.057). Sea star densities appeared highest between the 50m and 100m isobaths (Fig. 8). Table 2 indicates the area (km<sup>2</sup>) and percent area sampled where each species group was present or absent in the Elephant Trunk Closed Area. Sponges appeared most abundant at depths less than 50 m (Fig. 10), whereas hydrozoans/bryozoans were randomly distributed across the survey area (Fig. 11).

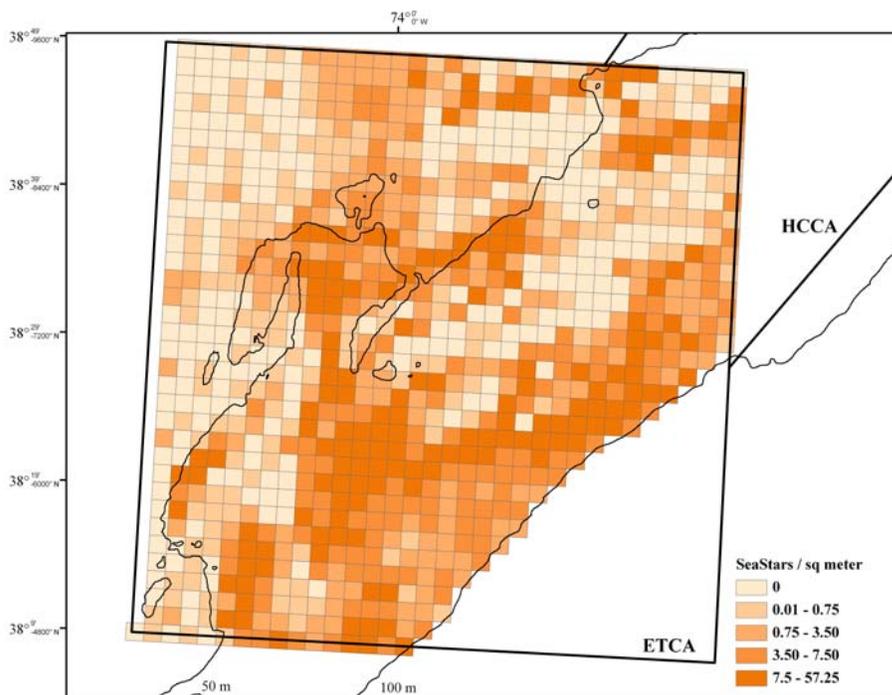


Figure 9. Sea star distributions (sea stars·m<sup>-2</sup>) in the Elephant Trunk Closed Area.

Table 2. Area (km<sup>2</sup>) and % Area sampled where sponges and hydrozoans/bryozoans were present or absent in the Elephant Trunk Closed Area.

<b>Sponges</b>		
	<b>km<sup>2</sup></b>	<b>% Area</b>
<b>Present</b>	622.82	13.3%
<b>Absent</b>	4053.45	86.7%

<b>Hydrozoans / Bryozoans</b>		
	<b>km<sup>2</sup></b>	<b>% Area</b>
<b>Present</b>	924.02	19.8%
<b>Absent</b>	3752.25	80.2%

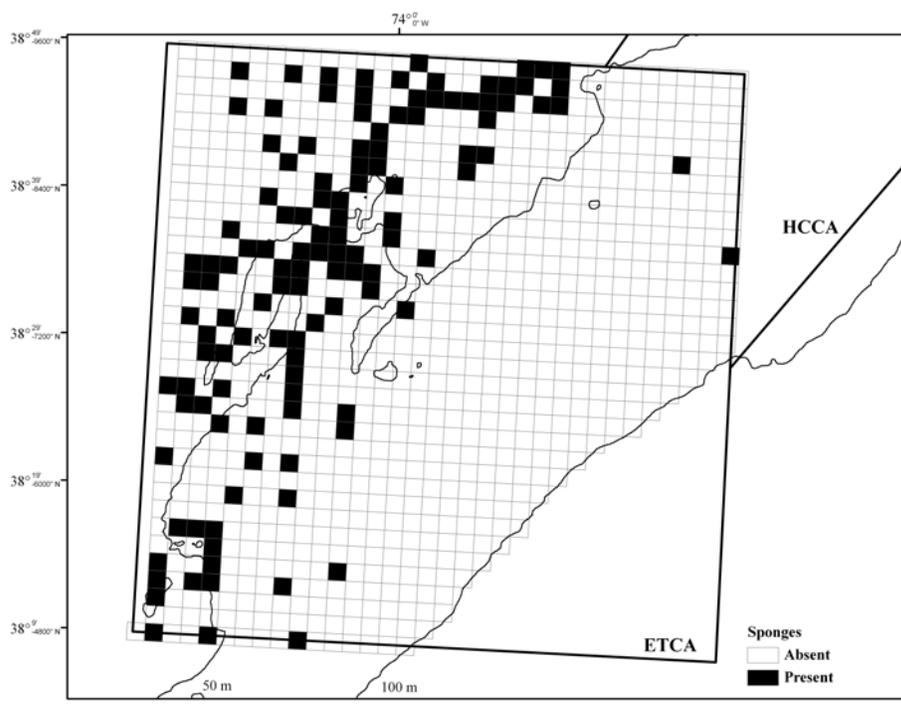


Figure 10. Sponge distributions (present or absent) in the Elephant Trunk Closed Area.

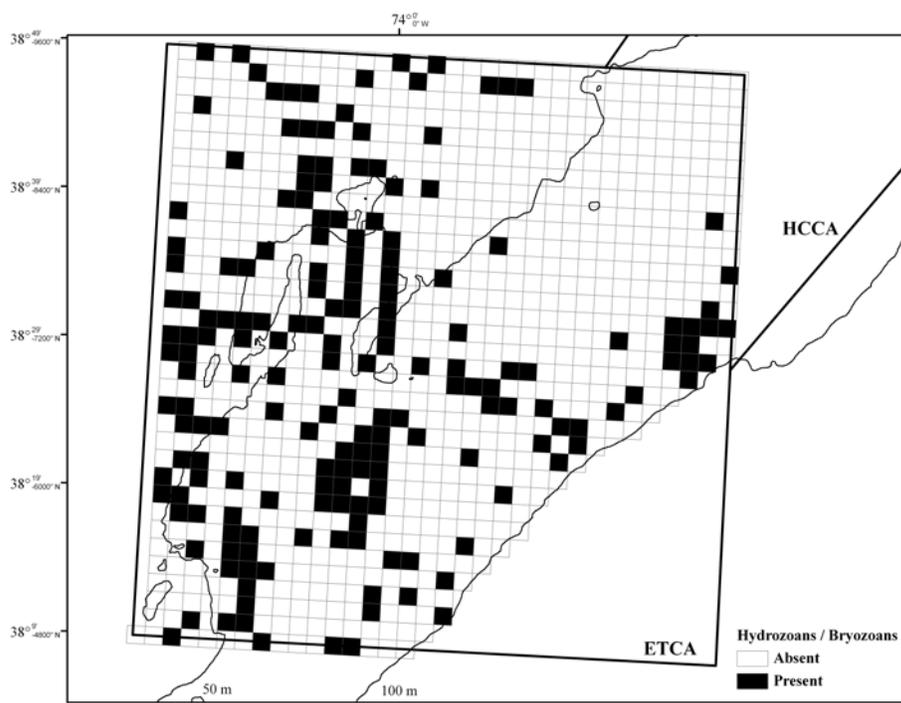


Figure 11. Hydrozoan / Bryozoan distributions (present or absent) in the Elephant Trunk Closed Area.

## **Evaluation**

The total sea scallop biomass in the Elephant Trunk Closed Area ranged from 97.1 to 126.4 million pounds and exploitable biomass ranged from 67.8 to 89.7 million pounds. The spatial distribution of scallops was highly aggregated with the highest densities occurring between the 50 and 100 m isobaths. Scallop distribution was also segregated by size, suggesting prosecution of a relatively high fishing mortality on large scallops and a relatively low fishing mortality on small scallops in the Elephant Trunk Closed Area to optimize yield from the resource.

Surficial substrates within the Elephant Trunk Closed Area were dominated by sand and shell debris. Sea stars were the most abundant macroinvertebrates observed in the Elephant Trunk Closed Area

We accelerated the time line and met the objectives of this proposal in advance due to the request for data from the NEFMC Scallop PDT for the NMFS Emergency Action Interim Rule on the Elephant Trunk Closed Area. As a result, we worked on examining the scallop population density and distribution using traditional and geostatistical analysis, and a manuscript (Adams et al., submitted, attached document) was submitted to ICES Journal of Marine Science.

Furthermore, we are preparing a manuscript, “Notes on the Dynamics of Scallops in the Elephant Trunk Area,” (Rothschild et al., in prep, attached document). The manuscript examines the growth potential of scallops, spatial distribution by size, estimates of abundance, and management implications. Preliminary results suggest prosecuting a relatively high fishing mortality on large scallops and a relatively low fishing mortality on small scallops, which will easily be accomplished because of spatial aggregation (Fig. 6).

## **Benefits and contributions to management decision making**

This work addresses the NEFMC and NMFS priorities of identifying the physical and biological variables associated with scallop and Essential Fish Habitat, improving information concerning scallop abundance estimates, and conducting high-resolution surveys examining distribution, recruitment, mortality and growth rate (such as size frequency) information. The 2006 Elephant Trunk Closed Area video survey data (Table 1 and Fig. 3) were presented to the Scallop PDT (25 Oct 06) and served as an integral factor in the decision to call for the Elephant Trunk Area Emergency Action measures (NOAA 2006).

The updated biomass estimate from this survey was consistent with dredge surveys conducted by the NMFS and the Virginia Institute of Marine Science (VIMS). Although the biomass in the Elephant Trunk Closed Area remained very high relative to the rest of the scallop resource, it was less abundant than was projected in Framework 18. As a result, even though the fishing mortality is expected to be lower than the target fishing mortality in the area, it would be high enough at the lower biomass to contribute to resource wide overfishing in the 2007 fishing year (NOAA 2006). While, none of the surveys suggested a reduction in the number of trips that should be allocated in the Elephant Trunk Closed Area based on the defined thresholds identified in Framework 18, the PDT stressed the need to use precaution when managing the scallop resource in this area. The New England Fisheries Management Council reduced the number of scallop harvest trips allocated in the Elephant Trunk Closed Area as a result of these recommendations (Attached document letter from PDT to Regional Administrator 30 Oct 2006).

### List of Supporting Documents:

Adams, CF, BP Harris, KDE Stokesbury. (submitted) Geostatistical comparison of two independent video surveys of sea scallop abundance in the Elephant Trunk Closed Area, USA. ICES J Mar Sci.

Rothschild, BJ, KDE Stokesbury, BP Harris, C Sarro. (in prep) Notes on the Dynamics of Scallops in the Elephant Trunk Area.

### References

Cochran W.G. 1977. Sampling Techniques. 3rd edition. New York: John Wiley & Sons.

Krebs C. J. 1989. Ecological methodology. Harper & Row, New York

Lincoln R. J, Boxshall G. A, Clark P. F. 1992. A dictionary of ecology, evolution and systematics. Cambridge University Press, Cambridge

NOAA (2006) 50 CFR Part 648; Fisheries of the Northeastern United States; Atlantic Sea Scallop Fishery; Interim Rule. US Fed Reg 71(246): 76945-76949.

Stokesbury, K. D. E. 2002. Estimation of sea scallop abundance in closed areas of Georges Bank, USA. Trans Am Fish Soc 131: 1081-1092.

Stokesbury, K. D. E., Harris, B. P., Marino, M. C., and Nogueira, J. I. 2004. Estimation of sea scallop abundance using a video survey in off-shore USA waters. J Shell Fish Res, 23: 33-44.

Stokesbury, K. D. E., Jacobson, L. D., Allard, M., Chute, A., Harris, B. P., Jaffarian, T., Marino, M. C. II, Nogueira, J. I., and Rago, P. (submitted) Quantification, effects and stock assessment modeling approaches for dealing with measurement error in body size data using Atlantic sea scallops as an example, Can J Fish Aquat Sci.

Wentworth, C. K. 1922. A Grade Scale and Class Terms for Clastic Sediments. J Geol, Vol. 30, p 377-392.

# Geostatistical comparison of two independent video surveys of sea scallop abundance in the Elephant Trunk Closed Area, USA

Charles F. Adams, Bradley P. Harris and Kevin D. E. Stokesbury

## Abstract

Geostatistical prediction at unsampled locations is done by kriging, an interpolation technique that minimizes the error variance. Our goal was to verify the technique by comparing kriged abundance estimates with true counts from an area containing the highest sea scallop (*Placopecten magellanicus*) densities offshore of the northeastern USA. In 2006 two independent video surveys of scallop abundance were done in the Elephant Trunk Closed Area, one using a  $5.6 \times 5.6$  km grid and the other using a  $2.2 \times 2.2$  km grid. We generated kriged surfaces of scallop abundance with the 5.6 km grid data using different combinations of semivariograms and theoretical models, then tested the null hypothesis of no difference between the predicted and true values (2.2 km grid data). There were significant differences between predicted and true values for three out four combinations of semivariogram–model fits to untransformed data, assuming isotropy. In contrast, there was no significant difference between kriged and true values for any combination of semivariogram–model fits to log–transformed, detrended data. Classical and robust semivariograms performed equally well. Kriging can be used to generate accurate maps of scallop abundance if the assumptions of geostatistics are met.

Keywords: geostatistics, kriging, *Placopecten magellanicus*, sea scallop, semivariogram

C.F. Adams, B.P. Harris and K.D.E. Stokesbury: School for Marine Science and Technology, University of Massachusetts Dartmouth, 838 South Rodney French Boulevard, New Bedford MA 02744–1223, USA. Correspondence to: C.F. Adams: tel +1 508 910 6386; fax +1 508 910 6396; email [cadams@umassd.edu](mailto:cadams@umassd.edu).

## Introduction

Geostatistics is a branch of spatial statistics that was developed by Matheron (1963, 1971) for the mining industry. The technique describes the spatial structure of a natural resource and can be used to predict values at unsampled locations. The strength of geostatistics is that it exploits the spatial correlation inherent in environmental data through the semivariogram. Geostatistics is now considered an important tool for estimating the distribution and abundance of fish stocks (Petitgas, 1993; Rivoirard *et al.*, 2000).

Geostatistical prediction at unsampled locations is done by kriging. The distinguishing feature of kriging, as compared with other interpolation techniques, is that it minimizes the error variance (Isaaks and Srivastava, 1989). Although simulation studies have examined the effect of methodological choices on kriging results (e.g., Rivoirard *et al.*, 2000; Rufino *et al.*, 2006), we aimed to verify the technique by comparing kriged abundance estimates with true counts from an area containing the highest sea scallop (*Placopecten magellanicus*) densities offshore of the northeastern USA (Stokesbury *et al.*, 2004). Kriging has been used to generate sea scallop abundance and biomass estimates (Conan, 1985; Ecker and Heltshe, 1994; Warren, 1998), as well as to estimate scallop dredge efficiency (Gedamke *et al.*, 2005).

The School for Marine Science and Technology (SMAST), University of Massachusetts Dartmouth, in cooperation with members of the USA commercial sea scallop industry,

developed a video survey to assess sea scallop abundance offshore of the northeastern USA. Since 2003, 7200 quadrat samples covering approximately 60 000 km<sup>2</sup> of continental shelf, including Georges Bank and the mid-Atlantic, have been examined annually using a 5.6 × 5.6 km grid (Stokesbury *et al.*, 2004). In 2006 an additional, independent video survey of sea scallop abundance was done in the Elephant Trunk Closed Area (ETCA) using a 2.2 × 2.2 km grid (Figure 1) to provide a precise estimate of scallop abundance prior to reopening the area to harvest in 2007.

We tested the null hypothesis of no difference between the mean predicted and actual scallop counts by generating a kriged surface of sea scallop abundance using the 5.6 km grid data, and then comparing these predicted values with the actual scallop counts observed at the 2.2 km grid stations. This analysis complements previous simulation studies by providing insights into the effects of methodological options, such as detrending data, choice of semivariogram, etc., on kriging results. As spatial models and spatially explicit management strategies, such as time-area closures and marine protected areas, are increasingly being used to manage fisheries (Jensen and Miller, 2005), our work offers some guidelines for the use of kriging in these important management decisions.

## Materials and methods

### Study area

The Elephant Trunk Closed Area (Figure 1) is a ca. 5400 km<sup>2</sup> of continental shelf that was closed to commercial sea scallop fishing in 2004 as part of an area–rotation management plan, with a scheduled reopening in 2007 (50 CFR Part 648, 2006). Sea scallop densities in the ETCA vary widely, ranging from areas with no scallops to areas with the highest densities of scallops offshore of the northeastern USA (Stokesbury *et al.*, 2004).

### Field sampling

The 5.6 km grid survey of the ETCA was conducted from 25 June to 18 July 2006, while the 2.2 km grid survey was conducted from 24 to 31 August 2006. Grid resolution for each survey was determined by balancing the logistics of covering the study area with maintaining low coefficients of variation, assuming a random or negative binomial distribution (Stokesbury *et al.* 2004).

Both surveys used the SMAST video survey pyramid (Stokesbury *et al.*, 2004). We used a centric systematic sampling design (Krebs, 1999). An element of randomization was added by incorporating a random starting point for the entire grid (Simmonds and Fryer, 1996; Rivoirard *et al.*, 2000). Sampling was done with a live–feed downward looking S–

VHS underwater video camera viewing a 3.235 m<sup>2</sup> quadrat. Four replicate quadrats were sampled at each station (Stokesbury *et al.*, 2004). Pooling the four replicates increases the sampled view area to 12.94 m<sup>2</sup> to allow detection of densities as low as 0.08 scallops m<sup>-2</sup>, which corresponds to minimum commercially viable levels (Brand, 1991; Stokesbury, unpublished data, 1992). After each cruise technicians reviewed the footage to verify the scallop counts recorded during the survey (Stokesbury *et al.* 2004).

#### Data preparation

Geographical referencing was done by setting the southwest corner of the ETCA as 0, 0 and converting all coordinates to km (Rivoirard *et al.*, 2000). The midpoint of the latitude in the ETCA (38.5°N) was used to convert longitude.

Geostatistical analysis was done on summed sea scallop counts for each station. Values can be expressed in density (scallops m<sup>-2</sup>) by simply dividing the counts by the station sampled view area (12.94 m<sup>2</sup>).

Normality of the 5.6 km grid data ( $n = 154$ ) was assessed using the Kolmogorov–Smirnov test, with critical values recomputed for tests of normality (Stephens, 1974). This test is more robust in the presence of autocorrelation than other tests of normality (Dutilleul and Legendre, 1992; Legendre and Legendre, 1998). There was a significant departure from normality ( $D = 0.29, p < 0.01$ ) that improved somewhat with log–transformation ( $D = 0.13, p < 0.01$ ). Although normality is not a requirement for kriging,

the procedure works best when the distribution is close to normal (Isaaks and Srivastava, 1989).

Post plots with row and column means (Webster and Oliver, 2001) suggested the presence of a slight longitudinal and/or latitudinal trend in the data (Figure 2) that, unless removed, would violate the assumption of intrinsic stationarity (i.e., constant mean and variance throughout the sampling space). However, the mean kriged estimate using the untransformed 5.6 km grid data, assuming isotropic conditions, was close to the true 2.2 km grid mean. Thus, we performed two concurrent geostatistical analyses (Figure 3): one using the actual 5.6 km grid scallop counts, assuming isotropic conditions (AB); and the other using the residuals of the log-transformed 5.6 km grid scallop counts, corrected for anisotropy (RE).

After log-transformation the RE data were fitted with a locally weighted regression, or loess (Cleveland and Devlin, 1988), using easting and northing as predictor variables (depth was dropped from the final equation due to lower multiple  $R^2$  values). Next, this trend was subtracted from the data and the geostatistical analysis was performed on the residuals (Kaluzny *et al.*, 1998; Lo *et al.*, 2001; Giannoulaki *et al.*, 2003; Mello and Rose, 2005). Finally, directional semivariograms revealed a constant sill with differing ranges, indicating a geometric anisotropy in the residuals. This was corrected by determining the direction of maximum spatial continuity, calculating the ellipse ratio, then rotating and rescaling the data (Isaaks and Srivastava, 1989). Note that the trend was added back to the kriged values for calculating summary statistics, plotting kriged maps, etc.

## Geostatistics

We constructed omnidirectional empirical semivariograms for the AB and RE data using the classical estimator of Matheron (1963) and the robust estimator of Cressie and Hawkins (1980). Data were binned at 5.6 km intervals, as the grid spacing is the appropriate lag spacing for data sampled on a grid (Isaaks and Srivastava, 1989). The maximum distance between pairs of points in the 5.6 km grid data was 100 km, so we initially calculated semivariograms with a maximum distance of 50 km, as only half the total distance measured in any direction may be legitimately represented in a semivariogram (Rossi *et al.*, 1992). Erratic behavior was observed in the last lag so the maximum distance was decreased to 44 km for all subsequent analysis.

We fit spherical and exponential models (Isaaks and Srivastava, 1989) to all empirical semivariograms using the method of weighted least squares (Cressie, 1985). The Gaussian model was not used because it can lead to unstable kriging equations in the absence of a nugget effect (Chilés and Delfiner, 1999), and some authors discourage the use of this function altogether (Webster and Oliver, 2001).

Theoretical model fits yielded nugget, sill and range estimates. The nugget  $C_0$  is a discontinuity from the origin at distance  $h = 0$ . It describes measurement error and/or microscale variation, the latter resulting from small scale variation not detected with the sampling grid. The sill is the asymptote of the semivariogram that occurs at range  $a$ . The

sill consists of the nugget and the partial sill  $C$ , the latter describing the spatial component of the semivariance  $\gamma$ . The range describes the extent of spatial correlation in the data (Isaaks and Srivastava, 1989) and the average patch diameter (Webster and Oliver, 2001).

Spatial correlation and patch diameter were also examined visually using standardized semivariograms and correlograms (Rossi *et al.*, 1992). The semivariogram can be an unreliable measure of spatial continuity if the assumption of intrinsic stationarity is not met. Correlograms, which are a plot of correlation coefficients by lag, filter out the lag means and variances, thereby providing a better estimate of the range in the absence of intrinsic stationarity. Standardized semivariograms and correlograms were also calculated for the 2.2 km grid to determine if a finer scale sampling grid would reveal different spatial structures.

Ordinary kriging was used to generate predicted scallop counts at 2.2 km grid stations ( $n = 852$ ).

#### Comparison of kriged vs. true values

The null hypothesis  $H_0$  of no difference between the mean predicted and true values was assessed qualitatively by comparing the predicted values  $\hat{z}$  with the “true” scallop counts  $z$  at the 2.2 km grid stations using methods outlined in Isaaks and Srivastava (1989).

Summary statistics were compared for  $\hat{z}$  and  $z$ , with the expectation that the distribution of  $\hat{z}$  should be similar to that of  $z$ . The univariate distribution of the error  $\hat{z} - z$  was also

examined, with the expectation that the mean, median and standard deviation would all be 0. Two additional statistics that incorporate both the bias and the spread of the error distribution are the mean absolute error:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{z} - z| \quad (1)$$

and the mean squared error:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{z} - z)^2 \quad (2)$$

A formal test of  $H_0$  was done with a Wilcoxon paired sample test by ranks (Zar, 1999).

The errors  $\hat{z} - z$  were not normally distributed (Table 1), which is an assumption of the paired  $t$ -test (Zar, 1999). Thus the nonparametric test was used, after we verified that the errors were symmetrical around the median (Table 1).

QQ plots (Wilk and Gnanadesikan, 1968) were used to compare the distributions of the kriged vs. true values.

We plotted colorscale maps of predicted values without contouring to avoid the additional interpolation that is introduced with the latter process to enable a more accurate comparison of kriged vs. true maps (Figure 4). Colorscale values were chosen to reflect minimum commercial densities, as well as departures from the 45° line  $\hat{z} = z$  observed in the QQ plots.

All analyses were done using S+ and the module S+SpatialStats (Insightful Corporation, Seattle WA, USA), except for the Kolmogorov–Smirnov tests of normality, which were done in PROC UNIVARIATE (SAS Institute, Cary, North Carolina, USA).

## Results

### Geostatistics

The standardized semivariogram and correlogram gave range estimates for the 5.6 km data of ca. 17.2 km (Figure 5). Similarly, the apparent range in the standardized semivariogram for the 2.2 km data was between 15.7 to 17.8 km. In contrast with these three plots, however, was the correlogram for the 2.2 km data, which showed an asymptote between 20.0 and 24.6 km. Another noteworthy trend in Figure 5 is the smoothness of the correlogram asymptotes relative to those observed in the standardized semivariograms. This confirmed the geographic trend suggested in Figure 2 and that the assumption of intrinsic stationarity was questionable.

The spherical model gave the best fit to the AB data, while the exponential model gave the best fit to the RE data (Table 2). For the AB data, spherical fits to the classical and robust semivariograms gave range estimates of 17.9 and 18.7 km, respectively. For the RE data, exponential fits to the classical and robust semivariograms gave range estimates of 21.0 and 20.0 km, respectively.

### Comparison of kriged vs. true values

For the AB data, the closest estimate of the true mean was given by a spherical fit to the classical semivariogram; the robust–spherical gave the closest approximation of the true median, maximum value and  $\sigma$ ; and the classical–exponential gave the minimum value closest to 0 (Table 3). For the RE data, the closest estimate of the true mean and median was given by the classical–exponential; the robust–exponential gave the closest approximation of the maximum value and  $\sigma$ ; and the classical–spherical gave the minimum value closest to 0. Although no combination of semivariogram–model gave consistently better estimates of the true distribution for either the AB or RE data, two general trends were apparent: classical semivariograms always gave the closest estimates of the true mean and minimum values closest to 0; while robust semivariograms gave the best approximations of the maximum value and  $\sigma$ .

Error distributions for the AB data were generally best with the classical–spherical, which gave a mean, maximum,  $\sigma$ , MAE and MSE closest to 0 (Table 4). The robust semivariogram yielded a minimum and median closest to 0 using an exponential and spherical model, respectively. Error distributions for the RE data were better for the robust–exponential, which gave  $\sigma$ , MAE and MSE closest to 0, but there was no consistent pattern for the other statistics.

The null hypothesis of no differences between the predicted and true values was rejected for the AB data with the exception of the robust–spherical, which was barely non–

significant (Table 5). In contrast,  $H_0$  was accepted for all combinations of semivariogram–model to the RE data.

QQ plots for the AB data showed three general trends (Figure 6): negative estimates for 0 true values (particularly for spherical models); a close fit to the 45° line  $\hat{z} = z$  for the majority of true values; and an underestimation of the highest values. The classical–spherical began to underestimate true values at ca. 35 scallops. It should be noted that these underestimates consisted of only 39 out of  $n = 852$  true values, or 4.6% of the total number of predictions. Similarly, the other three combinations of semivariogram–model began to underestimate true values at ca. 60 scallops, which amounts to only 0.9% of the total number of predictions.

QQ plots for the RE data revealed two additional trends (Figure 7). One was the range (i.e., columns) of predictions corresponding to  $\log(0 + 1)$ ,  $\log(1 + 1)$ , etc. Second, once the quantiles began to smooth ca.  $1.1 \log(\text{true})$  scallops, estimates resulting from the robust semivariograms were closer to the 45° line  $\hat{z} = z$  than those from the classical semivariogram.

Colorscale maps of kriged values for the AB data were less patchy than the distribution of true scallops (Figure 8). Amongst the kriged maps themselves, two trends were apparent. One, the classical–exponential had the fewest negative predictions (8), while the robust–spherical had the most (74). However, it should be noted that 171 of  $n = 852$  true values were zeroes: if the negative and 0–0.9 predictions are taken together then the

robust–spherical was actually the most realistic with 163 predictions  $< 1$  scallop. The other trend is that the number of actual values  $\geq 60$  was 9, which was matched only by the robust–spherical, which also had 9 values  $\geq 60$ .

Colorscale maps of kriged values for the RE data were also less patchy than the distribution of true scallops (Figure 9). In this case the classical–spherical had the fewest negative predictions (1), and the robust–exponential matched the true values exactly (171) when negative and 0–0.29 predictions were taken together. As for values  $\geq 1.79$ , the best approximation of the true values (9) was given by the robust–exponential (5).

A comparison of Figures 8 and 9 revealed a smoothing of the spatial structure in the latter along  $48^\circ$  (the direction of maximum spatial continuity) due to detrending and correcting for anisotropy. There were also fewer negative and red values in the RE maps.

## Discussion

### Geostatistics

The range estimates in Table 2, and the apparent ranges in Figure 5, show that data from stations within 15.7 to 24.6 km of each other are correlated. Although this means that classical statistical tests on these data would have an increased probability of committing a Type I error, parameter estimates would be unbiased, as each point has the same

probability of being included in the sample. Only the variance will increase if the distribution is patchy (Legendre and Legendre, 1998).

Our range estimates were larger than the 4.8 km reported by Conan (1985) for scallops in the Northumberland Strait off eastern Canada, and the 9.6 km reported by Gedamke et al. (2005) for scallops in Georges Bank Closed Area II. In contrast, our range estimate was much smaller than the 1° (i.e., 127.8 km) reported by Ecker and Heltshe (1994) for scallops in the New York Bight. These large differences are likely due to the type of data (i.e., abundance vs. biomass) as well as geographic variation in scallop populations. If useful inferences are to be made regarding patch size then what is clearly needed is a large scale analysis of sea scallop populations for both Georges Bank and the mid-Atlantic, utilizing absolute measures (i.e., abundance data) and consistent geostatistical methods.

The apparent range in the 2.2 km correlogram warrants further discussion. At first glance this appears to suggest that the 2.2 km grid would more accurately describe the spatial structure than the 5.6 km grid, once local means and variance have been filtered with a correlogram. But this is not the case. The parameters in Table 2 show that, in general, the RE data gave range estimates comparable to the apparent range in the 2.2 km correlogram. This is to be expected because the RE data were detrended to meet the assumption of intrinsic stationarity. In other words, the 5.6 km grid data can be used to describe the spatial structure, provided the data are properly detrended.

### Comparison of kriged vs. true values

Weighted least squares suggested that the classical–spherical would give the best results for the AB data, and that the classical–exponential would give the best results for the RE data. This was indeed the case in terms of the mean for the AB data, as well as the mean and median for the RE data, but not for other summary statistics. It is important to note that the seemingly small differences in the means reported in Table 3 would translate to tremendous differences in biomass when extrapolated to the entire ETCA. For example, the difference between the mean of the classical–spherical and –exponential for the AB data is 0.1, which would translate to a difference of ca. 42 million scallops over the entire ETCA.

The standard deviations for the kriged estimates reported in Table 3 are lower than the standard deviation for the 2.2 km grid data. This is to be expected: as more samples are incorporated in a weighted linear combination the resulting estimates generally become less variable (Isaaks and Srivastava, 1989). We emphasize that the standard deviations reported in Table 3 are based on summed quadrat counts and thus are not comparable with the two–stage sampling variances reported in Stokesbury *et al.* (2004).

A formal test of the null hypothesis of no difference between predicted and true values revealed that the robust–spherical was the only combination of semivariogram–model for the AB data that was not significantly different as compared with the true mean. This result was not anticipated by the least squares fits or the mean predicted value. However,

other summary statistics, such as the median, maximum and  $\sigma$  indicated that the robust-spherical most closely approximated the true distribution. The median was the only summary statistic indicating that the robust-spherical had the best error distribution. QQ plots for the AB data indicated poorest performance for the classical-spherical, but there was nothing to suggest that the robust-spherical was superior to either exponential fits. A tally of the number of cells  $< 1$  and  $\geq 60$  in the kriged maps would have suggested that the robust-spherical predictions would not have been significantly different from the true values. In short, there was no clear, consistent pattern indicating that the robust-spherical would have performed best for the AB data.

The null hypothesis was not rejected for any combination of semivariogram-model fit to the RE data. The highest  $p$ -value was given by the robust-exponential. Similar to the results for the AB data, this result would not have been anticipated by the least squares fit or the predicted mean, but it was predicted by the summary statistics maximum value and  $\sigma$ . In the case of the RE data, QQ plots indicated that the robust-exponential would perform best, as did the number of cells  $< 0.3$  and  $\geq 1.79$  in the kriged maps. As with the AB data, there was no consistent pattern, but the QQ plots were an additional tool that predicted that the robust-exponential would have performed best for the RE data.

## Conclusions

At first our results might seem to suggest that a spherical fit to a classical semivariogram on untransformed data would be a quick way to approximate the mean; or that an

exponential fit to a robust semivariogram would be the best way to test hypotheses on log-transformed, detrended data. Such an interpretation would be incorrect: just because a particular combination of semivariogram-model performs better with one data set does not mean the combination will always perform better with all fisheries data sets (Davis, 1987). Nevertheless, our analysis reveals some general guidelines for the typical application where the true values are not known.

The method of weighted least squares proposed by Cressie (1985) should be used with caution, particularly when working with robust semivariograms. This agrees with recent work that Cressie's least squares method should be considered a compromise solution, best used in conjunction with the classical semivariogram (Rufino *et al.*, 2006).

Recent work with simulated fisheries data has suggested that the classical semivariogram is superior to the robust semivariogram (Rufino *et al.*, 2006). Our results indicate that this is not the case. Indeed, the robust semivariogram has been applied successfully to a variety of other fisheries data sets (e.g., Sullivan, 1991; Maravelias *et al.*, 1996; Lo *et al.*, 2001; Jensen and Miller, 2005; Mello and Rose, 2005)

Recent work has also suggested the removal of outliers to help uncover the spatial structure of a stock (Rufino *et al.*, 2005). We did not employ this approach because the SMAST video survey is an absolute estimate and all scallop counts are known to be legitimate. Points should only be removed from a geostatistical analysis for valid physical

or ecological reasons, such as an erroneous measurement, an unusually large concentration of food, etc. (Rossi *et al.*, 1992; Rivoirard *et al.*, 2000).

Our results demonstrate that data transformation to approximate normality (if necessary), trend removal (if it exists) and correction for anisotropy (if it exists) will generate kriging predictions that are not significantly different from the (transformed) true values. This approach is recommended for describing the spatial structure of a stock, hypothesis testing, and answering other ecological questions.

The objective of this study was to generate kriged estimates of sea scallop abundance using data from a 5.6 km systematic grid, and then compare these predicted values with the actual scallop counts observed at a finer scale 2.2 km grid grid. There was no significant difference between kriged and true values for any combination of semivariogram–model fit to transformed, detrended data. In contrast, significant differences were found for three out four combinations of semivariogram–model fit to untransformed data. These empirical findings confirm the utility of the kriging technique to accurately predict scallop abundance at unsampled locations when the assumptions of geostatistics are met. We also found that the classical and robust semivariograms performed equally well. The ETCA contains the widest range of sea scallop densities (0 to 40 scallops  $\text{m}^{-2}$ ) offshore of the northeastern USA (Stokesbury *et al.*, 2004), so kriging of SMAST video data for other areas will be robust. The advantage of geostatistics over other statistical techniques is that it exploits the spatial correlation inherent in ecological data sets. These strengths will become increasingly important as the delineation of marine

protected areas, ecosystem based and other management strategies require spatially explicit interpretation of ecological data.

#### Acknowledgements

We thank B. Rothschild for his support and guidance. We thank the owners, Captains and crews who sailed with us. P. Christopher, D. Frie, R. Silva and L. Gavlin provided the Letters of Authorization. Aid was provided by SMAST, the Massachusetts Division of Marine Fisheries, NOAA award NA07NMF4540031, and the sea scallop fishery and supporting industries.

#### References

- 50 CFR Part 648. 2006. Fisheries of the Northeastern United States; Atlantic Sea Scallop Fishery; Framework 18. Federal Register 71:33211–33235.
- Brand, A.R. 1991. Scallop ecology: distributions and behavior. *In* *Scallops: Biology, Ecology and Aquaculture*, pp. 517–584. Ed. by S.E. Shumway. Elsevier, Amsterdam. 1095 pp.
- Chilés, J.-P., and Delfiner, P. 1999. *Geostatistics: Modeling Spatial Uncertainty*. Wiley–Interscience, New York. 695 pp.
- Cleveland, W.S, and Devlin, S.J. 1988. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:596–610.

- Conan, G.Y. 1985. Assessment of shellfish stocks by geostatistical techniques. ICES Document CM 1985/K:30. 24 pp.
- Cressie, N. 1985. Fitting variogram models by weighted least squares. *Mathematical geology* 17:563–586.
- Cressie, N., and Hawkins, D.M. 1980. Robust estimation of the variogram: I. *Mathematical Geology* 12:115–125.
- Davis, B.M. 1987. Uses and abuses of cross-validation in geostatistics. *Mathematical Geology* 19:241–248.
- Dutilleul, P., and Legendre, P. 1992. Lack of robustness in two tests of normality against autocorrelation in sample data. *Journal of Statistical Computation and Simulation* 42:79–91.
- Ecker, M.D., and Heltshe, J.F. 1994. Geostatistical estimates of scallop abundance. *In* Case Studies in Biometry, pp. 107–124. Ed. by N. Lange, L. Ryan, L. Billard, D. Brillinger, L. Conquest, and J. Greenhouse. John Wiley & Sons, New York. 529 pp.
- Gedamke, T., DuPaul, W.D., and 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.
- Giannoulaki, M., Machias, A., and Koutsikopoulos, C., Haralabous, J., Somarkis, S., and Tsimenides, N. 2003. Effect of coastal topography on the spatial structure of the populations of small pelagic fish. *Marine Ecology Progress Series*, 265:243–253.
- Isaaks, E.H., and Srivastava, R.M. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford. 561 pp.

- Jensen, O.P., and Miller, T.J. 2005. Geostatistical analysis of the abundance and winter distribution patterns of the blue crab *Callinectes sapidus* in Chesapeake Bay. Transactions of the American Fisheries Society 134:1582–1598.
- Kaluzny, S.P., Vega, S.C., Cardoso, T.P., and Shelly, A.A. 1998. S+SpatialStats: User's Manual for Windows<sup>®</sup> and UNIX<sup>®</sup>. Springer, New York. 327 pp.
- Krebs, C.J. 1999. Ecological Methodology, 2nd edn. Addison Wesley Longman, Reading, Massachusetts. 620 pp.
- Legendre, P., and Legendre, L. 1998. Numerical Ecology, 2nd edn. Elsevier, Amsterdam. 853 pp.
- Lo, N.C.H., Hunter, J.R., and Charter, R. 2001. Use of a continuous egg sampler for ichthyoplankton surveys: application to the estimation of daily egg production of Pacific sardine (*Sardinops sagax*) off California. Fishery Bulletin, 99:554–571.
- Maravelias, C.D., Reid, D.G., Simmonds, E.J., and Haralabous, J. 1996. Spatial analysis and mapping of acoustic survey data in the presence of high local variability: geostatistical application to North Sea herring (*Clupea harengus*). Canadian Journal of Fisheries and Aquatic Sciences 53:1497–1505.
- Matheron, G. 1963. Principles of geostatistics. Economic Geology 58:1246–1266.
- Matheron, G. 1971. The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphologie Mathématique 5. École Nationale Supérieure des Mines de Paris, Fontainebleau. 211 pp.
- Mello, L.G.S., and Rose, G.A. 2005. Using geostatistics to quantify seasonal distribution and aggregation patterns of fishes: an example of Atlantic cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences, 62:659–670.

- Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and acoustic application. *ICES Journal of Marine Science*, 50:285–298.
- Rivoirard, J., Simmonds, J., Foote, K.G., Fernandes, P., and Bez, N. 2000. *Geostatistics for Estimating Fish Abundance*. Blackwell Science, Oxford. 206 pp.
- Rossi, R.E., Mulla, D.J., Journel, A.G., and Franz, E.H. 1992. Geostatistical tools for modeling and interpreting spatial dependence. *Ecological Monographs*, 62:277–314.
- Rufino, M.M., Maynou, F., Abelló, P., Gil de Sola, L., and Yule, A.B. 2005. The effect of methodological options on geostatistical modeling of animal distribution: a case study with *Liocarcinus depurator* (Crustacea: Brachyura) trawl survey data. *Fisheries Research*, 76:252–265.
- Rufino, M.M., Stelzenmüller, V., Maynou, F., and Zauke, G.–P. 2006. Assessing the performance of linear geostatistical tools applied to artificial fisheries data. *Fisheries Research*, 82:263–279.
- Simmonds, E.J., and Fryer, R.J. 1996. Which are better, random or systematic acoustic surveys? A simulation using North Sea herring as an example. *ICES Journal of Marine Science* 53:39–50.
- Stephens, M.A. 1974. EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association* 69:730–737.
- Stokesbury, K.D.E. 2002. Estimation of sea scallop abundance in closed areas of Georges Bank, USA. *Transactions of the American Fisheries Society* 131:1081–1092.

- Stokesbury, K.D.E., Harris, B.P., Marino II, M.C., and Nogueira, J.I. 2004. Estimation of sea scallop abundance using a video survey in off-shore US waters. *Journal of Shellfish Research*, 23:33–40.
- Sullivan, P.J. 1991. Stock abundance estimation using depth-dependent trends and spatially correlated variation. *Canadian Journal of Fisheries and Aquatic Sciences* 48:1691–1703.
- Warren, W.G. 1998. Spatial analysis for marine populations: factors to be considered. *In* *Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management*, pp. 21–28. Ed. by G.S. Jamieson, and A. Campbell. *Canadian Special Publication of Fisheries and Aquatic Sciences*, 125. 462 pp.
- Webster, R., and Oliver, M.A. 2001. *Geostatistics for Environmental Scientists*. John Wiley & Sons, London. 271 pp.
- Wilk, M.B., and Gnanadesikan, R. 1968. Probability plotting methods for the analysis of data. *Biometrika* 55:1–17.
- Zar, J.H. 1999. *Biostatistical Analysis*, 4th edn. Prentice–Hall, Upper Saddle River, New Jersey. 663 pp.

Table 1. Kolmogorov–Smirnov tests for normality of the errors  $\hat{z} - z$ , and Wilcoxon signed rank tests for symmetry around the median error. *D*: Kolmogorov–Smirnov statistic, adjusted according to Stephens (1974).  $|Z|$ : Wilcoxon signed rank test statistic. AB: abundance data, assuming isotropic conditions. RE: residuals of the log–transformed abundance data, corrected for anisotropy. Clas: classical semivariogram. Rob: robust semivariogram. Sph: spherical model. Exp: exponential model.  $n = 852$ .

Data	Vario	Model	<i>D</i>	<i>p</i> –value	$ Z $	<i>p</i> –value
AB	Clas	Sph	0.18	<0.01	0.31	0.76
		Exp	0.17	<0.01	1.01	0.31
	Rob	Sph	0.16	<0.01	0.67	0.50
		Exp	0.17	<0.01	1.20	0.23
RE	Clas	Sph	0.04	<0.05	0.03	0.98
		Exp	0.04	<0.01	0.16	0.87
	Rob	Sph	0.03	>0.15	0.36	0.72
		Exp	0.03	<0.05	0.07	0.95

Table 2. Parameters of theoretical model fits to empirical semivariograms for scallop abundance at 5.6 km grid stations ( $n = 154$ ).  $C_0$ : nugget.  $C$ : partial sill.  $a$ : range (km). CLS: Cressie least squares fit. AB: abundance data, assuming isotropic conditions. RE: residuals of the log-transformed abundance data, corrected for anisotropy. Clas: classical semivariogram. Rob: robust semivariogram. Sph: spherical model. Exp: exponential model.

Data	Vario	Model	$C_0$	$C$	$a$	CLS
AB	Clas	Sph	71.92	188.71	17.92	16.07
		Exp	0.00	263.02	16.59	26.68
	Rob	Sph	0.00	65.07	18.73	49.32
RE	Clas	Exp	0.00	66.56	20.94	109.45
		Sph	0.09	0.05	22.78	0.84
	Rob	Exp	0.06	0.08	21.00	0.78
		Sph	0.04	0.09	20.42	1.84
		Exp	0.00	0.13	20.03	1.72

Table 3. Summary statistics for scallop abundance at 2.2 km grid stations and kriged estimates based on parameters of theoretical model fits presented in Table 1. Upper panel compares statistics for the true abundance with statistics resulting from kriging the AB data; lower panel compares statistics for the log-transformed true abundance with statistics resulting from kriging the RE data. AB: abundance data, assuming isotropic conditions. RE: residuals of the log-transformed abundance data, corrected for anisotropy. Clas: classical semivariogram. Rob: robust semivariogram. Sph: spherical model. Exp: exponential model.  $n = 852$ .

Data	Vario	Model	$\bar{x}$	$\sigma$	Min	$Q_1$	$M$	$Q_3$	Max
True	—	—	8.41	13.63	0	1	4	11	141
AB	Clas	Sph	8.72	10.37	-2.73	2.19	5.20	10.30	68.00
		Exp	8.82	11.80	-0.49	2.06	4.76	9.78	94.93
	Rob	Sph	8.82	12.60	-5.10	1.63	4.50	9.97	98.52
		Exp	8.86	12.10	-0.80	1.94	4.71	9.85	96.19
Log(true)	—	—	0.69	0.50	0	0.30	0.70	1.08	2.15
		Clas	Sph	0.68	0.34	-0.01	0.42	0.71	0.87
	Exp		0.69	0.36	-0.02	0.41	0.70	0.88	1.64
	Rob	Sph	0.68	0.39	-0.02	0.38	0.68	0.91	1.79
Exp		0.68	0.41	-0.03	0.37	0.68	0.92	1.89	

Table 4. Summary statistics for the error distributions from kriging estimates based on parameters of theoretical model fits presented in Table 1. MAE: mean absolute error. MSE: mean squared error. AB: abundance data, assuming isotropic conditions. RE: residuals of the log-transformed abundance data, corrected for anisotropy. Clas: classical semivariogram. Rob: robust semivariogram. Sph: spherical model. Exp: exponential model.  $n = 852$ . Note that values for the RE data have been extended to 3 decimal places due to small differences with log-transformed data.

Data	Vario	Model	$\bar{x}$	$\sigma$	Min	$M$	Max	MAE	MSE
AB	Clas	Sph	0.31	10.05	-85.85	0.75	44.93	5.80	101.02
		Exp	0.42	10.51	-80.23	0.43	64.92	5.82	110.48
	Rob	Sph	0.41	10.73	-80.29	0.24	68.86	6.08	115.14
		Exp	0.45	10.60	-80.18	0.32	66.37	5.87	112.36
RE	Clas	Sph	-0.001	0.343	-1.119	-0.007	1.076	0.277	0.118
		Exp	-0.001	0.339	-1.099	-0.009	1.081	0.271	0.115
	Rob	Sph	-0.001	0.343	-1.122	-0.004	1.122	0.272	0.118
		Exp	-0.001	0.348	-1.075	-0.008	1.167	0.274	0.121

Table 5. Wilcoxon paired sample tests by ranks of  $H_0$ : no difference between the predicted and true scallop counts.  $|Z|$ : Wilcoxon paired sample test statistic. AB: abundance data, assuming isotropic conditions. RE: residuals of the log-transformed abundance data, corrected for anisotropy. Clas: classical semivariogram. Rob: robust semivariogram. Sph: spherical model. Exp: exponential model.  $n = 852$ . Note that  $p$ -values have been extended to 3 decimal places to illustrate that the robust-spherical fit to the AB data was barely non-significant.

Data	Vario	Model	$ Z $	$p$ -value
AB	Clas	Sph	4.39	<0.001
		Exp	3.66	<0.001
	Rob	Sph	1.95	0.051
		Exp	3.18	<0.005
RE	Clas	Sph	0.60	0.551
		Exp	0.63	0.529
	Rob	Sph	0.68	0.498
		Exp	0.60	0.553

Figure 1. Eastern coastline of the USA with a close up of the Elephant Trunk Closed Area (solid line) showing 5.6 km grid stations (large open circles) and 2.2 km grid stations (small solid circles).

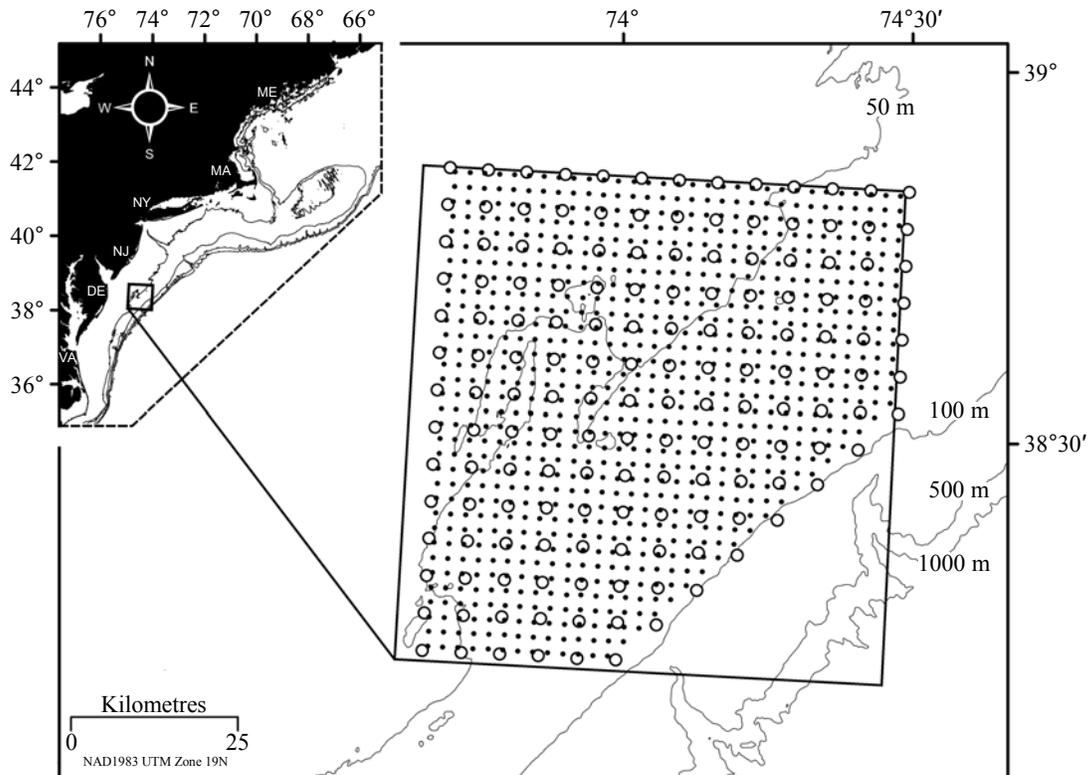


Figure 2. Post plot of sea scallop abundance (no. 12.94 m<sup>-2</sup>) for the 5.6 km grid survey in the Elephant Trunk Closed Area with column and row means on the top and right, respectively. Zero counts are denoted with a cross.

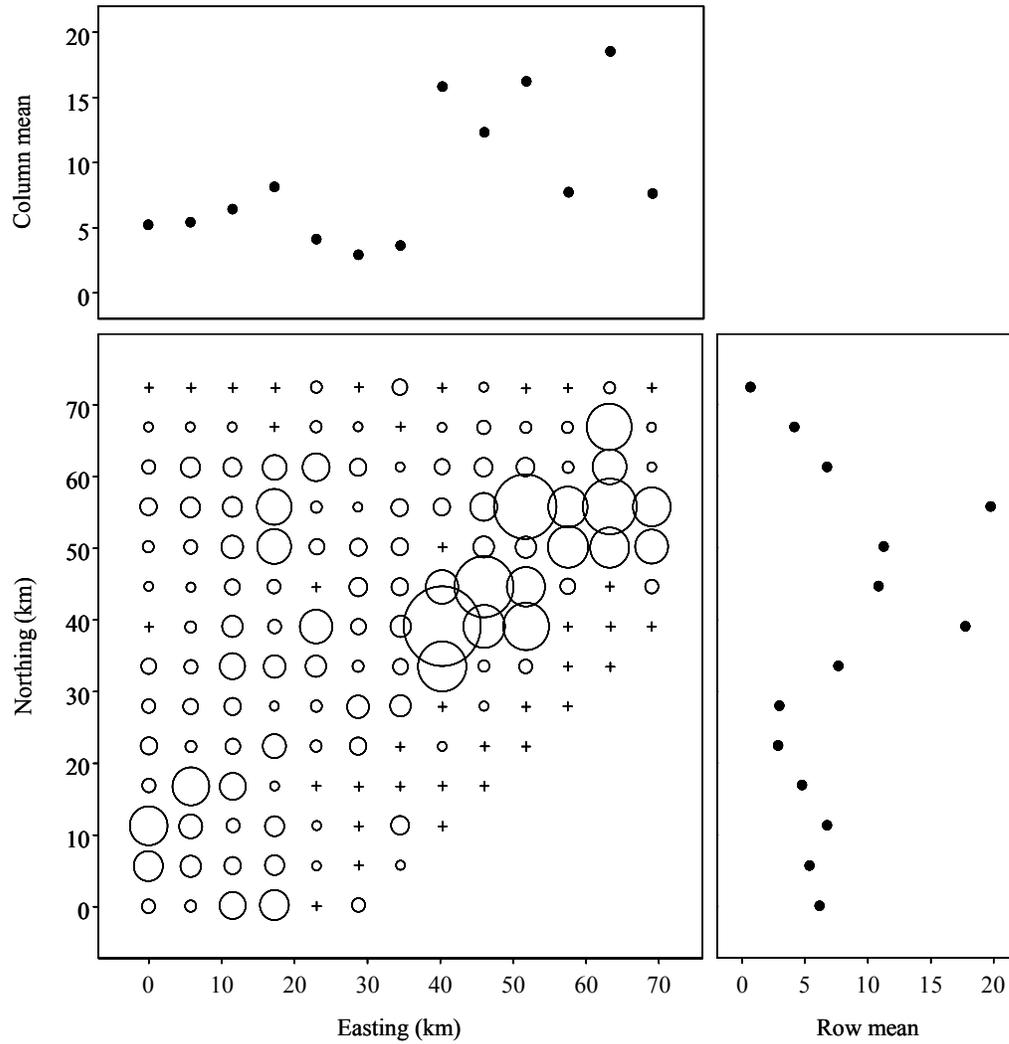


Figure 3. Flow chart showing the steps used to generate kriged estimates of sea scallop abundance. Note the boxes marked AB and RE, which summarize how the data were prepared for subsequent analyses.

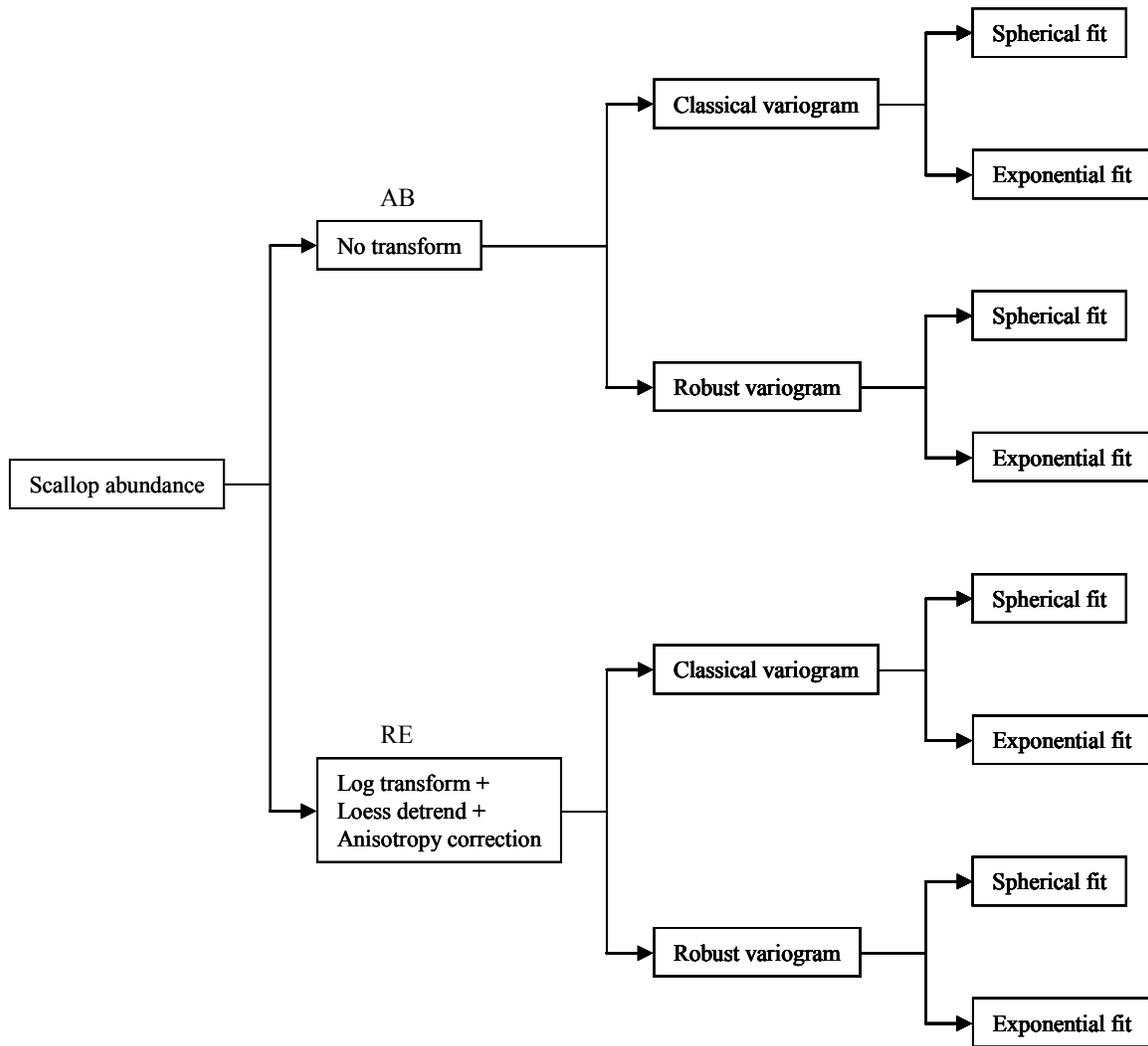


Figure 4. Comparison of contour (upper) vs. colorscale maps of predicted values (lower). Contour errors include the introduction of negative values (A), loss of a point  $\geq 60$  (B), and the incorrect representation of values between 1 to 4.9 (C).

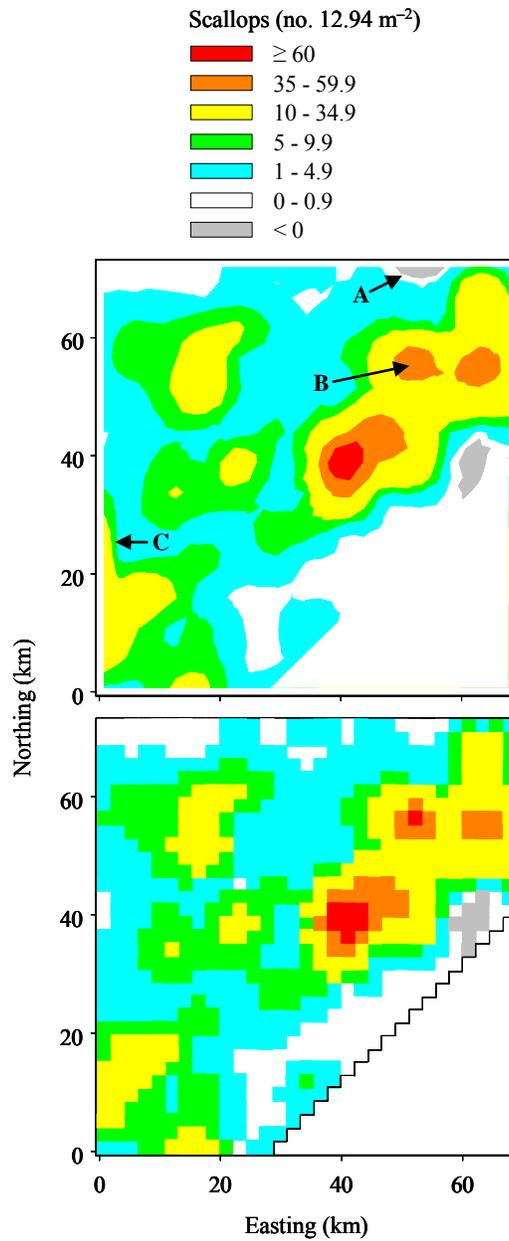


Figure 5. Standardized semivariograms (upper) and correlograms presented in semivariogram form (lower) for the 5.6 km grid (left) and 2.2 km grid (right).

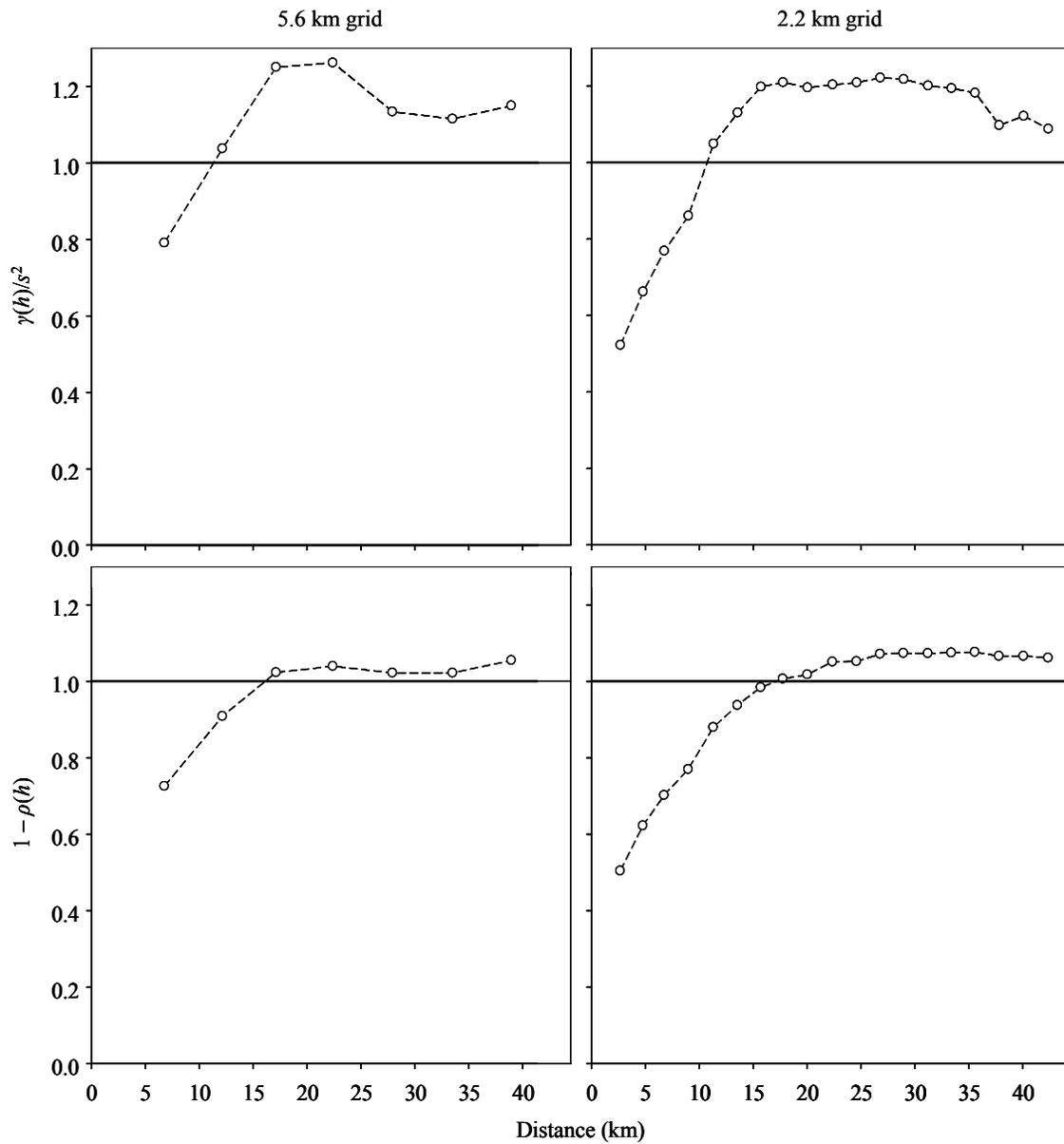


Figure 6. QQ plots comparing true scallop abundance with predicted scallop abundance based on kriging the AB data.

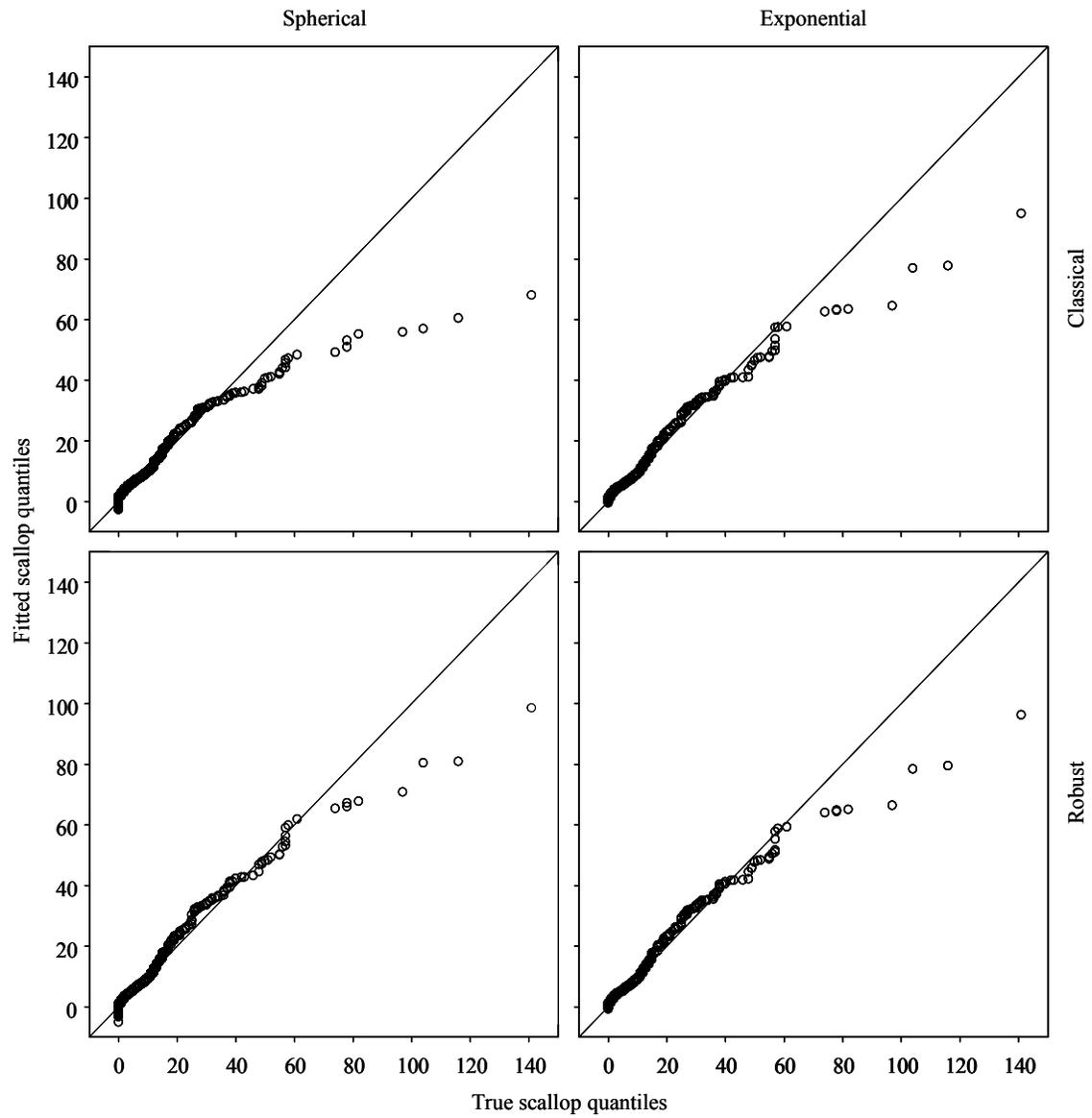


Figure 7. QQ plots comparing log(true) scallop abundance with predicted scallop abundance based on kriging the RE data.

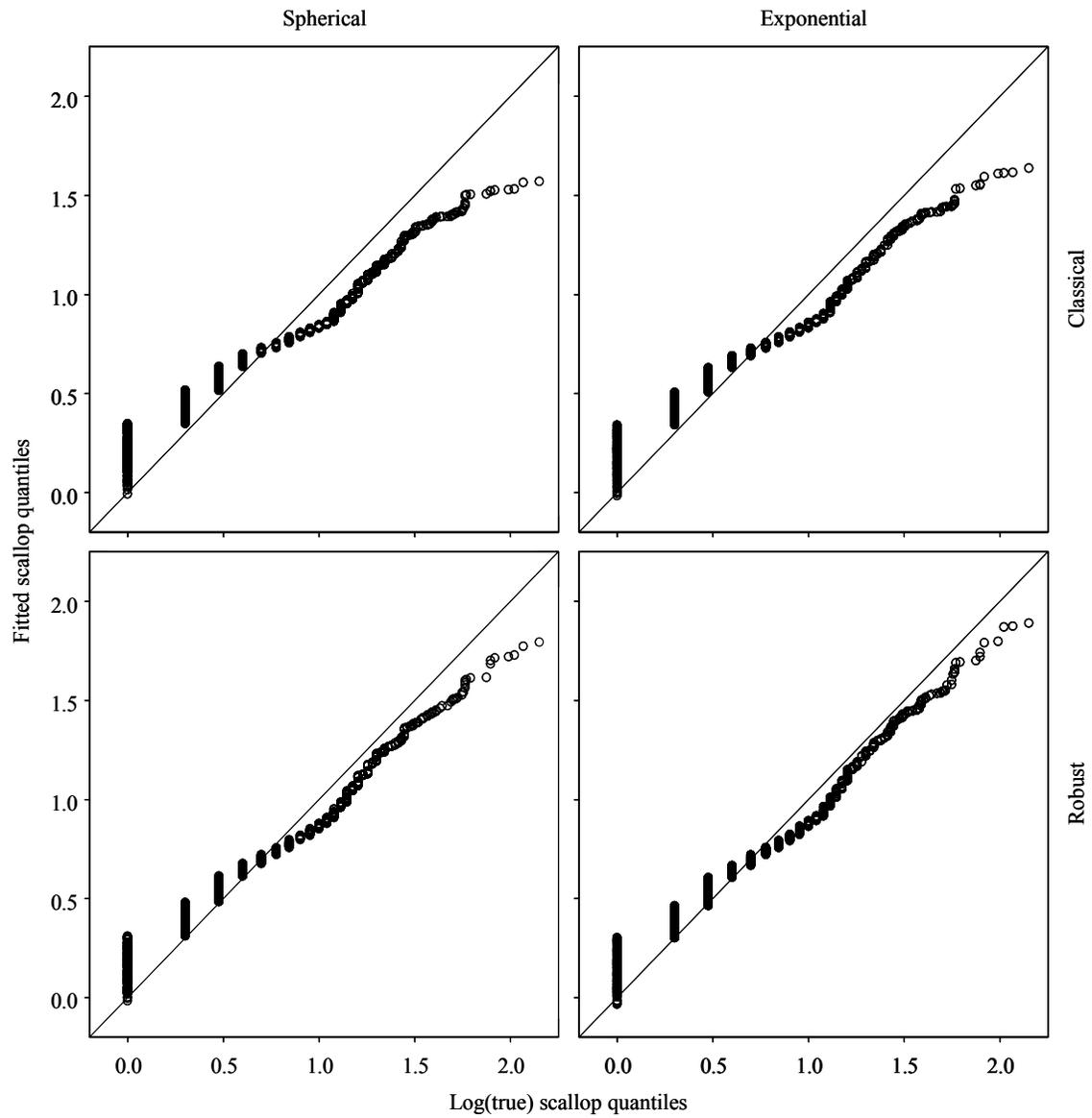


Figure 8. Colorscale maps of predicted scallop abundance in the Elephant Trunk Closed Area based on kriging the AB data. True scallop abundance is shown in the lower left panel.

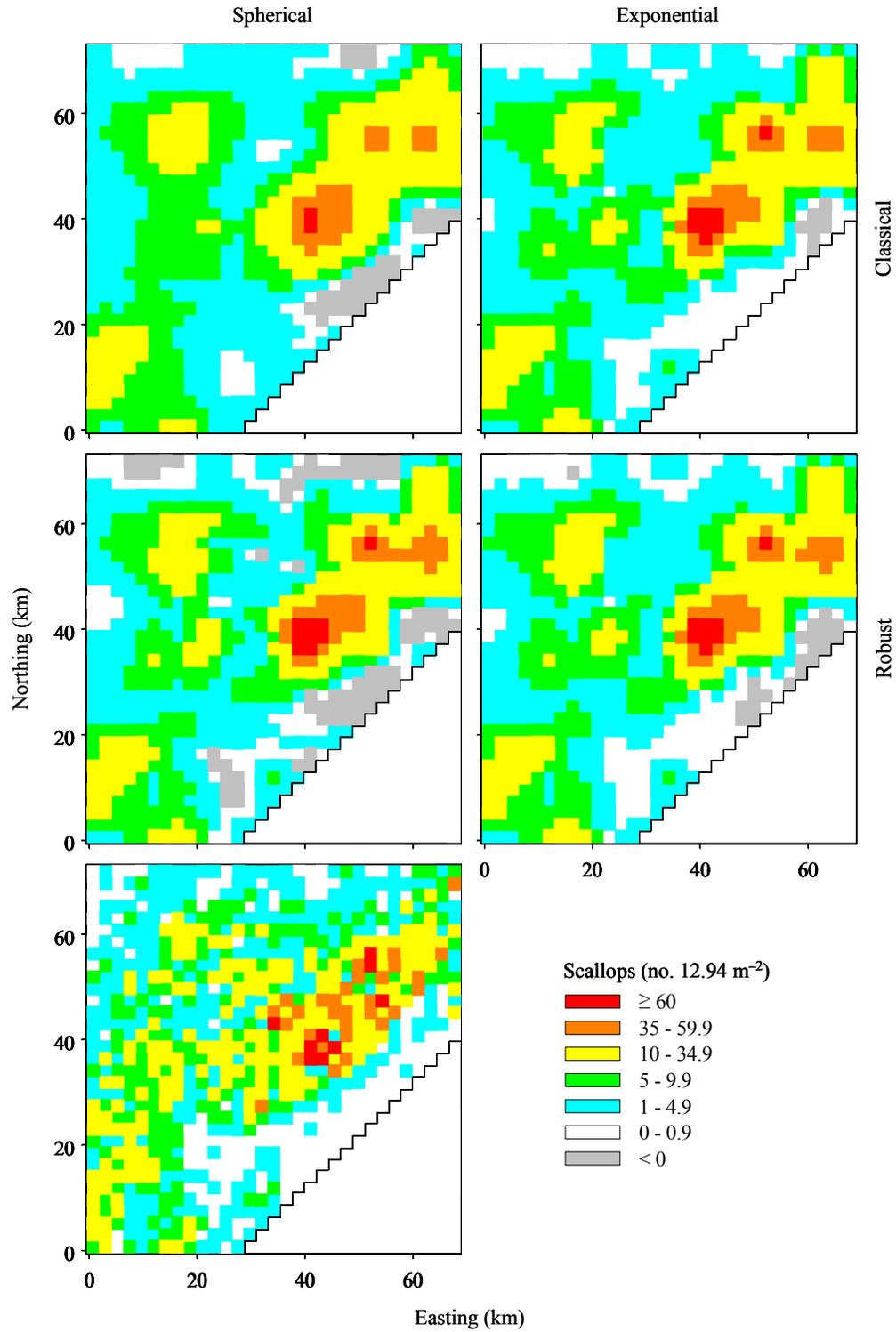


Figure 9. Colorscale maps of predicted scallop abundance in the Elephant Trunk Closed Area based on kriging the RE data. Log (true) scallop abundance is shown in the lower left panel.

