

## **B. Sea Scallop Assessment Report Appendixes**

Appendix B1 - Invertebrate Subcommittee meetings and participants

Appendix B2 - Sea Scallop Discard Estimates

Appendix B3 - Shell Height Meat Weight Relationships

Appendix B4 - Estimation of Dredge efficiency from paired dredge HabCam observations

Appendix B5 - Empirical Assessment

Appendix B6 - NEFSC HabCam survey for sea scallops: survey design, implementation, and data analysis

Appendix B7 - Assessment of the sea scallop resource in the Northern Gulf of Maine management area

Appendix B8 - Relationships between chlorophyll and scallop recruitment potentially useful for stock projections and assessment modeling

Appendix B9 - Technical documentation for the CASA length structured stock assessment model used in the SARC 59 sea scallop stock assessment

Appendix B10 – Forecasting methodology (SAMS Model)

## **Appendix B1. Invertebrate Subcommittee meetings and participants**

The Invertebrate Subcommittee met March 17-21, April 21-25, May 27-30, June 6, June 18 and June 23 during 2014 while preparing the SARC-59 stock assessment for Atlantic sea scallops. Meetings during March-May were held in the Stephen H. Clark Conference Room at the Northeast Fisheries Science Center in Woods Hole, MA with some participation by video conference. Meetings in June were exclusively by video conference. The following members participated in one or more meetings.

Larry Jacobson, NEFSC, chair  
Dvora Hart, NEFSC, Assessment Team Lead  
Burton Shank, NEFSC  
Jia-Han Chang, NEFSC  
Jiashen Tang, NEFSC  
Toni Chute, NEFSC  
Vic Nordahl, NEFSC  
Chris Legault, NEFSC  
Dan Hennen, NEFSC  
Mark Terciero, NEFSC  
Kevin Friedland, NEFSC  
Paul Rago, NEFSC  
Stephen Smith, DFO, Canada  
Mary Beth Tooley, NEFMC  
Dierdre Boelke, NEFMC  
David Rudders, VIMS  
Bill DuPaul, VIMS  
Carl Huntsberger, Coonamesset Farm Foundation  
Ron Smolowitz, Coonamesset Farm Foundation  
Katherine Thompson, Coonamesset Farm Foundation  
Daphne Munroe, Rutgers U.  
Kevin Stokesbury, SMAST  
Gregory DeCelles, SMAST  
Susan Inglis, SMAST  
Karen Bolles, HabCam Group  
Richard Taylor, HabCam Group  
Trish DeGraaf, Maine DMR  
Kevin Kelly, Maine DMR  
Matt Camisa, Massachusetts DMR  
Sam Truesdell, University of Maine

## Appendix B2. Sea Scallop Discard Estimates

Jessica Blaylock (NEFSC, Woods Hole, MA)

This paper presents discard estimates for Atlantic sea scallop (*Placopecten magellanicus*) for scallop dredge, scallop trawl and otter trawl fleets, calculated using the Standardized Bycatch Reporting Methodology (Wigley et al. 2007). This approach was also used in the previous assessment for this stock; however discard estimates were not included as input in the assessment model (NEFSC 2010).

### Methods

Estimates of Atlantic sea scallop discards (mt meats) were derived for seven fleets using Northeast Fishery Observer Program (NEFOP) and Northeast Fishery Science Center (NEFSC) commercial landings (i.e., dealer) data for the 1989 to 2013 time period: Georges Bank and Mid-Atlantic Bight scallop dredge, Mid-Atlantic Bight scallop trawl, Georges Bank and Mid-Atlantic Bight small-mesh otter trawl, and Georges Bank and Mid-Atlantic Bight large-mesh otter trawl. Additionally, sea scallop discard estimates were also derived for scallop dredge fleets at a finer stratification level using NEFOP and Vessel Trip Report (VTR) data for the 1994 to 2013 time period. This analysis considered the two scallop dredge fleets above as four fleets: Georges Bank open and closed scallop dredge, and Mid-Atlantic Bight open and closed scallop dredge,

A broad stratification scheme was used with trips partitioned into fleets using the following four classification variables: calendar quarter, gear type, area fished, and mesh. Trips were not partitioned by trip category ('limited' versus 'general', for scallop dredge and scallop trawl) due to small sample size over the time series. Calendar quarter was based on landed date and used to capture seasonal variations in fishing activity. Gear type was based on Northeast gear codes (scallop dredge: negear 132; scallop trawl: negear 052; otter trawl: negear 050). Trips for which gear was unknown were excluded. Two broad geographical regions are defined for area fished based on statistical area: areas 520-562 constituted the Georges Bank (GBK) area, and areas 600 and above constituted the Mid-Atlantic Bight (MAB) area. Two mesh size groups were formed for otter trawl: small (mesh less than 5.5 inches) and large (5.5 inch mesh and greater). The additional analysis considering scallop dredge at a finer scale included access area as another classification variable. Here, two access area categories were used: 'open' and 'closed', where 'closed' includes all trips fishing in one of the scallop access areas (Closed Area I, Closed Area II and Nantucket Lightship in the GBK region; Hudson Canyon, Virginia Beach, Elephant Trunk, and Delmarva in the MAB region). Observer trips were assigned to the access area category based on program code, and VTR trips were assigned based on latitude and longitude.

Discards were estimated using a combined  $d/k_{all}$  ratio estimator (Cochran 1963), where  $d$  is discarded pounds of sea scallops and  $k_{all}$  is kept pounds of all species, calculated from NEFOP data. Discard weight was derived by multiplying the  $d/k_{all}$  ratio of each fleet by the corresponding dealer or VTR landings (Wigley et al. 2007). Coefficients of variation (CV) were calculated as the ratio of the standard error of the discards divided by the discards.

In cases where limited observer data were available (i.e. two or less observed trips in a calendar quarter), an imputation approach was used to 'fill in' the missing (or

incomplete) information using data from adjoining strata. In this imputation procedure, the temporal stratification (i.e., calendar quarter) was relaxed to entire year, recognizing that seasonal variations may occur that will thus not be accounted for. Numbers of annual observed trips by fleet are summarized in Tables 1 and 2.

To evaluate the proportion of estimated sea scallop discards to landings, the sum of the current discard estimates for scallop dredge was compared to the sum of estimated landings from Georges Bank, Southern New England, and Mid-Atlantic Bight for the 1992 to 2013 time period.

## Results and Discussion

Annual Atlantic sea scallop discard estimates by fleet are presented in Tables 1, 2, and 3. Tables 1A-1D show estimates for the seven fleets without access area classification: Georges Bank and Mid-Atlantic Bight scallop dredge, Mid-Atlantic Bight scallop trawl, Georges Bank and Mid-Atlantic Bight small-mesh otter trawl, and Georges Bank and Mid-Atlantic Bight large-mesh otter trawl. Tables 2A-2B present discard estimates for the scallop dredge fleets at a finer scale that includes access area as a classification variable.

This analysis indicates that during the 1989 to 2013 time period, sea scallops were primarily discarded in the scallop dredge fleets (Tables 1A-1D, Table 3, Figure 1). For 2013, estimated discards from the Georges Bank and Mid-Atlantic Bight scallop dredge were 299 and 128 mt meats, respectively. Discard estimates for the other five fleets for the same year ranged from less than 1 mt meats (Georges Bank small-mesh otter trawl) to 10 mt meats (Mid-Atlantic Bight scallop trawl).

Discard estimates for scallop dredge at the access area classification level (Tables 2A-2B) suggest a higher discarding rate in the ‘open’ category fleets. For 2013, estimated discards from the Georges Bank open and closed scallop dredge fleets were 370 and 8 mt meats, respectively. Estimated discards from the Mid-Atlantic Bight open scallop dredge fleet were 46 mt meats; discards could not be estimated for 2013 for the Mid-Atlantic Bight closed scallop dredge fleet due to VTR trip misclassification.

The discard estimation presented here used a broad stratification approach. In addition, there are inherent limitations in the use of VTR data for trip assignment to the ‘access area’ category because of missing or inaccurate position data. Consequently, the discard estimates from scallop dredge at the access area classification level should be considered as preliminary.

Current estimates of discards and landings from scallop dredge fleets for 1994 to 2013 are presented in Figure 2. Total catch (discards plus landings) averaged 6,814 mt meats between 1993 and 1998. Catch increased in the following six years to peak at 31,435 mt meats in 2004, and averaged 26,560 mt meats from 2005 to 2012. Total catch in 2013 was 18,516 mt meats. Discards generally represent a small portion of total catch, with discard-to-landing ratios ranging from 0.010 in 1997 and 1998 to 0.1233 in 2000.

These results represent estimated sea scallop discards and landings in weight (mt meats). It is likely that discard-to-landing ratios of numbers would be higher because of the different size distribution of discarded scallops compared to that of landed scallops.

## Acknowledgements

I wish to thank all the NEFOP observers for their diligent efforts to collect the discard information used in this analysis. Additionally, I would like to thank Toni Chute for her assistance with the classification of VTR trips to access area categories.

## References

Cochran, W.L. 1963. Sampling Techniques. J. Wiley and Sons. New York.

Northeast Fisheries Science Center. 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report. US Dept Commer., Northeast Fish. Sci. Cent. Ref. Doc. 10-17; 844 p. Available online:  
<http://nefsc.noaa.gov/publications/crd/crd1017/>

Wigley, S.E., P.J. Rago, K.A. Sosebee, and D.L. Palka. 2007. The Analytic Component to the Standardized Bycatch Reporting Methodology Omnibus Amendment: Sampling Design, and Estimation of Precision and Accuracy (2nd Edition). US Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 07-09; 156 p. Available online:  
<http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0709/index.htm>

Table 1A. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) for the Georges Bank (GBK) scallop dredge and Mid-Atlantic Bight (MAB) scallop dredge fleets, 1989-2013. Discards were not estimated prior to 1992 due to small sample size.

GBK scallop dredge				MAB scallop dredge			
YEAR	Trips	Discards (mt meats)	CV	YEAR	Trips	Discards (mt meats)	CV
1989				1989			
1990				1990			
1991	1			1991	1		
1992*	11	464	0.48	1992*	7	121	0.00
1993*	12	345	0.32	1993*	10	12	0.80
1994*	7	3	0.89	1994	16	576	0.54
1995*	6	22	0.62	1995*	20	322	0.28
1996	15	116	0.36	1996	23	24	0.71
1997*	11	46	0.73	1997*	18	8	1.14
1998*	9	4	0.57	1998*	16	48	0.66
1999*	63	141	0.28	1999*	8	8	0.56
2000*	228	989	0.09	2000	28	779	0.33
2001*	18	529	0.17	2001*	88	1,955	0.11
2002*	11	105	0.58	2002	87	1,894	0.13
2003*	14	328	0.58	2003	108	2,225	0.10
2004*	46	58	0.20	2004	235	2,446	0.09
2005	107	228	0.27	2005	220	357	0.19
2006	135	347	0.20	2006*	93	78	0.49
2007	180	231	0.21	2007	177	260	0.20
2008	216	334	0.14	2008	425	414	0.15
2009	81	380	0.26	2009	408	923	0.12
2010	98	668	0.18	2010	238	688	0.21
2011	141	668	0.18	2011	251	482	0.14
2012	222	603	0.11	2012	201	237	0.12
2013	269	299	0.14	2013	182	128	0.22

\* Imputed data were used for discard estimation for these years.

Table 1B. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) for the Mid-Atlantic Bight (MAB) scallop trawl fleet, 1989-2013. Discards were not estimated prior to 2004 due to small sample size.

MAB scallop trawl			
YEAR	Discards		CV
	Trips	(mt meats)	
1989			
1990			
1991			
1992			
1993			
1994			
1995			
1996			
1997			
1998			
1999			
2000			
2001	4		
2002	1		
2003			
2004*	44	99	0.25
2005	137	61	0.13
2006*	30	150	0.33
2007	34	17	0.59
2008*	38	6	0.58
2009*	8	49	1.59
2010*	29	12	0.33
2011*	10	12	0.78
2012*	19	<1	0.75
2013*	20	10	0.35

\* Imputed data were used for discard estimation for these years.

Table 1C. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) for the Georges Bank (GBK) small-mesh otter trawl, and Mid-Atlantic Bight (MAB) small-mesh otter trawl fleets, 1989-2013.

GBK small-mesh otter trawl				MAB small-mesh otter trawl			
YEAR	Trips	Discards (mt meats)	CV	YEAR	Trips	Discards (mt meats)	CV
1989	65	2	0.53	1989	34	213	0.39
1990	31	<1	1.22	1990	47	8	0.44
1991	68	<1	0.80	1991	78	11	2.05
1992	42	<1	0.68	1992	47	6	0.53
1993	25	<1	0.57	1993*	16	8	0.81
1994*	18	7	1.88	1994*	15	29	0.78
1995*	11	<1	1.26	1995	63	71	0.23
1996*	10	0	0.00	1996	80	14	1.70
1997*	20	<1	0.87	1997*	48	1	2.76
1998*	6	<1	1.39	1998*	32	4	1.35
1999*	8	<1	2.62	1999	35	12	1.65
2000*	17	<1	0.49	2000	39	2	0.94
2001*	15	<1	0.64	2001	55	<1	8.75
2002*	33	<1	0.82	2002	32	68	0.34
2003	55	<1	1.11	2003	74	17	0.80
2004	109	2	0.96	2004	257	5	0.42
2005	194	<1	0.47	2005	172	4	0.32
2006	62	<1	0.56	2006	151	13	2.63
2007	60	<1	1.44	2007	218	5	0.56
2008	50	<1	0.49	2008	152	8	0.42
2009	199	<1	0.50	2009	286	23	0.52
2010	217	<1	0.54	2010	361	16	0.48
2011	168	<1	0.49	2011	365	5	0.33
2012	130	<1	0.83	2012	226	3	0.61
2013	186	<1	0.45	2013	395	5	0.35

\* Imputed data were used for discard estimation for these years.

Table 1D. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) for the Georges Bank (GBK) large-mesh otter trawl, and Mid-Atlantic Bight (MAB) large-mesh otter trawl fleets, 1989-2013. Discards were not estimated for MAB large-mesh otter trawl prior to 1992 due to small sample size.

GBK large-mesh otter trawl				MAB large-mesh otter trawl			
YEAR	Discards			YEAR	Discards		
	Trips	(mt meats)	CV		Trips	(mt meats)	CV
1989	27	1	0.88	1989	4		
1990	33	1	0.72	1990			
1991	34	4	0.54	1991	4		
1992	35	<1	1.10	1992*	14	4	0.40
1993	35	<1	1.30	1993*	12	3	1.54
1994	36	<1	1.21	1994*	21	99	0.53
1995	61	<1	0.36	1995	55	102	0.83
1996	38	<1	0.69	1996*	18	<1	0.62
1997	26	<1	1.00	1997*	9	1	0.62
1998*	10	<1	0.89	1998*	13	1	0.69
1999	20	<1	2.48	1999*	8	94	1.16
2000	30	2	0.66	2000*	26	32	0.57
2001	52	1	0.82	2001*	50	13	0.48
2002	83	2	0.61	2002*	39	8	2.36
2003	163	3	0.77	2003*	16	<1	2.26
2004	316	42	0.35	2004	109	9	0.43
2005	959	9	0.18	2005	93	1	0.94
2006	462	30	0.37	2006	71	3	2.39
2007	465	5	0.25	2007	160	12	0.59
2008	563	6	0.21	2008	132	29	0.88
2009	536	9	0.22	2009	167	19	0.22
2010	526	4	0.23	2010	274	9	0.73
2011	782	6	0.17	2011	253	9	1.00
2012	599	6	0.32	2012	169	4	0.78
2013	593	6	0.20	2013	251	7	0.53

\* Imputed data were used for discard estimation for these years.

Table 2A. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) by the Georges Bank (GBK) open scallop dredge and GBK closed scallop dredge fleets, 1994-2013. Discards were not estimated for the GBK open scallop dredge fleet in 2000 and 2001 due to small sample size.

GBK open scallop dredge				GBK closed scallop dredge			
YEAR	Discards		CV	YEAR	Discards		CV
	Trips	(mt meats)			Trips	(mt meats)	
1994*	7	2	0.82	1994	n/a		
1995*	6	23	0.63	1995	n/a		
1996	15	103	0.37	1996	n/a		
1997*	11	41	0.70	1997	n/a		
1998*	9	4	0.57	1998	n/a		
1999*	48	97	0.39	1999*	15	53	0.26
2000	2			2000	226	246	0.03
2001	2			2001	16	26	0.15
2002*	11	99	0.57	2002	n/a		
2003*	14	324	0.58	2003	n/a		
2004*	16	39	0.29	2004	30	25	0.19
2005	41	371	0.36	2005	66	40	0.27
2006*	56	783	0.25	2006	79	41	0.26
2007	53	194	0.30	2007	127	40	0.26
2008	73	202	0.23	2008	140	53	0.12
2009	58	295	0.33	2009*	23	24	0.30
2010	44	576	0.36	2010*	54	117	0.18
2011*	68	603	0.24	2011	71	84	0.20
2012	101	981	0.15	2012	119	48	0.11
2013	202	370	0.16	2013	30	8	0.07

\* Imputed data were used for discard estimation for these years.

n/a: not applicable

Table 2B. Number of observed trips, sea scallop discards (mt meats) and coefficient of variation (CV) by the Mid-Atlantic Bight (MAB) open scallop dredge and MAB closed scallop dredge fleets, 1994-2013. Discards were not estimated for the MAB open scallop dredge fleet in 2001 due to small sample size.

MAB open scallop dredge				MAB closed scallop dredge			
YEAR	Trips	Discards (mt meats)	CV	YEAR	Trips	Discards (mt meats)	CV
1994	16	276	0.59	1994	n/a		
1995*	20	341	0.28	1995	n/a		
1996	23	22	0.72	1996	n/a		
1997*	18	8	1.15	1997	n/a		
1998*	16	42	0.66	1998	n/a		
1999*	8	7	0.56	1999	n/a		
2000	28	749	0.33	2000	n/a		
2001	3			2001	85	301	0.09
2002*	13	1,446	0.19	2002	74	151	0.11
2003	62	2,253	0.14	2003	46	120	0.12
2004	143	1,869	0.13	2004	92	510	0.10
2005	166	368	0.29	2005	54	39	0.21
2006*	87	71	0.39	2006*	6	3	0.49
2007	84	65	0.41	2007	93	63	0.22
2008	89	215	0.54	2008	336	97	0.14
2009	118	597	0.15	2009	290	219	0.13
2010	130	583	0.30	2010	108	94	0.20
2011	145	489	0.20	2011	45	22	0.22
2012	100	143	0.20	2012^			
2013	137	46	0.25	2013^			

\* Imputed data were used for discard estimation for these years.

^ no discard estimation because of VTR missclassification

n/a: not applicable

Table 3. Summary of sea scallop discard estimates (mt meats) from Table 1 by region, 1989-2013.

Georges Bank (GBK)					Mid-Atlantic Bight (MAB)					
YEAR	scallop dredge	small-mesh otter trawl	large-mesh otter trawl	Total	YEAR	scallop dredge	scallop trawl	small-mesh otter trawl	large-mesh otter trawl	Total
1989	*	2	1	4	1989	*	*	213	*	213
1990	*	<1	1	1	1990	*	*	8	*	8
1991	*	<1	4	5	1991	*	*	11	*	11
1992	464	<1	<1	465	1992	121	*	6	4	131
1993	345	<1	<1	346	1993	12	*	8	3	22
1994	3	7	<1	10	1994	576	*	29	99	703
1995	22	<1	<1	23	1995	322	*	71	102	495
1996	116	0	<1	116	1996	24	*	14	<1	38
1997	46	<1	<1	46	1997	8	*	1	1	11
1998	4	<1	<1	4	1998	48	*	4	1	53
1999	141	<1	<1	142	1999	8	*	12	94	114
2000	989	<1	2	991	2000	779	*	2	32	813
2001	529	<1	1	531	2001	1,955	*	<1	13	1,969
2002	105	<1	2	107	2002	1,894	*	68	8	1,970
2003	328	<1	3	332	2003	2,225	*	17	<1	2,244
2004	58	2	42	102	2004	2,446	99	5	9	2,559
2005	228	<1	9	238	2005	357	61	4	1	424
2006	347	<1	30	378	2006	78	150	13	3	244
2007	231	<1	5	236	2007	260	17	5	12	294
2008	334	<1	6	341	2008	414	6	8	29	457
2009	380	<1	9	389	2009	923	49	23	19	1,013
2010	668	<1	4	672	2010	688	12	16	9	724
2011	668	<1	6	675	2011	482	12	5	9	508
2012	603	<1	6	610	2012	237	<1	3	4	245
2013	299	<1	6	306	2013	128	10	5	7	150

\* No discard estimate due to small sample size.

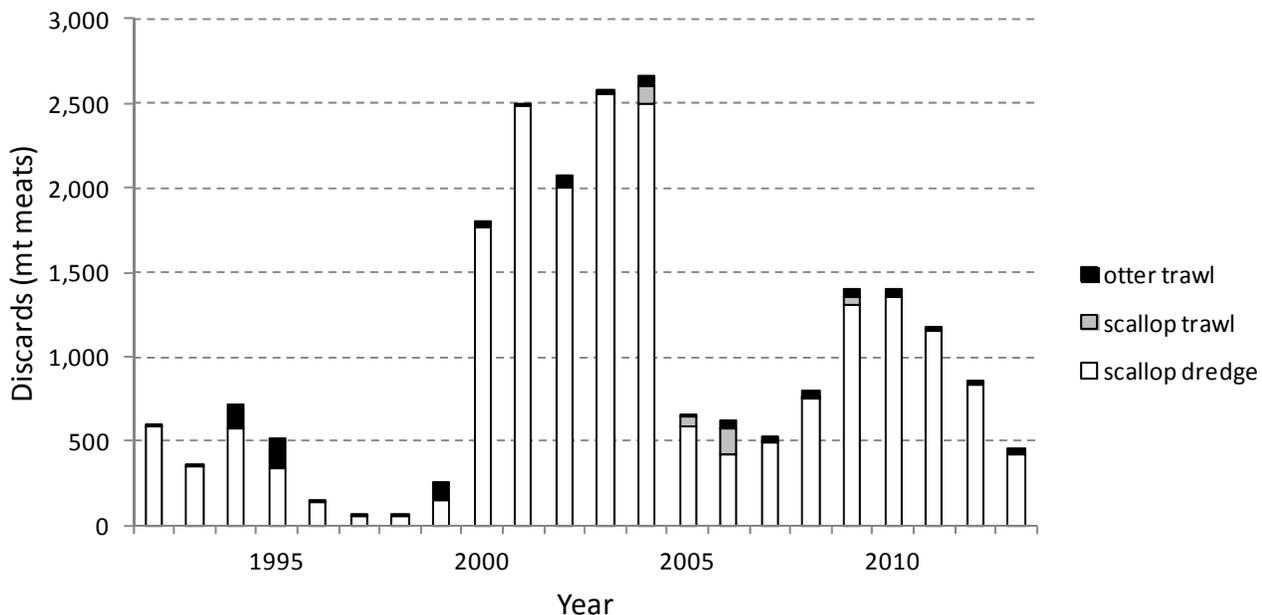


Figure 1. Sea scallop discard estimates (mt meats) from trips using scallop dredge, scallop trawl, and otter trawl gear presented in Table 1, 1992-2013. Discards from scallop trawl were not estimated prior to 2004 due to small sample size.

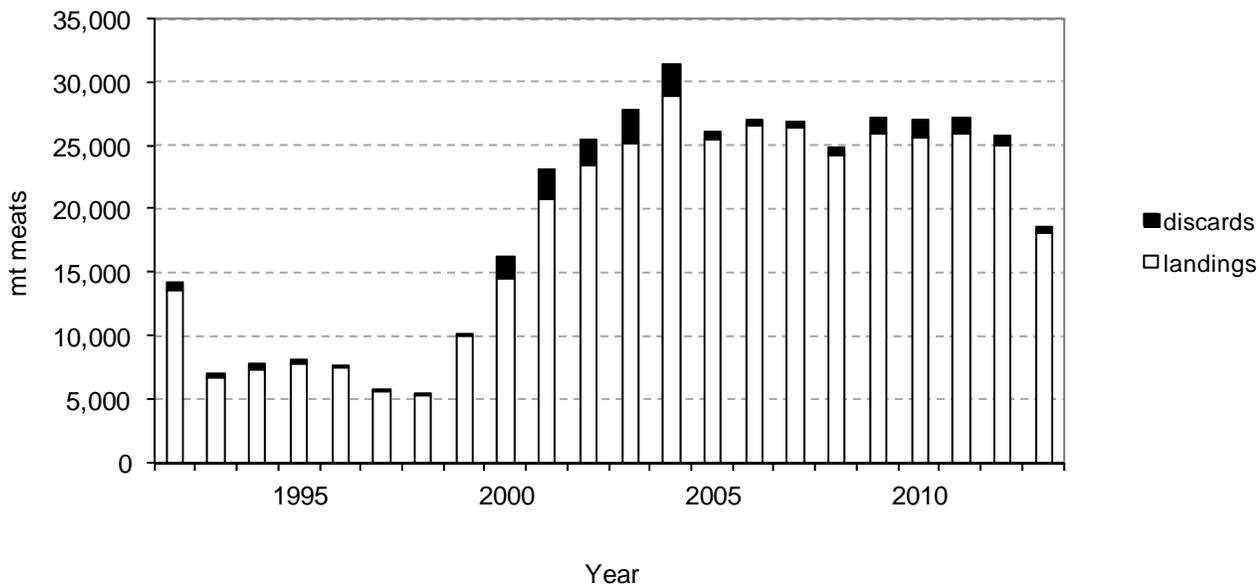


Figure 2. Estimated scallop landings and current estimated sea scallop discards from scallop dredge fleets (mt meats), 1992-2013.

## Appendix B3. Shell Height Meat Weight Relationships

Dan Hennen, NEFSC, Woods Hole, MA

### 1 Methods

Sea scallops (averaging about 6 per station) were selected for analysis on roughly half of all NEFSC survey stations from 2004 to 2013. The scallops were measured to the nearest millimeter, carefully shucked, excess water was removed from the meat, and the meat was weighed to the nearest gram.

Preliminary analysis indicated a residual pattern for those scallops with shell heights less than 70 mm. The small weights of these scallops (1-3 g) combined with the fact that meat weight could only be measured to the nearest gram resulted substantial measurement error. For this reason, the analysis was restricted to scallops that are at least 70 mm shell height (Figure A1).

A generalized linear mixed model (GLMM) with a log link was used to predict meat weight using shell height, depth, density, latitude, and subarea (a finer scale regional division within each broad region). The GLMM used the gamma likelihood with a log link which is appropriate for data (such as these) with "constant CV" error (McCullagh and Nelder [1989]). This method avoids log-transforming the response variable (meat weight) which can lead to biased estimates when the results are back-transformed. The best model was chosen by AIC (Tables 1-5; Burnham and Anderson [2002]). The grouping variable for the random effects was a combination of survey station number and the year in which the survey took place. Survey stations are chosen randomly (though stratified to fit NEFSC survey design specifications) and survey stations numbers are assigned sequentially so that a survey station number in one year does not have any particular relationship to the same station number in the next year. Thus, a grouping variable based on a combination of survey station number and year incorporates random variation in the data that is due to both time (year) and fine scale spatial differences (station number).

Several analyses using simplified versions of the best model were employed to explore the effects of year, subarea, and fishing regulations.

All data analysis was conducted using the R statistical program (v2.13.2).

#### 1.1 Seasonal variation and commercial meat anomalies

The NMFS Observer program provided meat weight estimates from commercial catches that occurred throughout the year. These meat weights are based on meats that are shucked by fishermen. Meats from the observer program are not weighed individually. They are packed into a graduated cylinder and a volume for a sample (typically 100 scallops) is recorded. The meat weight for a sample was calculated using a density estimate of  $1.05 \frac{\text{g}}{\text{ml}^3}$  (Caddy and Radley-Walters [1972]; Smolowitz et al. [1989]). These "observed" meat weights are therefore an average weight for all the meats in the cylinder, not a direct observation of the weight of a meat. The observer program does generate approximate shell heights for individual scallops, though they are binned by 5 mm increments. Therefore predicted meat weights can be generated for each shell height represented in the sample. Predicted meat weights were calculated using the best model (by AIC) from the analysis of survey meat weights described above.

It was noted this year that in many cases the number of shells measured was  $> 100$ . Because there were only 100 scallop meats packed into the cylinder and there is no way to determine which of the shells were associated with the meats in the cylinder, all observations in which the shell heights exceeded 100 in number were excluded from this analysis. This

correction reduced the sample size by approximately 52%, but reduced the error in predicted meat weights considerably (compare Figure 9 to Figure 10).

The best model was applied to predict meat weights for observer samples based on shell heights, latitude and longitude recorded for each sample during 2001-2013. Depth outliers were excluded by restricting maximum depths in the observed hauls from each subarea to the maximum depths observed in the survey for that subarea.

Predicted meat weights for each month were compared to the (observed) density derived meat weights for each month by  $\frac{\text{pred.}-\text{obs.}}{\text{pred.}}$  (Figure 10). The median of these ratios by month are referred to as the monthly meat weight “anomaly”. A positive anomaly indicates that the observed meat weight was greater than the expected meat weight, while a negative anomaly indicates the opposite is true. Annual meat weight anomalies for use in the CASA stock assessment model were computed by averaging the monthly values within a year using the landings during each month as weights.

## 2 Results and Discussion

In general, the observed meat weights (from observed volumes) should be less than the survey-based, predicted meat weights (a negative anomaly) because the commercially shucked scallops leave some meat on the shell, and because the surveys occur in late spring or summer (depending on the year), a time of typically high meat weight. The pattern in the anomaly calculated for MAB roughly follows this pattern in that the anomaly is negative in all months excluding April through July, a period that overlaps the survey (Figure 12). On Georges Bank, however, there were months of the year where the observed scallop meats were almost 15% heavier than the predicted meats, resulting in a positive anomaly (Figure 13). The positive anomaly appears in February through July. It is clear from examination of Figure 13 that either observed meat weights were heavier than expected and/or predicted meat weights lighter between January and May since 2009. In 2009, the timing of the survey was shifted to earlier in the year. Predicted meat weights have increased for scallops greater than about 130 mm since the last assessment (Figure 8). Therefore observed meat weights must have increased. In fact, observed meat weights have both increased and stabilized dramatically in the years since 2009 (Figure 14). It is possible that this reflects an increase in efficiency among fishers by selecting areas and time periods when meat weights were high. The early months of the year were not as well sampled by observers relative to the summer months and smaller sample sizes may be influencing this pattern as well (Table 6). There is also some indication of a systemic increase in meat weight for the region generally, based on the shell height to meat weight model estimates reflected in Figure 8, but this result is confounded with the shift in the timing of the survey.

The anomalies refine assessment model estimates of the total annual weight of meats removed by the fishing fleet, based on the lengths recorded by port-side samplers. To make the conversion from port-side shell height to meat weight, the median monthly meat weight anomalies were smoothed by a second order polynomial loess function with a span of 0.25 (months). This short smoothing span provided a modest smooth that allowed the data to strongly influence the model fit (Figures A15). The smooth was applied to a duplicated annual cycle (i.e. 24 months were fit, using identical data in each 12 month period) and the middle 12 months were selected and reordered so that January was the first month in the resulting model fit. This manipulation guaranteed that December and January produced linking estimates and minimized edge effects. The smoothed monthly anomalies were then weighted by the landings in each month in each year for which we have landings data (1975 – 2012) and annual median values

were calculated.

The annual values were somewhat different from similar values calculated for the last assessment (Figures A16 -A17). The anomalies are generally lower (~ 2%) in the MAB and higher (~ 15%) in the GBK. The difference in the GBK region is due to the large shift in the monthly anomalies between the last assessment and the current one, based primarily on the increase in observed meat weight (Figure 14). The shift in the MAB is relatively minor and is likely attributable to a combination of the various manipulations to the observer data and small changes in the shell height to meat weight model.

### **3 Literature Cited**

K.P. Burnham and D.R. Anderson. Model selection and multimodel inference: a practical information-theoretic approach. Springer, 2002.

J.F. Caddy and C. Radley-Walters. Estimating count per pound of scallop meats by volumetric measurement. Technical Report 1202, Fish. Res. Brd. Can. Man. Rep., 1972.

P. McCullagh and J.A. Nelder. Generalized Linear Models, 2nd Ed. Chapman and Hall, Boca Raton, FL, 1989. 511 pp.

R.J. Smolowitz, F.M. Serchuk, and R.J. Reidman. The use of a volumetric measure for determining sea scallop meat count. NOAA Tech. Mem. F/NER-1, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543-1026, 1989.

Table 1: AIC results from model fits to predict meat weight.

Formula	AIC	BIC	logLik	deviance
sh+d+sh*d+area+(sh+1)	101114.57	101267.34	-50537.28	101074.57
sh+d+lat+clop+area+(sh+1)	101123.48	101283.90	-50540.74	101081.48
sh+d+area+(sh+1)	101129.14	101274.28	-50545.57	101091.14
sh+d+lat+area+(sh+1)	101130.13	101282.90	-50545.06	101090.13
sh+d+clop+sh*d+(sh+1)	101166.05	101234.80	-50574.02	101148.05
sh+d+lat+clop+(sh+1)	101175.50	101244.25	-50578.75	101157.50
sh+d+clop+(sh+1)	101180.69	101241.80	-50582.35	101164.69
sh+d+sh*d+(sh+1)	101187.51	101248.63	-50585.76	101171.51
sh+d+lat+sh*d+(sh+1)	101188.53	101257.28	-50585.26	101170.53
sh+d+(sh+1)	101202.36	101255.83	-50594.18	101188.36
sh+area+(sh+1)	101288.53	101426.03	-50626.26	101252.53
sh+clop+(sh+1)	101359.04	101412.51	-50672.52	101345.04
sh+lat+(sh+1)	101363.62	101417.09	-50674.81	101349.62
d+(sh+1)	103485.29	103531.13	-51736.65	103473.29
sh+d+sh*d+(1)	105482.86	105528.69	-52735.43	105470.86
sh+d+area+(1)	105660.31	105790.17	-52813.16	105626.31
sh+d+clop+(1)	105750.75	105796.58	-52869.37	105738.75
sh+d+lat+(1)	105769.06	105814.89	-52878.53	105757.06
sh+d+(1)	105773.59	105811.78	-52881.79	105763.59
sh+area+(1)	105824.38	105946.60	-52896.19	105792.38
sh+clop+(1)	105915.93	105954.12	-52952.96	105905.93
sh+(1)	105923.56	105954.12	-52957.78	105915.56
sh+lat+(1)	105925.11	105963.31	-52957.56	105915.11
d+(1)	119777.65	119808.20	-59884.82	119769.65

Table 2: Results from model fits to predict meat weight. The coefficients estimated are: the intercept(int), ln(shell height) (sh), ln(depth) (d), latitude (lat), an interaction between ln(shell height) and ln(depth)(shXd) and an Identifier which is either a marker for a model with subarea coefficients (see Tables 3 and4) or a coefficient for closed vs. open (clop). Random effects are either on the shell height coefficient and intercept (sh+1) or intercept alone (1). The models are listed in order of increasing AIC (lowest AIC model is in the top row).

formula	int	sh	d	lat	shXd	Identifier
sh+d+sh*d+area+(sh+1)	-16.98(0.013)	4.6(0.021)	1.93(0.018)		-0.48(0.087)	1
sh+d+lat+clop+area+(sh+1)	-6.43(0.016)	2.61(0.022)	-0.38(0.019)	-0.02(0.012)		0.09(0.019)
sh+d+area+(sh+1)	-7.45(0.013)	2.61(0.021)	-0.38(0.018)			2
sh+d+lat+area+(sh+1)	-6.55(0.016)	2.61(0.022)	-0.39(0.019)	-0.02(0.012)		3
sh+d+clop+sh*d+(sh+1)	-17.08(0.006)	4.59(0.021)	1.94(0.016)		-0.48(0.087)	-0.06(0.008)
sh+d+lat+clop+(sh+1)	-8.02(0.006)	2.61(0.021)	-0.38(0.016)	0.01(0.003)		-0.07(0.008)
sh+d+clop+(sh+1)	-7.56(0.006)	2.61(0.021)	-0.36(0.016)			-0.06(0.008)
sh+d+sh*d+(sh+1)	-17.38(0.004)	4.64(0.021)	2.01(0.016)		-0.49(0.087)	
sh+d+lat+sh*d+(sh+1)	-17.56(0.004)	4.64(0.021)	2.01(0.016)	0.005(0.003)	-0.49(0.087)	
sh+d+(sh+1)	-9.09(0.004)	2.61(0.021)	-0.34(0.016)			
sh+area+(sh+1)	-9.07(0.013)	2.61(0.022)				4
sh+clop+(sh+1)	-9.04(0.006)	2.61(0.022)				-0.04(0.008)
sh+lat+(sh+1)	-8.63(0.004)	2.61(0.022)		-0.01(0.003)		
d+(sh+1)	4.96(0.005)		-0.36(0.019)			
sh+d+sh*d+(1)	-28.64(0.004)	6.98(0.015)	4.94(0.017)		-1.1(0.064)	
sh+d+area+(1)	-6.38(0.014)	2.38(0.016)	-0.38(0.019)			5
sh+d+clop+(1)	-6.64(0.006)	2.4(0.016)	-0.34(0.017)			-0.06(0.008)
sh+d+lat+(1)	-7.18(0.004)	2.4(0.016)	-0.34(0.017)	0.01(0.003)		
sh+d+(1)	-6.76(0.004)	2.4(0.016)	-0.32(0.017)			
sh+area+(1)	-7.99(0.014)	2.38(0.016)				6
sh+clop+(1)	-8.02(0.006)	2.39(0.016)				-0.04(0.009)
sh+(1)	-8.05(0.004)	2.39(0.016)				
sh+lat+(1)	-7.91(0.004)	2.39(0.016)		-0.003(0.003)		
d+(1)	4.69(0.007)		-0.31(0.028)			

Table 3: Results from model fits to predict meat weight in MAB subareas.

Identifier	VB	DMV	DMV.VB	ET	HC	NYB
1	-0.13(0.023)	-0.06(0.018)	-0.14(0.028)	-0.17(0.022)	-0.08(0.019)	-0.07(0.019)
2	-0.14(0.023)	-0.06(0.018)	-0.15(0.028)	-0.17(0.022)	-0.08(0.019)	-0.07(0.019)
3	-0.14(0.023)	-0.12(0.041)	-0.22(0.05)	-0.23(0.039)	-0.12(0.031)	-0.11(0.028)
4	-0.06(0.024)	0.04(0.018)	-0.03(0.028)	-0.07(0.022)	0.002(0.02)	0.04(0.019)
5	-0.14(0.024)	-0.05(0.019)	-0.2(0.029)	-0.24(0.023)	-0.11(0.02)	-0.07(0.02)
6	-0.07(0.025)	0.05(0.019)	-0.08(0.029)	-0.13(0.023)	-0.03(0.021)	0.04(0.02)

Table 4: Results from model fits to predict meat weight in GBK subareas.

Identifier	NLS	SCH	CA1	SEP	NEP	CA2
1	0.07(0.021)	-0.13(0.018)	0	-0.07(0.023)	-0.13(0.017)	0.004(0.017)
2	0.07(0.021)	-0.13(0.018)	0	-0.07(0.023)	-0.13(0.017)	0.005(0.017)
3	0.06(0.022)	-0.13(0.018)	0	-0.08(0.024)	-0.12(0.018)	0.008(0.017)
4	0.14(0.021)	-0.07(0.019)	0	-0.08(0.024)	-0.12(0.017)	0.05(0.018)
5	0.08(0.021)	-0.12(0.019)	0	-0.06(0.024)	-0.14(0.018)	0.001(0.018)
6	0.14(0.022)	-0.06(0.02)	0	-0.07(0.025)	-0.12(0.018)	0.04(0.018)

Table 5: Results from model fits to predict meat weight. Predictors are ln(shell height) (sh) ln(depth) (d), region (reg) and open vs. closed to fishing (clop). MAB and open coefficients are shown. GBK and closed are assumed to have coefficients equal to 0.

formula	int	sh	d	reg	clop	AIC	BIC
sh+d+reg+clop+(sh+1)	-7.35(0.012)	2.61(0.03)	-0.4(0.028)	-0.05(0.014)	-0.06(0.013)	101171	101240
sh+d+reg+(sh+1)	-7.46(0.009)	2.61(0.03)	-0.38(0.029)	-0.04(0.014)		101195	101256
sh+reg+clop+(sh+1)	-9.07(0.012)	2.61(0.03)		0.04(0.014)	-0.04(0.014)	101353	101414
sh+reg+(sh+1)	-9.09(3e-04)	2.61(4e-04)		0.04(0.01)		101361	101414

Table 6: Sample sizes for observed meat weights by month in GBK.

month	pre2010	post2009	Total
1	142	82	224
2	86	38	124
3	18	62	80
4	32	88	120
5	84	149	233
6	431	333	764
7	433	404	837
8	356	404	760
9	269	174	443
10	201	151	352
11	249	138	387
12	167	58	225
Total	2468	2081	4549

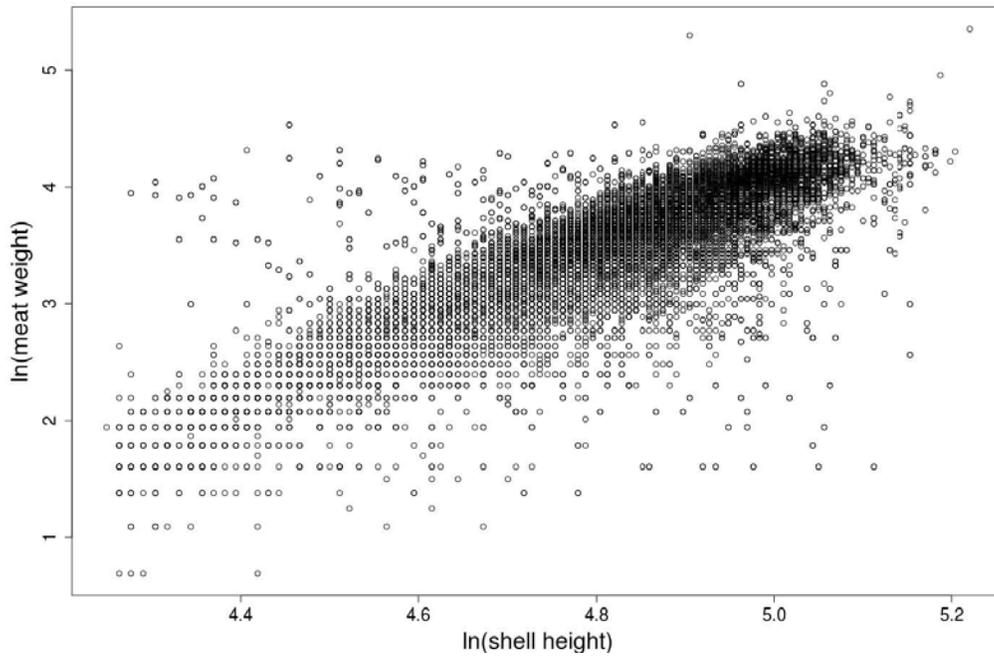


Figure 1: Natural log of shell height against the natural log of meat weights measured on NEFSC scallop surveys between 2003 and 2013.

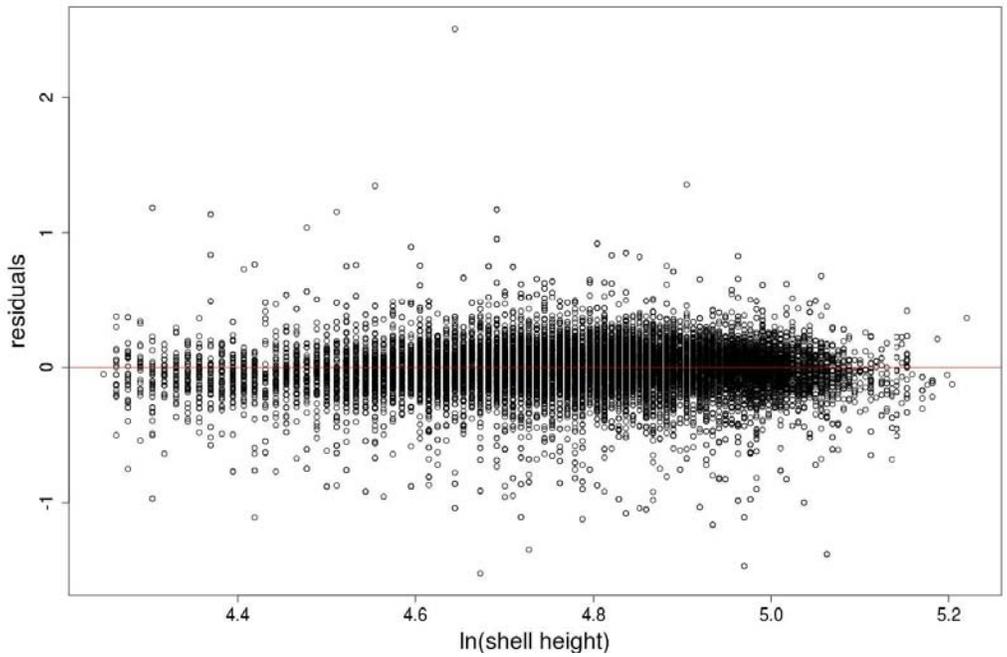


Figure 2: Residuals from the fit of best model predicting meat weight by the natural log of shell height.

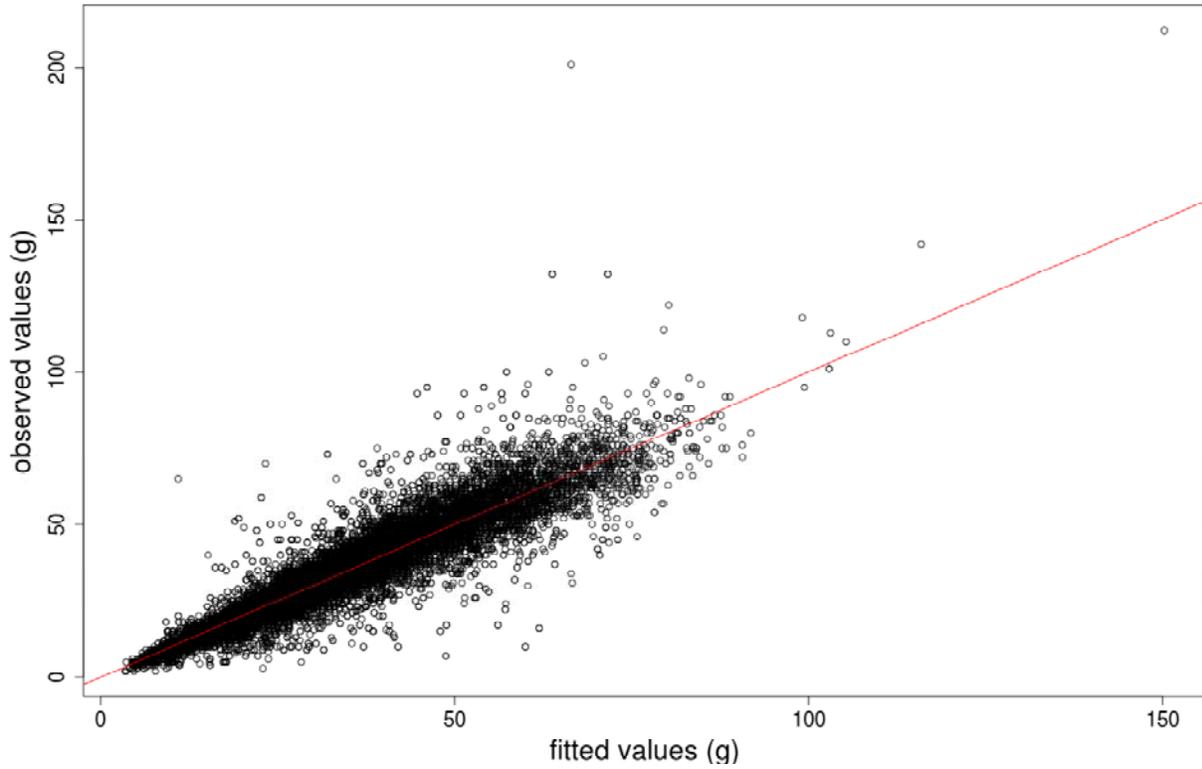


Figure 3: Observed vs. predicted meat weight using the best model by AIC.

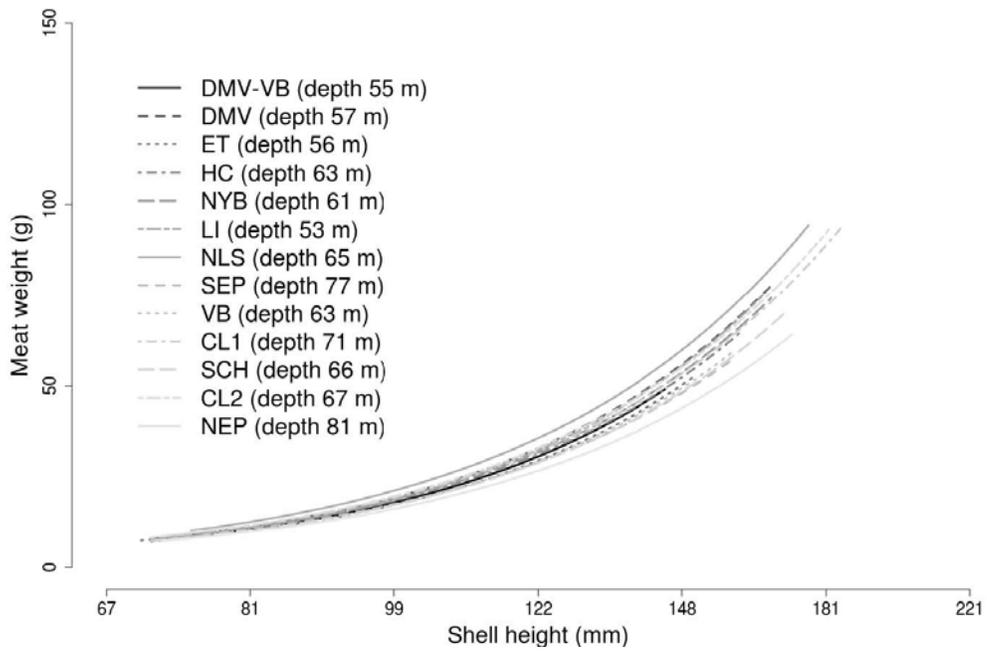


Figure 4: Meat weight curves by subarea. The depths used are the median depths observed in each subarea during all available years of the survey.

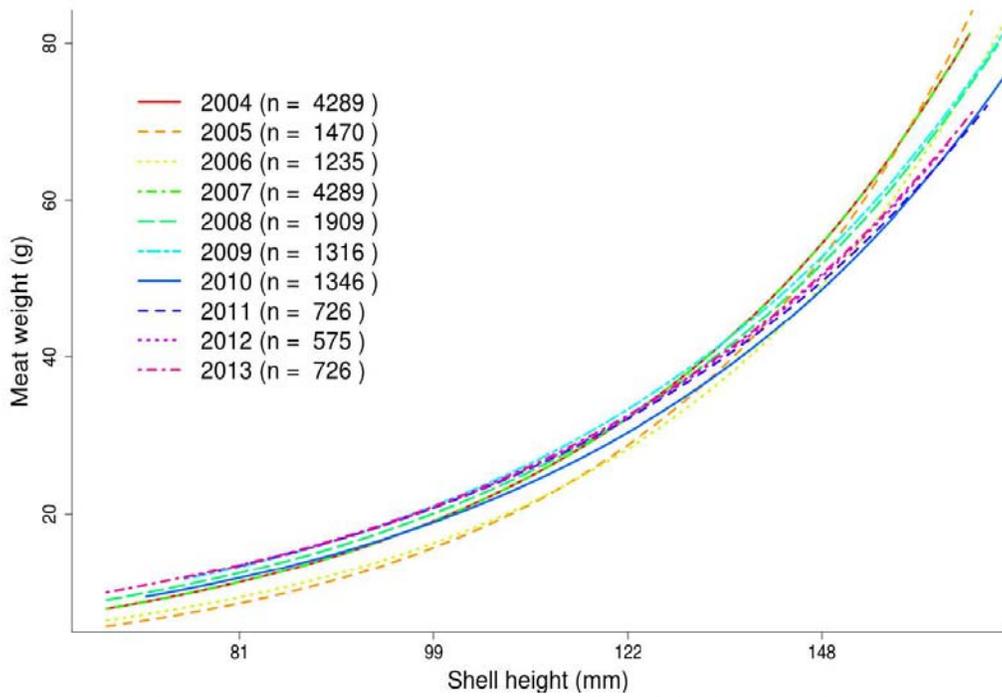


Figure 5: Meat weight curves by year. The curves are fits of the best model to annual subsets of the data. The sample size of each subset are shown in the legend.

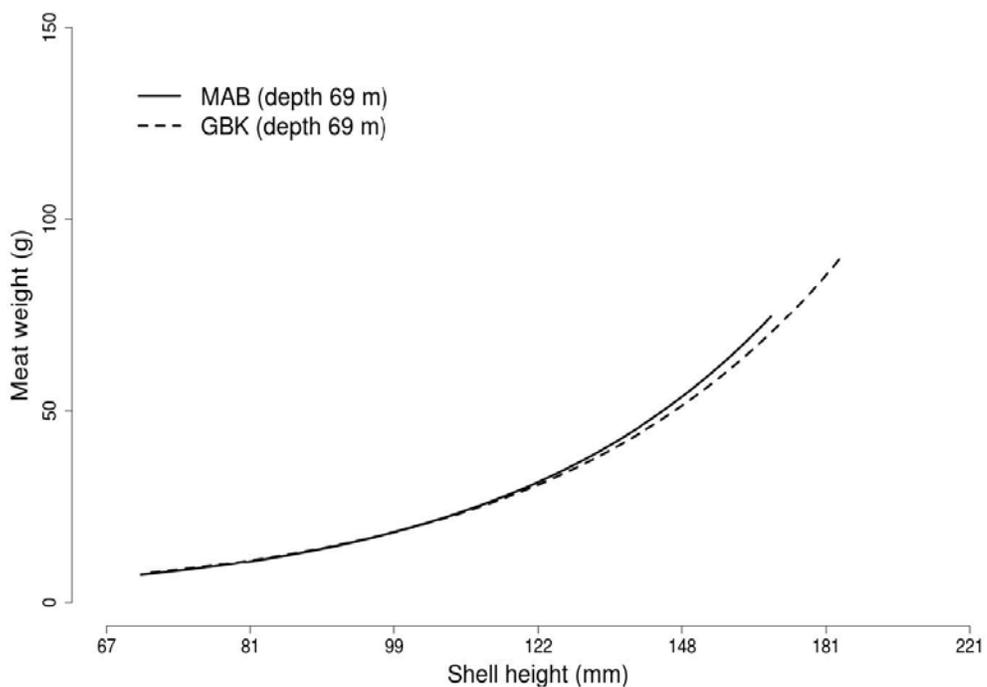


Figure 6: Shell height to meat weight relationship for each region based NEFSC survey data from 2003 -2013. The length of the curves represents the range of shell heights observed in each region.

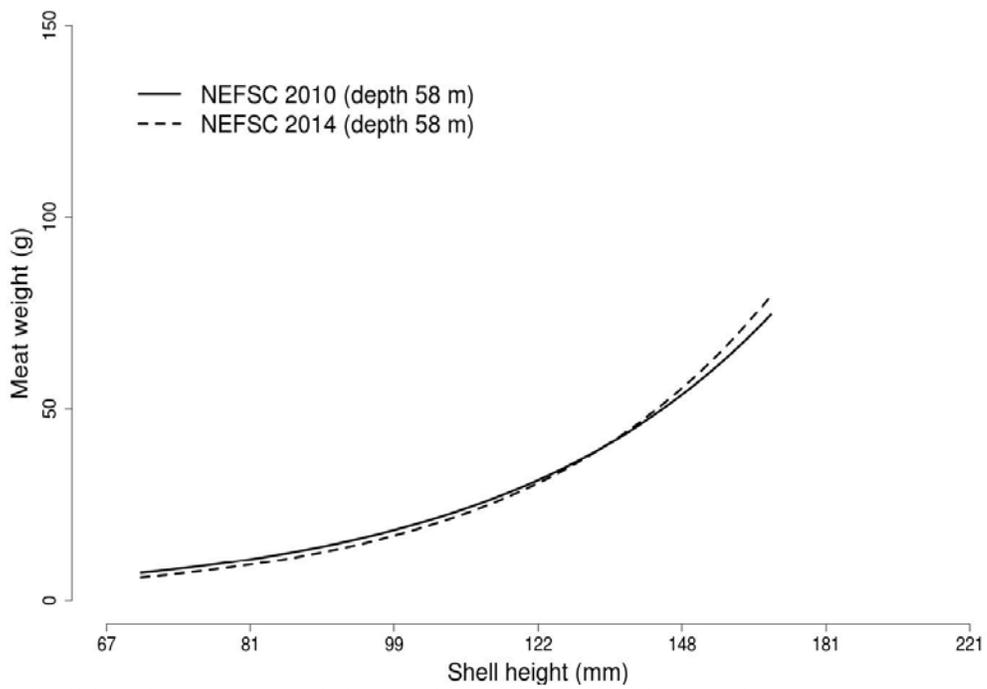


Figure 7: Shell height to meat weight relationship for two time periods in MAB. The length of the curves represents the range of shell heights observed in each period.

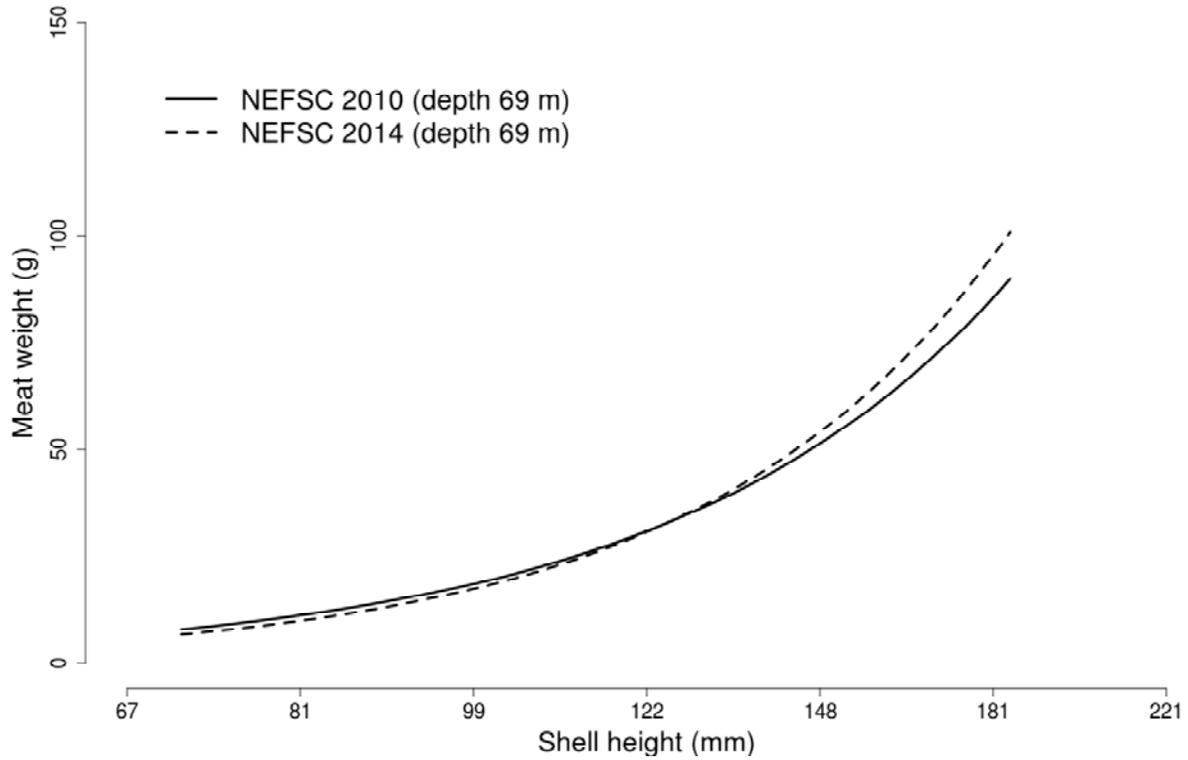


Figure 8: Shell height to meat weight relationship for two time periods in GBK. The length of the curves represents the range of shell heights observed in each period.

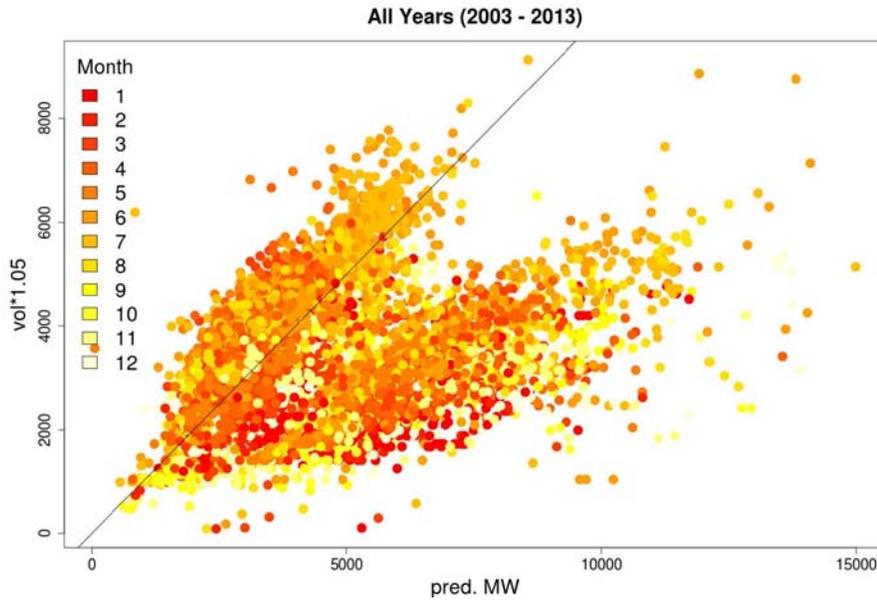


Figure 9: Meat weights estimated using data from the observer program compared to those expected based on NEFSC survey data. The solid line shows one to one correspondence and is for illustrative purposes only. The large cluster of points below the one to one line is an artifact of many more shells being measured for height than were packed into the cylinder for volume determination.

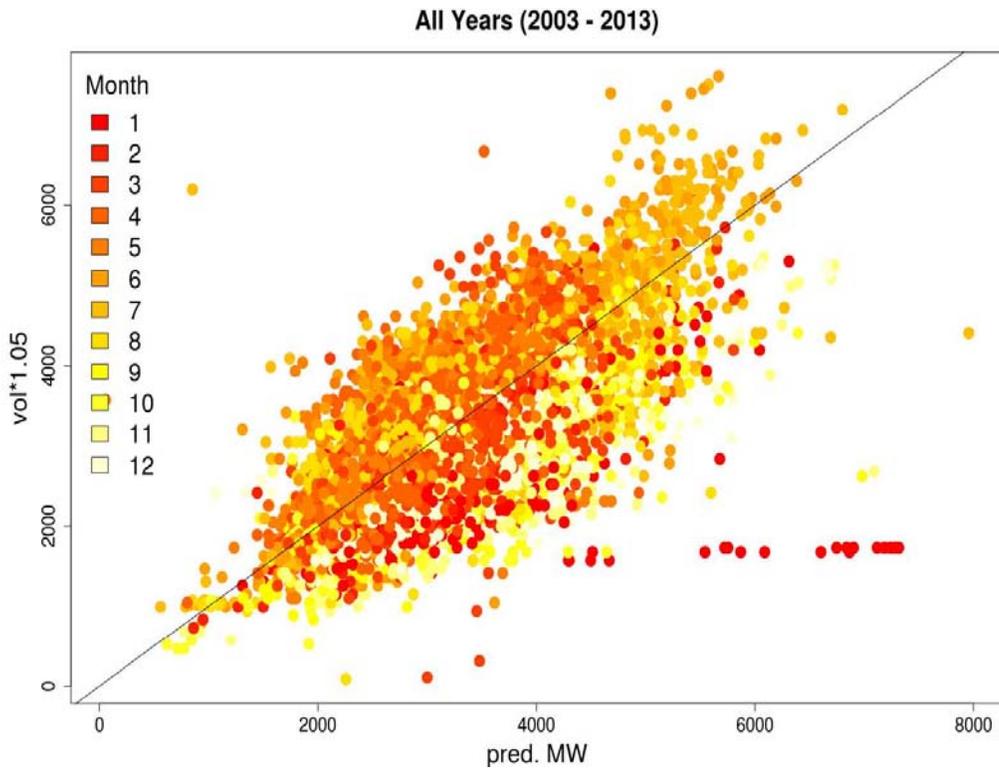


Figure 10: Meat weights estimated using data from the observer program compared to those expected based on NEFSC survey data. The solid line shows one to one correspondence and is for illustrative purposes only. Observations including more than 100 measured shells were excluded.

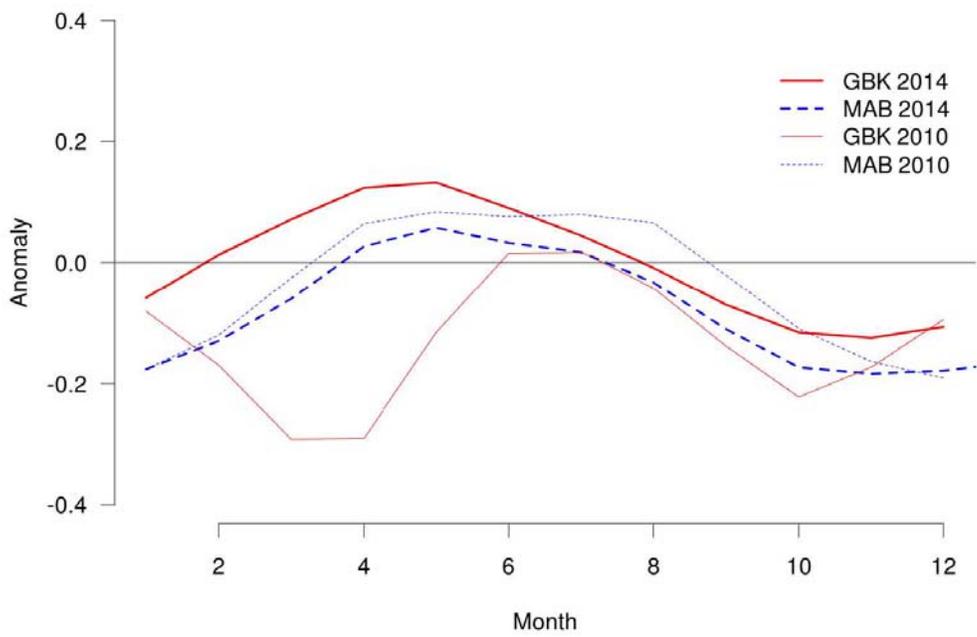


Figure 11: The anomalies estimated in the last assessment compared to the current anomalies.

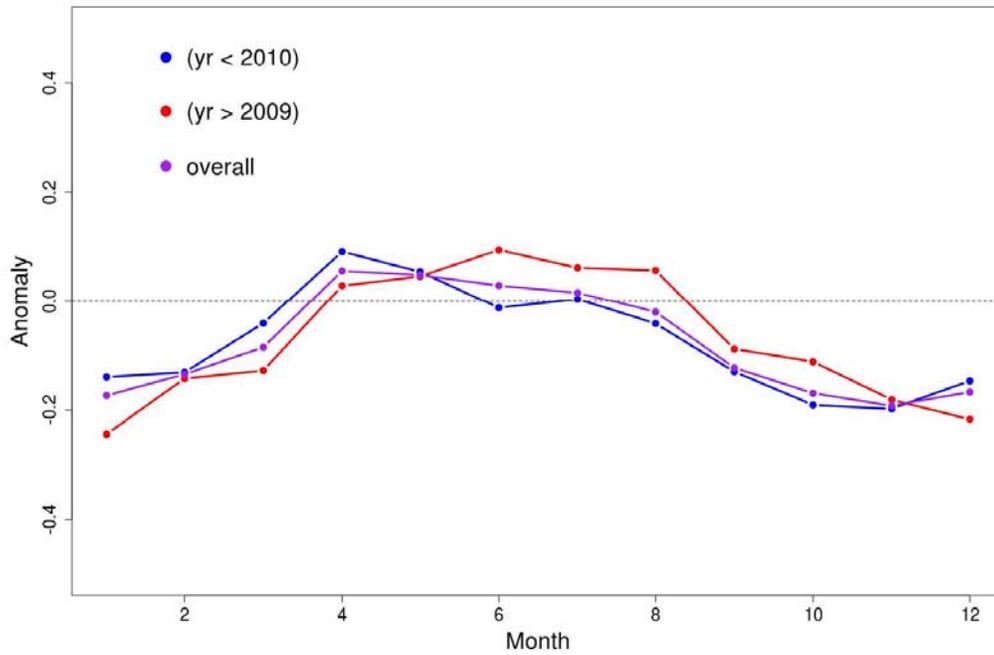


Figure 12: Monthly meat weight anomalies for the period prior to 2010, the period after 2010 and overall in the MAB.

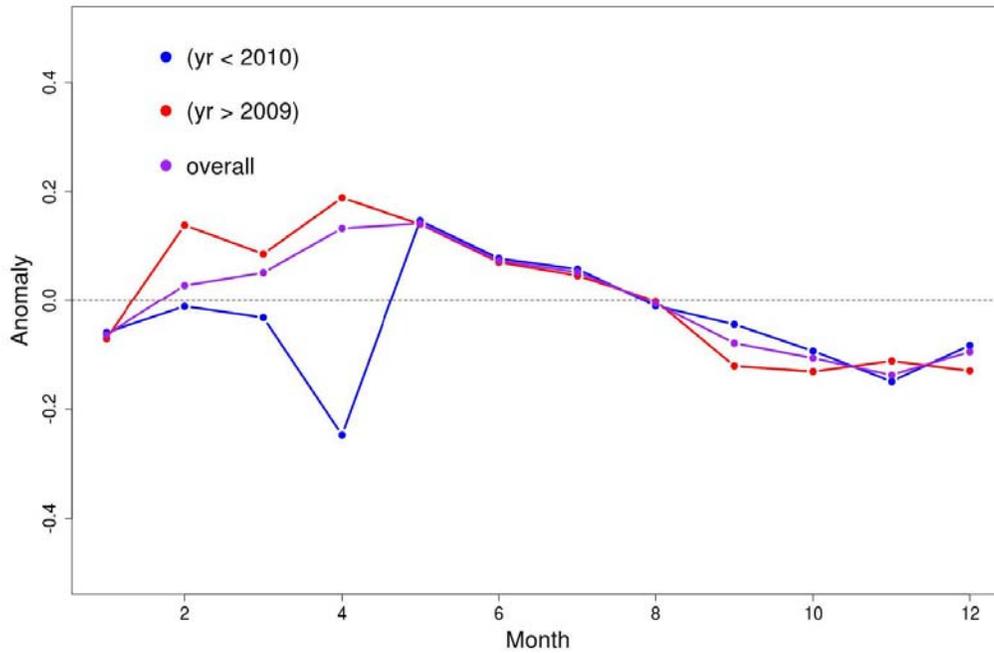


Figure 13: Monthly meat weight anomalies for the period prior to 2010, the period after 2010 and overall on GBK.

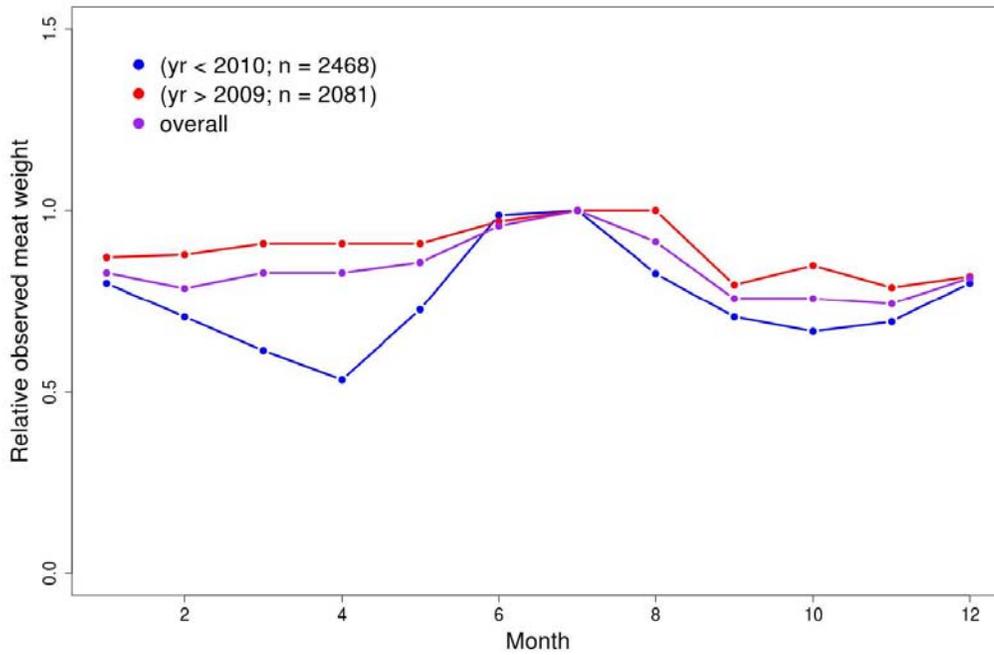


Figure 14: Relative monthly meat weight in observed commercial catches for the period prior to 2010, the period after 2010 and overall on GBK.

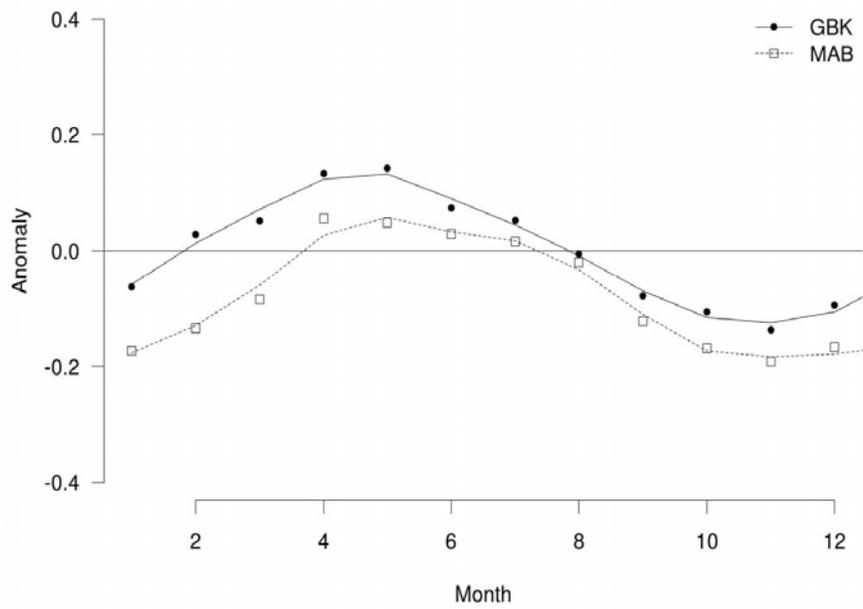


Figure 15: Smoothed anomalies for MAB and GBK.

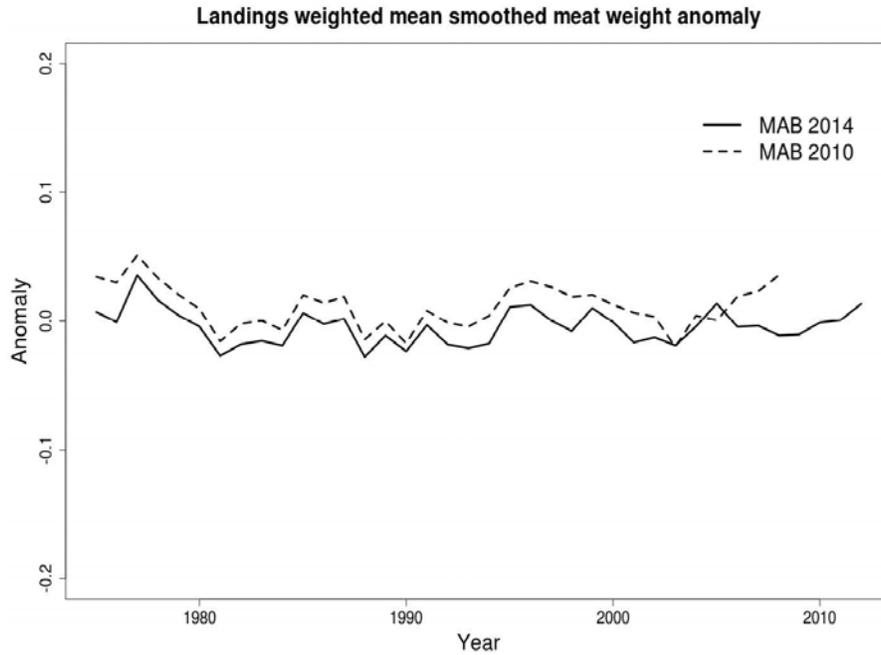


Figure 16: Landings weighted annual anomaly for MAB.

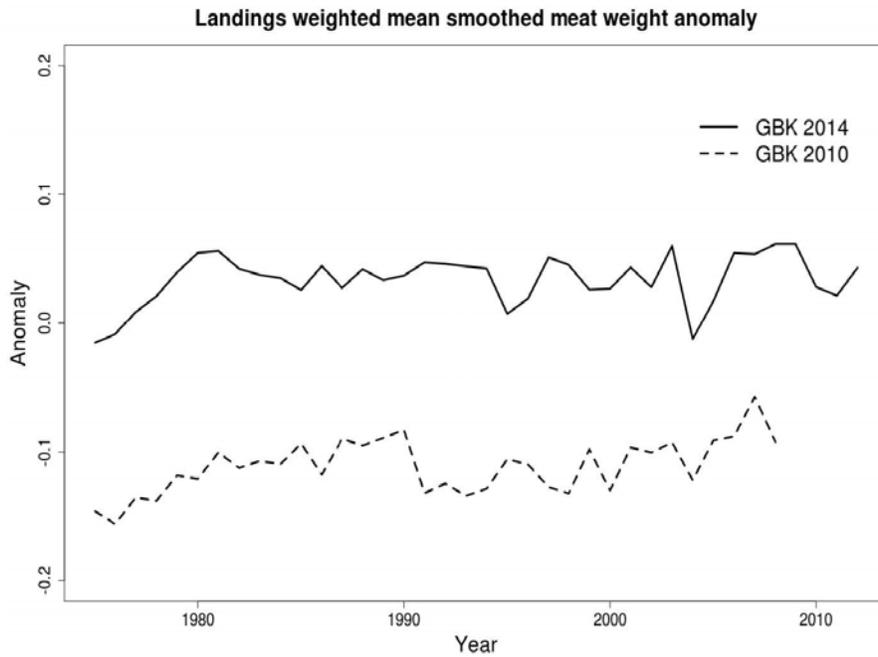


Figure 17: Landings weighted annual anomaly for GBK.

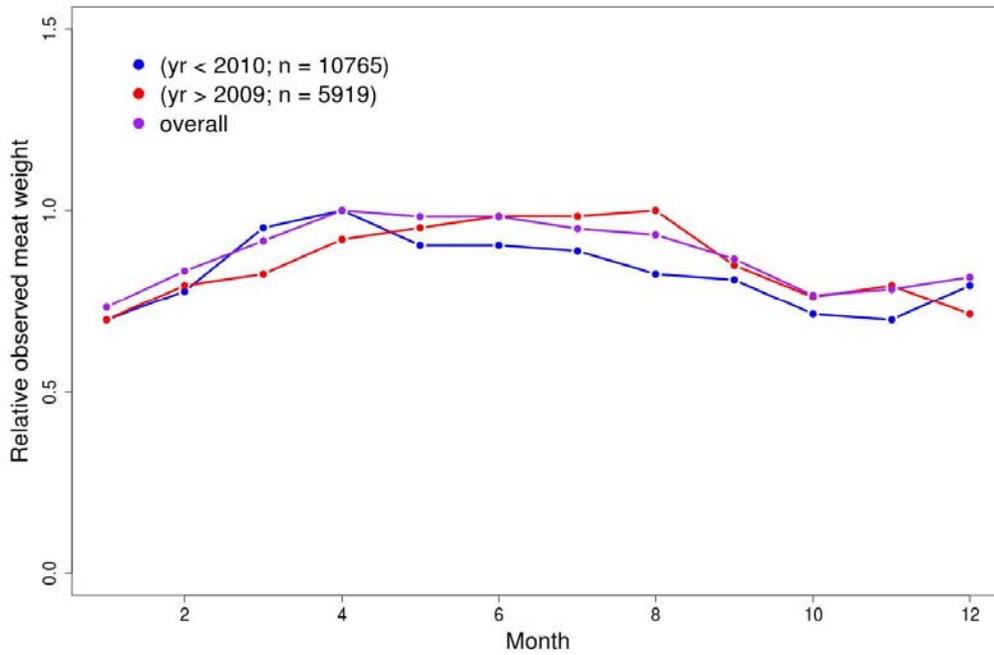


Figure 18: Relative monthly meat weight in observed commercial catches for the period prior to 2010, the period after 2010 and overall for MAB.

## Appendix B4. Estimation of Dredge efficiency from paired dredge-HabCam observations

Timothy J. Miller, NEFSC, Woods Hole, MA

We use HabCam optical survey data to estimate capture efficiency of the NEFSC scallop survey dredge where capture efficiency is the probability of capture for a scallop in the path of the dredge. The literature on methods for analysis of comparative gear studies is extensive, but an alternative observation model is used here because HabCam provides hundreds or thousands of observation for each dredge tow. We develop a general hierarchical model for the dredge and HabCam observations, compare relative performance of a set of specific models, assess the statistical behavior of the estimators to determine the best model, and provide relatively precise estimates of the efficiency of the scallop survey dredge on sand and gravel/cobble substrates.

### Materials and Methods

A dredge survey is conducted annually by the Northeast Fisheries Science Center to obtain relative abundance indices and other data for sea scallops. The dredge tows are conducted at stations according to a stratified random design. At a subset of these stations in 2008 and 2009, the HabCam optical survey device was also deployed. The HabCam captures images continuously along its track, but a thinned set were used in our analyses to make correlation between successive images within a station analyzed negligible. In all, we had 110 dredge stations where the number of sea scallops and swept area were recorded and where HabCam data including area searched, shell heights and number of scallops observed was recorded. There were 95-1,669 HabCam images used for each station.

The density of scallops differs by substrate type as based on HabCam as may the efficiency of the dredge. Sea scallop density is generally higher in sand than gravel substrates. We observe the substrate in each HabCam image, but the dredge track may cover various substrates which are not directly observed. The lack of these observations for the dredge makes estimation of relative efficiency for specific substrates impossible. However, sand and gravel/cobble substrates are more prevalent in particular survey strata. Sandy bottom is predominant in the Mid-Atlantic strata 6130, 6140, 6150, 6180, and 6190 and Georges Bank strata 6460, 6470, 6530, 6540, 6550, 6610, 6621, and 6670. Rock and gravel substrates are more common in Georges Bank strata 6490, 6500, 6510, 6520, 6651, 6652, 6661, 6662, and 6710. We therefore used stratum to establish proxies for substrate type when estimating dredge efficiency. In all there were 22 stations classified gravel (G) and 88 classified as sand (S).

### *Observation model*

At station  $i$  out of  $n$  total stations, we have the numbers captured by the dredge  $ND_i$  and the total number of sea scallops counted in associated HabCam images  $n_i$ . For HabCam, we assume all scallops are observed in each image and that the surface area  $A_{ij}$  of the substrate in the field of view is known. We also assume that the area swept by the dredge (determined using inclinometer sensors) is known. Conditional on the density of scallops in the image  $j$  at station  $i$

$\delta_{Hij}$ , we assume the number of scallops observed in the image is Poisson distributed with mean

$$(1) \quad E(N_{Hij}|\delta_{Hij}, A_{Hij}) = \delta_{Hij}A_{Hij}.$$

Conditional on the density of scallops  $\delta_{Di}$  and the known area swept by the dredge at station  $A_{Di}$ , we assume the number of captured scallops is Poisson distributed with mean

$$(2) \quad E(N_{Hij}|\delta_{Di}, A_{Di}) = q\delta_{Di}A_{Di}$$

where  $q$  is the efficiency of the dredge (cf. Paloheimo and Dickie 1964). Note that HabCam images is assumed to be 100% efficient at detecting scallops. More generally,  $q$  in Eq. 2 can be viewed as a relative efficiency when the HabCam is less than fully efficient.

We consider two different models for densities in each HabCam image  $\delta_{Hij}$ . The first simply assumes that the densities within a station are equal  $\delta_{Hij} = \delta_{Hi}$  and the second assumes that the densities are gamma distributed with station-specific mean  $\delta_{Hi}$  and shape  $\sigma_{Hi}$  parameters,

$$f(\delta_{Hij}|\delta_{Hi}, \sigma_{Hi}) = \frac{\delta_{Hij}^{\sigma_{Hi}-1} \exp\left(-\delta_{Hij} \frac{\sigma_{Hi}}{\delta_{Hi}}\right)}{\Gamma(\sigma_{Hi}) \left(\frac{\delta_{Hi}}{\sigma_{Hi}}\right)^{\sigma_{Hi}}}.$$

In the former model the counts in the HabCam images  $N_{Hij}|\delta_{Hi}$ , are still conditionally Poisson distributed. In the latter model, they are negative binomial distributed in the with mean

$$E(N_{Hij}|\delta_{Hi}) = \delta_{Hi}A_{Hij}$$

and variance

$$V(N_{Hij}|\delta_{Hi}) = E(N_{Hij}|\delta_{Hi}) \left[ 1 + \frac{E(N_{Hij}|\delta_{Hi})}{\sigma_{Hi}} \right].$$

For models where we assume the HabCam densities are gamma distributed we also consider variants where the shape parameter is constant across stations  $\sigma_{Hi} = \sigma_H$  and where the shape parameter is itself gamma distributed with mean  $\sigma_H$  and shape parameter  $\sigma_{\sigma_H}$ . The former corresponds to an assumption that the variability of the densities observed in each image is constant across stations and the latter allows the variability to change from station to station. For stations where  $\sigma_{Hi}$  is large, the distribution of HabCam image observations is closer to Poisson.

The dredge efficiency  $q$  and densities  $\delta_{Di}$  resulting in the dredge observations and the average densities  $\delta_{Hi}$  for HabCam observations at a given station are not all estimable as fixed parameters. Estimation of dredge efficiency requires some assumption about the relationship of dredge and HabCam densities both within and across stations. We use a bivariate gamma distribution described by Moran (1969) to relate the densities producing the HabCam and dredge observations at each station (see Attachment B4-1). The distribution is a function of the mean and shape parameters for the marginal gamma distributions and a correlation parameter ( $-1 <$

$\rho_\delta < 1$ ) that defines the relationship of dredge and HabCam densities within a station. The densities are independent when  $\rho_\delta = 0$  and identical when  $\rho_\delta = 1$ . We assume the means of the dredge and HabCam densities are the same, but that these means are a function of the substrate type at a given station. The details for the different components of five plausible models we consider are provided in Table 1.

The general likelihood that we maximize for parameter estimation is

$$(3) \quad L = \prod_{i=1}^n \left\{ \int_0^\infty \int_0^\infty f(N_{Di}|\delta_{Di})f(\delta_{Di}, \delta_{Hi}) \prod_{j=0}^{n_i} \left[ \int_0^\infty f(N_{Hij}|\delta_{Hij})f(\delta_{Hij}|\delta_{Hi})d\delta_{Hij} \right] d\delta_{Di}d\delta_{Hi} \right\}.$$

Unobserved densities are treated as random effects and integrated out to obtain the marginal model likelihood. Models such as  $M_1$  where HabCam densities within stations are assumed constant do not require the corresponding integration in Eq. 3. When densities within stations are gamma distributed, the numbers in the HabCam images conditional on  $\delta_{Hi}$  are negative binomial distributed. The closed form for this marginal sub-model is computationally more efficient. Because the densities are marginally gamma distributed and the dredge counts are Poisson distributed conditional on the realized densities at each station, dredge observations  $N_{Di}|\delta$  are marginally negative binomial distributed. The HabCam observations are also marginally negative binomial when the densities within a station are constant. In all models, the correlation of HabCam and dredge observations is defined by  $\rho$ .

We used AD Model Builder (Fournier et al. 2012) and the random effects library (Skaug and Fournier 2006) to maximize the marginal likelihood for all models. Parameters  $\theta$  were estimated on log scale except  $\rho_\delta$  which was defined as  $\rho_\delta = -1 + 2/(1 + e^{-\theta})$ . Standard errors were approximated using the delta method and asymmetric 95% confidence intervals were calculated by making the appropriate transformation of  $\hat{\theta} \pm z_{1-\frac{\alpha}{2}}SE(\hat{\theta})$  where  $\alpha = 0.05$  and  $z_{1-\frac{\alpha}{2}}$  is the quantile of the standard normal distribution with cumulative probability  $1 - \frac{\alpha}{2}$ .

### Simulation study

Because the methods were new, we used simulation to evaluate the reliability of the parameter estimates in the best model chosen by AIC. Using the parameter estimates from the best model, we simulated 1000 data sets and fit the same model to each data set. We calculated bias of parameter and standard error estimators and 95% confidence interval coverage.

## Results

The best performing model M5 demonstrated that the efficiency of the dredge differed substantially in gravel (0.24) and sandy (0.40) substrates (Table 2). There were dramatic reductions in AIC between  $M_1$  and  $M_2$  and between  $M_2$  and  $M_3$ . The reduction for  $M_2$  implies strong evidence of variability in densities among HabCam observations within stations. The reduction in AIC for model  $M_3$  implies strong evidence of variation among stations in the variance of HabCam observations. The very small difference in AIC values for  $M_3$  and  $M_4$  implies, implies that there is little evidence for differences in variability in mean densities among

stations for both HabCam and dredge observations.

Mean densities were much greater in gravel substrates ( $> 0.5 m^2$ ) than sand substrates ( $< 0.5 m^2$ ) for all models. Because there were fewer stations in the gravel substrate than sand, the relative precision of mean density estimates for gravel was lower for all models (CV about 0.3 for gravel vs. about 0.1 for sand). The precision of the dredge efficiency estimate was lower in gravel also (CV about 0.14 for gravel vs. about 0.06 for sand) for the best performing model  $M_5$ . The correlation of mean densities for dredge and HabCam observations was high ( $\rho_\delta > 0.9$ ) in all models.

### *Statistical behavior*

Seventy file out of 1000 simulations with model M5 did not converge. However, average parameter estimates for the unconverged fits were similar to averages for simulations where the model did converge. The relative bias for estimates from converged model fits was negligible for most parameters except that the shape parameter  $\sigma_{\sigma_H}$  which determines the variability of HabCam densities at each station was biased high by about 12% (Table 3). Standard error estimates were negligible for most parameters except  $\sigma_{\sigma_H}$  (SE approximately -15%) and the efficiency of the dredge in gravel substrates (SE approximately 6%). Bias of coverage for 95% confidence intervals was also small with the exception of the parameter  $\sigma_{\sigma_H}$  (bias about -9%).

### References

- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* 27(2): 233-249.
- Moran, P. A. P. 1969. Statistical inference with bivariate gamma distributions. *Biometrika* 56(3): 627-634.
- Paloheimo, J. E. and Dickie, L. M. 1964. Abundance and fishing success. *Rapports et Procès-Verbaux des Réunions. Conseil Permanent Internationale pour l'Exploration de la Mer* 155: 152-163.
- Skaug, H. J. and Fournier, D. A. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics & Data Analysis* 51(2): 699-709.

Table 1. Details of the fixed effects and random effects sub-models in the hierarchical models we fitted to paired HABCAM and dredge data.

Model	$E(N_{Di} \delta_{Di}, q)$	$E(N_{Hij} \delta_{Hij})$	$\delta_{Hij}$	$\sigma_{Hi}$	$\delta_{Di}, \delta_{Hi}$	$n_p$
M <sub>1</sub>	$q\delta_{Di}A_{Di}$	$\delta_{Hij}A_{Hij}$	$\delta_{Hi}$	—	BGamma( $\delta(G, S), \sigma_\delta, \rho_\delta$ )	5
M <sub>2</sub>	$q\delta_{Di}A_{Di}$	$\delta_{Hij}A_{Hij}$	Gamma( $\delta_{Hi}, \sigma_{Hi}$ )	$\sigma_H$	BGamma( $\delta(G, S), \sigma_\delta, \rho_\delta$ )	6
M <sub>3</sub>	$q\delta_{Di}A_{Di}$	$\delta_{Hij}A_{Hij}$	Gamma( $\delta_{Hi}, \sigma_{Hi}$ )	Gamma( $\sigma_H, \sigma_{\sigma_H}$ )	BGamma( $\delta(G, S), \sigma_\delta, \rho_\delta$ )	7
M <sub>4</sub>	$q\delta_{Di}A_{Di}$	$\delta_{Hij}A_{Hij}$	Gamma( $\delta_{Hi}, \sigma_{Hi}$ )	Gamma( $\sigma_H, \sigma_{\sigma_H}$ )	BGamma( $\delta(G, S), \sigma_{D\delta}, \sigma_{H\delta}, \rho_\delta$ )	8
M <sub>5</sub>	$q(G, S)\delta_{Di}A_{Di}$	$\delta_{Hij}A_{Hij}$	Gamma( $\delta_{Hi}, \sigma_{Hi}$ )	Gamma( $\sigma_H, \sigma_{\sigma_H}$ )	BGamma( $\delta(G, S), \sigma_\delta, \rho_\delta$ )	8

Table 2. AIC and parameter estimates for each fitted model. Parameters denoted with (G) and (S) are specific to observations from gravel and sand substrates, respectively and (D) and (H) denote parameters specific to dredge and HABCAM observations, respectively.

Model	$\Delta(AIC)$	$q$	$\delta$	$\sigma_\delta$	$\rho_\delta$	$\sigma_H$	$\sigma_{\sigma_H}$
M <sub>1</sub>	6724.0	0.376 (0.020)	5.048 (1.476) (G) 0.470 (0.056) (S)	0.621 (0.084)	0.905 (0.022)	—	—
M <sub>2</sub>	986.6	0.376 (0.020)	5.041 (1.480) (G) 0.469 (0.056) (S)	0.622 (0.084)	0.905 (0.022)	1.576 (0.044)	—
M <sub>3</sub>	8.20	0.376 (0.020)	5.085 (1.505) (G) 0.469 (0.056) (S)	0.620 (0.084)	0.906 (0.022)	3.419 (0.625)	0.880 (0.207)
M <sub>4</sub>	8.60	0.383 (0.021)	5.448 (1.653) (G) 0.461 (0.054) (S)	0.586 (0.084) (D) 0.647 (0.091) (H)	0.910 (0.021)	3.418 (0.624)	0.880 (0.207)
M <sub>5</sub>	—	0.243 (0.034) (G) 0.400 (0.022) (S)	5.771 (1.708) (G) 0.458 (0.054) (S)	0.630 (0.085)	0.912 (0.020)	0.630 (0.624)	0.880 (0.207)

Table 3. Relative bias of parameter and standard error estimators and coverage probability of approximate 95% confidence interval for 925 simulated data sets with parameters specified from the best performing model  $M_5$ .

Parameter	Value	Relative Bias	SE	Relative Bias of SE	95% CI coverage
$q$ (G)	0.24	0.01	0.03	-0.06	0.93
$q$ (S)	0.40	0.00	0.02	-0.01	0.94
$\delta$ (G)	5.77	-0.01	1.33	0.00	0.94
$\delta$ (S)	0.46	0.00	0.06	-0.03	0.93
$\sigma_\delta$	0.63	0.03	0.09	-0.03	0.94
$\rho_\delta$	0.91	0.00	0.02	-0.04	0.94
$\sigma_H$	3.42	-0.03	0.53	-0.02	0.91
$\sigma_{\sigma_H}$	0.88	0.12	0.19	-0.15	0.86

Attachment B4-1. Bivariate gamma distribution.

This is the same formulation described by Moran (1969). Let  $Y_1$  and  $Y_2$  be bivariate standard normal distributed with correlation parameter  $\rho$ ,

$$f(Y_1, Y_2) = \frac{1}{2\pi(1 - \rho^2)^{\frac{1}{2}}} \exp \left[ -\frac{1}{2(1 - \rho^2)}(y_1^2 - 2\rho y_1 y_2 + y_2^2) \right].$$

Then letting the marginal distributions  $F(X_1) = F(Y_1)$  and  $F(X_2) = F(Y_2)$ , where

$$F(X_i) = \int_0^{X_i} \frac{1}{\Gamma(\sigma_i)\beta_i^{\sigma_i}} w_i^{\sigma_i-1} \exp(-w_i\beta_i^{-1}) dw_i,$$

$X_1$  and  $X_2$  have a bivariate gamma distribution with means  $\sigma_i\beta_i$  and marginal variances  $\sigma_i\beta_i^2$ , but correlation defined by  $\rho$ . When  $\rho = 0$ ,  $X_1$  and  $X_2$  are independent and when  $\rho = 1$ ,  $X_1$  and  $X_2$  are identically distributed.

## Appendix B5. Empirical assessment

Larry Jacobson, NEFSC, Woods Hole, MA

### Introduction

The empirical assessment used simple techniques to estimate sea scallop stock abundance, biomass and fishing mortality in the MAB, GBK and combined stock areas. The purpose was to evaluate the accuracy of CASA estimates as independently as possible. However, empirical results could be used in place of CASA model estimates if the later were unavailable. The data and various parameters used in the empirical analysis are a subset of those also used in the CASA model and were all obtained independently in field studies or other analyses rather than from a stock assessment model.

### Materials and methods

Survey swept-area abundance data used in the empirical analysis were the best available estimates of total 40+ mm stock abundance and considered reliable. Abundance from the dredge and optical surveys (HabCam and SMAST large camera) were the same as used in CASA except that SMAST data were adjusted for logistic size selectivity using externally estimated selectivity curves (Appendix B7 in NEFSC 2007). In CASA, the same selectivity curves are applied in the model after data input. In addition abundance estimates were not rescaled for comparison to a prior distribution as in CASA although this had no impact on results. Size selectivity was assumed to be flat in the dredge and HabCam surveys.

Updated capture efficiency estimates were used in expansion of the dredge survey to calculate swept-area abundance prior to their use in this analysis (Appendix B4). Additional variance due to uncertainty about dredge efficiency was included (see below). Capture efficiency was assumed to be 100% in the dredge and HabCam surveys for scallops 40+ mm SH in calculating swept-area abundance for this analysis. Thus, capture efficiency was factored in to all of the survey abundance data prior to use here.

As in the CASA model analysis, dredge survey abundance estimates were adjusted to account for scallops in deep or shallow water areas not sampled by the dredge but no adjustments were made for areas of poor habitat within the survey area. Survey abundance at length data were not adjusted for errors in measuring shell height as in the CASA model although such errors are appreciable in the optical surveys because the adjustment requires information available in a simulation based stock assessment model. These type of errors smooth size composition estimates making modes lower, valleys higher and proportions in the largest and smallest length groups larger (Jacobson et al. ???).

Five mm length groups (40-45, 45-50 ...) were used and the last length group was always a plus group. Intermediate calculations included all of the size groups in the original data but results are summarized using a 140+ mm size group, which is roughly the same as von Bertalanfy  $L_{max}$  (asymptotic mean size) estimates. Only years 2003-2013 were included because at least two surveys (dredge+SMAST, dredge+HabCam, or dredge+SMAST+HabCam) were conducted each year. Using multiple independent surveys helps smooth estimates without using a population dynamics model like CASA.

Total abundance in each year and for each size group ( $N_{y,L}$ ) was estimated by averaging swept-area abundance estimates from each survey:

$$N_{y,L} = \frac{\sum_s N_{s,y,L}}{n_{s,y}}$$

where  $N_{s,y,L}$  was swept area abundance data for year  $y$  and survey  $s$  while  $n_{s,y}=2$  or  $3$  was the number of surveys. Total survey stock abundance was  $N_y = \sum_L N_{y,L}$ . Stratified random CVs for mean total number per tow and the number of positive tows by year in the dredge survey provide some information about precision of abundance data (Table 1 and Figure 1).

Variances for  $N_{y,L}$  were calculated from length specific average CVs for mean number per tow in the dredge survey. Length specific variances were not easily available for the SMAST and HabCam surveys. In particular:

$$Var(N_{s,y,L}) = (CV_L N_{s,y,L})^2$$

where  $CV_L$  is the average CV at length in the dredge survey for either Georges Bank or the Mid-Atlantic (Figure 2). CVs for total abundance  $N_y$  were from the CVs for total catch per tow in each survey (Figure 1):

$$Var(N_{s,y}) = CV_{s,y} N_{s,y}$$

and:

$$Var(N_y) = \sum_s Var(N_{s,y})/n_s^2.$$

Dredge survey abundance CVs were increased to account for uncertainty in capture efficiency. CVs for dredge survey capture efficiency were  $0.034/0.243=0.14$  (gravel/cobble) and  $0.022/0.4=0.05$  (sand, Appendix B4). Therefore, the adjusted CV for a dredge survey abundance estimate was  $\sqrt{CV_{s,y}^2 + 0.1^2}$  where  $0.1$  is close to the average CV for gravel/cobble and sand.

Uncertainty about stock area, area sampled, and other factors were ignored in calculating survey abundance. However, variance from these factors was probably modest relative to the variance in mean catch per tow and capture efficiency for the dredge survey. Uncertainty about stock area is relatively small because scallops are sessile with a static spatial distribution that is well defined by the optical surveys and covered effectively by each survey after the dredge data are adjusted for area not surveyed. Uncertainty about size selectivity in the experimentally derived size selectivity curve for the SMAST survey was ignored for lack of time but could have been included.

For plotting, mean abundance at length estimates were smoothed with GAM models fit assuming gamma errors using the mgcv library in the R programming language (Wood 2006):

$$\text{gam}(y \sim s(x), \text{family}=\text{Gamma}(\text{link}=\text{log}), \text{weights}=\text{wts})$$

The variances used for weights were, for example,  $Var(N_{y,L}) = [N_{y,L} CV(N_{y,L})]^2$ . Assuming predicted values were gamma distributed, 95% percent confidence intervals were calculated for means equal to the fitted values and variances  $Var(\hat{N}_{y,L}) = [\hat{N}_{y,L} CV(F_{y,L})]^2$ . The variance of the fitted values calculated in the GAM was not used because it grossly underestimated uncertainty. Better confidence intervals might have been obtained by combining the CV above with the CV for uncertainty in the smooth trend calculated by the GAM software.

Fishing mortality rates by year and length ( $F_{y,L}$ ) were approximated by dividing catch numbers by estimated abundance:

$$F_{y,L} = \frac{C_{y,L}}{N_{y,L}}$$

Where  $C_{y,L}$  is catch number at length. This approximation is reasonable because the instantaneous rate of fishing mortality is exactly  $F = C/\bar{N}$  (Ricker 1975) and because scallop surveys tend to occur near the middle of the year when abundance may be similar to average abundance (Table 2).

Catch numbers at length in each year ( $C_{y,L}$ ) were calculated:

$$C_{y,L} = \frac{W_y}{m_y} p_{y,L}$$

where  $W_y$  is total meat weight for landings,  $m_y$  is mean weight of scallops in the catch and  $p_{y,L}$  is a size-specific proportion of the total commercial catch. The mean weight ( $m_y$ ) was calculated from commercial size composition data, survey shell height-meat weight parameters and annual commercial meat weight anomalies as in the CASA model.

Variances for fishing mortality were approximated based on CVs for average survey abundance and an assumed CV=10% for catch to give  $CV(F_{y,L}) = \sqrt{CV(N_{y,L}) + 0.1^2}$ .

Abundance weighted fishing mortality (all sizes combined) was approximated  $F_y = C_y/N_y$  with  $(F_y) = \sqrt{CV(N_y) + 0.1^2}$ .

CASA models include a correction for incidental mortality which is highest on the smallest size groups. This adjustment was not made in the empirical analysis because it requires an a-priori estimate of fishing mortality and fishery selectivity not available in the empirical analysis. Therefore, fishing mortality  $F_{y,L}$  and  $F_y$  are underestimated relative to total fishery mortality. Fishing mortality attributable to landings and fully recruited fishing mortality are unaffected.

GAM models were used to smooth fishing mortality at size estimates and confidence intervals were estimated in a manner similar to abundance at size. The variances used for weights were  $Var(F_{y,L}) = [F_{y,L} CV(F_{y,L})]^2$  and the variances used to calculate confidence intervals were  $Var(\hat{F}_{y,L}) = [\hat{F}_{y,L} CV(F_{y,L})]^2$ . Fully recruited fishing mortality was estimated using the gam model to predict  $F_{y,L}$  over a wide range of narrowly spaced shell height values and selecting the largest value of predicted  $F_{y,L}$ .

Commercial size selectivity estimates are useful although not required in the empirical assessment or in projections which are handled independently in the SAMS model. However, for illustration, size selectivity by year and size  $s_{y,L}$  was estimated by rescaling fishing mortality at size:

$$s_{y,L} = \frac{F_{y,L}}{\max(F_{y,L})}$$

and then smoothing the rescaled estimates using a model for proportions:

$$\text{gam}(y \sim s(x), \text{family} = \text{quasibinomial}, \text{weights} = \text{wts})$$

The weights were one when estimating selectivity at size in individual years. Weights equal  $n_s$  were used when selectivity estimates for multiple years were combined to estimate average fishery selectivity. After the GAM model was fit, predicted selectivity were rescaled again to a maximum value of one. Fishable abundance (available to the fishery) in each year  $A_y$  can be calculated using abundance at size and a fishery selectivity estimate although the estimates are not required for this empirical assessment. For example:

$$A_y = \sum_L s_L N_{y,L}$$

## Results

Empirical abundance at size estimates appear reasonably precise and smooth although the smoothness is due partly to measurement errors in survey size data (Figure 3). The progression of two large year classes is clear during 2003-2006 in the Mid-Atlantic and during 2012-2013 in

both regions. There are clear differences between the two regions in population size composition (e.g. the 140+ mm size group) seem clear. Important aspects of the fishery (relatively low exploitation rates and targeting large animals) are evident in comparing abundance and catch numbers at size (Figure 4).

Empirical fishing mortality at length data show that fishing pressure is higher in the Mid-Atlantic than on Georges Bank (Figure 5). The working group concluded that the variation over time in fishery selectivity between domed and ascending patterns could be explained in terms of management measures that: 1) increased the minimum ring size on commercial vessels and decreased selectivity of small scallops during 1994-1995, 2) recruitment events, and 3) management measures that opened and closed rotational harvest areas where large scallops were common. Average fishery selectivity curves for 2003-2013 illustrate how selectivity for particular time periods can be estimated as needed for management related or other analyses (Figure 6).

Empirical abundance and fishing mortality for the combined Mid-Atlantic and Georges Bank regions were calculated by summing catch numbers and abundance for the Mid-Atlantic and Georges Bank regions and then computing approximate fishing mortality rates from the ratio of the sums. CVs and were calculated using standard formulas for sums of random variables.

Empirical and CASA model estimates of abundance and fishing mortality show similar trends in all regions (Tables 3-4 and Figure 7). However, empirical abundance estimates were generally higher reflecting the tension in CASA models between matching the scale of the abundance data (matching the prior on Q) versus fitting the survey and fishery data. As expected, fishing mortality show the inverse pattern with empirical generally lower than CASA estimates.

Fully recruited fishing mortality estimates from empirical calculations were usually lower than from CASA the CASA model as well (Figure 8). However, the comparison may not be very useful because of fully recruited F depends on fishery selectivity assumptions which differed in the two assessment approaches.

#### Status determination and catch advice

No special provisions are necessary for providing catch advice to the scallop fishery using the empirical methods. Catch advice is generated using a simulation models (SAMS) which is initialized using best estimates of abundance at length from surveys (i.e. using the empirical method).

Reference points used to determine if the scallop stock is overfished or if overfishing is occurring are more difficult. For this assessment, it would be reasonable to compare empirical fishing mortality estimates to reference points calculated in terms of landings divided by 1 July abundance from the SYM reference point model. The CASA model may be problematic due to the tension between scale of the model estimates and general fit to the data. However, the current condition of the stock (not overfished and overfishing not occurring) is clear based on both sets of models and common sense. Empirical and CASA results are broadly similar. If the trend in  $B/B_{MSY}$  estimates from the CASA and SYM models are roughly correct, then the ratio for 2013 should be sufficient to determine if the stock is overfished despite uncertainty about scale.

### Advantages and disadvantages

It was advantageous to use both empirical and the complex CASA modeling approach for CASA, if only for comparison and to determine if the CASA model results were plausible. Empirical estimates depend almost entirely on data while the CASA model depends on data, biological assumptions (e.g. about growth and natural mortality) and modeling techniques. The empirical approach requires fewer assumptions about growth, natural mortality, size selectivity, etc. and uses most of the data also used in CASA. However, the empirical approach is sensitive to survey measurement errors which are relatively high in the Georges Bank area. It is therefore necessary to have multiple surveys each year for empirical estimation. The empirical approach cannot be applied in all years and the CASA model may give a clearer long term perspective on stock size and productivity.

In theory, the CASA model should do a better job of balancing goodness of fit to survey, catch and size composition data to estimate realistically smooth population trends. However, experience with many real stocks and models indicates that stock assessment models often have pathological problems that may be difficult to resolve due to many potential causes including inaccurate catch data, changes in natural mortality, etc..

An assessment model like CASA makes it easier to calculate reference points. Empirical reference point methods were not evaluated in this assessment but there are a number of methods that could be applied.

Empirical estimates do not suffer from retrospective patterns, which are usually blamed on model structure or assumptions about the data which may remain hidden in empirical analyses. CASA model results did not show retrospective error in this assessment but this was probably due to the proximity of the estimates to priors for survey capture efficiency with tension in the model pulling abundance estimates low enough so that implied capture efficiency estimates were trapped near the upper prior bound. The empirical estimates in this assessment for 2003-2013 are less sensitive to errors in historical catch which are often suspected when modeling problems occur.

### Reference:

NEFSC. 2007. 45th Northeast Regional Stock Assessment Workshop (45th SAW): 45th SAW assessment report. US Dep Commer, Northeast Fish Sci Cent Ref Doc 07-16.

Table 1. Numbers of tows in which at least one scallop was caught in the MAB and GBK areas during dredge surveys during 2003-2013 by size group. For example, the 40 mm size group is 40-44.9 mm SH. The last size bin (140+ mm SH) is a plus group. The number of positive tows is a lower bound estimate for the effective sample size in each year/size group category..

Year	Size group (mm)																				
	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140+
<b>MAB</b>																					
2003	110	113	120	127	145	146	145	151	147	145	152	158	160	159	156	157	135	122	91	56	39
2004	124	132	145	137	150	146	154	170	187	192	191	188	187	186	192	187	169	150	120	84	41
2005	157	160	170	161	147	152	142	168	188	205	215	217	220	224	224	223	216	210	194	164	127
2006	111	139	160	176	196	232	222	231	242	235	239	240	246	250	248	252	246	234	211	163	117
2007	70	97	130	148	150	172	186	204	209	218	237	249	250	257	250	249	244	237	203	168	131
2008	168	183	179	178	176	158	154	159	172	180	199	214	217	222	215	217	207	202	175	149	125
2009	77	88	104	97	114	108	121	147	152	151	160	152	157	153	162	157	156	150	130	103	86
2010	141	156	156	135	131	117	122	134	171	199	219	227	240	236	234	241	233	227	196	132	100
2011	119	149	151	146	123	111	96	117	165	191	223	214	219	225	232	230	238	238	225	187	163
2012	155	165	158	141	131	120	126	149	156	174	185	187	192	211	208	201	213	217	204	171	119
2013	99	129	164	167	213	216	229	227	222	232	231	238	224	220	220	216	213	214	203	161	140
<b>GBK</b>																					
2003	64	72	76	84	99	92	95	99	96	110	115	116	124	137	137	131	131	139	128	114	122
2004	83	94	96	105	102	92	95	108	120	141	140	145	148	156	153	164	166	169	163	141	140
2005	46	57	98	94	108	101	106	109	133	142	164	177	205	229	245	254	267	277	276	256	248
2006	67	74	88	103	108	96	103	96	93	112	127	138	138	144	154	154	170	172	173	165	172
2007	153	181	217	215	240	222	204	189	190	185	199	202	210	212	208	246	271	276	277	274	284
2008	111	114	129	146	156	145	131	141	138	148	158	174	178	183	168	170	159	176	169	180	196
2009	95	107	135	132	128	126	119	117	130	145	158	160	156	162	164	162	160	161	152	148	168
2010	81	77	92	88	111	108	117	130	152	150	170	161	185	193	214	215	219	223	224	206	216
2011	44	44	43	50	68	72	85	92	119	132	146	138	154	148	155	176	177	184	180	176	184
2012	61	86	100	105	100	94	107	107	125	133	144	155	151	157	168	174	176	181	181	178	177
2013	81	106	115	123	138	139	118	108	112	116	122	126	134	133	141	142	155	153	163	156	161

Table 2. Dates (Julian) for sea scallop surveys during 2003-2013 in the MAB and GBK regions.

Survey	Mid-Atlantic			Georges Bank			Comment
	Min	Max	Mid	Min	Max	Mid	
Dredge	130	215	173	163	230	197	1979-2013
SMAST	130	194	162	165	233	199	2003-2009
HabCam	153	201	177	159	210	184	2011-2012 for the Mid-Atlantic and 2011-2013 for Georges Bank

Table 3. Abundance and fishing mortality (estimates from the empirical approach and CASA model for the Georges Bank (top) and Mid-Atlantic (bottom) regions).

Year	Empirical					CASA	
	Abundance (Mid-year, 40+ mm, 10 <sup>6</sup> )	CV	Landings	Aprox. F	CV	Abundance (1 July, 40+ mm, 10 <sup>6</sup> )	Landings/ Abundance
Georges Bank							
2003	4,145	0.10	173	0.04	0.14	3,517	0.05
2004	3,788	0.12	133	0.04	0.15	3,159	0.04
2005	3,660	0.11	267	0.07	0.15	3,132	0.09
2006	3,216	0.11	448	0.14	0.15	2,769	0.16
2007	3,979	0.11	249	0.06	0.15	3,219	0.08
2008	3,941	0.10	179	0.05	0.14	3,300	0.05
2009	5,332	0.12	221	0.04	0.15	3,690	0.06
2010	4,883	0.17	170	0.03	0.19	3,801	0.04
2011	4,169	0.12	217	0.05	0.15	4,194	0.05
2012	3,498	0.08	316	0.09	0.13	4,607	0.07
2013	4,073	0.14	365	0.09	0.17	5,620	0.06
Mid-Atlantic							
2003	13,601	0.31	807	0.06	0.33	5,511	0.15
2004	7,324	0.21	918	0.13	0.23	4,036	0.23
2005	6,154	0.15	545	0.09	0.18	4,811	0.11
2006	6,261	0.15	272	0.04	0.18	4,226	0.06
2007	5,521	0.15	503	0.09	0.18	4,310	0.12
2008	6,340	0.13	463	0.07	0.16	4,647	0.10
2009	5,312	0.11	664	0.13	0.15	3,202	0.21
2010	3,794	0.11	687	0.18	0.15	2,458	0.28
2011	2,747	0.10	598	0.22	0.14	1,606	0.37
2012	4,617	0.10	365	0.08	0.14	3,387	0.11
2013	4,163	0.14	219	0.05	0.17	2,648	0.08

Table 4. Abundance and fishing mortality (estimates from the empirical approach and CASA model to the combined Georges Bank plus Mid-Atlantic regions (whole stock).

Year	Empirical			CASA			
	Abundance (Mid-year, 40+ mm, 10 <sup>6</sup> )	CV	Landings	Aprox. F	CV	Abundance (1 July, 40+ mm, 10 <sup>6</sup> )	Landings/ Abundance
Whole stock							
2003	17,746	0.24	980	0.06	0.26	9,028	0.11
2004	11,112	0.14	1,051	0.09	0.17	7,195	0.15
2005	9,814	0.11	812	0.08	0.15	7,942	0.10
2006	9,477	0.11	720	0.08	0.15	6,994	0.10
2007	9,500	0.10	752	0.08	0.14	7,529	0.10
2008	10,281	0.09	643	0.06	0.13	7,946	0.08
2009	10,644	0.08	885	0.08	0.13	6,891	0.13
2010	8,677	0.11	857	0.10	0.15	6,259	0.14
2011	6,915	0.08	815	0.12	0.13	5,799	0.14
2012	8,115	0.07	681	0.08	0.12	7,995	0.09
2013	8,237	0.10	584	0.07	0.14	8,269	0.07

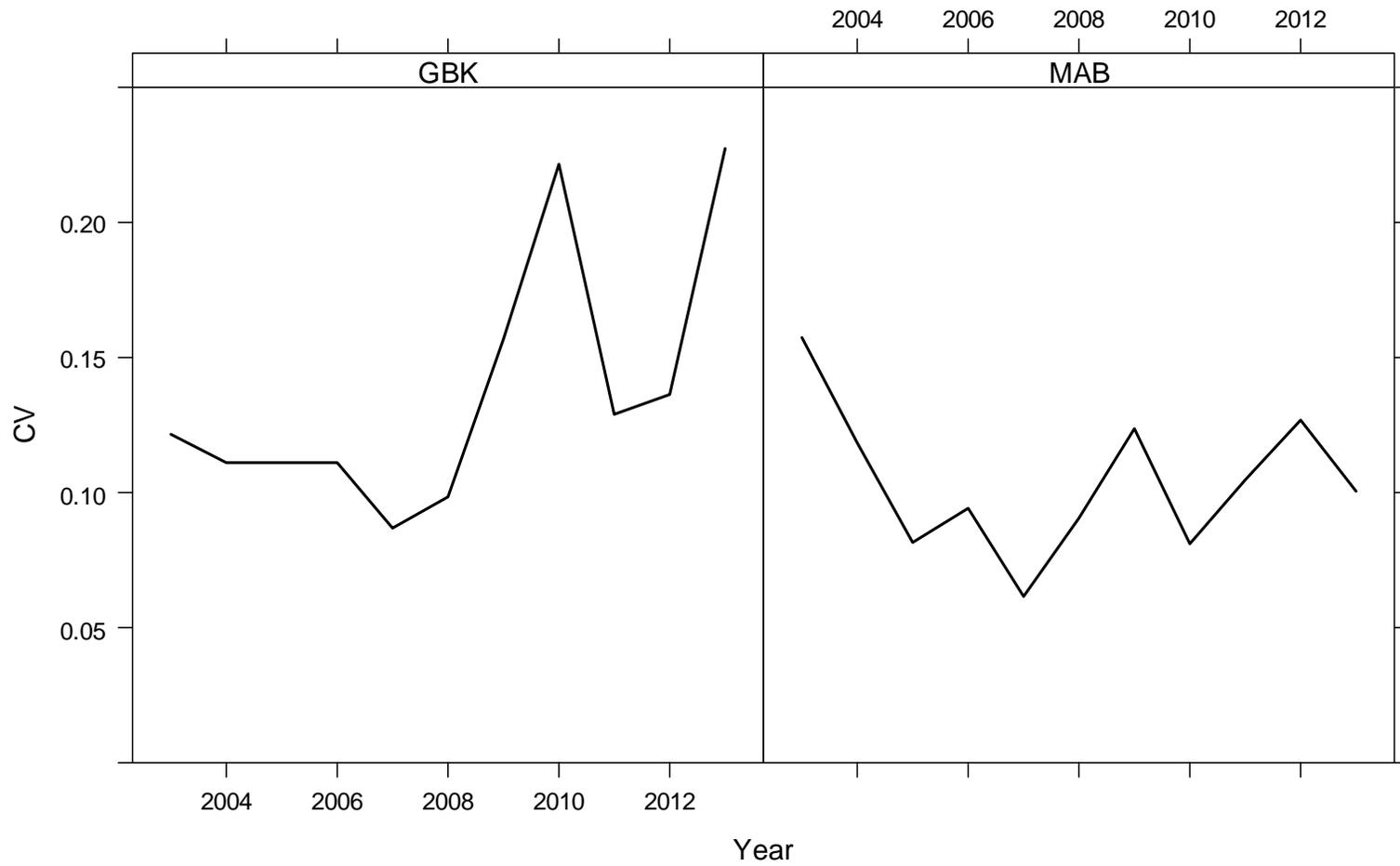


Figure 1. CVs for total mean catch per tow (all sizes) in the dredge survey during 2003-2013.

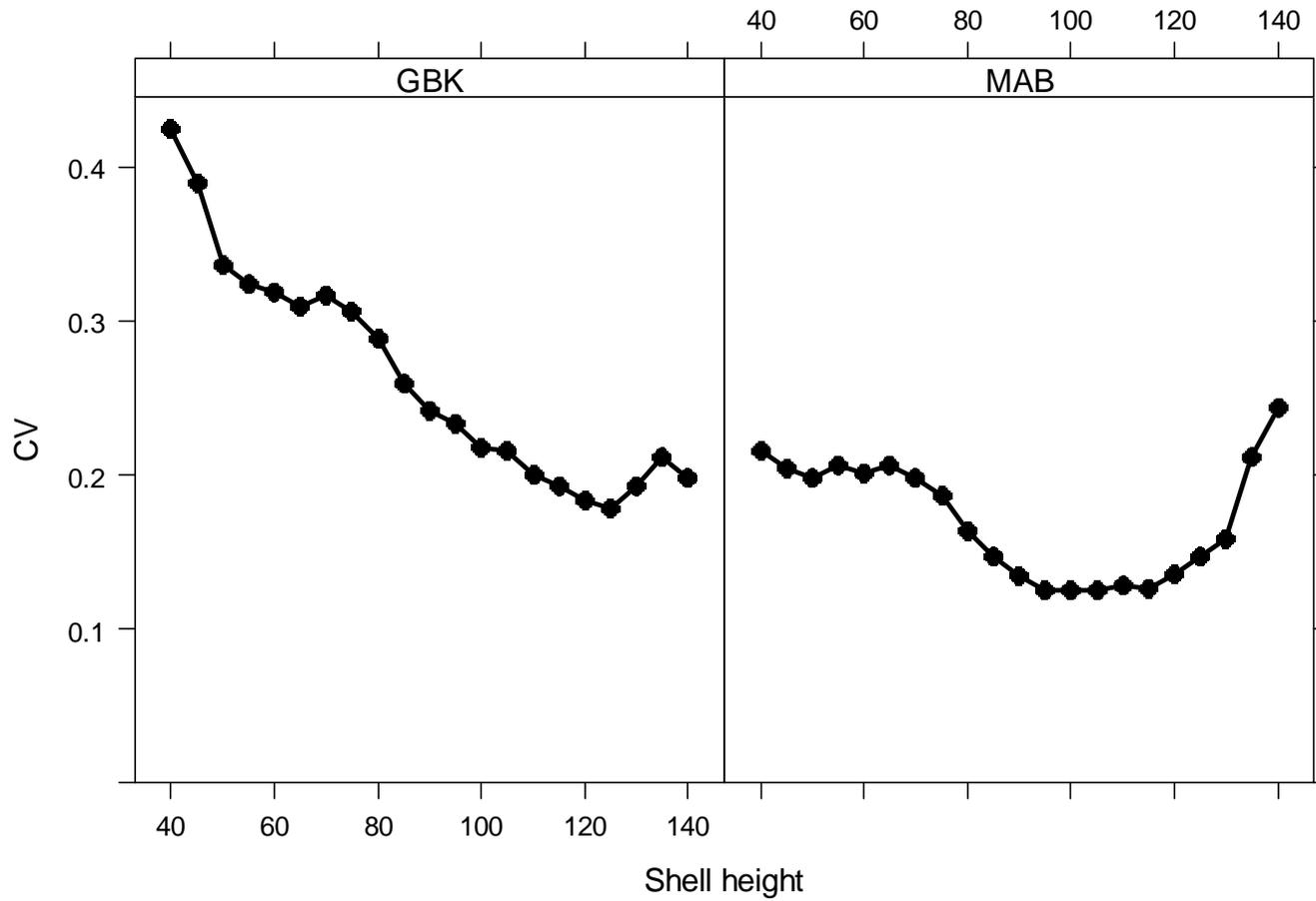


Figure 2. Average CVs for mean scallop catch per tow in the dredge survey during 1978-2013 by shell height size group and stock area.

MAB empirical population abundance and 95% CI (y-axis varies)

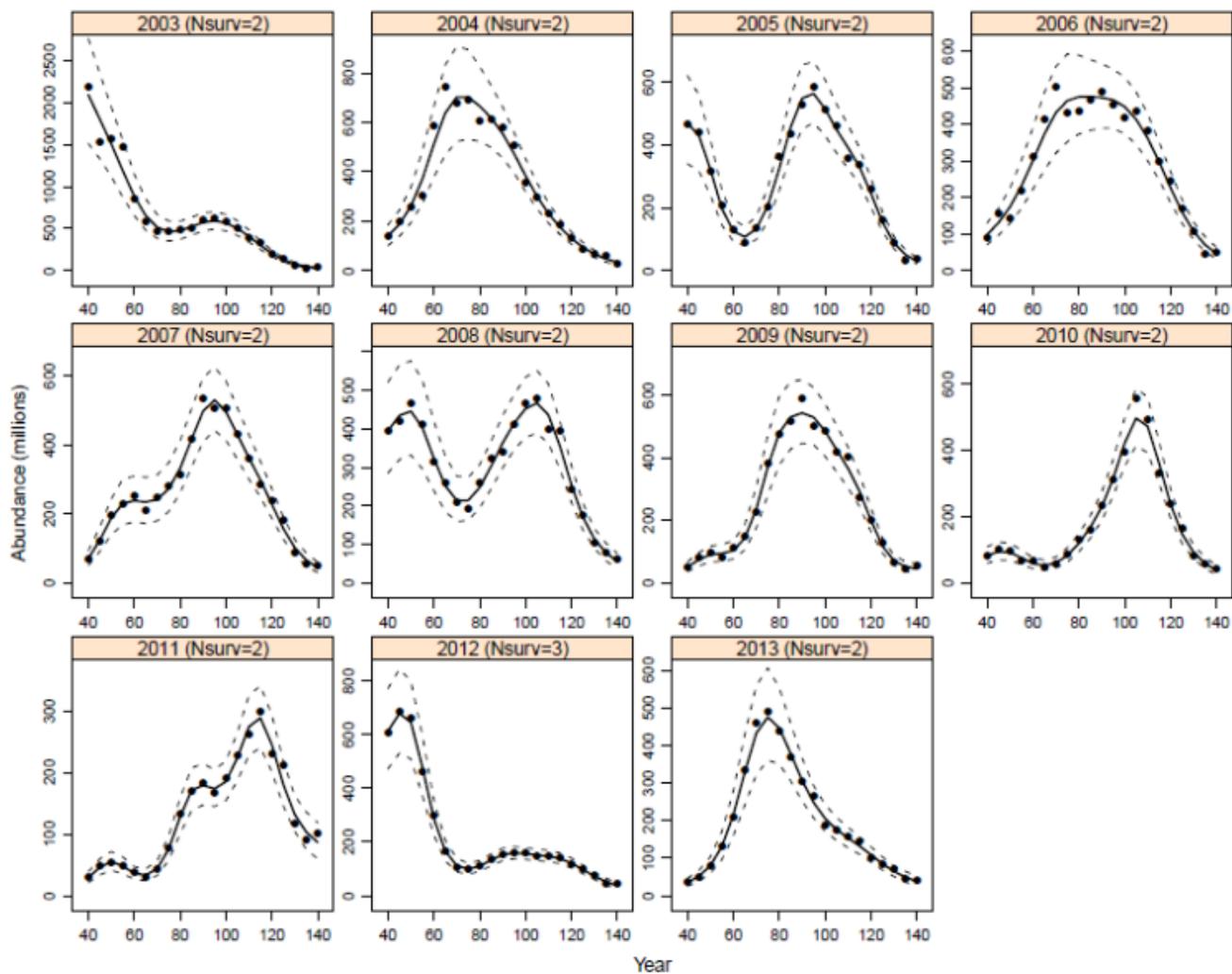


Figure 3a. Empirical abundance at length during 2003-2013 in the Mid-Atlantic region with approximate 95% confidence intervals. Note that the scales on the y-axis vary.

GBK empirical population abundance and 95% CI (y-axis varies)

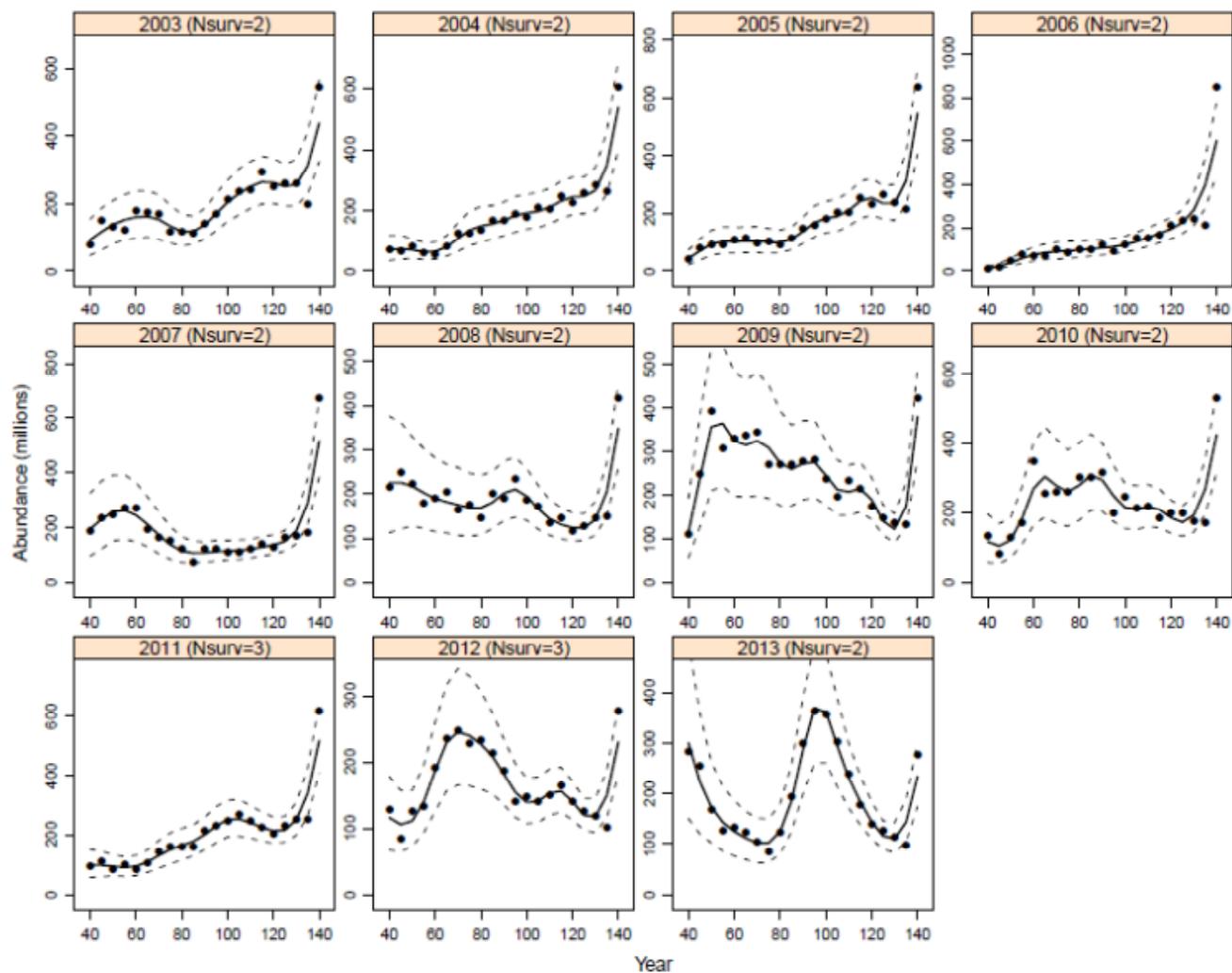


Figure 3b. Empirical abundance at length during 2003-2013 in the Georges Bank region with approximate 95% confidence intervals. Note that the scales on the y-axis vary.

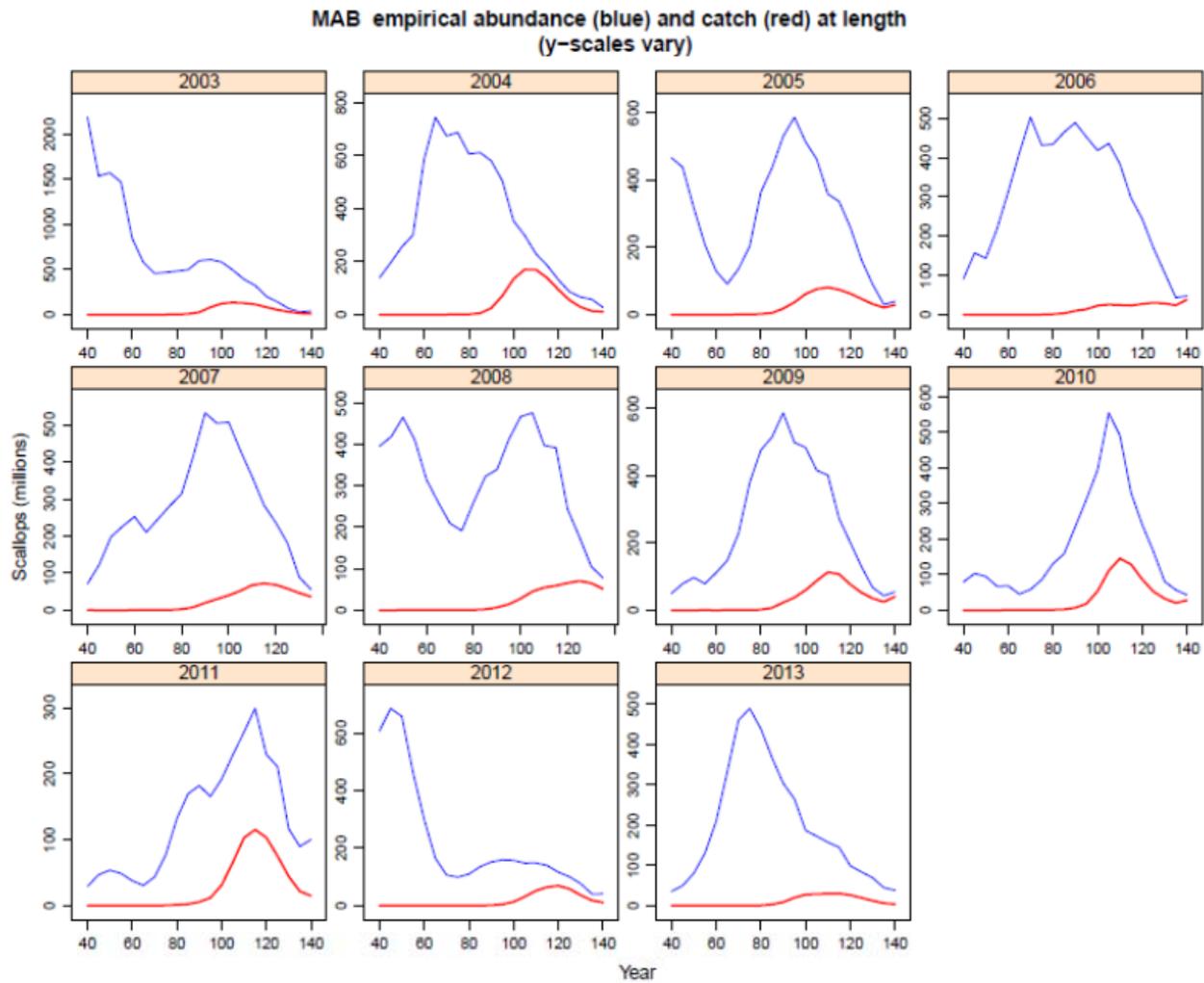


Figure 4a. Empirical abundance and catch at length during 2003-2013 in the Mid-Atlantic region. Note that the scales on the y-axis vary.

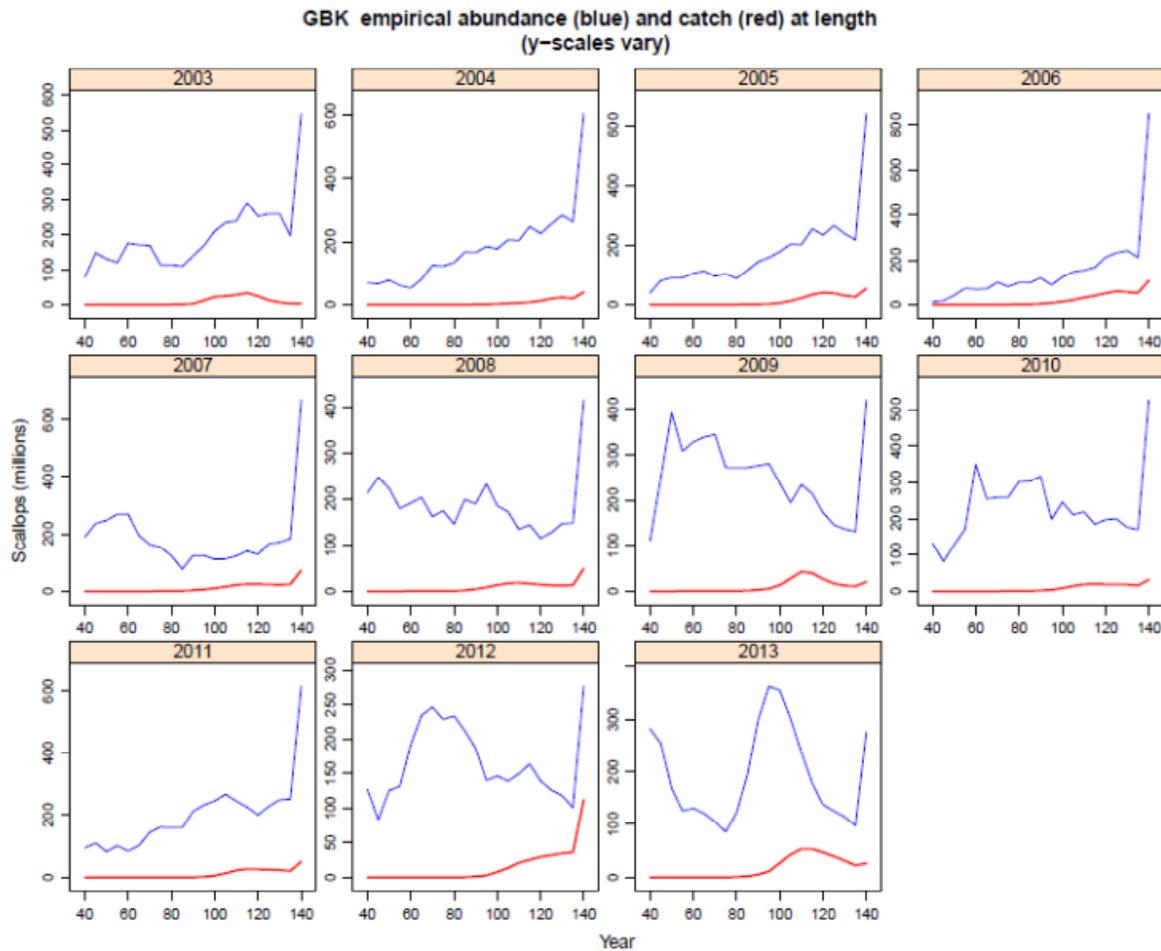


Figure 4b. Empirical abundance and catch at length during 2003-2013 in the Georges Bank region. Note that the scales on the y-axis vary.

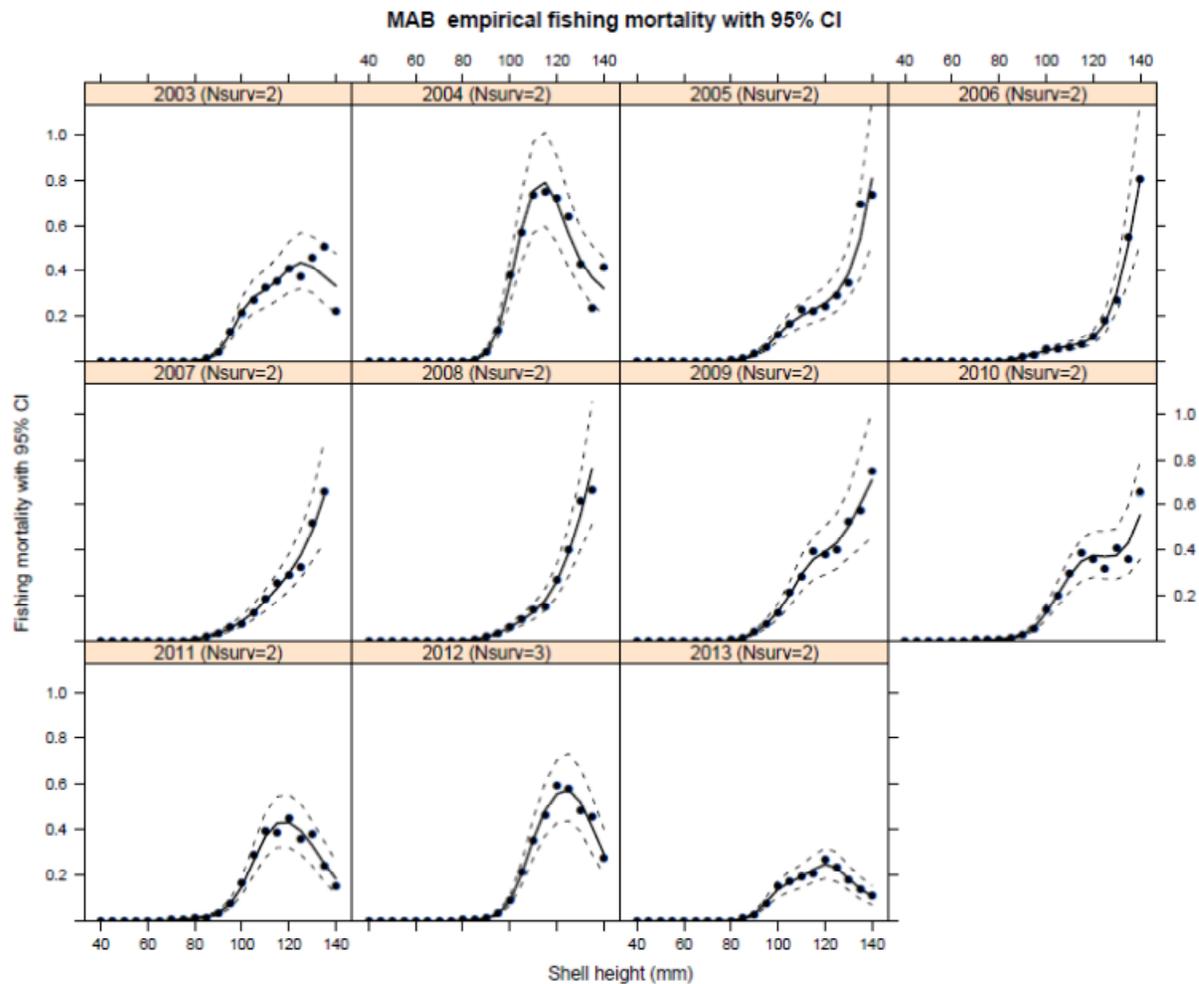


Figure 5a. Empirical fishing mortality at length during 2003-2013 in the Mid-Atlantic region with approximate 95% confidence intervals. Note that the scales on the y-axis differ (fishing mortality was typically higher in the Mid-Atlantic region).

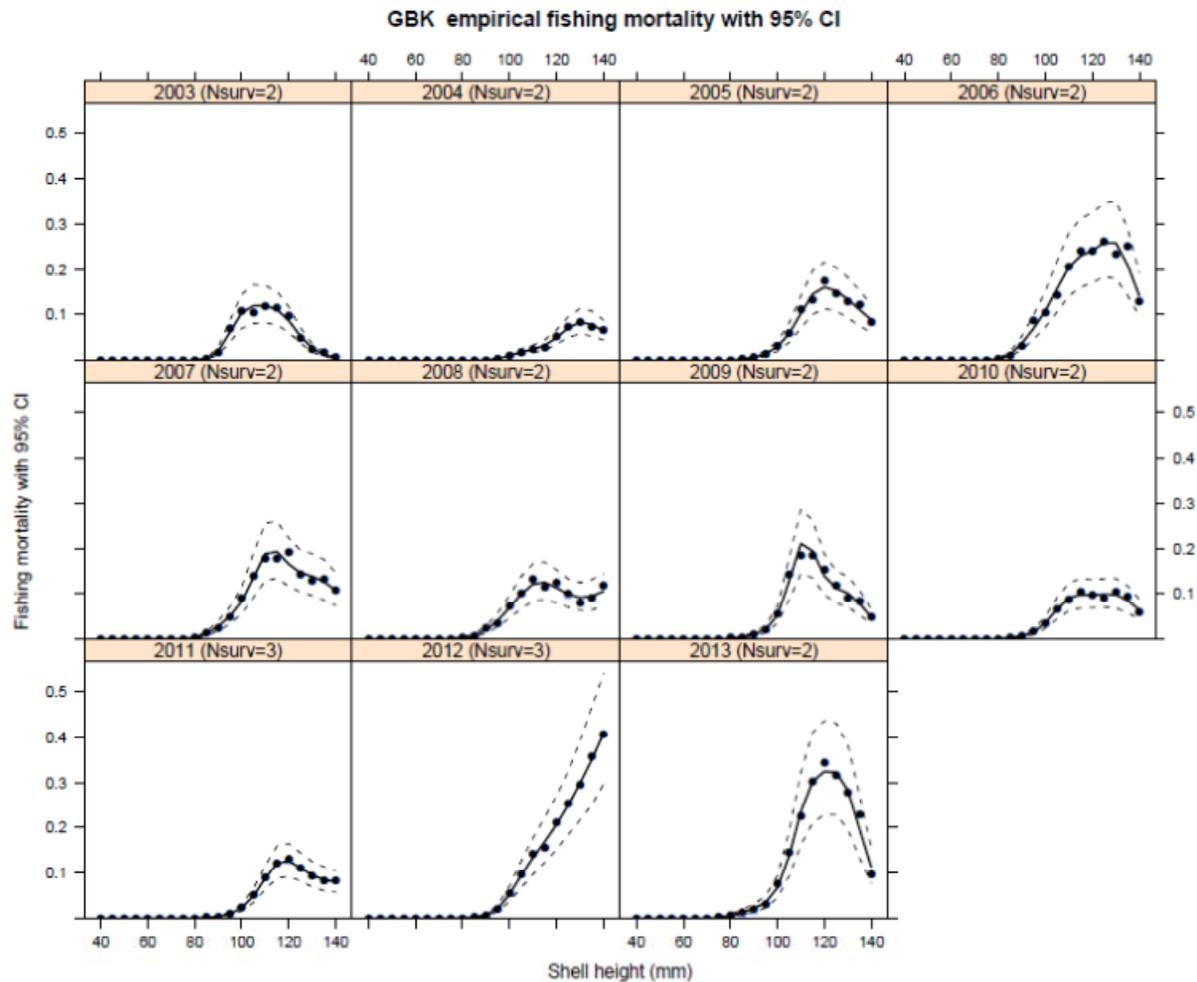


Figure 5b. Empirical fishing mortality at length during 2003-2013 in the Georges Bank region with approximate 95% confidence intervals. Note that the scales on the y-axis differ (fishing mortality was typically higher in the Mid-Atlantic region).

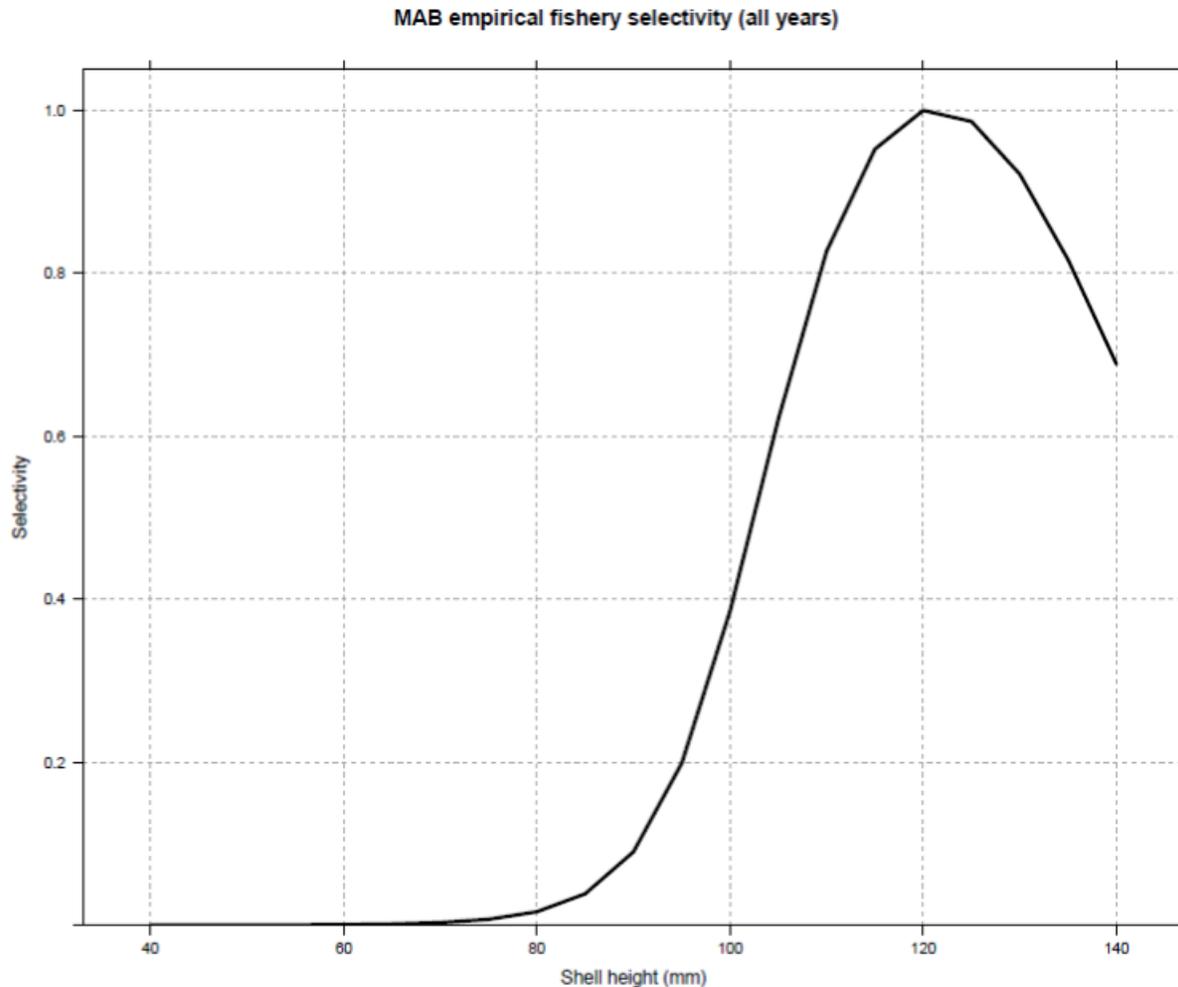


Figure 6a. Empirical estimates of average size selectivity for the scallop fishery during 2003-2013 in the Mid-Atlantic region. This curve was calculated by pooling data for different years and fitting a single line to show the trend. Another approach is to average the fitted selectivity curves for each year.

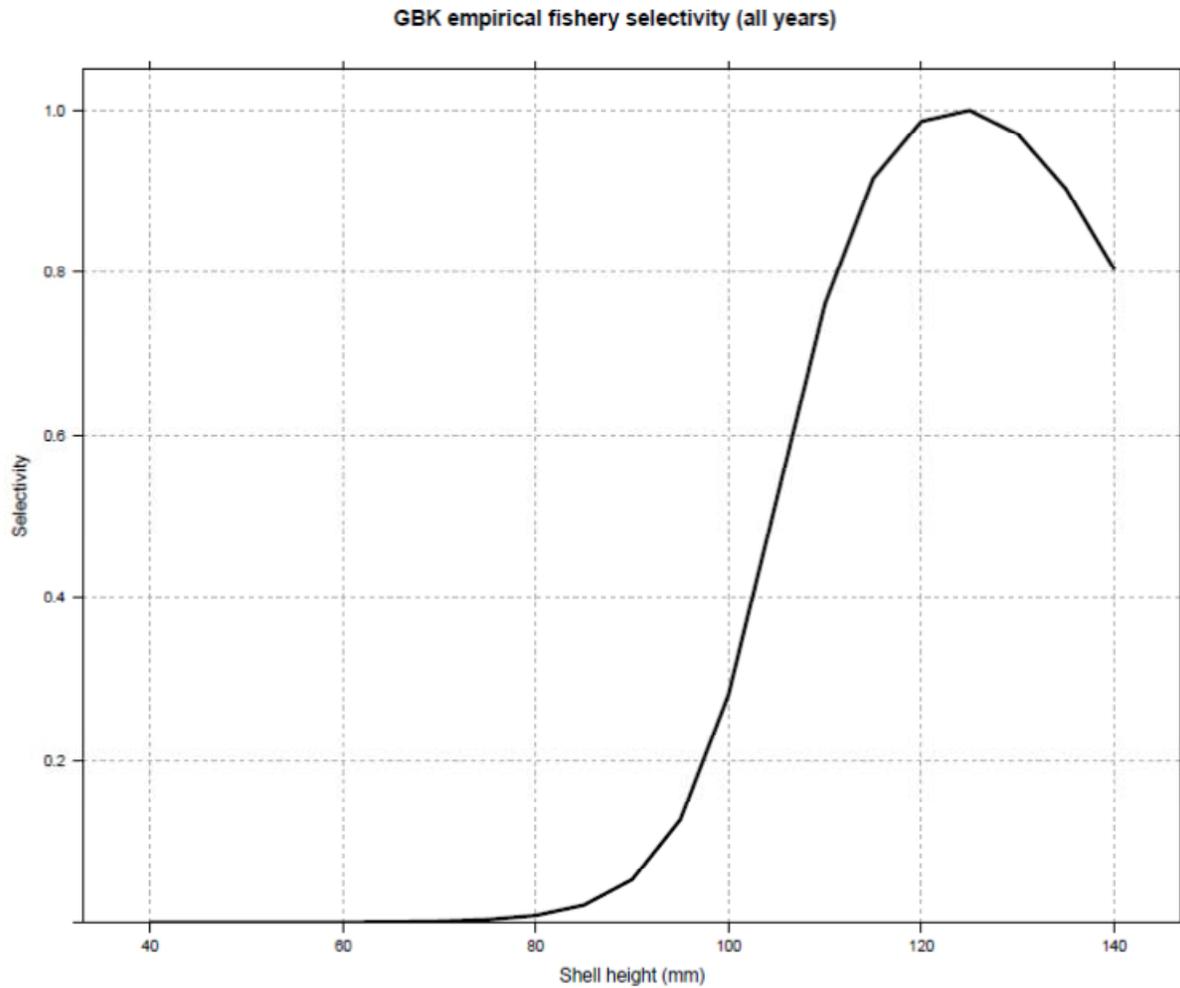


Figure 6b. Empirical estimates of average size selectivity for the scallop fishery during 2003-2013 in the Georges Bank region. This curve was calculated by pooling data for different years and fitting a single line to show the trend. Another approach is to average the fitted selectivity curves for each year.

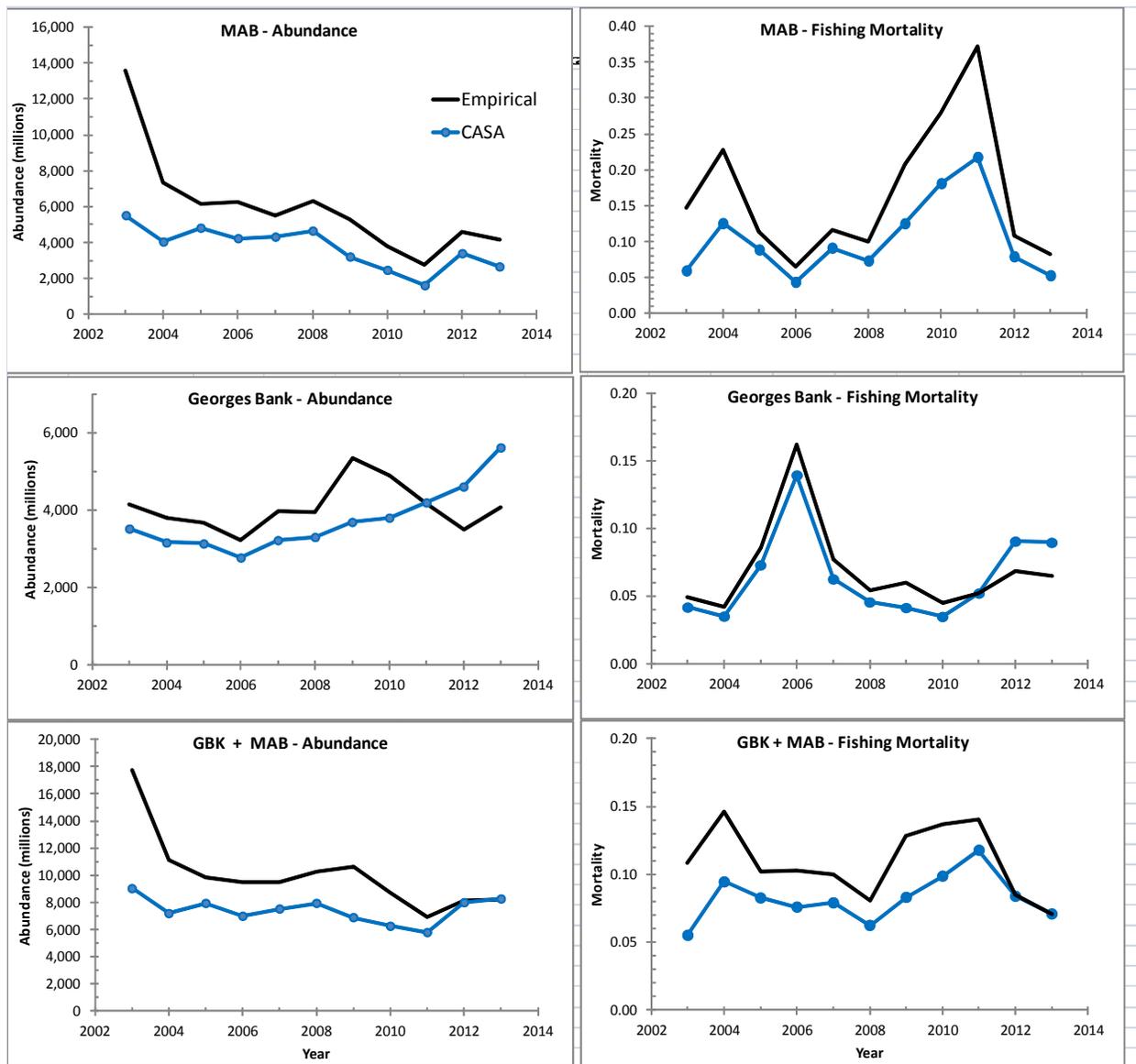


Figure 7. Abundance (left) and fishing mortality estimates (right) from the empirical method and the CASA model during 2003-2013 for the Georges Bank (top), Mid-Atlantic (middle) and combined (bottom) regions.

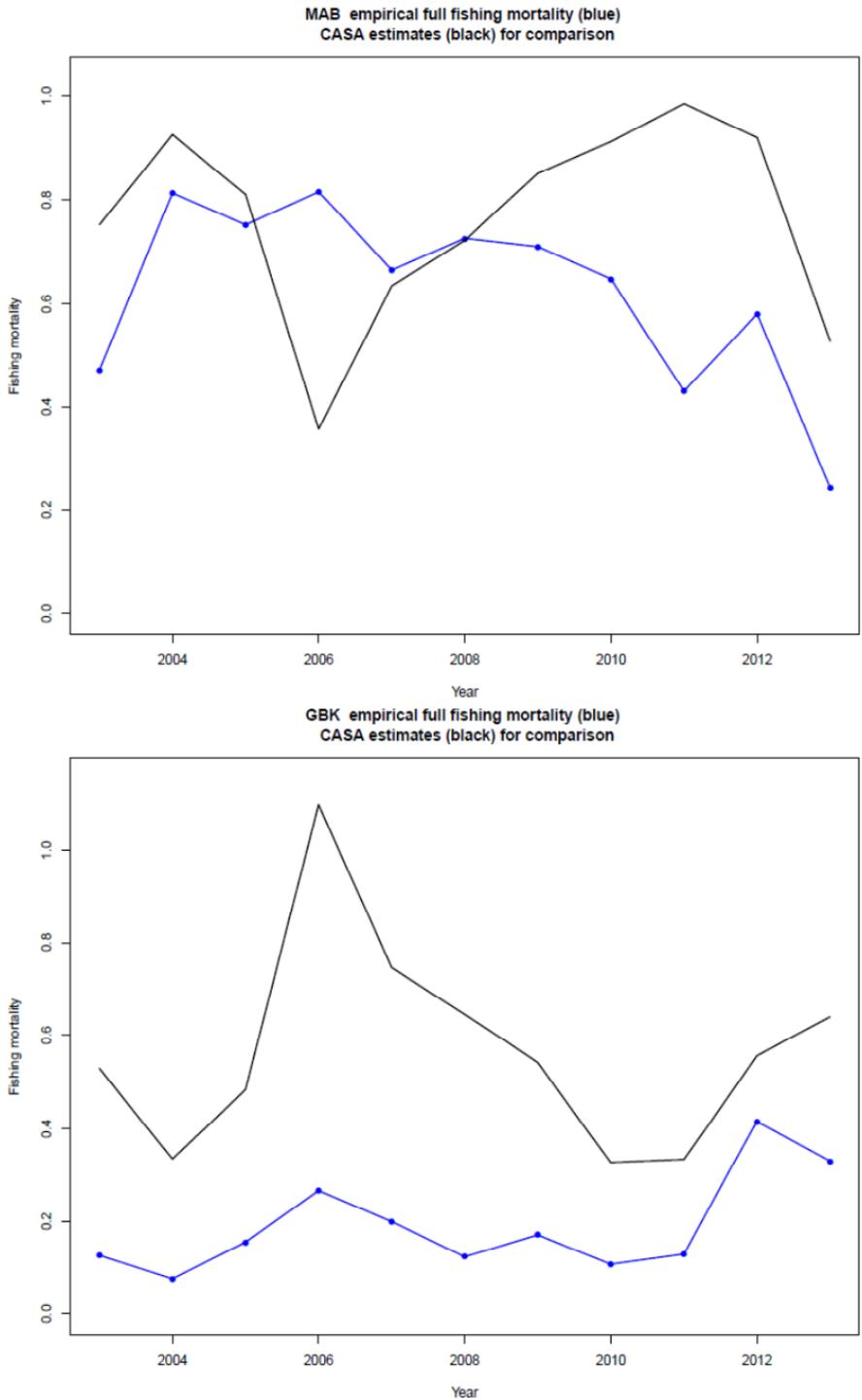


Figure 8. Fully recruited fishing mortality estimates for the Mid-Atlantic (top) and Georges Bank (bottom) regions. The empirical estimates are in blue, CASA estimates are black.

## **Appendix B6. NEFSC HabCam survey for sea scallops: survey design, implementation, and data analysis.**

Jui-Han Chang, Burton Shank, and Dvora Hart. Northeast Fisheries Science Center, Woods Hole MA.<sup>4</sup>

This report contains five stand-alone sections that together describe HabCam gear and operations, simulation work used to develop and test survey designs, how the actual surveys during 2011-2013 were carried out, and how abundance estimates and size composition for 2011-2013 used in assessment models were made.

### **1. Introduction to the HabCam survey**

HabCam is an underwater vehicle that was originally developed through collaboration of commercial fishermen, independent scientists, and staff at the Woods Hole Oceanographic Institute as a vehicle for documenting the size and abundance of benthic / demersal organisms and mapping sea floor habitats. The vehicle is towed behind a vessel, while actively “flown” ~2m off the bottom by a pilot. It collects overlapping, downward facing digital still imagery. Between 2005 and 2010, the HabCam group developed and improved this technology and successfully performed a number of surveys on the Mid-Atlantic continental shelf, Georges Bank and in the Gulf of Maine, primarily using the HabCam\_V2 vehicle which preceded the current design. The development of the vehicle and many of these surveys were supported by the Sea Scallop Research Set-Aside program and the vehicle proved to be appropriate technology for assessing sea scallops (NEFSC 2010). In 2009 a paired HabCam / dredge experiment was conducted to determine the capture efficiency of the NEFSC survey dredge (probability of catching a scallop in the path of the dredge) and in 2011 the HabCam\_V2 was used in the NEFSC scallop survey to get an estimate of the entire scallop resource on Georges Bank.

With an interest in making a HabCam-type survey a standard part of the sea scallop assessment survey, NEFSC secured funds from NOAA Office of Science and Technology and contracted WHOI to build a vehicle for NEFSC’s use. This vehicle, (HabCam\_V4 or NOAA HabCam), completed resource-wide surveys in 2012 and 2013, beginning a new assessment time series for sea scallops that is used for the first time in this assessment. The HabCam\_V4 vehicle is equipped with stereo digital still cameras, altimeters, and a compliment of oceanographic sensors including temperature, salinity, water spectrometer, 3D side-scan sonar, and optical sensors for dissolved oxygen, cdom, and turbidity.

### **2. Survey Design**

Because the HabCam vehicle collects a constant track of images, data derived from the images are autocorrelated and not appropriate for analysis as a random or stratified survey. Resource assessments from such data are typically use spatial models including Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs) or geostatistical methods

---

<sup>4</sup> First and second coauthors alphabetical.

such as kriging (Rivoird et al. 2008). Literature on sampling designs for this type of survey comes primarily from literature on acoustic surveys. With geostatistical methods, the uncertainty in the estimate at any given location increases with distance from the survey track. As a result, evenly spaced grids are optimal for acoustic surveys as the distance from the survey track is minimized with even spacing. A second common survey design is a two-stage approach where a low resolution survey is first performed to determine the location of high-density aggregations and a second high-resolution survey is conducted on the aggregations. The results from two-stage surveys are post-stratified to account for spatial heterogeneity in survey effort. In both cases, geostatistical methods assume that the mean and variance is homogeneous throughout a survey stratum.

The HabCam sea scallop survey differs from these situations because adult sea scallops are relatively easy to detect and intensively surveyed, do not move long distances, and because spatial heterogeneity is primarily driven by management measures and known habitat affinities. While geostatistical methods assume a landscape with a stationary mean (Figure 1a), a landscape with a higher mean along the center of the landscape (Figure 1b) is more realistic for sea scallops because densities typically decrease in habitats deeper and shallower than the optimal habitat for a region (Figure 1c). In this case, it may be advantageous to increase sampling effort in the core habitats along the center of the survey area. Given a survey track of evenly-spaced transects of equal length (Figure 2a) and assuming an underlying variogram model, we can derive a map of kriging variances for the survey at each location in the landscape (Figure 2b). If the mean density is higher in the center of the landscape instead of stationary, we may assume that the standard deviation of the mean is proportional to the mean (similar to a Gamma distribution) and calculate an adjusted kriging variance for each location as:

$$\text{AdjVar}_{x,y} = \text{KrVar}_{x,y} * [e^{(\text{CE})}]^2 \quad (2.1)$$

Where  $\text{AdjVar}_{x,y}$  is the adjusted variance of the estimate at a given location,  $\text{KrVar}_{x,y}$  is the unadjusted variance at the location and  $e^{(\text{CE})}$  is the magnitude of the center effect from Figure 1c.

As a proof of concept, we used geostatistical simulation to examine the effect of allowing the mean (and variance) to vary across the shelf and longitudinally along the shelf. We first simulated varying the mean across the shelf and examined how the survey variances were affected by varying (1) the proportion of the effort concentrated along the center of the survey area and (2) the length of the survey track. We modeled the cross-shelf gradient as a double-logistic with higher densities along the center of the study area and the amplitude of the center effect varying from 0 (no effect) to 1 (variance is  $e^2$  or 7.38 times higher along the center of the study area (Figure 3). To assess the effect of increasing sampling intensity along the center of the study area, we decreased the length of alternating transects (range from 0 – 100% of the total width of the study area) and increased the total number of transects to keep the total survey track length constant (Figure 4). We then varied the total survey track length from 1,000 to 4,000 pixels. For each simulation, we examined the resulting variance maps (i.e. Figure 5) and used the sum of the adjusted kriging variance (eq. 2.1) as a relative proxy for the variance of the survey. While this is not the true variance of the survey, as the variances are correlated across the landscape, we are not aware of established methods for calculating a kriging variance for survey areas with non-stationary variances and this should be an effective relative measure for

comparison purposes.

The adjusted kriging variances varied across center effects and transect lengths (Figure 6). Optimal short transect lengths decreased as center effects increased and increased as total track length increased. The center effect and total track length interacted to produce an optimal short transect length. With a track length of 1,000 pixels, increasing the center effects from 0 and 1 decreased optimal short transect length from 67% to 30%. However, for track lengths of 4,000 pixels, varying the center effect from 0 to 1 only decreased optimal short transect length from 92% to 85%.

For a second simulation, we examined the effect of the mean and variance varying longitudinally along the survey area (i.e. zonal anisotropy, Figure 7). The zonal effect was implemented by dividing the landscape into two zones (upper and lower) and adding an additional, longitudinally-oriented logistic trend to the landscape. We then varied the amplitude of the longitudinal effect (Zone effect) the spacing of adjacent transects between the two zones, and total track length (Figure 8).

The optimal solutions for landscapes with Zone effects placed more transects in the zone with higher underlying means and variances (Figure 9). The effect was most notable for shorter total track lengths, increasing transect density in the higher mean zone by as much as 300% over the lower mean zone.

### **3. Survey Area and Design for Actual HabCam surveys**

The above simulations indicate that the variance of a survey can be decreased by alternating the length of survey transects and increasing transect density in areas with known higher abundances. These simulation results are used informally in the design of each year's survey but actual survey design is based on researchers' knowledge of where the current stock biomass and incoming cohorts are.

The two stock areas (MAB and GB) are each divided into multiple subregions, based on changes in habitat type, habitat orientation (anisotropy), and management boundaries (Figure 10 and 11). These subregions are used both for designing the survey and for abundance estimation from the resulting survey data.

The extent of the survey area is based on an updated analysis of biomass patterns from the NMFS dredge and RSA surveys, Vessel Trip Reports, sea scallop observer trips, and Vessel Monitoring System data. In general, the current extent of the dredge survey was found to be very adequate for covering the scallop resource, though small areas were added to the extent of the HabCam survey to capture areas where there was evidence of adequate scallop densities or commercial activity.

The survey tracks are constructed in one long track for the MAB and three separate tracks for GB. Each track is bounded by a set of subregions. A midline, drawn along the center of biomass, runs through each set of subregions. Survey transects are centered around and oriented orthogonal to the midline.

### 3.1 Software and procedures used in designing HabCam surveys

In designing actual surveys, specialized software prompts a user to enter the total effort (survey days) to allocate to a track, the relative lengths of the short transects on the track, and the transect density offset for each subregion along the track. The software varies the relative transect densities and provides a number of alternative tracks of similar lengths for the user to choose among, based on appropriate allocation of effort across the subregions, how well each track works around complex bathymetric structures, and other logistical considerations.

## **4. Image Acquisition, Processing and Annotation.**

The HabCam vehicle is towed along the survey track at speeds from 6 – 7 knots while a pilot maintains the unit at an altitude of ~2m off the bottom. Digital still images are generally collected at a sufficiently high frequency that ~35% of adjacent images overlap. Collected images are initially stored as raw TIFF-formatted images. The raw TIFFs are then light-field and color corrected to improve image quality and saved in processed PNG format. Each image is named with a unique identifier and metadata for each image is recorded including longitude, latitude, time, vehicle depth, bottom depth, and vehicle altitude, roll, and pitch as well as the data from the oceanographic sensors. The altitude of each image is critical for determining the field of view of the image and measuring objects in the images. As altitude can be measured in multiple ways, the value used for a particular image is based on the following list ordered by expected accuracy:

1. Altitude as measured via disparity mapping (parallax) from the stereo images
2. Altitude as measured by the altimeters on the vehicle
3. Altitude inferred from the side-scan sonar

The metadata associated with each image is then stored in a PostgreSQL database and used for selecting images for annotation.

We select blocks of images for annotation, termed “assignments”, based on the spatial extent of the image set and a target image density. Based on the desired density of images to be annotated, we break the survey track into equal length segments and select one image from each segment. Individual image selection is biased towards preferred vehicle heights (Gaussian-weighted, based on known issues with water turbidity or other factors that affect image quality) but image selection is otherwise random within each segment.

The selected image list is uploaded to the Postgres database for direct observation and annotation using a web-based annotation tool. Additional assignments may be created once an assignment is completed if additional images are desired from the same region. In such cases, we first remove all images from a buffered region around each image that has already been annotated from the pool of available images before the next random subset of images is selected. The goal of this is to keep the density of annotated images consistent within subregions along the track.

Data on the abundance, size and behavior of scallops are extracted from each image using an online annotation tool developed by collaborators at WHOI (Figure 12). Only scallops where

the center of the scallop is judged to be inside the image are enumerated. Scallops larger than about 35mm (age 2+) are measured by drawing a line over the shell while smaller scallops are only marked with a point and counted. Additional data are recorded including confidence in identification, swimming, dead, clappers, etc. Image quality may be poor due to turbid waters, extremely high or low altitudes, image corruption, or other objects obscuring the bottom. In this case, the image can be noted as poor quality and data from this image excluded from derived data sets. All annotations, as well as comments on image quality and sediment types, are recorded directly to the Postgres database by the annotation tool.

Because scallops are not always oriented normal to the camera or may be partially obscured, scallops measurements are either shell heights (umbo to opposite margin) or widths (lateral margins), whichever is judged to be more accurate. Shell widths are converted to shell heights using a statistical model derived from paired measurements of scallops that were well oriented to the camera:

$$\text{Shell\_height} = 3.538 + 1.034 * (\text{Shell\_width}) - 0.0003502 * (\text{Shell\_width})^2 \quad (4.1)$$

Shell height is calculated in pixels based on the start and end coordinates of the annotated line. The size of each pixel in an image is calculated from the altitude of the associated image, based on tank calibration experiments, and this pixel size is used to convert the shell height to actual millimeters. The altitude is also used to calculate the field of view for each image for density calculations.

For estimating size frequency distributions and abundance for each year, we constructed standardized data sets from the database and posted them to a common location on a network drive. The annual data sets include data from both the NEFSC HabCam surveys and from the HabCam group RSA surveys, which have to be drawn from multiple databases and corrected individually for problems in altitude measurements or other issues. The data sets include the metadata from all annotated images of acceptable quality, plus the classification of all scallops observed in each image and calculated lengths of for any scallop measured with a line segment.

## **5. Model-based estimation of sea scallop abundance and biomass**

### **5.1 Introduction and summary**

The goal of this section is to assess different model-based methods for estimating total abundance and biomass from HabCam and then apply these methods to HabCam data for 2011 - 2013 data to estimate abundance, biomass and size composition of sea scallops in the Georges Bank (GB) and Mid-Atlantic Bight (MAB) assessment regions (Figures 14 and 15). We also present design-based method (stratified mean) for this data set as an alternative to model-based methods and use it to validate the model-based estimates and CV's.

Scallop abundance or biomass data from HabCam are highly spatially autocorrelated and zero inflated, reflecting the patchiness of scallop distributions and the continuous nature of the observations. Thus, model-based estimation methods might be required to extrapolate observations along the observed track to larger areas. We used 2013 HabCam biomass data to

test 3 geostatistical models: (1) ordinary kriging on spatially averaged data (OK), (2) zero-inflated Generalized Additive Models on spatially averaged data with kriged model residuals (GAM+OK), and (3) zero-inflated Generalized Additive Mixed Models where small scale variations are treated as random effects, combined with kriged model residuals (GAMM+OK). Effects of scale (neighborhood) size to average the data or scale of random effects was also evaluated. Co-located survey data from other gear types (dredge surveys from NEFSC and VIMS and video surveys from SMAST) were used for model validation. No single modeling approach and scale was consistently superior but GAM+OK performed better than OK and GAMM+OK in general.

We then conducted a simulation to evaluate performance of the 3 model-based methods along with a design-based method (stratified mean method, SM) and effects of scale size for data averaging and random effects. The GAM+OK method with small scale size outperformed the other 2 model-based methods and scale sizes in the simulation in terms of accuracy and precision of estimating mean and CV in most cases. SM estimates were more accurate and precise than the model-based estimates but only when the study region was stratified more correctly than might be expected in practice.

Based on the results of 2013 HabCam biomass data analysis and simulations, we selected the GAM+OK method to estimate scallop abundance and biomass for the GB and MAB stock for 2011 to 2013. SM estimates estimated with careful stratifications are also provided to back up the model-based estimates. Following are detailed descriptions of the simulation design, model- and design-based methods, simulation results, and procedures to estimate GB and MAB scallop abundance and biomass for 2011 to 2013.

## 5.2 Simulation Design

The area covered (domain) of simulated scallop populations was 50 km longitude and 100 km latitude (roughly the size of Hudson Canyon subregion, Figure 2) with a 100 m grid size. The scallop spatial distributions are non-stationary due to the influences of physical and biological environment including current, depth, and predator distributions (Brand, 1991). The simulated scallop population is therefore assumed to be heterogeneous in global trend (first-order effect), combined with stationary second-order effects. We simulated different first-order and second-order effects in order to test whether the abundance and biomass estimation methods are robust to the type of spatial distributions of the underlying population.

Variations in global mean quantity were simulated using a double logistic function

$$p_{i,j} = \frac{1}{1+\exp(-a(i-b))} + \frac{1}{1+\exp(a(i-b+\frac{\max(i)}{2}))}, \quad (5.1)$$

where  $a$  and  $b$  parameters determine the shape of the logistic curve, and  $i$  and  $j$  are the longitude and latitude, respectively. The simulated first-order effects are high in the middle and decrease logistically toward the left and right edge of the simulation domain (Figure 16). Two types of first-order effects were simulated, one narrow but highly dense and the other wide and less dense (Figure 16).

Second-order effects were simulated as stationary Gaussian random fields with a spherical isotropic covariance structure (Cressie 1993)

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ c_0 + c_1 \left\{ \frac{3h}{2r} - \frac{1}{2} \left( \frac{h}{r} \right)^3 \right\} & 0 < h \leq r, \\ c_0 + c_1 & h \geq r \end{cases} \quad (5.2)$$

where  $c_0$ ,  $c_1$ , and  $r$  are the nugget, partial sill, and range parameter, respectively. The nugget/sill ratio ( $\frac{c_0}{c_0+c_1}$ ) determines randomness and  $r$  determines aggregations size of the second-order effects. We simulated combinations of 2 levels of nugget/sill ration and 2 levels of the range parameter resulting in 4 types of second-order effects: small aggregation, large aggregation, small aggregation with a large random noise, and large aggregation with a large random noise (Figure 17). We chose the parameter values based estimates from actual HabCam data.

Scallop distributions are patchy, resulting in HabCam data being highly zero-inflated (Table 1). To reflect the patchiness of scallop distribution, for each second-order realization, densities smaller than 90th percentile were set to zero. The zero-inflated second-order effects were combined with first-order effects to produce realistic simulated scallop distributions (Figure 18).

We simulated combinations of 2 first-order and 4 second-order effects resulting in 8 types of simulated population distributions. Thirty realizations were generated for each population type. Total abundance and biomass of each realization was scaled to equality across realizations. Each realization was surveyed using 30 different tracks. Shape and direction of tracks was designed to mimic the actual HabCam survey design.

Model-based and designed-based methods were used to estimate total biomass and abundance for the simulated populations. These estimation methods were evaluated using percent bias and percent root mean square error (RMSE)

$$\% \text{ Bias} = \frac{\sum_{i=1}^n (\hat{T}_i - \mu)}{n \mu} \quad (5.3)$$

$$\% \text{ RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{T}_i - \mu)^2}{n \mu^2}}, \quad (5.4)$$

where  $\hat{T}_i$  is the estimated total biomass or abundance for sample set  $i$ ,  $\mu$  is the true population size, and  $n$  is the total number of sample sets analyzed. Percent bias and percent RMSE of CVs for the precision of model estimates were also evaluated. The method that produced the least biased and most precise estimates was selected to analyze the actual HabCam data.

### 5.3 Model-Based Estimation

Kriging is one of the most widely used geostatistical method for spatial interpolation (Webster and Oliver 2001). We tested performance of 3 different kriging methods including

OK, GAM+OK, and GAMM+OK on the simulated scallop populations. OK is a standard version of the kriging models with the assumption of a constant mean and consideration of variation and distance between sample points (Hengl 2009, Webster and Oliver 2001). Although the constant mean assumption might not be reasonable for scallops, the simulation tests are necessary to determine whether the observed non-stationary pattern can be modeled as an autocorrelation among errors with a constant mean or a trend with mean changing with variance.

Isotropy and anisotropy is the variation of scallop abundance or biomass being identical or directionally dependent. It is not clear whether the samples are isotropic or anisotropic although actual observations indicate that first-order effects the simulated populations should have the largest variations along the horizontal axis. Therefore, we built both the isotropic and anisotropic models and selected the final OK model using RMSE

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{z}_i - z_i)^2}{n}} \quad (5.5)$$

Total abundance or biomass ( $T$ ) and its variance were estimated as

$$\hat{T} = A \sum_{i=1}^n \hat{z}_i \quad (5.6)$$

$$\text{Var}(\hat{T}) = A^2 \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\hat{z}_i, \hat{z}_j), \quad (5.7)$$

where  $\hat{z}_i$  is the kriging estimates at location  $i$  and  $A$  is the grid size.

Regression kriging (RK) extends the OK to account for a global trend, which can be estimated by apply a regression model (e.g. GAM or GLM) to a series of ancillary variables (e.g. depth, latitude or longitude) then applying OK to the residuals of the regression model (Hengl 2009, Odeh et al. 1995). The final predictions of RK are obtained by summing the regression predicted values and the kriged residuals. This approach was criticized by Cressie (1993) and Lark et al. (2006) because the variogram estimates of the random component of spatial variation are theoretically biased. Generalized least squares and residual maximum likelihood-empirical best linear unbiased predictor are two potential solutions (Lark et al. 2006). However, Kitanidis (1993) and Minasny and McBratney (2007) showed that although these methods are theoretically preferable to RK, they did not substantially improve model predictions. We therefore used the RK approach.

Scallop data from the HabCam survey are highly spatially autocorrelated and zero inflated, reflecting the patchiness of scallop distributions. Therefore, we estimated the first order effects (over relatively large geographic areas) using a two-stage hurdle model which models the probability that scallops are found in a sample (presence/absence) separately from the density given that at least one scallop was found (Barry and Welsh, 2002). Predictions from the two models are combined to make the complete estimates of abundance and biomass. Hurdle model results were usually modified further to account for second order effects over smaller geographic areas as described below. We tested a hurdle GAM on data averaged within segments along the tracks (to reduce the autocorrelation and zero-inflation) and a hurdle GAMM where the fine-scale variations within track segments were treated as random effects. A quasi-binomial distribution was assumed for the presence/absence model and a quasi-Poisson distribution for the

positive model. The first-order effects were estimated using an interaction term of latitude and longitude for both GAM and GAMM. OK was performed on the residuals using the same algorithm described above. Total abundance and biomass of GAM+OK and GAMM+OK model estimates were estimated using

$$\hat{T} = A \sum_{i=1}^n \hat{x}_i \hat{y}_i + \hat{z}_i, \quad (5.8)$$

where  $\hat{x}_i$  is the probability of presense estimate,  $y_i$  is the positive estimate,  $\hat{z}_i$  is the kriged residual at location  $i$ . By assuming that  $\hat{x}$  and  $\hat{y}$  are independent, the variance of the  $\hat{T}$  was calculated using

$$\text{Var}(\hat{T}) = A^2 (\sum_{i=1}^n E^2 \hat{x}_i \text{Var}(\hat{y}_i) + E^2 \hat{y}_i \text{Var}(\hat{x}_i) + \text{Var}(\hat{x}_i) \text{Var}(\hat{y}_i) + \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\hat{x}_i, \hat{y}_i)) \quad (5.9)$$

Effects of segment length to average the data or determine random effects along the tracks was evaluated. The dense scallop aggregations occurred at approximately 400 to 900 m (NESFC 2010) and therefore we tested 3 segment lengths, 750, 1500, and 2,250 m. These segment lengths were also used to define the grid size A.

#### 5.4 Design-Based Estimation

We tested a SM method to estimate total abundance and biomass from the simulated data. Only horizontal transects were used in the SM estimation because variance of these transects were different from the vertical transects. Horizontal transects were post-stratified into 2 strata based on high and low first-order effects (Figure 19). Mean and its variance of the simulated scallops ( $t$ ) by segment ( $j$ ) and stratum ( $i$ ) were calculated by

$$\bar{t}_{i,j} = \frac{\sum_{k=1}^{n_{i,j}} t_{i,j,k}}{n_{i,j}} \quad (5.10)$$

$$\text{Var}(\bar{t}_{i,j}) = \frac{\text{Var}(t_{i,j,k})}{n_{i,j}}, \quad (5.11)$$

where  $n_{i,j}$  is the number of images by segment and stratum. Total abundance and biomass estimates ( $\hat{T}$ ) and variance were estimated as

$$\hat{T} = A \sum_{i=1}^2 S_i \frac{\sum_{j=1}^{n_i} \bar{t}_{i,j}}{n_i} \quad (5.12)$$

$$\text{Var}(\hat{T}) = A^2 \sum_{i=1}^2 S_i^2 \sum_{j=1}^{n_i} \frac{\text{Var}(\bar{t}_{i,j})}{n_i^2}, \quad (5.13)$$

where  $n_i$  is the number of segments by stratum  $i$ , and  $S_i$  is the size of stratum  $i$ .

The simulation domain was well-stratified based on the first-order trend; however, we do not have the same information when dealing with the real data which tend to complicated as

shown below using real data. We tested whether the SM estimates are sensitive to the stratification by enlarging (Stratified Mean Wide, SMW) and shrinking (Stratified Mean Narrow, SMN) the central high density stratum by 20% (Figure 19) and estimated total abundance and biomass under the original (incorrect) assumptions about stratum size.

### 5.5 Simulation Results

Proportion of converged model runs was 99% for GAM+OK and OK but 80-93% for GAMM+OK (Table 2). Percent bias and percent RMSE showed that GAM+OK with data averaged by 750 m (scale) is the best way to estimate the scallop biomass among all the model-based methods. For abundance, the method that produce the least biased estimates is GAM+OK with 1500 m scale, though it only outperformed the GAM+OK with 750 m scale by 0.006%. When both the bias and precision of the estimates are taking into account, GAM+OK with 750 m is the best way to estimate the scallop abundance (lowest percent RMSE, Table **Error! Reference source not found.**). The GAM+OK with 750 m segments also produce the least biased CV estimates for both biomass and abundance estimates (Table 2).

Percent bias and percent RMSE of the SM estimates are smaller than all the model-based estimates (except for the percent RMSE of the abundances estimated using GAM+OK with 750 m) but the CVs were highly underestimated. Beside the problems of estimating CVs, SM estimates were sensitive to the quality of post stratification. SMW and SMN estimates were biased and worse than all the model-based estimates.

Based on the simulation results, we concluded that GAM+OK method with data averaged over 750 m segments was the best way to estimate total abundance and biomass using HabCam data. SM estimates with careful stratifications were also provided in order to validate the model-based estimates although variances for the SM method are probably understated.

### 5.6 Analysis of actual HabCam data for 2011-2013

The HabCam data were collected during 2011-2013 in GB and during 2012-2013 in MAB. We divided the GB and MAB stock region into 14 subregions based on geographic characteristics and management areas and analyzed them separately because their topology, orientation and covariance structures differ (Figures 14 and 15).

Images with altitudes higher than 4 m and scallops with measured shell heights smaller than 40 mm were excluded for estimating scallop abundance and biomass. The shell height ( $SH$ ) measures were converted to meat weights (g) ( $MW$ ) based Hennen and Hart (2012)

$$\text{MAB: } MW = -16.88 + 4.64 \log(SH) + 1.57 \log(D) - 0.43 \log(SH) \log(D) \quad (5.14)$$

$$\text{GB: } MW = 14.38 + 2.826 \log(SH) - 0.529 \log(D) - 5.98 \log(L), \quad (5.15)$$

where  $D$  is depth and  $L$  is latitude. The counts and weight data were summed by image and standardized into abundance and biomass per  $m^2$  by field of view of the image. A summary of the HabCam data used by subregion for 2011-2013 is listed in Table 1.

As described above and based on the simulation results, the GAM+OK method with 750 m segments was used to estimate total abundance and biomass for each subregion. For estimation purposes, we constructed a 1-km buffer zone around each subregion and used the data within the buffered region to build the subregional models. An average of weight or count ( $t$ ) by image ( $j$ ) and distance group ( $i$ ) weighted by field of view ( $f$ ) was calculated for every 750 m segment along the tracks

$$\bar{t}_i = \sum_{j=1}^{n_i} \frac{f_{i,j} t_{i,j}}{\sum_{j=1}^{n_i} f_{i,j}}, \quad (5.16)$$

The  $\bar{t}_i$  was weighted by both variation ( $s$ ) and number of images ( $n$ ) in the hurdle GAM using

$$w_i = \frac{s_i - s_{(1)}}{2(s_{(n_i)} - s_{(1)})} + \frac{n_i - n_{(1)}}{2(n_{(n_i)} - n_{(1)})} \quad (5.17)$$

A hurdle GAM with a quasi-binomial distribution for the presence/absence model and quasi-Poisson distribution for the positive model was used to estimate the first-order trend with respect to latitude, longitude, and depth. Depth is correlated with latitude and / or longitude in some of the subregions. To prevent potential problems cause by collinearity, latitude and longitude were transformed into composite variables: latitude plus longitude and half of the latitude or longitude plus longitude/latitude. A list of models with the different combination of covariates is supplied in Table 3. Depth is included in all of the candidate models because it is one of the most important variables that affecting scallop distributions. The maximum amount of knots for interactions between covariates in GAM models was limited to 15 (reduced to 10 for some of the subregions) and 10 for the single terms to prevent over-fitting. We selected the final first-order model using the RMSE from a 10-fold cross validation.

OK were performed on the GAM residuals. We tested isotropic and a series of anisotropic (from 0 to 180 by 20 degrees) residual OK models and selected the final OK model using the median standard error (MedSE).

$$\text{MedSE} = \sum_{i=1}^n \text{Median}(\hat{\bar{t}}_i - \bar{t}_i) \quad (5.18)$$

GAM and OK final models by subregion and year are listed in Table 4.

For the SM analysis we used only the data within the subregion. The transects were separated into segments based on the following criteria: parallel or perpendicular to depth contour, distance between points (2 km), depth strata, and distance along the transect (10 km). We first separated transects into segments at locations where the direction of the transects changed between parallel and perpendicular to the depth contour. These segments were further separated into smaller ones by depth strata or any location where the distance of any two points in the segment was greater than 2 km. The resulting segments were again broken into smaller ones if length where segments were longer than 10 km. An example of segmentations of the HabCam data (abundance data for 2013) is in Figure 20.

Thresholds for the depth strata were estimated using a maximum likelihood based change

point analysis (Killick et al. 2010). A GAM with a quasi-Poisson error distribution was built for each subregion. The depth partial residuals from the GAM were used in the change point analysis to estimate the depth thresholds. The thresholds were detected based on changes in mean or variance or both mean and variance of the partial residuals. Each subregion is post-stratified into a maximum of 3 depth strata. The depth stratification was done for each year by subregion and separately for abundance and biomass data.

The mean count or weight and its variance was estimated by segment and stratum using equations 10 and 11 and weighted by total field of view ( $f$ ) and length of the segment ( $d$ ) to estimate the total abundance or biomass and its variance

$$\hat{T} = A \sum_{i=1}^3 S_i \sum_{j=1}^{n_i} w_{i,j} \bar{t}_{i,j} \quad (5.19)$$

$$\text{Var}(\hat{T}) = A^2 \sum_{i=1}^3 S_i^2 \sum_{j=1}^{n_i} w_{i,j}^2 \text{Var}(\bar{t}_{i,j}), \quad (5.20)$$

where  $n_i$  is number of segments within depth stratum  $i$ ,  $S_i$  is the size of depth stratum  $i$ , and  $w_{i,j}$  is the weighting factor

$$w_{i,j} = \frac{d_{i,j} - d_{i,(1)}}{2(d_{i,(n_i)} - d_{i,(1)})} + \frac{f_{i,j} - f_{i,(1)}}{2(f_{i,(n_i)} - f_{i,(1)})} \quad (5.21)$$

The resulting GAM+OK and SM abundance and biomass estimates and CV's by subregion are listed in Table 5 and by stock in Table 6 for 2011-2013.

### 5.7 Size composition data for assessment modeling

Calculating scallop size frequency distributions from HabCam data for use in this assessment required re-stratifying Georges Bank for each year for appropriate spatial expansions because inclusion of the RSA surveys resulted in very high densities of annotated images in localized areas (Figure 13). A simple union of the sea scallop strata and HabCam estimation areas was sufficient for the Mid Atlantic in 2012 and 2013 as there were no RSA surveys in this region. Based on these stratifications, we derived stratified size frequency distributions by calculating the density of scallops within each strata and size class, weighted these densities by strata area, and averaging across the region. No adjustments for measurement errors were made although such measurement errors in the two optical surveys for sea scallops (HabCam and SMAST) may have standard deviations on the order of 1 cm. Instead, this type of error is accommodated in the CASA stock assessment model as predicted population length distributions are transformed into predicted length composition observations (Jacobson et al. 2010).

## References

- Barry, S. C., Welsh, A. H. (2002) Generalized additive modeling and zero inflated count data. *Ecological Modeling*, 157(2), 179–188
- Brand, A. R. 1991. Scallop ecology: distributions and behavior. In: Shumway, S. (Ed.), *Scallops: Biology, Ecology, and Aquaculture*, Elsevier, Amsterdam, p. 517-584.
- Cressie, N. A. C. 1993. *Statistics for Spatial Data*, revised edition. 68 John Wiley and Sons, New York, p. 416.
- Hengl, T. 2009. *A practical guide to geostatistical mapping*. University of Amsterdam, Amsterdam, p. 291.
- Hennen, D.R., Hart, D.R. 2012. Shell height-to-weight relationships for Atlantic sea scallops (*Placopecten magellanicus*) in offshore U.S. waters. *J. Shellfish Res.* 31:1133-1144.
- Jacobson, L.D., Stokesbury, K.D.E., Allard, M.A., Chute, A., Harris, B.P., Hart, D., Jaffarian, T., Marino, M.C., Nogueira, J.I., and Rago, P. 2010. Quantification, effects, and stock assessment modeling approaches for measurement errors in body size data using sea scallops (*Placopecten magellanicus*) as an example. *Fish. Bull.* 108: 233–247.
- Killick, R., Eckley, I. A., Jonathan, P., Ewans, K. 2010. Detection of changes in the characteristics of oceanographic time-series using statistical change point analysis. *Ocean Engineering*, 37(13), 1120-1126.
- Kitanidis, P. K. 1993. Generalized covariance functions in estimation. *Mathematical Geology*, 25(5), 525-540.
- Lark, R. M., Cullis, B. R., Welham, S. J., 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science*, 57(6), 787-799.
- Minasny, B., McBratney, A. B. 2007. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma*, 140(4), 324-336.
- Northeast Fisheries Science Center (NEFSC) 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW): Assessment Report. Northeast Fisheries Science Center Reference Document 10-17.
- Odeh, I. O. A., McBratney, A. B., Chittleborough, D. J. 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* 67(3), 215-226.
- Rivoirard, J., Simmonds, J., Foote, K. G., Fernandes, P., Bez, N. 2008. *Geostatistics for estimating fish abundance*. John Wiley & Sons.
- Webster, R., Oliver, M. A. 2001. *Geostatistics for Environmental Scientist*. John Wiley and Sons, Chichester, England, p. 149.

Table 1: Sample size, percent zero, mean weight and count per m<sup>2</sup> of for HabCam data by regions during 2011-2013.

Stock	Year	Subregion	Sample Size	% Zero	Meat Wt (g/m <sup>2</sup> )	Meat Ct (m <sup>2</sup> )
GB	2011	CA1	1942	0.86	15.09	0.39
GB	2011	CA2_N	213	0.91	26.93	1.11
GB	2011	CA2_S	614	0.96	3.22	0.1
GB	2011	GSC_NW	1022	0.83	21.31	0.56
GB	2011	GSC_SE	677	0.97	24.67	0.42
GB	2011	NF	797	0.96	7.77	0.24
GB	2011	NLS	349	0.94	8.48	0.25
GB	2011	SF	554	0.99	2.34	0.08
GB	2012	CA1	660	0.91	6.18	0.35
GB	2012	CA2_N	1382	0.52	27.95	0.91
GB	2012	CA2_S	1415	0.93	3.34	0.12
GB	2012	GSC_NW	735	0.77	8.5	0.47
GB	2012	GSC_SE	276	0.94	5.42	0.23
GB	2012	NF	1486	0.82	22.84	0.75
GB	2012	NLS	298	0.87	5.85	0.26
GB	2012	SF	982	0.96	3.83	0.14
GB	2013	CA1	2054	0.95	1.54	0.07
GB	2013	CA2_N	1015	0.61	21.83	0.51
GB	2013	CA2_S	476	0.86	2.14	0.29
GB	2013	GSC_NW	953	0.86	2.99	0.15
GB	2013	GSC_SE	676	0.95	1.77	0.07
GB	2013	NF	1818	0.93	11.34	0.28
GB	2013	NLS	322	0.85	2.8	0.13
GB	2013	SF	491	0.84	2.05	0.3
MAB	2012	DMV_VB	753	0.9	0.84	0.11
MAB	2012	ET	665	0.85	1.28	0.19
MAB	2012	HC	1159	0.9	1.66	0.15
MAB	2012	HCnr	732	0.93	1.45	0.1
MAB	2012	HCsr	619	0.92	1.86	0.14
MAB	2012	LI	486	0.95	1.24	0.07
MAB	2013	DMV_VB	561	0.91	1.93	0.17
MAB	2013	ET	922	0.87	4.25	0.35
MAB	2013	HC	1114	0.96	2.02	0.18
MAB	2013	HCnr	657	0.95	1.55	0.08
MAB	2013	HCsr	585	0.96	1.7	0.14
MAB	2013	LI	608	0.96	1.55	0.08

Table 2: Percent bias, CV, percent RMSE, estimated CV and number of converged sample runs for biomass and abundance estimates by segment sizes and estimation methods.

Model Type	Scale	% Bias	CV	% RMSE	Estimated CV	# Runs	% Bias	CV	% RMSE	Estimated CV	# Runs
GAM	750	0.048	0.194	0.209	0.191	7194	0.038	0.167	0.177	0.16	7187
GAMM	750	0.088	0.19	0.225	0.308	6699	0.069	0.165	0.189	0.249	6106
OK	750	0.136	0.195	0.26	0.289	7196	0.098	0.173	0.214	0.241	7182
GAM	1500	0.052	0.276	0.295	0.173	7198	0.033	0.188	0.197	0.154	7195
GAMM	1500	0.088	0.192	0.227	0.465	6305	0.066	0.167	0.19	0.507	5774
OK	1500	0.173	0.385	0.484	0.272	7184	0.113	0.288	0.34	0.225	7194
GAM	2250	0.056	0.227	0.246	0.16	7199	0.036	0.206	0.198	0.156	7199
GAMM	2250	0.09	0.193	0.228	0.559	6342	0.063	0.206	0.19	0.651	5953
OK	2250	0.178	0.339	0.438	0.259	7199	0.126	0.206	0.415	0.213	7199
SM		-0.002	0.193	0.193	0.09	7200	0.001	0.206	0.181	0.064	7200
SMN		0.219	0.233	0.359	0.091	7200	0.168	0.206	0.294	0.068	7200
SMW		0.13	0.201	0.262	0.094	7200	0.085	0.206	0.216	0.067	7200

Table 3: List of GAMs tested in the 10-fold cross validation.

GAM Models
s(Longitude, Latitude, k=15)+s(Depth)
s(Latitude, Depth, k=15)
s(Longitude, Depth, k=15)
s(LatPlusHalfLong, Depth, k=15)
s(HalfLatPlusLong, Depth, k=15)
s(LatPlusLong, Depth, k=15)
s(Latitude)+s(Depth)
s(Longitude)+s(Depth)
s(LatPlusHalfLong)+s(Depth)
s(HalfLatPlusLong)+s(Depth)
s(LatPlusLong)+s(Depth)
s(Latitude)+Depth

Table 4: List of first-order and second-order models for the biomass and abundance estimates of GB and MAB subregions for 2011 to 2013.

Stock	Year	Subregion	GAM (Biomass)	GAM (Abundance)	OK (Biomass)	OK (Abundance)
GB	2011	CA1	s(HalfLatPlusLong) + s(Depth)	s(HalfLatPlusLong) + s(Depth)	No angle	Angle: 100
GB	2011	CA2_N	s(LatPlusLong) + s(Depth)	s(LatPlusLong) + s(Depth)	Angle: 0	No angle
GB	2011	CA2_S	s(Longitude, Depth, k = 15)	s(Longitude, Depth, k = 15)	No angle	Angle: 160
GB	2011	GSC_NW	s(Longitude, Latitude, k = 15) + s(Depth)	s(LatPlusLong) + s(Depth)	No angle	No angle
GB	2011	GSC_SE	s(Latitude, Depth, k = 10)	s(Latitude, Depth, k = 15)	No angle	Angle: 140
GB	2011	NF	s(LatPlusLong) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 60	No angle
GB	2011	NLS	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 120	Angle: 160
GB	2011	SF	s(Latitude) + Depth	s(LatPlusHalfLong, Depth, k = 15)	Angle: 160	No angle
GB	2012	CA1	s(HalfLatPlusLong) + s(Depth)	s(HalfLatPlusLong) + s(Depth)	Angle: 160	No angle
GB	2012	CA2_N	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 100	Angle: 60
GB	2012	CA2_S	s(Latitude, Depth, k = 15)	s(Latitude, Depth, k = 15)	Angle: 40	No angle
GB	2012	GSC_NW	s(Latitude) + s(Depth)	s(LatPlusLong) + s(Depth)	No angle	Angle: 0
GB	2012	GSC_SE	s(HalfLatPlusLong) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 60	Angle: 20
GB	2012	NF	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	No angle	No angle
GB	2012	NLS	s(HalfLatPlusLong) + s(Depth)	s(HalfLatPlusLong) + s(Depth)	Angle: 120	Angle: 80
GB	2012	SF	s(LatPlusLong, Depth, k = 15)	s(Longitude, Latitude, k = 15) + s(Depth)	No angle	Angle: 40
GB	2013	CA1	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Depth, k = 15)	Angle: 120	No angle
GB	2013	CA2_N	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	No angle	No angle
GB	2013	CA2_S	s(Longitude, Latitude, k = 15) + s(Depth)	s(Latitude, Depth, k = 10)	Angle: 0	Angle: 160
GB	2013	GSC_NW	s(Latitude) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 0	Angle: 0
GB	2013	GSC_SE	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 160	Angle: 20
GB	2013	NF	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 160	Angle: 160
GB	2013	NLS	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 0	Angle: 160
GB	2013	SF	s(LatPlusLong, Depth, k = 15)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 20	No angle
MAB	2012	DMV_VB	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 60	Angle: 60
MAB	2012	ET	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 80	No angle
MAB	2012	HC	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 160	No angle
MAB	2012	HCnr	s(LatPlusHalfLong, Depth, k = 15)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 140	Angle: 100
MAB	2012	HCsr	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 0	Angle: 60
MAB	2012	LI	s(Latitude) + s(Depth)	s(Latitude) + s(Depth)	Angle: 100	Angle: 0
MAB	2013	DMV_VB	s(LatPlusLong) + s(Depth)	s(LatPlusLong) + s(Depth)	No angle	Angle: 20
MAB	2013	ET	s(Longitude, Latitude, k = 15) + s(Depth)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 60	No angle
MAB	2013	HC	s(Longitude, Latitude, k = 15) + s(Depth)	s(LatPlusHalfLong) + s(Depth)	No angle	Angle: 120
MAB	2013	HCnr	s(Latitude) + s(Depth)	s(Latitude) + s(Depth)	Angle: 100	Angle: 40
MAB	2013	HCsr	s(Longitude, Latitude, k = 15) + s(Depth)	s(LatPlusHalfLong, Depth, k = 10)	No angle	No angle
MAB	2013	LI	s(LatPlusLong, Depth, k = 15)	s(Longitude, Latitude, k = 15) + s(Depth)	Angle: 0	Angle: 40

Table 5: Abundance and biomass and its CVs estimated using GAM+OK and SM methods by subregions for 2011 to 2013.

Stock	Year	Subregion	Number (million)				Weight (mt)			
			SM	GAM+OK	SM CV	GAM+OK CV	SM	GAM+OK	SM CV	GAM+OK CV
GB	2011	CA1	1151.70	1220.70	0.02	0.92	41772.14	42648.48	0.01	0.05
GB	2011	CA2_N	406.92	409.21	0.05	0.07	8325.85	12797.17	0.05	0.06
GB	2011	CA2_S	215.35	338.48	0.08	0.35	8882.94	10237.32	0.07	0.29
GB	2011	GSC_NW	1480.93	1289.01	0.04	0.17	32578.17	21675.43	0.04	0.15
GB	2011	GSC_SE	79.21	75.00	0.12	0.77	3578.14	2051.35	0.14	1.50
GB	2011	NF	336.35	201.78	0.10	0.09	5002.90	4631.38	0.08	1.70
GB	2011	NLS	218.19	159.66	0.07	0.06	7285.42	6224.57	0.07	0.17
GB	2011	SF	103.55	138.22	0.12	1.86	2778.85	2553.07	0.20	3.05
GB	2012	CA1	489.46	763.04	0.08	0.13	10102.43	11744.98	0.08	0.29
GB	2012	CA2_N	659.49	568.81	0.02	0.09	19660.00	21527.78	0.02	0.02
GB	2012	CA2_S	257.40	372.81	0.09	0.07	9803.77	9590.06	0.08	0.16
GB	2012	GSC_NW	1401.52	1721.65	0.05	0.04	25584.05	26266.07	0.05	0.22
GB	2012	GSC_SE	97.12	65.23	0.30	0.23	2390.65	4359.93	0.60	0.30
GB	2012	NF	375.65	259.75	0.05	0.09	8809.68	5919.12	0.05	0.23
GB	2012	NLS	275.23	256.81	0.14	0.44	8139.02	7111.74	0.16	0.14
GB	2012	SF	447.59	634.37	0.11	1.00	9534.90	7519.81	0.12	0.17
GB	2013	CA1	223.26	434.47	0.07	0.05	4479.75	6313.61	0.09	1.09
GB	2013	CA2_N	358.69	279.35	0.03	0.03	15818.66	12027.82	0.03	0.04
GB	2013	CA2_S	545.50	1026.61	0.04	0.09	5594.88	5445.98	0.10	0.05
GB	2013	GSC_NW	471.15	501.50	0.05	0.47	8518.39	8875.60	0.06	0.31
GB	2013	GSC_SE	78.82	57.64	0.16	0.90	1934.28	2281.77	0.21	0.08
GB	2013	NF	135.35	175.20	0.06	1.40	3413.10	4206.02	0.09	2.87
GB	2013	NLS	227.46	188.51	0.12	0.07	4519.21	4039.83	0.11	0.03
GB	2013	SF	1521.91	1385.35	0.06	0.05	10405.12	6480.77	0.09	0.18
MAB	2012	DMV_VB	487.11	340.30	0.06	0.09	3563.73	2657.57	0.08	0.08
MAB	2012	ET	1069.29	1431.26	0.06	0.02	7872.55	7455.85	0.06	0.68
MAB	2012	HC	1056.73	1417.64	0.05	0.02	12865.32	13196.17	0.07	0.10
MAB	2012	HCnr	497.09	616.72	0.11	0.99	8320.79	8607.06	0.12	0.03
MAB	2012	HCsr	418.46	435.87	0.13	0.15	6398.27	6531.35	0.12	0.03
MAB	2012	LI	637.03	660.37	0.11	0.04	11553.18	10748.32	0.11	0.25
MAB	2013	DMV_VB	594.70	529.23	0.07	0.09	5928.37	5742.01	0.05	0.05
MAB	2013	ET	1607.36	1555.18	0.04	0.04	20500.36	19429.08	0.04	0.05
MAB	2013	HC	1324.67	1091.30	0.08	0.16	9953.54	10758.67	0.09	0.05
MAB	2013	HCnr	644.33	502.73	0.26	0.48	8899.89	9953.83	0.14	0.78
MAB	2013	HCsr	262.77	266.72	0.27	0.40	5107.50	4946.65	0.26	0.11
MAB	2013	LI	630.57	665.43	0.09	0.10	11925.31	10655.17	0.10	0.06

Table 6: Abundance and biomass and its CVs estimated using GAM+OK and SM methods by stocks for 2011 to 2013.

Stock	Management Area	Year	Number (million)				Weight (mt)			
			SM	GAM+OK	SM CV	GAM+OK CV	SM	GAM+OK	SM CV	GAM+OK CV
GB	Close	2011	1992.16	2128.05	0.02	0.53	66266.35	71907.54	0.02	0.06
GB	Close	2012	1681.58	1961.46	0.04	0.08	47705.22	49974.57	0.04	0.08
GB	Close	2013	1354.91	1928.94	0.03	0.05	30412.50	27827.24	0.03	0.25
GB	Open	2011	2000.04	1704.01	0.04	0.20	43938.06	30911.23	0.03	0.39
GB	Open	2012	2321.88	2681.00	0.04	0.24	46319.28	44064.93	0.05	0.14
GB	Open	2013	2207.23	2119.68	0.04	0.17	24270.90	21844.16	0.05	0.57
GB	Total	2011	3992.20	3832.06	0.02	0.31	110204.42	102818.77	0.02	0.12
GB	Total	2012	4003.46	4642.46	0.03	0.14	94024.50	94039.50	0.03	0.08
GB	Total	2013	3562.13	4048.62	0.03	0.09	54683.40	49671.39	0.03	0.29
MAB	Total	2012	4165.70	4902.15	0.03	0.13	50573.84	49196.34	0.04	0.12
MAB	Total	2013	5064.39	4610.57	0.05	0.07	62314.98	61485.41	0.04	0.13

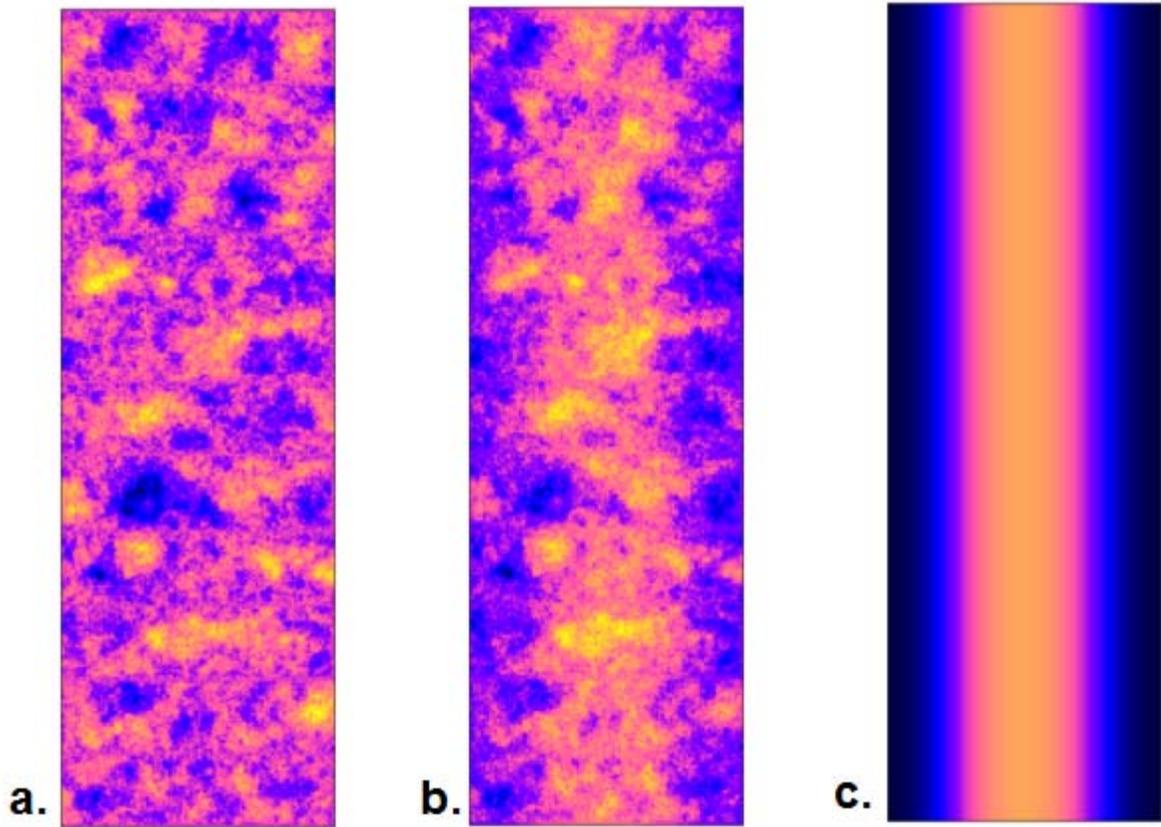


Figure 1. Hypothetical landscapes for (a) a landscape with a stationary mean, (b) a biased landscape with a higher mean along the center, and (c) the bias applied to landscape a to produce landscape b. In all plots, densities are higher in lighter-colored pixels.

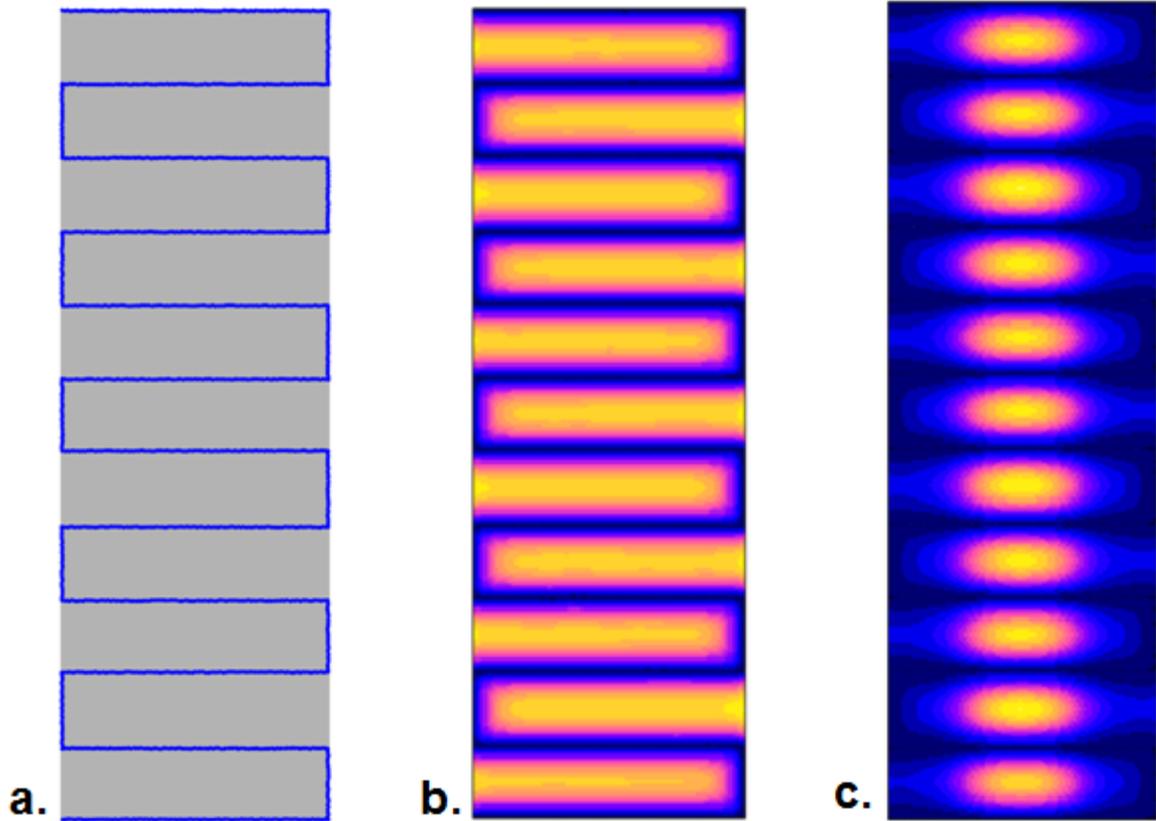


Figure 2 (a) A regularly spaced survey track across a rectangular survey area, (b) map of kriging variance derived from the survey track and an assumed underlying variogram model describing the data, (c) adjusted kriging variances resulting from applying the underlying trend from Figure 1c to (b).

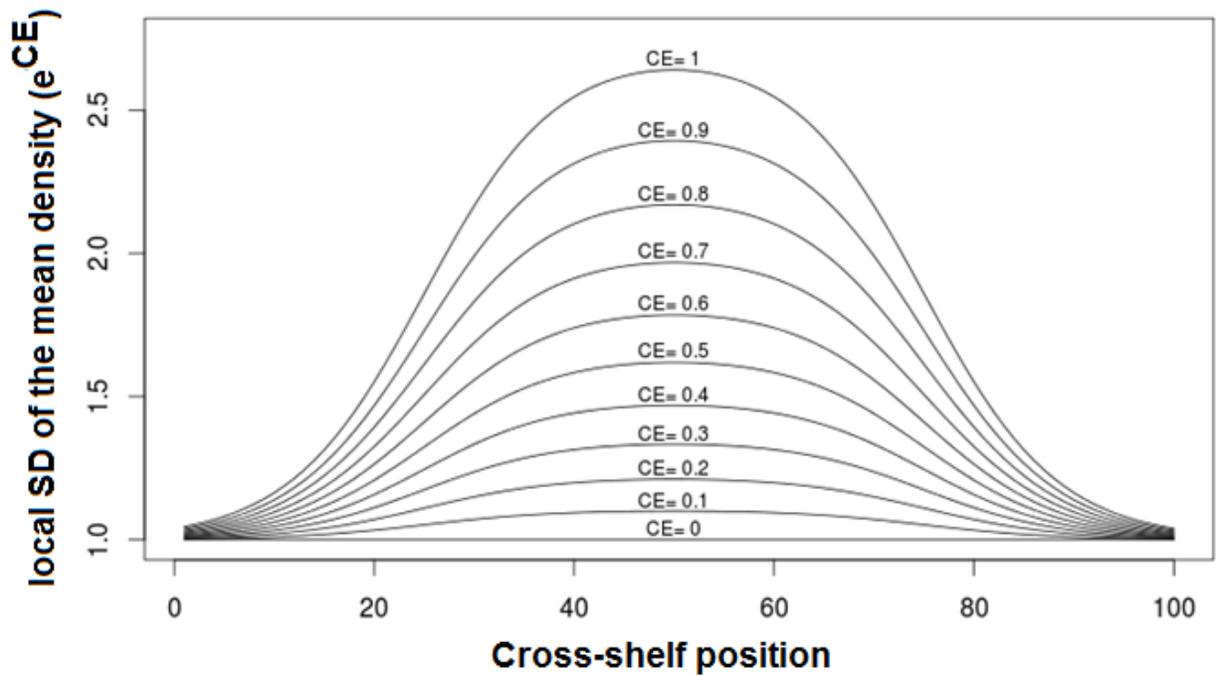


Figure 3. Levels of center effect used in simulations and resulting effects on the local standard deviation of the mean.

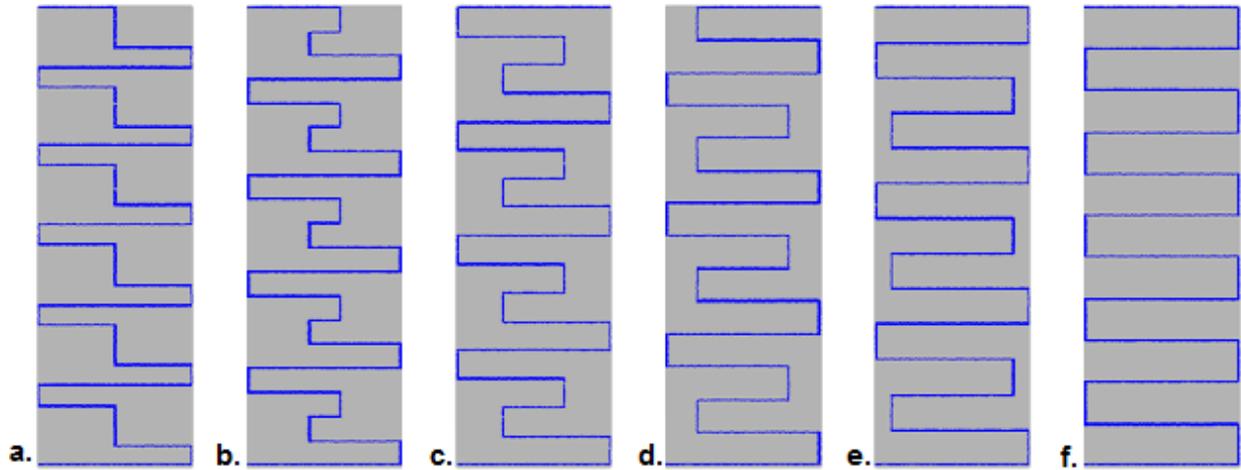


Figure 4a. Alternative survey configurations, alternating the length of transects along the track and keeping total survey length constant. Short transect lengths are (a) 0%, (b) 20%, (c) 40%, (d) 60%, (e) 80%, and (f) 100% of the length of the long transects.

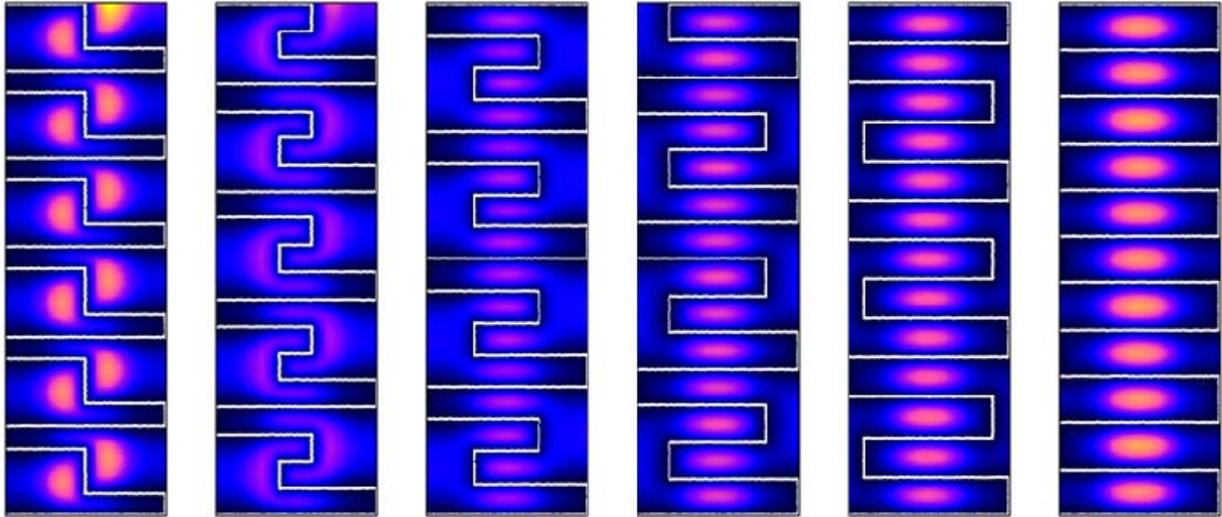


Figure 5. Variance maps for the survey tracks in Figure 4 with an applied center effect. Lighter colors indicate higher areas of variance.

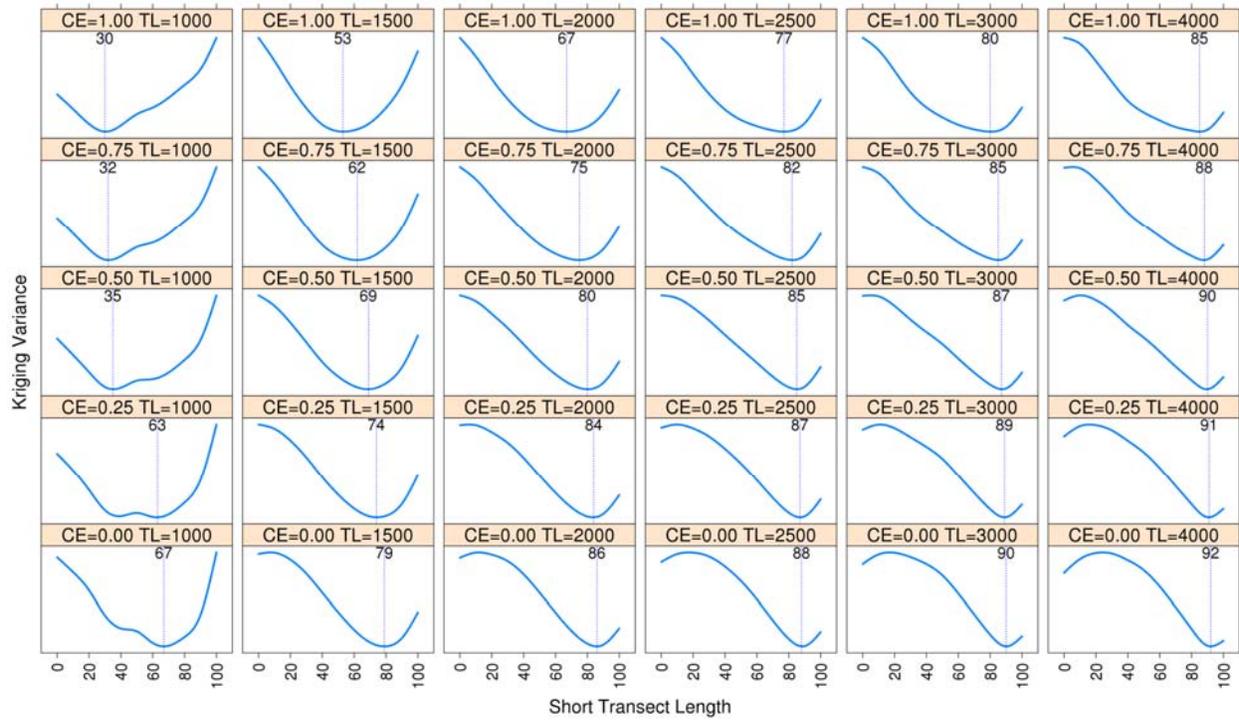


Figure 6. Adjusted kriging variances for different center effects (CE), short transect lengths and total survey track lengths (TL). Optimal solutions for each combination are marked with a dotted vertical line and labeled.

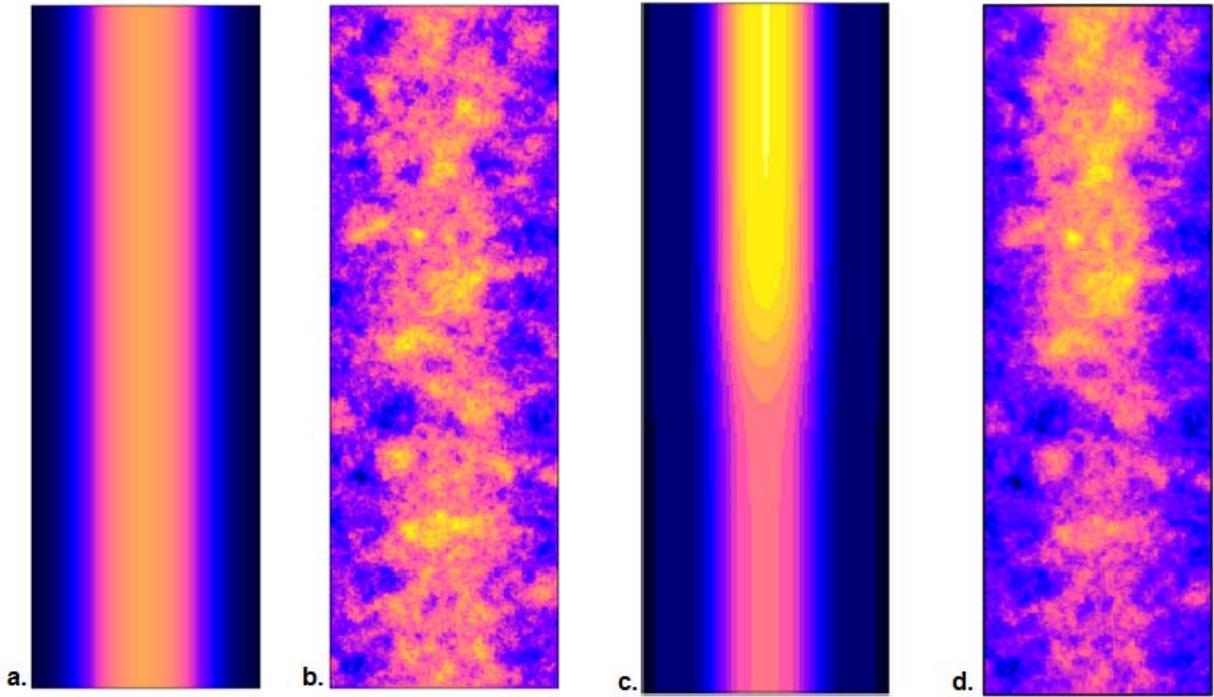


Figure 7. Comparison of survey landscapes without (a and b) and with (c and d) zonal anisotropy effects. Figures a and c represent the underlying trend in the mean while b and d represent the resulting simulated landscapes.

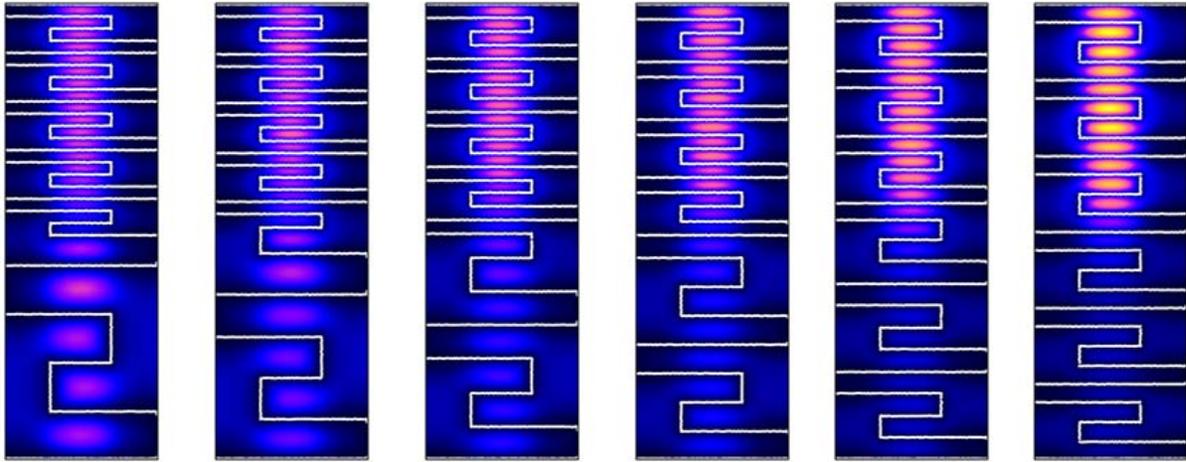


Figure 8. Example of varying transect density between zones and resulting variance maps for a simulated landscape with an underlying trend similar to C-7c. The survey track is represented in white. Lighter colors indicate higher variances.

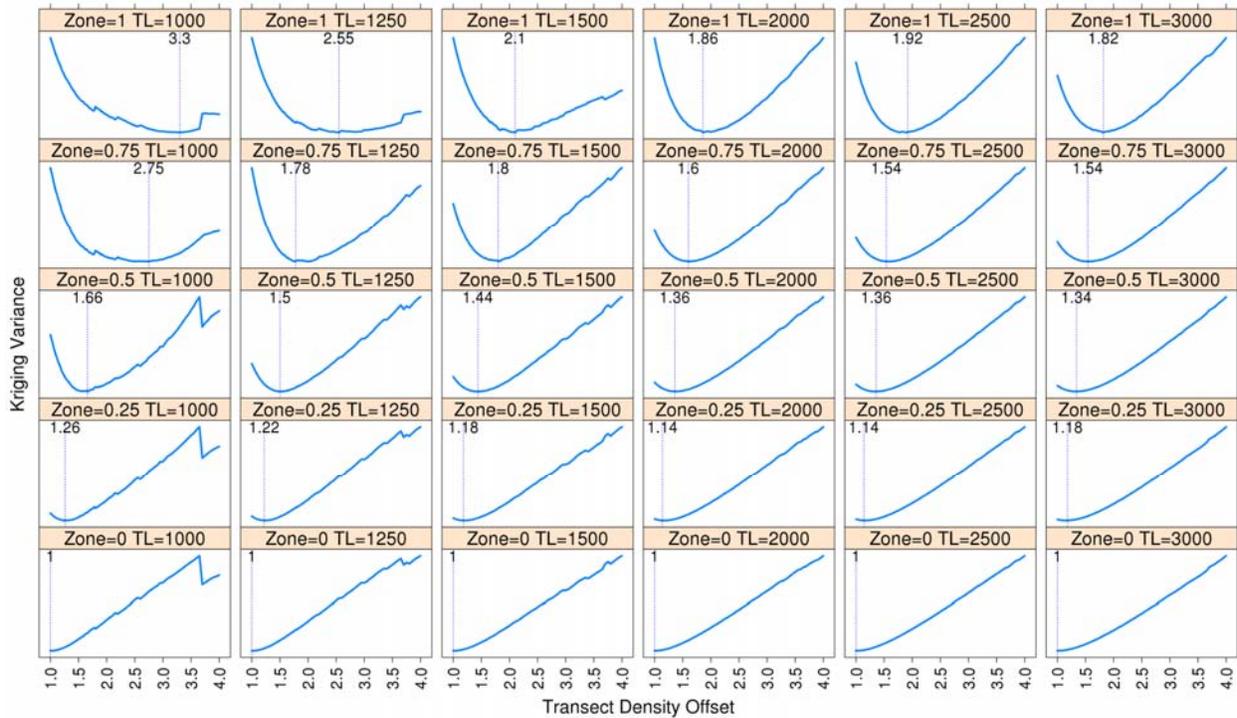


Figure 9. Zonal effects on transect density allocation. Higher “Transect Density Offsets” represent the placement of proportionally more transects in the high density zones. Optimal solutions for each simulation set are labeled and marked with a dotted line.

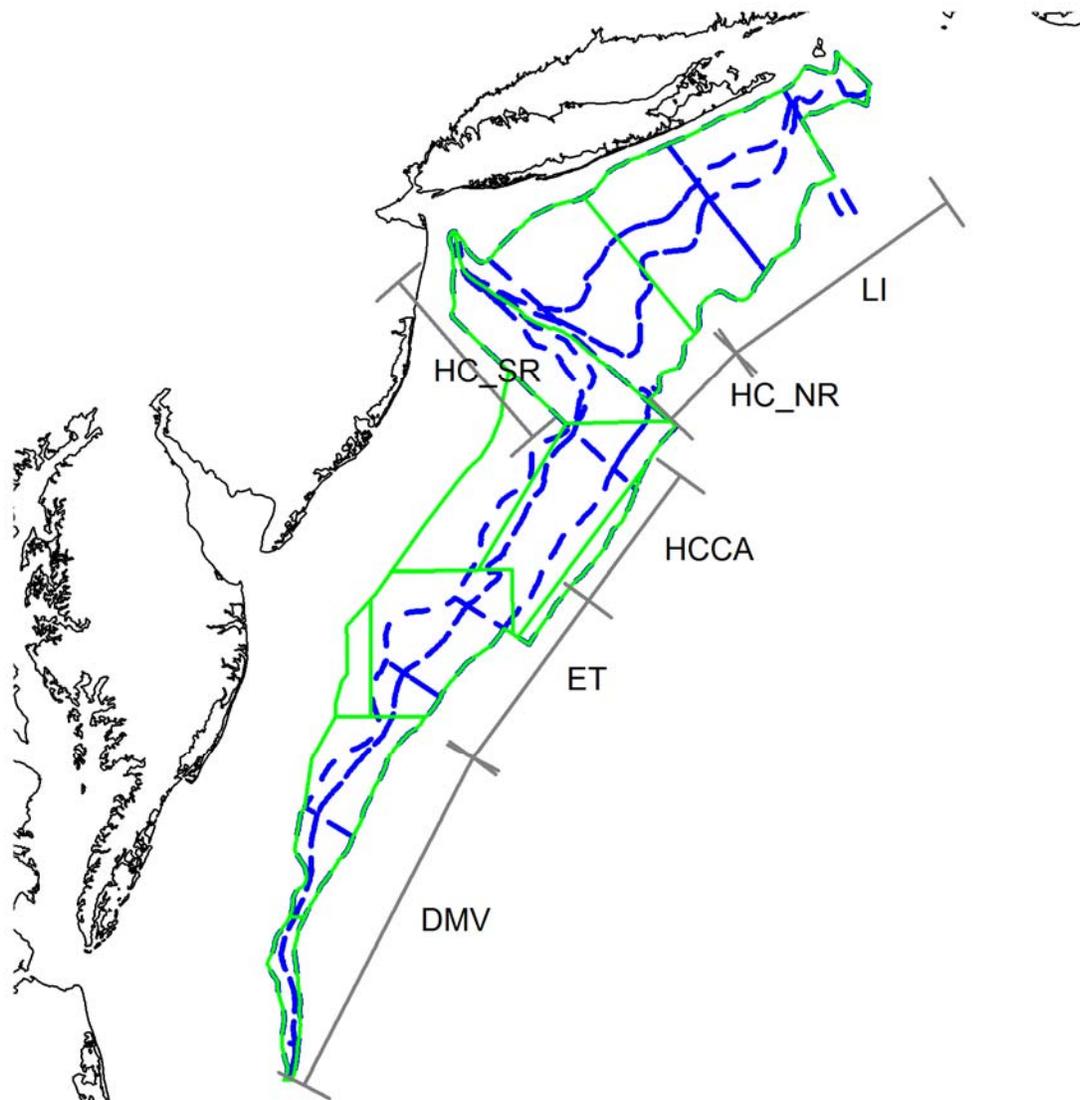


Figure 10. HabCam survey area (solid green line) compared to NEFSC scallop survey core strata (dashed blue line) in the MAB region. Subregions used for allocating survey effort and abundance estimation are: LI – Long Island, HC\_NR – Hudson Canyon North Rim, HC\_SR Hudson Canyon South Rim, HCCA – Hudson Canyon Closed Area, ET – Elephant Trunk, and DMV – DelMarVa.

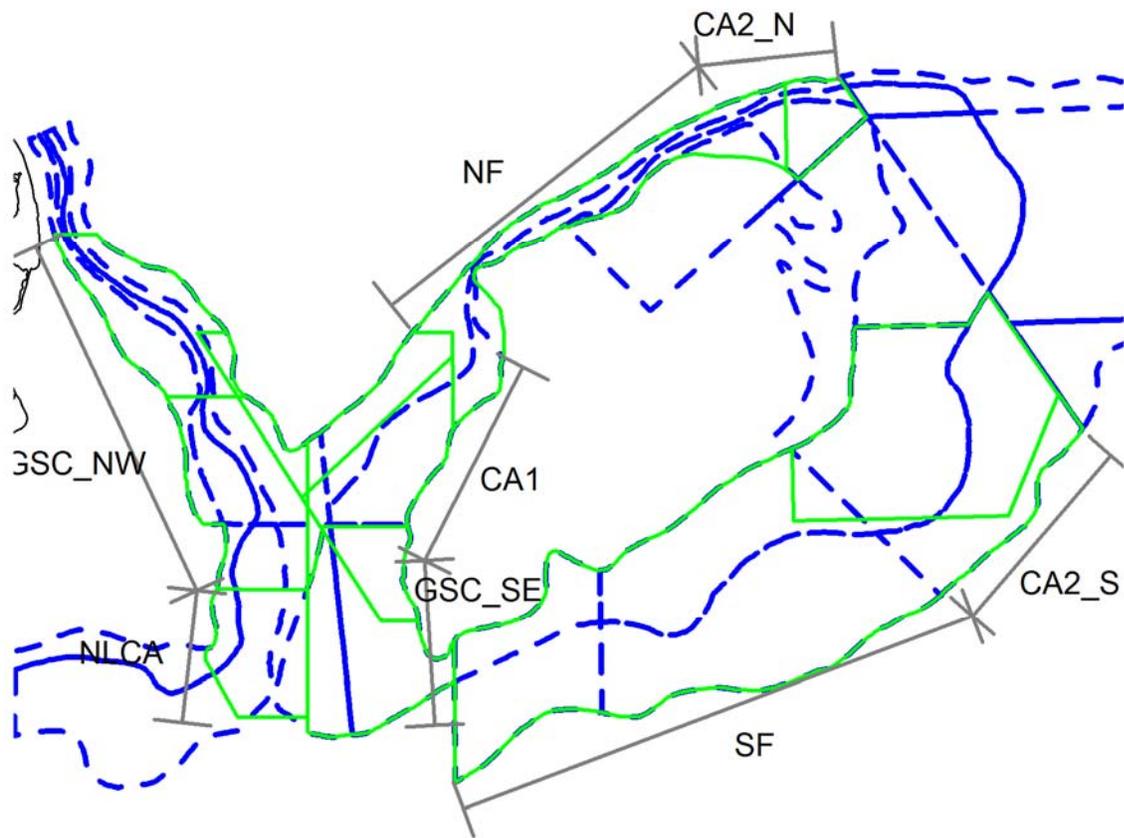


Figure 11. HabCam survey area (solid green line) compared to NEFSC scallop survey core strata (dashed blue line) for Georges Bank. Subregions used for allocating survey effort and abundance estimation are: GSC\_NW – Great South Channel Northwest, NLCA – Nantucket Lightship Closed Area, GSC\_SE – Great South Channel Southeast, CA1 – Closed Area 1, NF – Northern Flank, CA2\_N – Closed Area 2 North, CA2\_S – Closed Area 2 South, and SF – Southern Flank.

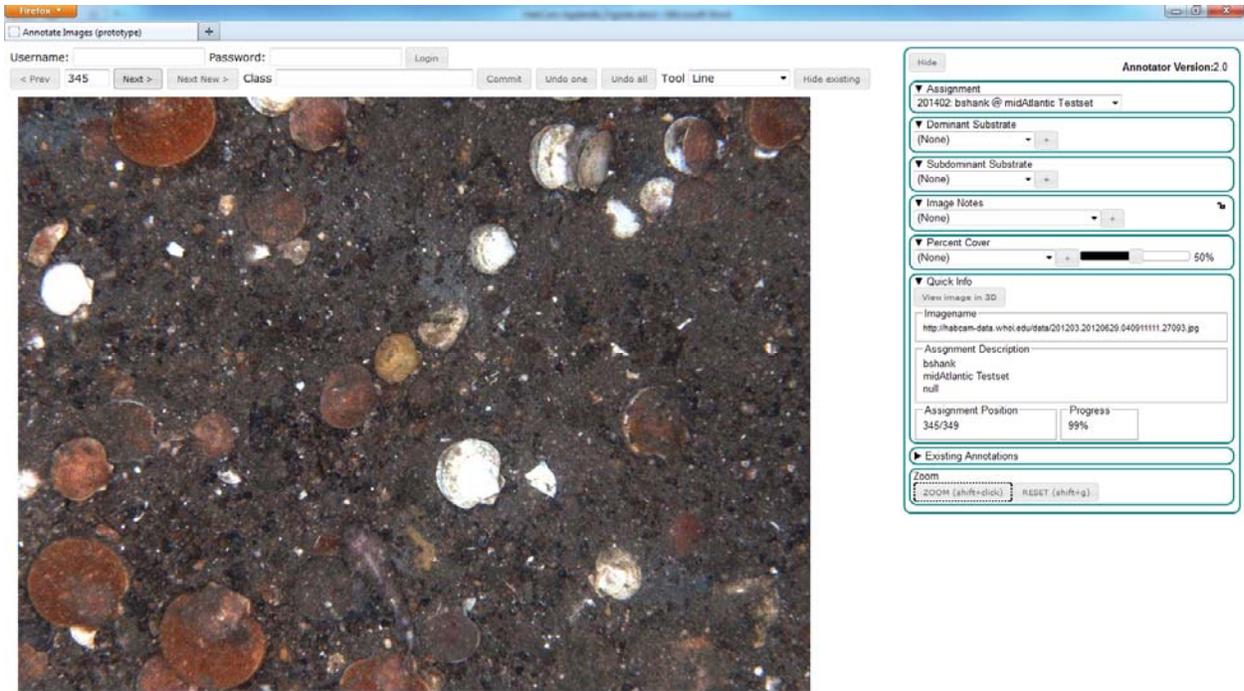


Figure 12. Screen image showing the web-based annotation tool for counting and measuring scallops from HabCam images.

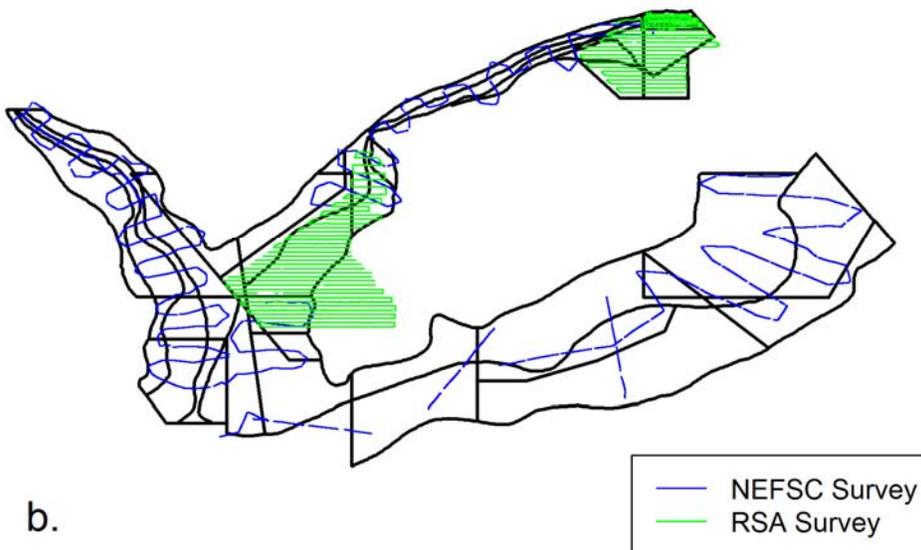
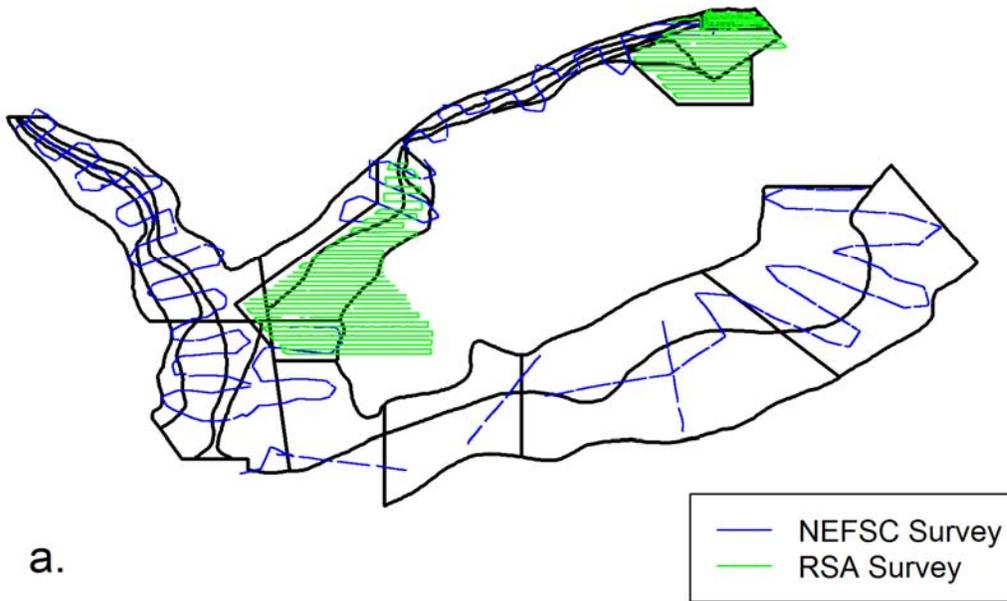


Figure 13 Example re-stratification of Georges Bank used for 2013 size frequencies: (a) open and closed areas combined and (b) open and closed areas separate.

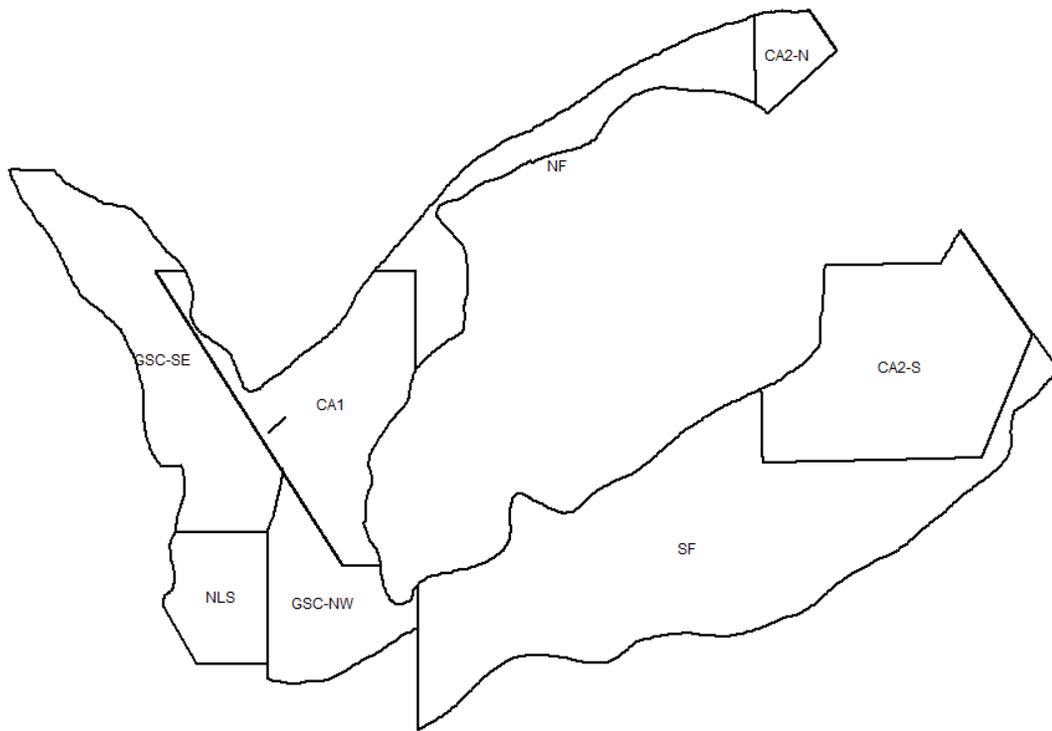


Figure 14: Subregions of the GB scallop stock area used in the HabCam survey.

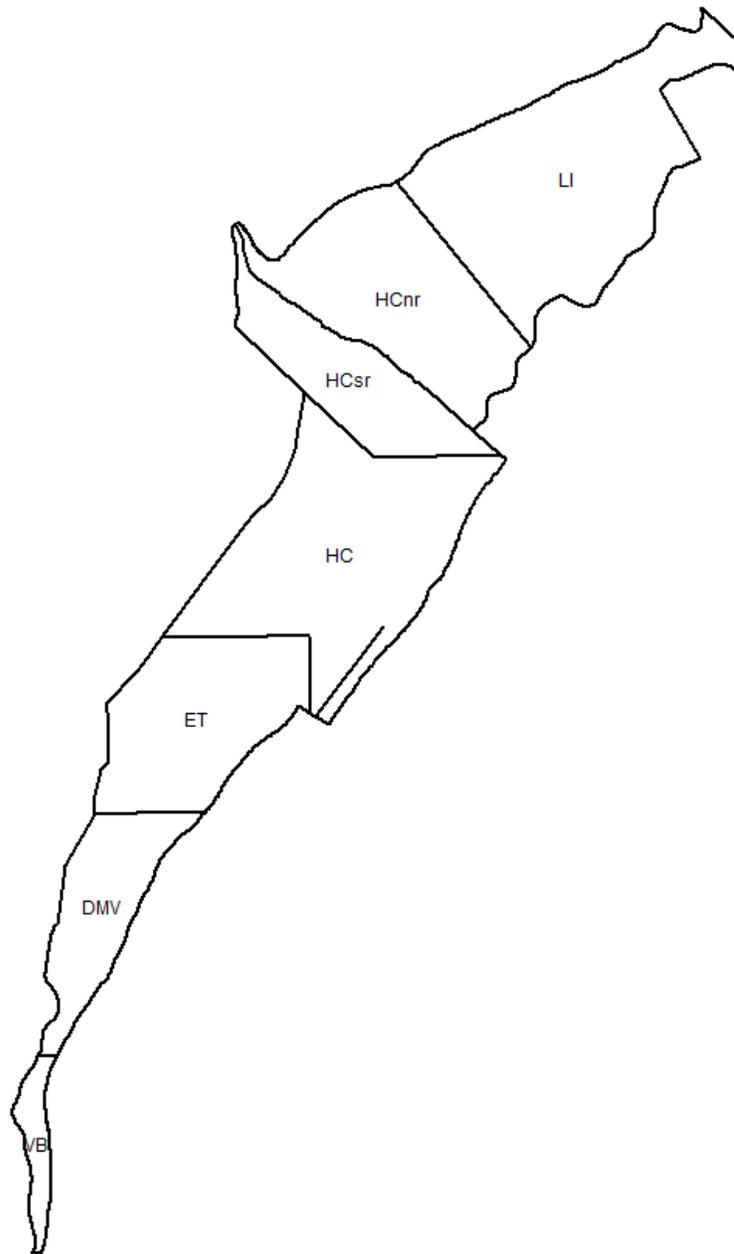


Figure 15: Subregions of the MAB scallop stock area used for the HabCam survey.

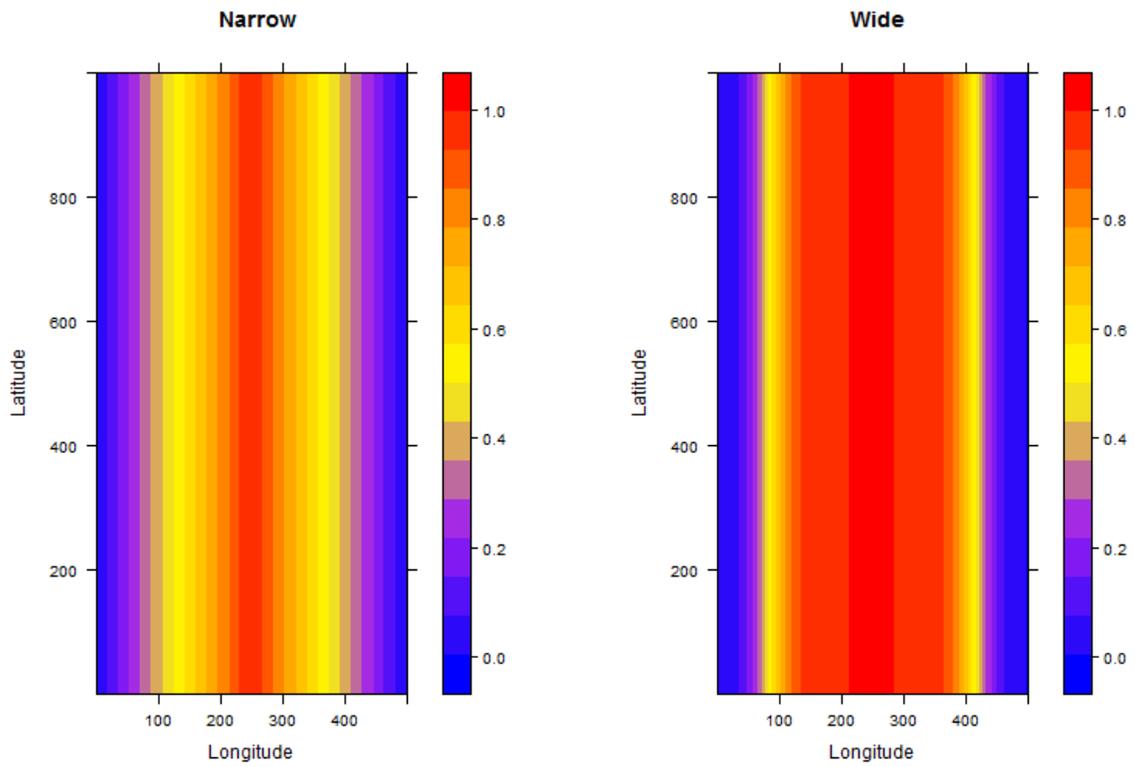


Figure 16: The two types of first-order effects used to simulate scallop populations: a narrow but highly dense first-order effect (left) and a wide but relatively less dense first-order effect (right).

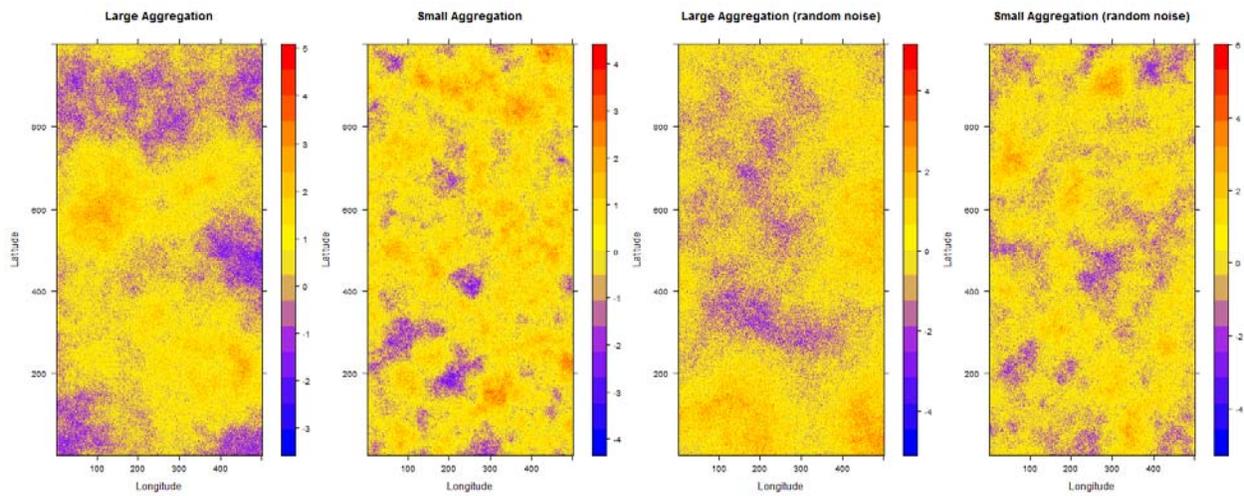


Figure 17: The four types of second-order trends tested to simulate scallop populations: large aggregations, small aggregations, large aggregations with a high random noise, and small aggregations with a high random noise (from left to right).

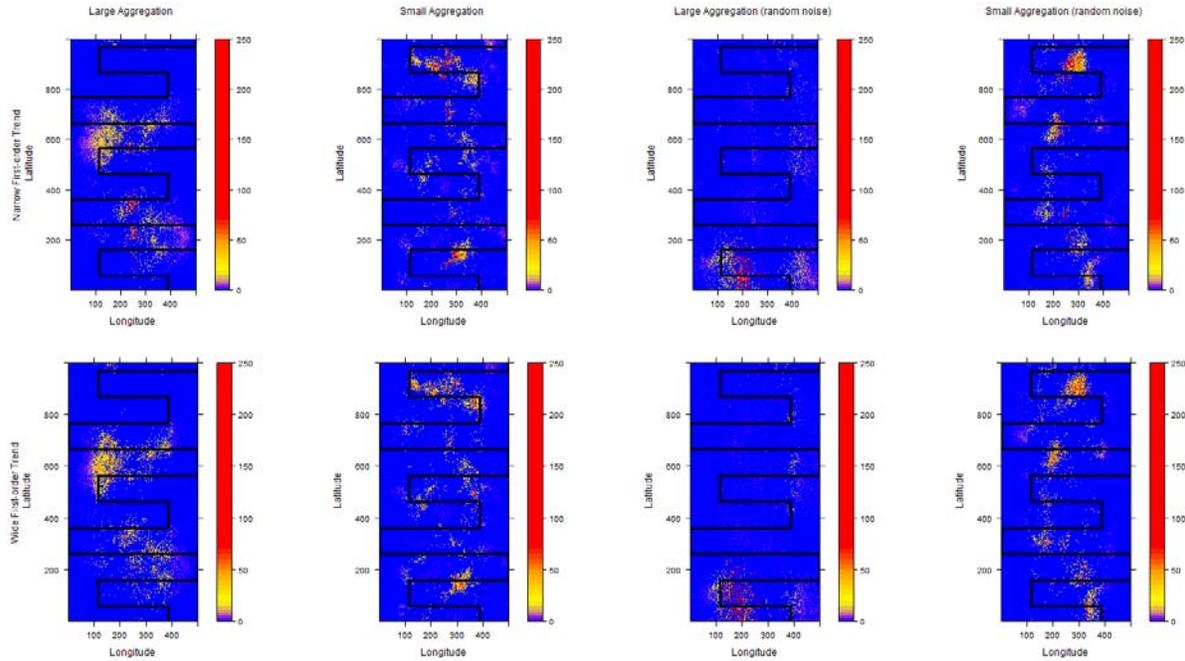


Figure 18: Example simulated scallop population distributions with an over-layed sampling track.

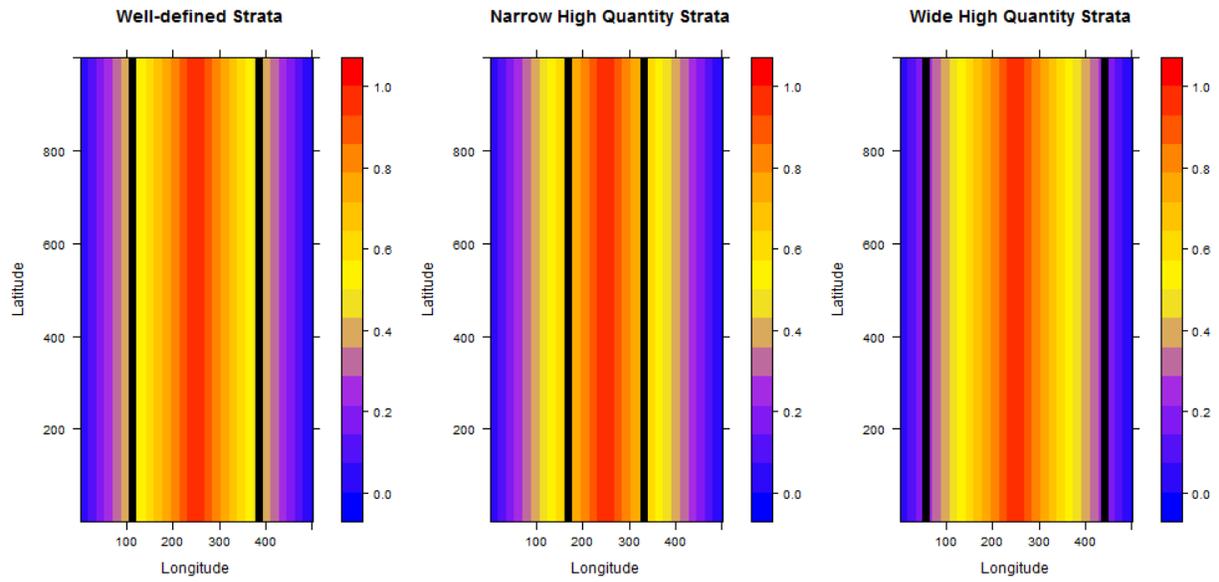


Figure 19: Alternative types of stratifications used for stratified mean estimations.

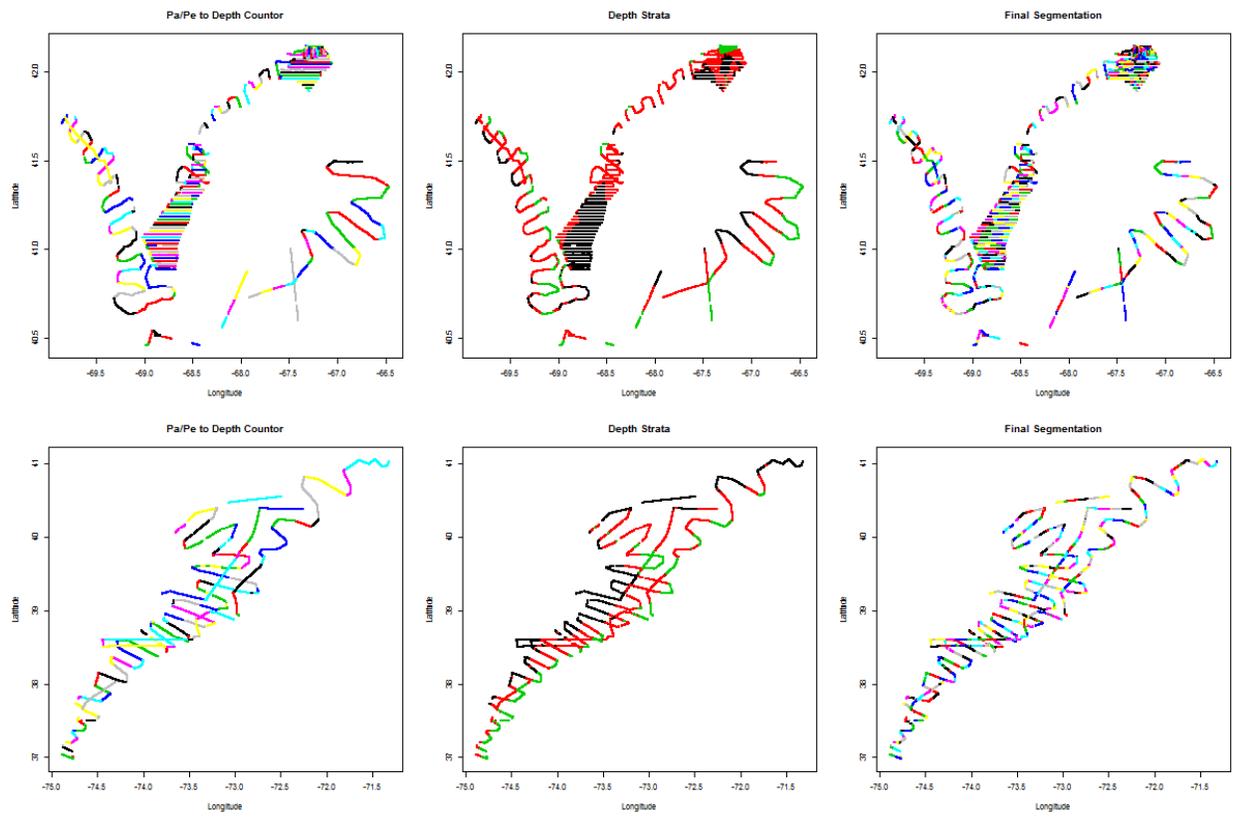


Figure 20: Transect segmentation for stratified mean estimation and the 2013 survey based on orientation to depth contours and distance between points (2 km) (left) and depth strata (center). The final combined segmentation is on the far right for GB (upper panels) and MAB (lower panels).

## **Appendix B7. Assessment of the sea scallop resource in the Northern Gulf of Maine management area**

Samuel B. Truesdell (University of Maine, Orono, ME); Kevin H. Kelly (Maine Department of Marine Resources, W. Boothbay Harbor, ME); and Yong Chen (University of Maine, Orono, ME).

### **Summary**

The sea scallop (*Placopecten magellanicus*) fishery in the Northern Gulf of Maine management area (NGOM) occurs in federal waters and is managed by the New England Fishery Management Council. The NGOM resource and associated fishery are locally important but amount to a small portion of the total stock and less than 0.1% of total landings. The fishery is managed by a TAC independently of the rest of the EEZ sea scallop stock. Management of the NGOM fishery does not involve biological reference points as targets or thresholds.

A cooperative survey was carried out by the Maine Department of Marine Resources and University of Maine during May-June of 2012. Based on survey results, estimated biomass of NGOM sea scallops targeted by the fishery (102+ mm or 4+ in shell height) was approximately 164.19 MT (90% confidence interval from 74.35 to 278.91), an increase from 115.40 MT (66.05 to 173.31) in 2009. These estimates are based on density estimates from the survey assuming a capture efficiency of 43.6%. The previous survey in 2009 noted a large year class of 10-50 mm scallops on Platts Bank; this year class was still evident in 2012 and had grown to approximately 65-90 mm.

Based on these biomass estimates the exploitation rate in weight (landings/stock biomass, assuming harvested scallops greater than 102 mm shell height and a dredge efficiency of 43.6%) during 2012 was 2.1% with a 90% confidence interval from 1.3% to 4.7%.

Several analyses were performed to determine how representative the survey was of the NGOM to determine applicability of survey results to management of the NGOM. The fraction of the NGOM covered by the survey area is 0.11, however using information regarding habitat preferences of scallops, the fraction of the suitable habitat area for the stock within the NGOM covered by the survey is 0.37. The survey extent was designed to ensure coverage of the primary fishing areas, and the fraction of fishing locations within the survey bounds was greater than 50% since 2006 and greater than 70% since 2011. Thus, the survey probably encompasses most of the areas with scallop concentrations high enough to support fishing activity indicating that survey results should be useful information for management of the NGOM scallop stock.

### **Introduction**

The Gulf of Maine scallop fishery that occurs in federal waters is managed by the New England Fishery Management Council. Amendment 11 to the New England Fishery Management Council Sea Scallop Fishery Management Plan (NEFMC 2008) created a separate limited entry program for general category fishing in the Northern Gulf of Maine management area (NGOM; Fig. 1). The area is managed under an annual total allowable catch (TAC; currently 31.75 MT) and a daily possession limit of 90.7 kg (NEFMC 2008). Scallop dredge ring size must be greater than 102 mm, but there are currently no regulations regarding shell size (as in Maine state waters) or meat count.

Landings in the NGOM are low relative to the rest of the scallop stock, averaging just over 7 MT from 2008 to 2013 (total sea scallop landings have been over 20,000 MT in recent

years). In 2013 the most landings since the management area's inception in 2008 (over 18 MT) were reported, more than double any other year.

The region has limited fishery-independent data available. There was an offshore survey administered by the Maine Department of Marine Resources in 1974 (Spencer 1974), and in 1983 and 1984 NMFS sampled some areas in this region on their annual survey (Serchuk 1983; Serchuk and Wigley 1984), but no broad-scale surveys were completed between the early 1980s and 2008 when the region was first managed under a TAC. Given the lack of recent fishery-independent data, the initial allowable catch was determined using historical federal Gulf of Maine landings (NEFMC 2008). More recently, Maine Department of Marine Resources/University of Maine scallop surveys in 2009 and 2012, along with UMass Dartmouth video scallop surveys that occasionally sample in this area (e.g., Stokesbury et al. 2010) have offered fishery-independent sources of information to aid in generating the TAC.

The only management area-wide biomass estimate previously available was based on the Maine Department of Marine Resources/University of Maine scallop survey in 2009. This was a point estimate that used swept area to expand the survey results to a subset of the NGOM (this subset is discussed below; Fig. 1). This analysis estimated 103 MT of scallops greater than 102 mm shell height, with a confidence interval that ranged from 53 to 186 MT (Truesdell et al. 2010). This estimate was revised (see Results/Discussion section) during the current analysis and the new estimate for 2009 is 115.40 MT (90% confidence interval from 66.05 to 173.31). The best estimate based on the 2012 survey results indicates that the biomass of NGOM sea scallops over 102 mm shell height was approximately 164.19 MT of meats with a 90% confidence interval ranging from 74.35 to 278.91 MT.

## Methods

### *Survey area identification and delineation*

The NGOM management area is bounded by Cape Ann, Massachusetts in the west and the Canadian border in the east (Fig. 1). Prior to 2009 when the first survey was conducted, the NGOM had limited fishery dependent and no recent fishery-independent data available to help design the survey. Scallops are not found uniformly throughout this region so sampling efforts were focused on a subset of areas in the NGOM. To determine this subset, fishing locations from National Marine Fisheries Service vessel trip reports (VTRs) from 2000 to 2008 were reviewed as well as three historical surveys of the region from the 1970s and 1980s (Spencer 1974; Serchuk 1983; Serchuk and Wigley 1984). In addition to the information available, two fishermen with a history of scalloping in the Gulf of Maine were interviewed to help identify current and historical fishing grounds. These sources of information were used qualitatively to determine the five sampling areas: Machias-Seal Island (MSI), Mount Desert Rock (MDR), Platts Bank (PB), Northeast of Cape Ann (NCA) and Northern Stellwagen Bank (NSB; Fig. 1).

To increase sampling precision, the two western strata off the Massachusetts coast (where most fishing occurs), NCA and NSB, were further divided into substrata of expected high, medium and low scallop density.

### *Survey coverage area*

Although the survey is intended to represent the NGOM scallop management area, the entirety of the NGOM was not sampled (Fig. 1); as such it is necessary to document the survey coverage area relative to total stock area, total stock biomass and the area fished. This was accomplished most simply by calculating the ratio of the sampling area ( $A_{SURVEY}$ ) to the area of the NGOM ( $A_{NGOM}$ )

$$R_{BASE} = \frac{A_{SURVEY}}{A_{NGOM}} \quad \text{Eqn. 1}$$

where  $R_{BASE}$  is the proportion of survey coverage. This baseline ratio is only one approach to estimating the coverage area of the survey, and it assumes that scallops are as likely to be found within the survey area as they are outside. However, the survey was designed specifically to sample the areas where scallops are distributed within the NGOM, so  $R_{BASE}$  is likely to be an underestimate of the survey's coverage of the scallop stock. Two additional methods were used to arrive at a more realistic approximation: one based on the depths at which scallops are typically found and another based on fishing effort data.

Sea scallops are typically more abundant at shallower depths (Merrill 1972; Posgay 1979; Serchuk et al. 1979); a depth threshold was employed as one way to estimate the effective coverage proportion of the survey. Serchuk et al. (1979) note that most commercial quantities of scallops are found in depths less than 100 m; this is corroborated by analyses of the NMFS bottom trawl survey from 1982 to 2010 and the NMFS bottom trawl survey from 2010 to 2012.

Employing a depth threshold  $DTh$  to determine an effective coverage proportion for the survey can be

$$R_{DTh} = \frac{A_{SURVEY}^*}{A_{DTh}} \quad \text{Eqn. 2}$$

where  $R_{DTh}$  stands for the ratio at a particular depth threshold (100 m in this analysis),  $A_{SURVEY}^*$  is the survey area shallower than the threshold and  $A_{DTh}$  is the area of the NGOM shallower than the depth threshold.

Alternatively, an effective coverage proportion can be estimated using fishing effort data. This assumes that the Gulf of Maine scallop fleet follows an ideal free distribution (Fretwell and Lucas 1969; i.e., fishing activity is directly related to abundance). Vessel monitoring system (VMS) data from 2006 to 2013 were used to determine the effort-based effective coverage proportion  $R_{VMS}$  as

$$R_{VMS} = \frac{P_{SURVEY}}{P_{NGOM}} \quad \text{Eqn. 3}$$

where  $R_{VMS}$  is the coverage proportion with respect to VMS observations (satellite location records),  $P_{SURVEY}$  is the number of VMS observations within the survey areas and  $P_{NGOM}$  is the total number of VMS observations within the NGOM.

Two resolutions of VMS data were considered: 1km and 3km. The advantage to the finer resolution is that the locations were more accurate, which is important near the boundaries of the areas. The disadvantage is that for confidentiality reasons less VMS data was available at higher resolutions. At the 1 km resolution 83% of VMS observations were available and at 3 km resolution 91% were available.

#### *Survey design*

Surveys were carried out in June and July of 2009 and in May and June of 2012. Dredge tow stations were selected from a grid overlying each stratum. The dimensions of each grid unit were 1 km<sup>2</sup>. Each survey followed a two-stage random stratified design in the NCA and NSB strata. Station allocation in the first stage was based on fishing intensity from 2000-2008 vessel trip report (VTR) data and the size of each substratum. Forty stations in each stratum were assigned to the first stage and distributed among substrata according to the formula

$$N_s = 40 * \frac{M(V_s)A_s}{\sum_{s=1}^S M(V_s)A_s} \quad \text{Eqn. 4}$$

where  $N_s$  (rounded) is the number of stations to be sampled in substratum  $s$ ,  $M()$  is the median function,  $V_s$  is the VTR landings from 2000 to 2008 for substratum  $s$  and  $A_s$  is the area of substratum  $s$ . VTR data is assumed to be a proxy for scallop density and so was used to help allocate survey sample size. Such commercial data have also been used in the design of Canadian scallop surveys (Robert and Jamieson 1986; Serchuk and Wigley 1986). Area size was included in the weighting to ensure sufficient effort in the larger substrata.

In the NCA and NSB strata, which were further divided into substrata, a two-stage survey was employed. The approach taken by Francis (1984) was used to allocate tows to the second survey stage. His formula to assign one additional station among strata is:

$$G'_s = \frac{A_s^2 M_s^2}{n_s(n_s + 1)} \quad \text{Eqn. 5}$$

where  $G'_s$  is the assumed reduction in variance from adding a single station to a particular substratum  $s$ ,  $A_s$  is the area,  $M_s$  is the mean catch rate (when squared, a proxy for the variance suggested by Francis (1984)) and  $n_s$  is the number of additional stations. Twenty stations were available for the second stage and were apportioned among the substrata. They were assigned one-by-one (by repeated use of Eqn. 5) according to whichever substratum would gain the most in terms of reduced variance from receiving one additional station. As such, the assignment of the  $j^{\text{th}}$  station can be written

$$G'_s = \frac{A_s^2 M_s^2}{(n_s + j - 1)(n_s + j)} \quad \text{Eqn. 6}$$

A single stage design was used for the remaining three strata in the eastern GOM.

In 2012 206 stations were sampled using a 2.13 m New Bedford style dredge with 51 mm rings, 4.4 cm head bale, 8.9 cm twine top, 25.4 cm pressure plate and rock chains. This gear was identical to that used in the 2009 survey. The target tow duration in 2009 was 7 minutes at a speed of 6.5km/h (a distance of approximately 750m). This was reduced to 5 minutes and about 540m in 2012, though fixed gear in some locations forced shorter tows.

#### *Data Analyses*

Historically, meat count by shell height has been found to vary regionally within the Gulf of Maine (Serchuk and Rak 1983), so separate models predicting meat weight using shell height were employed for each stratum. Depth was included because it has been shown to influence many aspects of scallop life history (Naidu and Robert 2006) and has been used in this type of analysis by Hennen and Hart (2012). These models also included a random effect (as in Hennen and Hart 2012) to account for repeated sampling within a station. The mixed effects models were produced using R (v. 2.15.1, R Core Team 2012) with the package lme4 (Bates et al. 2013). The form of the model within each stratum was

$$\ln(W_{i,t}) = \beta_1 \ln(H_{i,t}) + \beta_2 D_t + R_t + \varepsilon_{i,t} \quad \text{Eqn. 7}$$

where  $W_{i,t}$  is the meat weight of individual  $i$  at station  $t$ ,  $H_{i,t}$  is its respective shell height,  $D_t$  is the depth,  $R_t$  is a random effect term associated with each station,  $\beta_1$  and  $\beta_2$ , are the coefficients of the explanatory variables, and  $\varepsilon_{i,t}$  is the error term for each sample. Depth was important to include as a covariate because although meat weights were sampled whenever possible, the samples were not always evenly distributed throughout the depth range of a stratum, though the results were extrapolated across all depths. PB had a low number of meat weight samples in 2009 so the 2009 samples were combined with those from 2012 for the 2009 PB meat weight

model.

Prior to analyzing length frequency distributions, the number of scallops in each 5 mm size class belonging to a particular station was standardized to the mean swept area per station in the relevant stratum or substratum according to the formula:

$$Z_{l,s,c} = \frac{R_{l,s}}{\bar{R}_s} N_{l,s,c} \quad \text{Eqn. 8}$$

where  $Z_{l,s,c}$  is the standardized count for scallops at station  $l$  within stratum (or substratum for strata 4 and 5)  $s$  in 5 mm size class  $c$ ,  $R_{l,s}$  is the swept area of the station tow,  $\bar{R}_s$  is the mean swept area for samples in area  $s$ , and  $N_{l,s,c}$  is the number of scallops of size class  $c$  in tow  $l$  of area  $s$ . In these analyses the middle of the size bin was always used as the reference size for estimation.

The mean number of scallops within each stratum was estimated and uncertainty was addressed using bootstrapping and percentile confidence limits. Survey sample counts were bootstrapped 50,000 times. Bootstrapping was chosen to estimate confidence bounds because it requires few distributional assumptions (Efron and Tibshirani 1986) and avoids unrealistic confidence bounds that drop outside the range of observation (such as below zero).

To estimate the biomass and confidence limits for each stratum, the predicted meat weights from the mixed effects models at each location ( $1 \text{ km}^2$ ) within each stratum were estimated by size class and combined with the (sub)stratum length frequency distribution and the number of scallops per station to calculate the overall biomass per stratum such that

$$B_s = \sum_{g=1}^G \sum_{c=1}^C W_{s,c,g} P_{s,c} N_s \quad \text{Eqn. 9}$$

where  $B_s$  is the estimated biomass in stratum  $s$ ,  $G$  is the number of  $1 \text{ km}^2$  grids in stratum  $s$ ,  $c$  is the number of 5 mm size classes over 102 mm (4 in; assumed to be harvestable size),  $W$  is the expected weight per scallop from Eqn. 7,  $P_{s,c}$  is the proportion of scallops in stratum  $s$  within size class  $c$ , and  $N$  is the bootstrapped standardized mean count per station in stratum  $s$ . The upper and lower confidence limits were estimated by substituting the upper and lower percentile estimates for  $N$  in each substratum.

The dredge efficiency (vulnerability coefficient) used in this study was 43.6% which was estimated experimentally in Maine state waters (Kelly 2007). The Maine value was used because it was generated near the survey area and is close to other estimates of dredge efficiency (e.g., Gedamke et al. 2004).

Weight-based exploitation rates for the NGOM were estimated for 2009 and 2012 as

$$E = \frac{L}{B} \quad \text{Eqn. 10}$$

where  $E$  is the exploitation rate,  $L$  is the landings in weight and  $B$  is the total estimated biomass in the NGOM of scallops larger than 102 mm shell height. A 90% confidence interval for the exploitation rate was computed using the 5<sup>th</sup> and 95<sup>th</sup> percentiles for biomass, derived from the bootstrapping. Landings were assumed to be error-free.

## Results and Discussion

### *Survey coverage area*

The Northern Gulf of Maine management area encompasses a region of 23,470 km<sup>2</sup>. Although this entire management area is under the regulations outlined in Amendment 11 to the sea scallop Fishery Management Plan (NEFMC 2008), scallops are not found throughout the region. The survey region (Fig. 1) has an area of 2,652 km<sup>2</sup> and the ratio of this region to the total area in the Northern Gulf of Maine regulatory region is 0.11 (Eqn. 1). While this areal coverage appears low, the effective survey coverage is larger in terms of both potential scallop habitat with respect to depth as well as the realized fishery area according to vessel monitoring system data.

The coverage proportion assuming a depth threshold of 100 m (Eqn. 2) is 0.37 (Fig. 2). This represents an estimate of the survey's coverage of the NGOM stock area, assuming depth is related to the probability of scallop occurrence. Using VMS data to determine the fraction of the fishery that occurs within the survey extent, the effective coverage proportion (Eqn. 3) was greater than 0.9 using either low or high resolution VMS data (Table 1). The proportion of total VMS observations by year (with no data excluded for confidentiality) was calculated by Burton Shank (NMFS NEFSC) for 2006-2013. Since 2009, the first year of the survey, the minimum coverage proportion was 0.69 (in 2010) and in 2013, the most recent year available, it was 0.87 (Table 2).

#### *Scallop demographics within the NGOM*

The most heavily fished area within the NGOM is the southwestern part, within survey areas NCA and NSB. In the NCA area, most scallops were found north of Cape Ann near the state waters boundary in both 2009 and 2012 (Fig. 3A). In the NSB area there were some scallops at the northern boundary (especially in 2012) and in both years scallops were found on the northern part of Stellwagen Bank near the southern-central part of the NSB area (Fig. 3A). Both NCA and NSB had wide, multimodal shell height distributions in both years (Figs. 4A-B). NSB was noteworthy in 2012 because it had signs of recent recruitment as well as some of the largest scallops seen on the survey.

In both years scallops were found on the southwest part of PB (Fig. 3B). The growth of the cohort first observed in 2009 was evident (Fig. 4C); the mode shell height grew from 32.5 mm in 2009 to 72.5 mm in 2012. In both years there was a small proportion of scallops that were between 125 and 150 mm. The survey in MDR encountered almost no scallops in both years (Fig. 3C). There were scallops to the south of this area near Mount Desert Rock in both years, but this small region is within Maine state waters and not part of the NGOM. The scallops that were caught in 2009 were mainly less than 100 mm (Fig. 4D). In 2012 only a single scallop was caught in this area. In the MSI area there was no obvious coherence between the spatial distribution of catch in the 2009 and 2012 surveys; scallops within this area appear from these surveys to be fairly evenly distributed relative to the patchiness observed in NCA and NSB to the south (Fig. 3D). The only persistent aggregation was near Machias Seal Island, again within state waters. Little signs of recruitment were seen in this region in either 2009 or 2012 and most scallops were between 110 and 150 mm (Fig. 4E).

The relationship between shell height and meat weight varied by area, as in 2009 (Fig. 5). The best condition meats were in NSB and NCA, while the meats in MSI were clearly smaller for their size. Few samples of larger scallops were taken on PB, but those greater than 100 mm were of similarly poor condition to the scallops sampled in MSI.

#### *Biomass and exploitation rate estimates*

Analysis of the surveys produced estimates indicating that the NGOM had overall harvestable biomass in 2009 and 2012 of 115.40 MT (90% confidence interval from 66.05 to

173.31) and 164.19 MT (74.35 to 278.91), respectively. The 2009 estimate was revised slightly since 2010 because of the new meat weight estimates for Platts Bank, a slightly different approach to the bootstrapping (previously a bootstrapping with replacement method was used along with bias corrected confidence intervals; see Truesdell et al. 2010), and the correction of an error that was found in the standardization of length frequencies. The original estimate given was 103 (53 to 186) MT (Truesdell et al. 2010). In addition, the assumed dredge efficiency in 2009 was 40%, but this was changed to 43.6%, which is based on a dredge efficiency study by the Maine Department of Marine Resources.

Harvestable biomass was distributed disproportionately across the areas surveyed. In the eastern half of the NGOM, the MSI area was found to have consistently high biomass (Fig. 6), though the density of biomass was lower than in some regions of the western NGOM (Fig. 7). Further west and offshore, PB was estimated to contain 5.6 harvestable MT in 2009 and 2.1 MT in 2012. However, this assumes that none of the large year class on PB is yet available to fishing since these scallops are under the assumed harvestable size of 102 mm used in this study (Fig. 4C). Given the increased activity on Platts Bank evident in VMS data however, it is likely that some fishermen are targeting this year class though its biomass is not included in the calculations presented here. Still further west in the two strata where most of the fishing currently occurs, NCA and NSB, the mean biomass available for harvest was 17.0 and 43.55 MT in 2009 and 55.6 and 67.2 MT in 2012. Despite their relatively small areas (Fig. 1), the high expected density strata within these regions supported considerable biomass of harvestable scallops in both survey years relative to the other areas surveyed (Fig. 6).

These biomass estimates are dependent on some fixed parameters. Survey dredge efficiency was assumed to be 43.6%, which was determined experimentally in Maine waters. No uncertainty is attached to this estimate however. Gadamke et al. (2004) estimated the efficiency of a dredge with 89 mm rings to be 42.7%, with a potential range based on sensitivity analyses from 35.5 to 52.5%. The gear was different (this study used 51 mm rings), but the mean estimate was similar to the Maine study. The approximate sensitivity range from Gadamke et al.'s study was used as a sensitivity range for the 2012 biomass estimates presented here. If dredge efficiency is assumed to be 35% the 2012 estimate is 207.51 MT (with a 90% confidence interval ranging from 93.35 to 353.29; Table 3). If dredge efficiency is 50% the estimate is 143.14 MT (65.00 to 242.88). No uncertainty was considered for the shell height to meat weight relationships or the length frequency distributions. These sources of uncertainty should be considered in subsequent analyses, though they are probably better estimated than sampling variability which is likely the main source of uncertainty and was quantified by bootstrapping.

Landings were low from 2008 to 2012, though increased notably in 2013 (Fig. 8). To determine the source of this change it would be necessary to examine vessel trip report data; however that information is not currently available to the authors. One possible reason for the higher landings, is the increased fishing effort on PB as the year class first observed in 2009 may have become targeted by the fishery.

The estimated exploitation rate during 2012 was 2.14% (90% confidence interval from 1.26% to 4.72%; Table 3), which is lower than the 6.1% (4.1% to 10.7%) estimated during 2009. The reduced exploitation rate was a function of both a decrease in landings (Figure 8) and the increase in estimated biomass from 115.40 MT in 2009 to 164.16 MT in 2012.

#### *Characterization of scallops in the Gulf of Maine*

The Maine Department of Marine Resources mid-1970s survey report (Spencer 1974) noted that most scallops encountered were older and there was no evidence of recent recruitment,

leading Spencer to conclude that “only in widely separated years do scallops set in these offshore waters.” The report stated that only near-shore fishing was tenable at present, though it was noted that in the 1960s the beds around Jeffereys Ledge were commercially viable. In the early 1980s scallop sets were recorded in the GOM: Serchuk (1984) and Serchuk and Wigley (1984) reported large quantities of small scallops offshore on Fippennies Ledge and Jeffereys Ledge. High densities of commercial size scallops, however, were not found in either of these surveys (Serchuk and Wigley 1984). The Maine Department of Marine Resources/University of Maine 2009 survey identified a large set of scallops on Platts Bank. Another 2009 Gulf of Maine survey corroborated these findings and also observed small scallops on Fippennies Ledge, Jeffereys Ledge and Cashes Ledge (Stokesbury et al. 2010). No such recruitment event was seen in the Maine Department of Marine Resources/University of Maine 2012 survey however.

While the fishery-independent data are not extensive for this region, it is clear that scallop sets in the NGOM are intermittent. In most years recruitment is limited or non-existent, but occasionally large recruitment events do occur. This is supported by the history of the commercial fishery in the region, which is highly variable (Dow 1971; Kelly 2012). The exception may be the western NGOM, in particular the NSB and NCA areas. It is evident from the length frequency distributions (Figs. 4A-B) that recruitment is more stable in this region than to the east where not all size classes are evident. This discrepancy may indicate environmental differences between the eastern and western NGOM, in particular how local oceanography interacts with the early life history of scallops.

## **Conclusions**

Scallops in the NGOM represent a small but locally important fishery. Landings have been low since the inception of the NGOM management area, though they more than doubled in 2013. The best estimates from 2009 and 2012 indicate that scallop biomass increased by about 40% over that period. The exploitation rate in weight (landings/stock harvestable biomass) during 2009 was 6.1% with a 90% confidence interval from 4.1% to 10.7%, and during 2012 was 2.1% (1.3% to 4.7%). Given the region’s low biomass relative to the rest of the stock along with its intermittent recruitment in eastern areas, it is probably not necessary to survey the NGOM every year. However, periodic surveys that provide point biomass estimates are likely to be helpful to managers for determining a TAC.

## **Acknowledgements**

We would like to thank Wallace and Steve Gray and the late Wallace Gray Jr. and Wayne Young of the *F/V Foxy Lady II*, and Glenn Nutting, Anna Henry, and Mike Kersula who assisted as scientific observers. The National Marine Fisheries Service Sea Scallop Research Set-Aside Program and the NMFS-Sea Grant fellowship program in population dynamics funded this study.

## References

- Bates, Douglas, Maechler, M., Bolker, B. and Walker, S. 2013. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-4. <http://CRAN.R-project.org/package=lme4>
- Dow, R.L. 1971. Periodicity of Sea Scallop Abundance Fluctuations in the Northern Gulf of Maine. Bulletin of the Department of Sea and Shore Fisheries. Research Bulletin No. 31.
- Efron, B., and Tibshirani, R. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science* 1: 54–75.
- Francis, R.I.C.C. 1984. An adaptive strategy for stratified random trawl surveys. *New Zealand Journal of Marine and Freshwater Research* 18: 59–71.
- Fretwell, S.D., and Lucas, H.L. 1969. On territorial behavior and other factors influencing habitat distribution in birds. *Acta biotheoretica* 19: 16–36.
- Gedamke, T., DuPaul, W.D., and Hoenig, J.M. 2004. A Spatially Explicit Open-Ocean DeLury Analysis to Estimate Gear Efficiency in the Dredge Fishery for Sea Scallop *Placopecten magellanicus*. *North American Journal of Fisheries Management* 24: 335–351.
- Hennen, D.R., and Hart, D.R. 2012. Shell height-to-weight relationships for Atlantic sea scallops (*Placopecten magellanicus*) in offshore US waters. *Journal of Shellfish Research* 31: 1133–1144.
- Kelly, K.H. 2007. Results from the 2006 Maine Sea Scallop Survey. Maine Department of Marine Resources; West Boothbay Harbor, Maine.
- Kelly, K.H. 2012. Results from the 2011 Maine Sea Scallop Survey. Maine Department of Marine Resources; West Boothbay Harbor, Maine.
- Merrill, A.S. 1960. Abundance and distribution of sea scallops off the Middle Atlantic coast. *Proceedings of the National Shellfisheries Association* 51: 74–80.
- Naidu, K.S., and Robert, G. 2006. Chapter 15 Fisheries sea scallop, *Placopecten magellanicus*. *In* *Scallops: Biology, Ecology and Aquaculture*. Elsevier. pp. 869–905.
- New England Fisheries Management Council (NEFMC). 2008. Fisheries of the Northeastern United States; Atlantic Sea Scallop Fishery; Amendment 11; Final Rule.
- Posgay, J.A. 1979. Sea scallop *Placopecten magellanicus* (Gmelin). *In* *Fish distribution*. Edited by M.D. Grosslein and T. Azarovitz. New York Sea Grant Institute. pp. 130–133.
- R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

- Robert, G., and Jamieson, G.S. 1986. Commercial fishery data isopleths and their use in offshore sea scallop (*Placopecten magellanicus*) stock evaluations. *In* Proceedings of the North Pacific Workshop on Stock Assessment and Management of Invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. pp. 76–82.
- Serchuk, F.M. 1983. Results of the 1983 USA Sea Scallop research survey: distribution and abundance of Sea Scallops in the Georges Bank, Mid-Atlantic and Gulf of Maine Regions and biological characteristics of Iceland Scallops off the coast of Massachusetts. Northeast Fisheries Science Center Reference Document 83-37, Woods Hole, Massachusetts.
- Serchuk, F.M., and Rak, R.S. 1983. Biological characteristics of offshore Gulf of Maine sea scallop populations: size distributions, shell height meat weight relationships and relative fecundity patterns. Northeast Fisheries Science Center Reference Document 83-07, Woods Hole, Massachusetts.
- Serchuk, F.M., and Wigley, S.E. 1984. Results of the 1984 USA sea scallop research vessel survey: status of sea scallop resources in the Georges Bank, Mid-Atlantic and Gulf of Maine regions and abundance and distribution of Iceland scallops off the southeastern coast of Cape Cod. Northeast Fisheries Science Center Reference Document 84-34, Woods Hole, Massachusetts.
- Serchuk, F.M., and Wigley, S.E. 1986. Evaluation of USA and Canadian Research Vessel Surveys for Sea Scallops (*Placopecten magellanicus*) on Georges Bank. *Journal of Northwest Atlantic Fisheries Science*. 7: 1–13.
- Serchuk, F.M., Wood, P.W., Posgay, J.A., and Brown, B.E. 1979. Assessment and status of sea scallop (*Placopecten magellanicus*) populations off the northeast coast of the United States. *Proceedings of the National Shellfisheries Association* 69: 161–191.
- Spencer, F. 1974. Final report: offshore scallop survey - Cape Ann, Massachusetts to Maine-Canadian border. Maine Department of Marine Resources.
- Stokesbury, K.D., Carey, J.D., Harris, B.P., and O’Keefe, C.E. 2010. High densities of juvenile Sea Scallop (*Placopecten magellanicus*) on banks and ledges in the central Gulf of Maine. *Journal of Shellfish Research* 29: 369–372.
- Truesdell, SB, KH Kelly, CE O’Keefe and Yong Chen. 2010. An assessment of the sea scallop resource in the Northern Gulf of Maine management area. Appendix B6 to the 50<sup>th</sup> Northeast Regional Stock Assessment Workshop (50<sup>th</sup> SAW) assessment report. Northeast Fisheries Science Center Reference Document 10-17, Woods Hole, Massachusetts.

Table 1: Survey coverage proportion calculated using three methods (See Eqns. 1-3). Num. stands for numerator, den. stands for denominator and prop. stands for proportion.

Type	Num. descriptor	Num. value	Den. descriptor	Den. value	Prop. inside
Area	Total survey area (km <sup>2</sup> )	2,652	NGOM	23,470	0.11
Depth thresh.	Survey area < 100m	2,652	NGOM < 100m	7,132	0.37
VMS – low res.	VMS inside survey area	26,661	Total VMS within NGOM	27,217	0.98
VMS – high res.	VMS inside survey area	21,901	Total VMS within NGOM	23,555	0.93

Table 2: Proportion of VMS observations within the NGOM survey area. All VMS observations were included (i.e., none were excluded for confidentiality). Table provided by Burton Shank (NMFS NEFSC).

Year	Proportion inside
2006	0.94
2007	0.94
2008	0.55
2009	0.84
2010	0.69
2011	0.74
2012	0.81
2013	0.87
Overall	0.86

Table 3: Best estimates for 2012 NGOM harvestable biomass (HB) and corresponding exploitation rates (ER) under three assumptions of dredge efficiency.

Assumed Dredge Efficiency	5 <sup>th</sup> percentile		Mean		95 <sup>th</sup> percentile	
	HB	ER %	HB	ER %	HB	ER %
35%	93.35	3.76	207.51	1.69	353.29	0.99
43.6%	74.35	4.72	164.19	2.14	278.91	1.26
50%	65.00	5.40	143.14	2.45	242.88	1.44

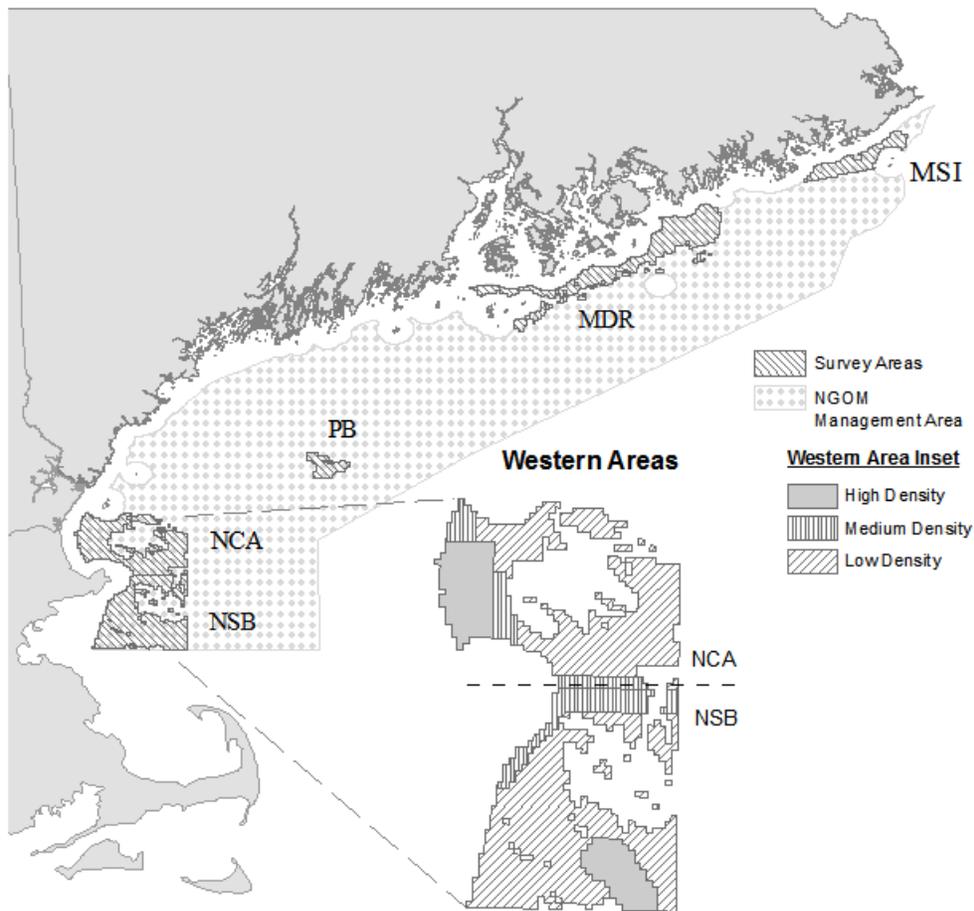


Figure 1: The NGOM and the 5 strata selected for the survey, with substrata of differing expected scallop density appearing in the western areas inset. MSI: Machias-Seal Island; MDR: Mount Desert Rock; PB: Platts Bank; NCA: Northeast of Cape Ann; NSB: Northern Stellwagen Bank.

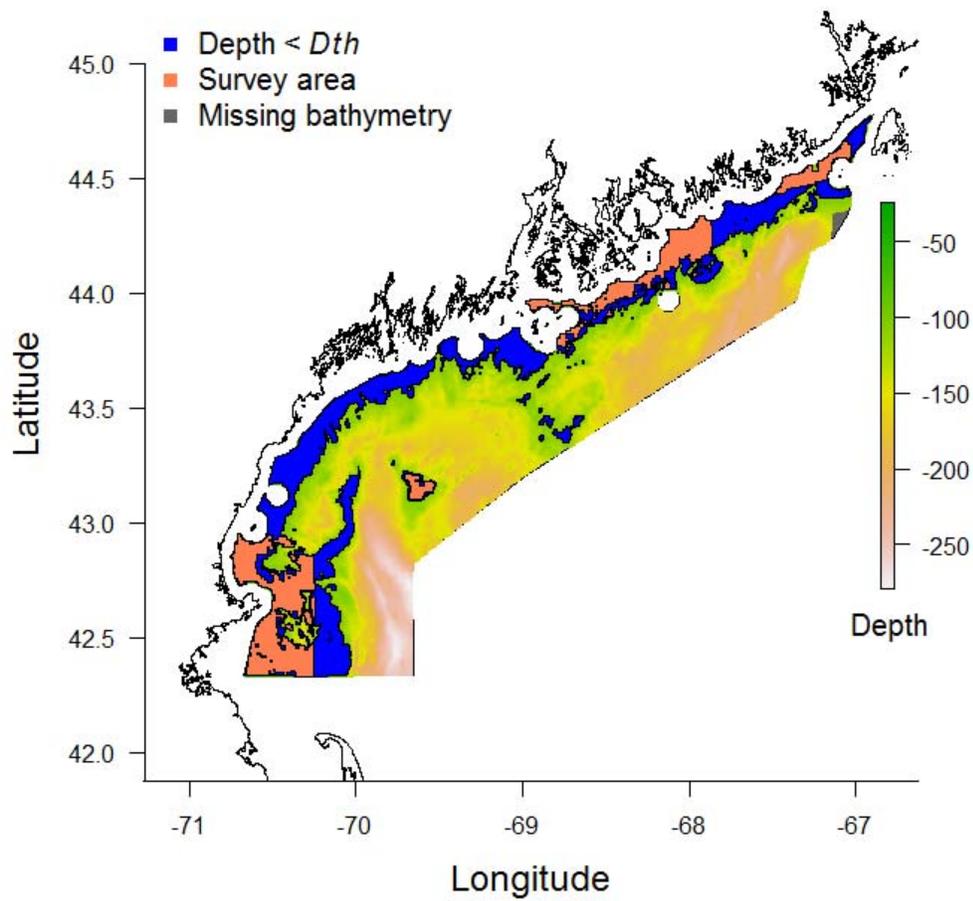


Figure 2: Survey area (pink) relative to the NGOM shallower than 100 m (*Dth*; blue). The survey area accounts for 37% of the NGOM shallower than 100 m.

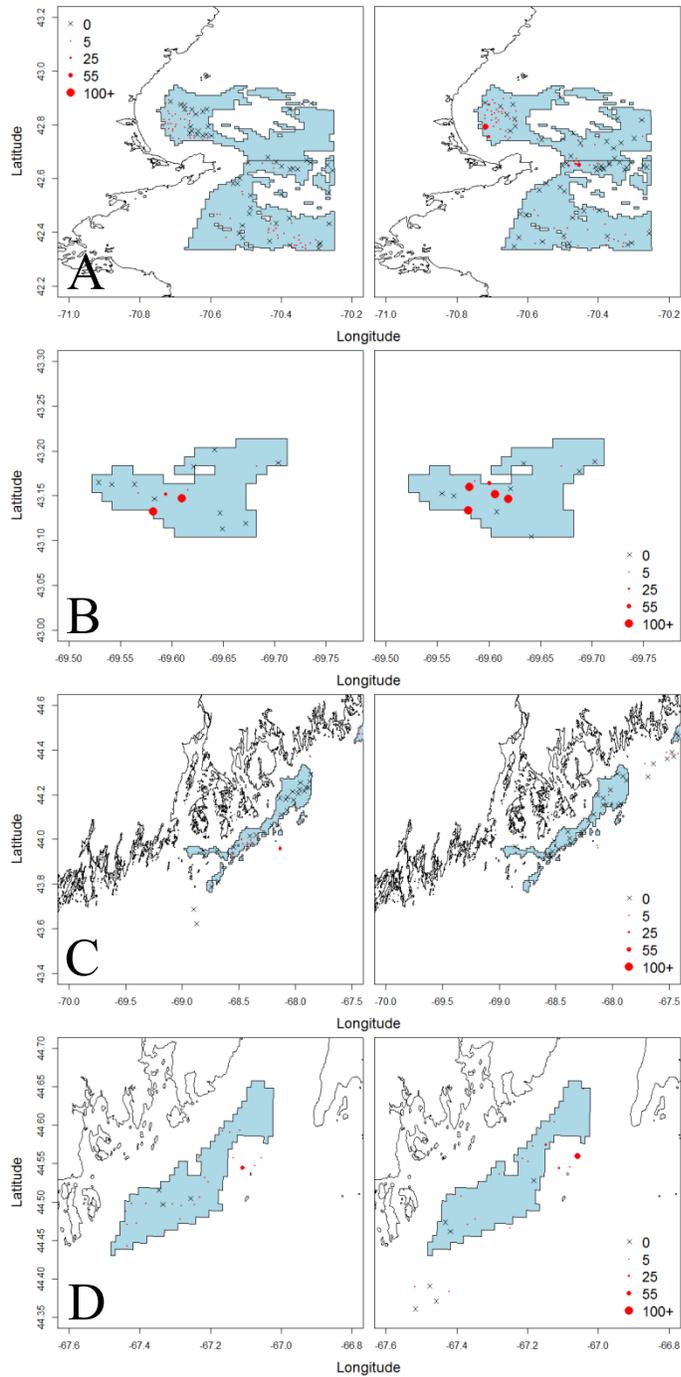


Figure 3: Distribution of survey scallop catch (all sizes) in 2009 (left panels) and 2012 (right panels). A: NCA and NSB; B: PB; C: MDR; and D: MSI.

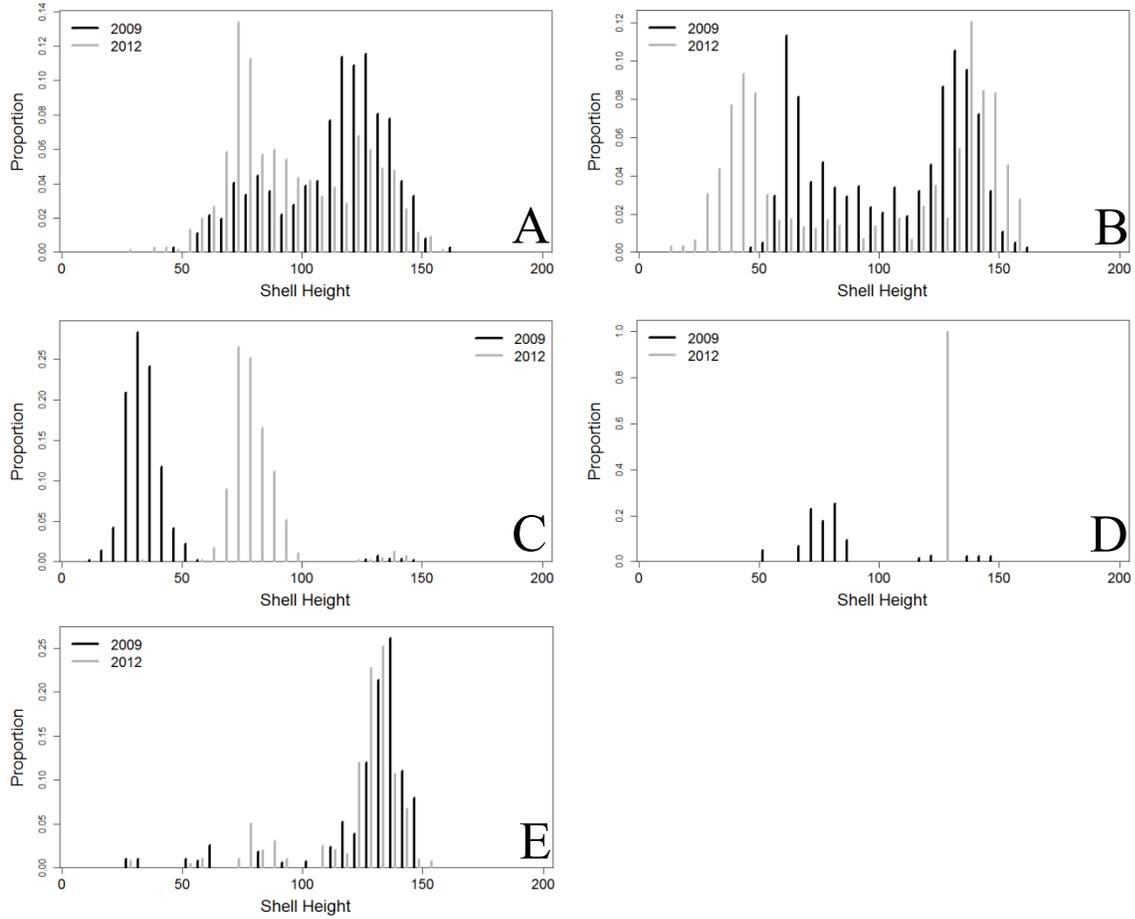


Figure 4: Shell height distribution in mm for each of the areas in 2009 and 2012. A: NCA; B: NSB; C: PB; D: MDR; E: MSI.

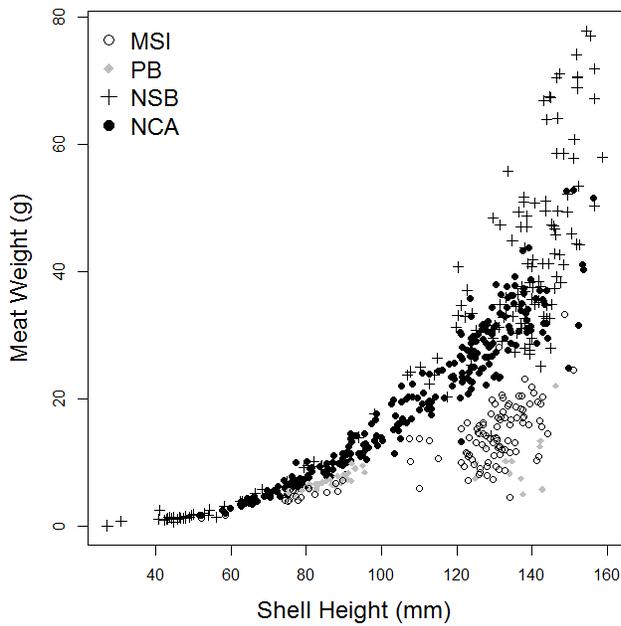


Figure 5: Relationship between shell height and meat weight in 2012 for the survey areas (excluding MDR).

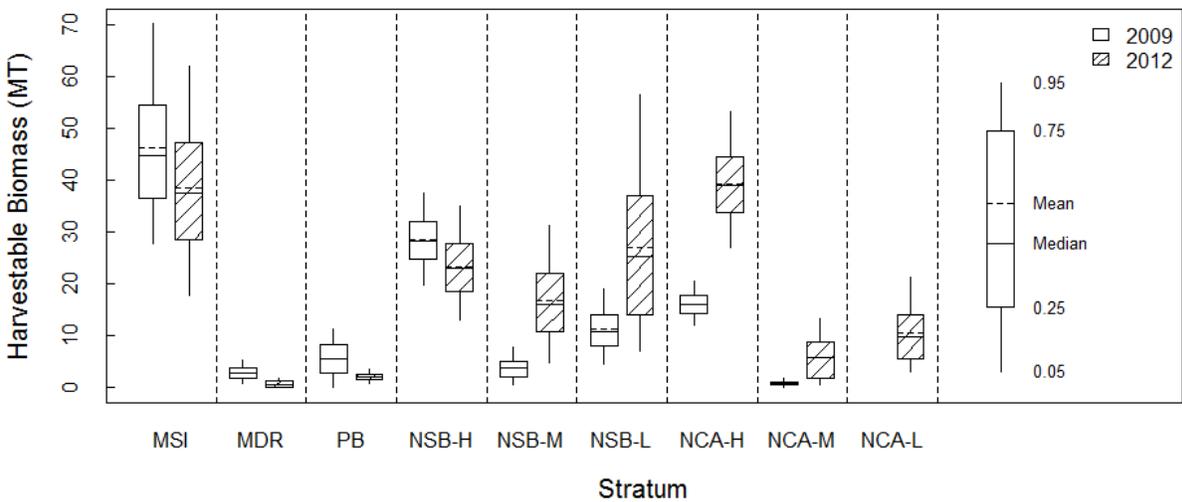


Figure 6: 2009 and 2012 harvestable biomass in NGOM survey strata (and substrata in the western region). H, M and L indicate expected high, medium and low density substrata.

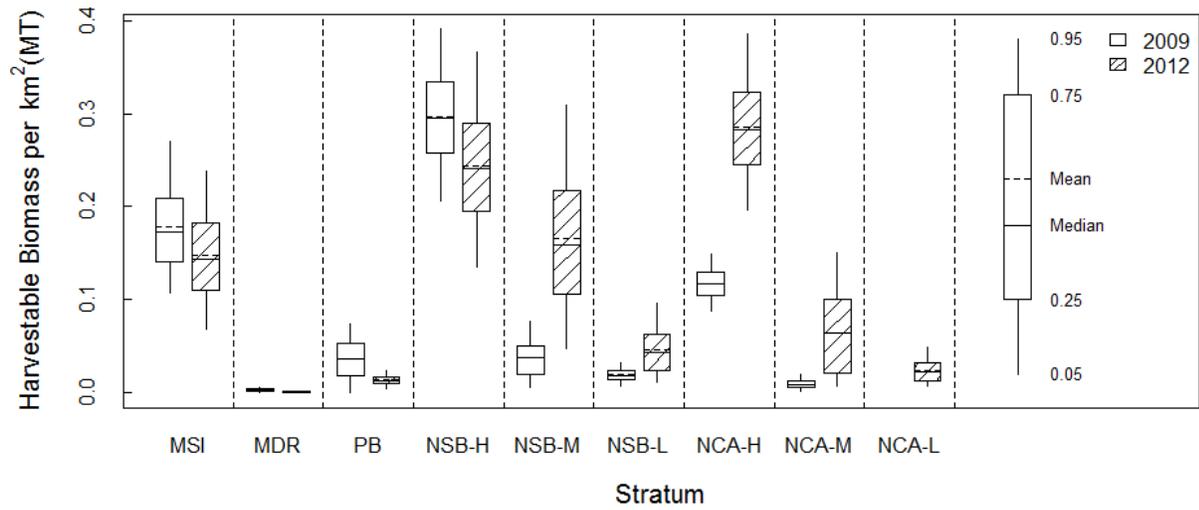


Figure 7: 2009 and 2012 harvestable density (in biomass per km<sup>2</sup>) in NGOM survey strata (and substrata in the western region). H, M and L indicate expected high, medium and low density substrata.

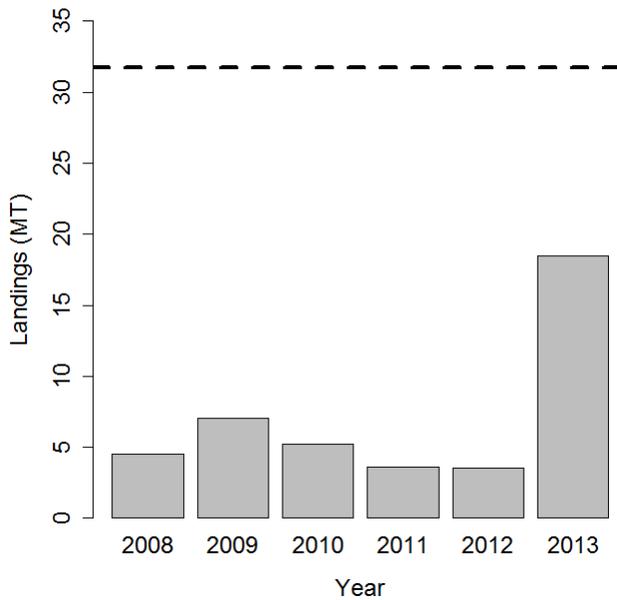


Figure 8: Landings history for the NGOM management area since its inception in 2008. Dashed line is the 31.75 MT quota.

## **Appendix B8. Relationships between chlorophyll and scallop recruitment potentially useful for stock projections and assessment modeling**

Kevin Friedland (NEFSC, Narragansett, RI), Deborah Hart and Burton Shank (NEFSC, Woods Hole, MA)

### Summary

Preliminary analyses of remote sensing and scallop dredge data suggest that recruitment to the yearling stage is influenced by summer phytoplankton bloom activity. Blooms in areas likely to influence Middle Atlantic spawning aggregations occur just prior to spring and summer spawning periods with larger bloom levels associated with high yearling settlement. The results of this analysis are encouraging and indicate further work developing techniques for predicting regional recruitment patterns based on chlorophyll concentrations is warranted. Such predictions are at spatial scales of interest to managers (e.g. rotational management areas) and might be used to improve management and profitability of the fishery.

### Introduction

This appendix describes an analysis of spring and summer bloom activity and scallop recruitment in the Middle Atlantic Bight during 1998 to 2012. The topic is important because uncertainty about recent and near-term scallop recruitment reduces the accuracy of stock projection analyses used to set harvest levels and to open rotational fishing areas. Recruitment of scallops in the region was represented by two indices based on survey data: i) a yearling index based on the abundance of 1-year old scallops, and ii) a 2-year old index. The two indices generally agree but there are notable disagreements for some year classes, indicating potential measurement errors in the survey data and/or variable survival between age-1 and 2. For the purpose of this summary, we will concentrate on the results of modeling recruitment to the yearling stage.

There are two spawning periods for Middle Atlantic Bight scallops. Spring spawning occurs mostly during May and fall spawning occurs in September. In line with these putative spawning periods, the spring and summer bloom dynamics of the Middle Atlantic Bight were characterized using chlorophyll *a* concentrations based on remote sensing data. The distribution of blooms was evaluated over a 0.5° spatial grid. Chlorophyll *a* concentrations were based on remote-sensing measurements made with the Sea-viewing Wide Field of View (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. The level-3 processed data, at 9 km and 8-day spatial and temporal resolutions, respectively, were obtained from the Ocean Color website ([oceancolor.gsfc.nasa.gov](http://oceancolor.gsfc.nasa.gov)). These two sensors provide an overlapping time series of chlorophyll *a* concentrations during the period 1998 to 2013. An analysis restricted to the overlapping period of data from both sensors revealed a systemic and consistent difference (relative bias) between them. We corrected for this bias with simple correction factors applied to MODIS data to approximate the mean levels of the SeaWiFS data. Chlorophyll *a* concentrations ( $\text{mg m}^{-3}$ ) were calculated by taking the average of the constituent pixel elements for each spatial-temporal cell.

The sequential averaging algorithm called STARS or “sequential *t*-test analysis of regime shifts” (Rodionov, 2004, 2006) was used to find the beginning and end of blooms (change

points) in the chlorophyll time series. A detected bloom could not exceed nine sample periods (approximately 2.4 months) based on analyses of climatological bloom patterns. Periods bracketed by positive and negative change points exceeding nine 8-day periods were considered to be ecologically different from discrete blooms. This method has been used in previous analyses of Northeast Shelf bloom patterns (Friedland et al., 2008, 2009) and elsewhere (Friedland and Todd, 2012).

We extracted statistics to characterize timing and magnitude of each bloom. Bloom start was defined as the day of initiation, which was the first day of the 8-day bloom period that exhibited bloom conditions. Bloom magnitude was the integral of the chlorophyll concentrations during the bloom period. In some years and locations, no distinct bloom period was detected by the STARS algorithm; when this occurred, bloom magnitude was taken as the integral of chlorophyll concentrations during the climatological (long-term average) bloom period based on average start and end dates for years with blooms.

## Results

Yearling scallop recruitment appears to be related to spring and summer phytoplankton blooms in the Middle Atlantic Bight. The area of highest correlation between spring chlorophyll concentrations and yearling recruitment was on the continental shelf off Long Island (Fig. 1a). In contrast, the area of the greatest correlative density between summer chlorophyll concentrations and yearling recruitment was off the New Jersey coast (Fig. 1b). Mean seasonal surface currents suggest that these blooms contributed to both water column chlorophyll and depositional particulate organic carbon in the areas of spawning scallops. These observations are consistent with the hypotheses that blooms either stimulate scallop spawning or support larval survival. Recruitment to age two was not related to the same spring and summer bloom patterns as yearling scallops due primarily to the change in population size of the 2011 year class between year-1 and 2.

## Future research

Refine models that predict scallop recruitment based on chlorophyll and predator data to improve estimates from stock assessment and projection models. Investigate statistical approaches to refine yearling recruitment indices. Develop complimentary models of bloom driven settlement and spatio-temporal predation pressure to ultimately stimulate recruitment of scallops to the fishery.

## References

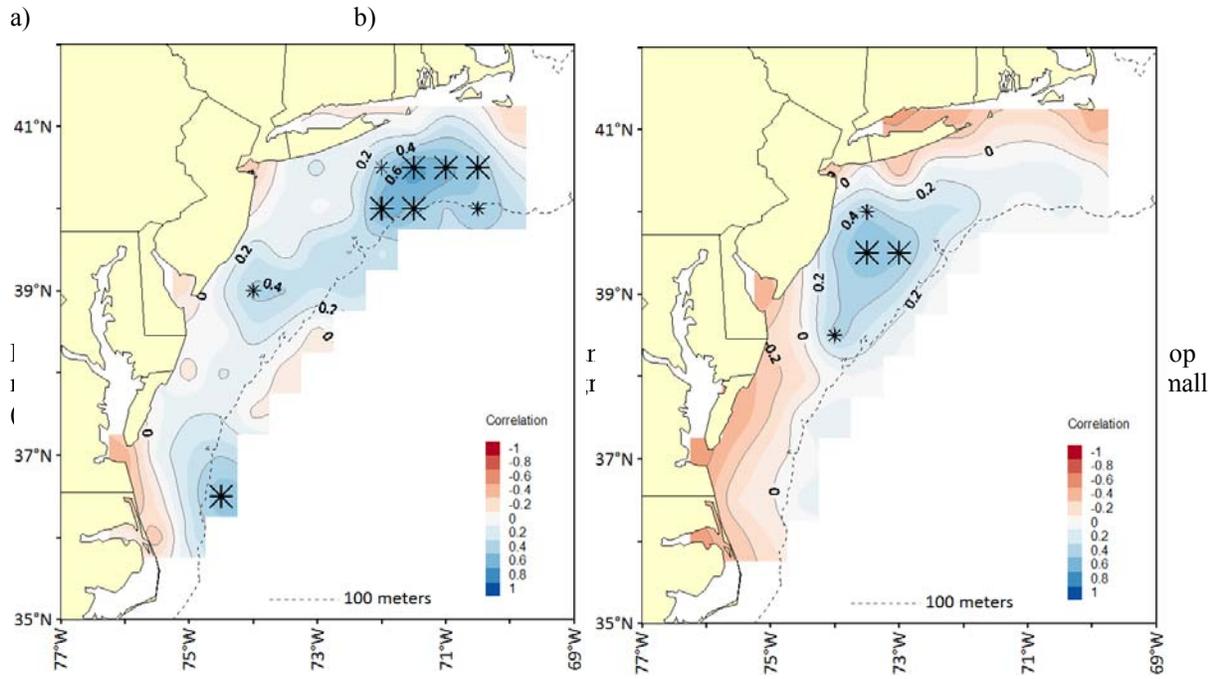
- Friedland, K.D., Hare, J.A., Wood, G.B., Col, L.A., Buckley, L.J., Mountain, D.G., Kane, J., Brodziak, J., Lough, R.G., Pilskaln, C.H., 2008. Does the fall phytoplankton bloom control recruitment of Georges Bank haddock, *Melanogrammus aeglefinus*, through parental condition? *Can J Fish Aquat Sci* 65, 1076-1086.
- Friedland, K.D., Hare, J.A., Wood, G.B., Col, L.A., Buckley, L.J., Mountain, D.G., Kane, J., Brodziak, J., Lough, R.G., Pilskaln, C.H., 2009. Reply to the comment by Payne et al. on "Does the fall phytoplankton bloom control recruitment of Georges Bank haddock, *Melanogrammus aeglefinus*, through parental condition?". *Can J Fish Aquat Sci* 66, 873-

877.

Friedland, K.D., Todd, C.D., 2012. Changes in Northwest Atlantic Arctic and Subarctic conditions and the growth response of Atlantic salmon. *Polar Biology* 35, 593-609.

Rodionov, S.N., 2004. A sequential algorithm for testing climate regime shifts. *Geophys Res Lett* 31, Doi 10.1029/2004gl019448.

Rodionov, S.N., 2006. Use of prewhitening in climate regime shift detection. *Geophys Res Lett* 33, Doi 10.1029/2006gl025904.



## Appendix B9. Technical documentation for the CASA length structured stock assessment model used in the SARC-59 sea scallop stock assessment.

Larry Jacobson, Northeast Fisheries Science Center, Woods Hole, MA.

[*This technical description is current through CASA version nc246.*]

The stock assessment model described here is based on Sullivan et al.'s (1990) CASA model.<sup>5</sup> CASA is entirely length-based with population dynamic calculations in terms of the number of individuals in each length group during each year. Age is almost completely irrelevant in model calculations. Unlike many other length-based stock assessment approaches, CASA is a dynamic, non-equilibrium model based on a forward simulation approach. CASA incorporates a very wide range of data with parameter estimation based on maximum likelihood. CASA can incorporate prior information about parameters such as survey catchability and natural mortality in a quasi-Bayesian fashion and MCMC evaluations are practical. The implementation described here was programmed in AD-Model Builder (Otter Research Ltd.).<sup>6</sup>

### Population dynamics

Time steps in the model are years, which are also used to tabulate catch and other data. Recruitment occurs at the beginning of each time step. All instantaneous rates in model calculations are annual ( $y^{-1}$ ). The number of years in the model  $n_y$  is flexible and can be changed easily (e.g. for retrospective analyses) by making a single change to the input data file. Millimeters are used to measure body size (e.g. sea scallop shell heights). Length-weight relationships should generally convert millimeters to grams. Model input data include a scalar that is used to convert the units for length-weight parameters (e.g. grams) to the units of the biomass estimates and landings data (e.g. mt). The units for catch and biomass are usually metric tons.

The definition of length groups (or length “bins”) is a key element in the CASA model and length-structured stock assessment modeling in general. Length bins are identified in CASA output by their lower bound and internally by their ordinal number. Calculations requiring information about length (e.g. length-weight) use the mid-length  $\ell_j$  of each bin. The user specifies the first length ( $L_{min}$ ) and the size of length bins ( $L_{bin}$ ). Based on these specifications, the model determines the number of length bins to be used in modeling as  $n_L = 1 + \text{int}[(L_\infty - L_{min})/L_{bin}]$ , where  $L_\infty$  is maximum asymptotic size supplied by the user, and  $\text{int}[x]$  is the integer part of  $x$ . The last length bin in the model is always a “plus-group” containing individuals  $L_\infty$  and larger. Specifications for length data used in tuning the model are separate (see below).

---

<sup>5</sup> Original programming in AD-Model Builder by G. Scott Boomer and Patrick J. Sullivan (Cornell University), who bear no responsibility for errors in the current implementation.

<sup>6</sup> AD-Model Builder can be used to calculate variances for any estimated or calculated quantity in a stock assessment model, based on the Hessian matrix with “exact” derivatives and the delta method.

## Growth

In population dynamics calculations, individuals in each size group grow (or not) at the beginning of the year, based on the annual growth transition matrix  $P_0(b,a)$  which measures the probability that a survivor in size bin  $a$  at the beginning of the previous year will grow to bin  $b$  at the beginning of the current year (columns index initial size and rows index subsequent size).<sup>7</sup> Growth probabilities do not include any adjustments for mortality and are applied to surviving scallops based on their original size in the preceding year.

There are two options for growth transition matrices. Under Option 1, a single annual growth matrix is calculated internally based on raw shell increment data:

$$P_0(b,a) = \frac{n(b|a)}{\sum_{j=a}^{n_L} n(j|a)}$$

where  $n(b|a)$  is the number of individuals that started at size  $a$  and grew to size  $b$  after one year in the raw size increment data.

Under option 2, the user specifies the number of transition matrices to be supplied in the input file and then assigns one of the matrices to each year in the model. All such growth matrices must have the same number of length groups. The number and size groups in the model and in the growth matrices should be large enough to accommodate the largest maximum size in any year. If growth varies such that maximum size in some time period is lower the maximum value, then the growth transition probabilities for that period of maximum size are set to one along the diagonal. For example, if there were five length groups in the model: [20,25), [25,30), [30,35), [35,40) and [40,45+] mm SH and the maximum size was 34 mm SH in period one and 44 mm SH in period two, the growth transition matrices might look like:

*Growth matrix for period 1*

		Starting size				
		[20, 25)	[25, 30)	[30, 35)	[35, 40)	[40, 45)
Ending size	[20, 25)	0.7	0	0	0	0
	[25, 30)	0.2	0.7	0	0	0
	[30, 35)	0.1	0.3	1	0	0
	[35, 40)	0	0	0	1	0
	[40, 45)	0	0	0	0	1

*Growth matrix for period 2*

		Starting size				
		[20, 25)	[25, 30)	[30, 35)	[35, 40)	[40, 45)
Ending size	[20, 25)	0.7	0	0	0	0
	[25, 30)	0.2	0.7	0	0	0
	[30, 35)	0.1	0.2	0.7	0	0
	[35, 40)	0	0.1	0.2	0.7	0
	[40, 45)	0	0	0.1	0.3	1

## Abundance, recruitment and mortality

Population abundance in each length bin during the first year of the model is:

<sup>7</sup> For clarity in bookkeeping, mortality and annual growth calculations are always based on the size on January 1.

$$N_{1,L} = N_1 \pi_{1,L}$$

where  $L$  is the size bin, and  $\pi_{1,L}$  is the initial population length composition expressed as

proportions so that  $\sum_{L=1}^{n_L} \pi_L = 1$ .  $N_1 = e^\eta$  is total abundance at the beginning of the first modeled

year and  $\eta$  is an estimable parameter. It is not necessary to estimate recruitment in the first year because recruitment is implicit in the product of  $N_1$  and  $\pi_L$ . The current implementation of CASA takes the initial population length composition as data supplied by the user, typically based on survey size composition data and a preliminary estimate of survey size-selectivity.

Abundance at length in years after the first is calculated:

$$\vec{N}_{y+1} = P_0 (\vec{N}_y \circ \vec{S}_y) + \vec{R}_{y+1}$$

where  $\vec{N}_y$  is a vector (length  $n_L$ ) of abundance in each length bin during year  $y$ ,  $P_0$  is the matrix ( $n_L \times n_L$ ) of annual growth probabilities  $P_0(\mathbf{b}, \mathbf{a})$ ,  $\vec{S}_y$  is a vector of length-specific survival fractions for year  $y$ ,  $\circ$  is the operator for an element-wise product, and  $\vec{R}_y$  is a vector holding length-specific abundance of new recruits at the beginning of year  $y$ .

Survival fractions are:

$$S_{y,L} = e^{-Z_{y,L}} = e^{-(M_{y,L} + F_{y,L} + I_{y,L})}$$

where  $Z_{y,L}$  is the total instantaneous mortality rate and  $M_{y,L}$  is the instantaneous rate for natural mortality (see below). Length-specific fishing mortality rates are  $F_{y,L} = F_y s_{y,L}$  where  $s_{y,L}$  is the size-specific selectivity<sup>8</sup> for fishing in year  $y$  (scaled to a maximum of one at fully recruited size groups),  $F_y$  is the fishing mortality rate on fully selected individuals. Fully recruited fishing mortality rates are  $F_y = e^{\phi + \delta_y}$  where  $\phi$  is an estimable parameter for the log of the geometric mean of fishing mortality in all years, and  $\delta_y$  is an estimable “dev” parameter.<sup>9</sup> The instantaneous rate for “incidental” mortality ( $I_{y,L}$ ) accounts for mortality due to contact with the fishing gear that does not result in any catch on deck (see below).<sup>10</sup> The degree of variability in dev parameters for fishing mortality, natural mortality and for other variables can be controlled by specifying variances or likelihood weights  $\neq 1$ , as described below.

Natural mortality rates are calculated:

$$M_{y,L} = u_L e^{\zeta + \xi_y} + p_L \psi_y g$$

where  $\vec{u}$  holds length-specific adjustments to the natural mortality rate for each length group (input by the user and assumed constant over time),  $\zeta$  is an estimable parameter measuring the mean log natural mortality rate during all years and  $\xi_y$  is an estimable year-specific dev parameter. The r.h.s. deals with density-dependent natural mortality which may be important in the population dynamics of small scallops after large recruitment events. In particular,  $p_L$  is a

8 In this context, “selectivity” describes the combined effects of all factors that affect length composition of catch or landings. These factors include gear selectivity, spatial overlap of the fishery and population, size-specific targeting, size-specific discard, etc.

9 Dev parameters are a special data type for estimable parameters in AD-Model Builder. Each set of dev parameters (e.g. for all recruitments in the model) is constrained to sum to zero. Because of the constraint, the sums  $\phi + \delta_y$  involving  $n_y + 1$  terms amount to only  $n_y$  parameters.

10. See the section on per recruit modeling below for formulas used to relate catch, landings and incidental mortality.

descending logistic function based on size (larger size groups experience less density dependent mortality),  $\psi_y$  is abundance of sea scallops used to calculate density dependent natural mortality,  $g=e^\alpha$  is a multiplier that converts from units of abundance to units of instantaneous mortality, and  $\alpha$  is an estimable scaling parameter. The logistic function is used to calculate the abundance that controls maximum density dependent mortality while reducing the importance of large individuals:

$$\psi_y = \sum_l p_L N_{y,L}$$

Where  $N_{y,L}$  is on January 1.

The logistic function in density dependent mortality calculations is calculated:

$$p_L = 1 - \frac{1}{1 + e^{-b(L-a)}}$$

where  $b$  is the slope parameter and  $a$  is the  $L_{50}$  parameter. The logistic curve is flat or decreasing with size because  $b=e^\alpha$  is  $> 0$  where  $\alpha$  is an estimable parameter. The  $L_{50}$  parameter is parameterized so that it automatically falls between the first and last sizes in the model:

$$a = L_{min} + (L_{max} - L_{min}) * \frac{e^\alpha}{1 + e^\alpha}$$

where  $L_{min}$  is the size at the bottom of the first size bin in the population model,  $L_{max}$  is the top of the last size bin, and  $\alpha$  is an estimable parameter.

Incidental mortality  $I_{y,L} = F_y u_L i$  is the product of fully recruited fishing mortality ( $F_y$ , a proxy for effective fishing effort, although nominal fishing effort might be a better predictor of incidental mortality), relative incidental mortality at length ( $u_L$ ) and a scaling parameter  $i$ , both of which are supplied by the user and not estimable in the model. Incidental mortality at length is supplied by the user as a vector ( $\vec{u}$ ) containing a value for each length group in the model. The model rescales the relative mortality vector so that the mean of the series is one.

Given abundance in each length group, natural mortality, and fishing mortality, predicted fishery catch-at-length in numbers is:

$$C_{y,L} = \frac{F_{y,L} (1 - e^{-Z_{y,L}}) N_{L,y}}{Z_{y,L}}$$

Total catch number during each year is  $C_y = \sum_{j=1}^{n_L} C_{y,L}$ . Catch data (in weight, numbers or as

length composition data) are understood to include landings ( $L_y$ ) and discards ( $d_y$ ) but to exclude losses to incidental mortality (i.e.  $C_y=L_y+d_y$ ).

Discard data are supplied by the user in the form of discarded biomass in each year or a discard rate for each year (or a combination of biomass levels and rates). In the current model, discards have the same selectivity as landed catch and size composition data for discards are not included in the input file.<sup>11</sup> It is important to remember that discard rates in CASA are defined the ratio of discards to landings ( $d/L$ ). The user may also specify a mortal discard fraction between zero and one if some discards survive. If the discard fraction is less than one, then the discarded biomass and discard rates in the model are reduced correspondingly. See the section on per recruit modeling below for formulas used to relate catch, landings and incidental mortality.

<sup>11</sup> The model will be modified in future to model discards and landing separately, and to use size composition data for discards.

Recruitment (the sum of new recruits in all length bins) at the beginning of each year after the first is calculated:

$$Ry = e^{\rho + \gamma_y}$$

where  $\rho$  is an estimable parameter that measures the geometric mean recruitment and the  $\gamma_y$  are estimable dev parameters that measure inter-annual variability in recruitment. As with natural mortality devs, the user specified variance or likelihood weight  $\neq 1$  can be used to help estimate recruitment deviations (see below).

Proportions of recruits in each length group are calculated based on a beta distribution  $B(w, r)$  over the first  $n_r$  length bins that is constrained to be concave down.<sup>12</sup> Proportions of new recruits in each size group are the same from year to year. Beta distribution coefficients must be larger than one for the shape of the distribution to be unimodal. Therefore,  $w=1+e^\omega$  and  $r=1+e^\rho$ , where  $\omega$  and  $\rho$  are estimable parameters. It is presumably better to calculate the parameters in this manner than as bounded parameters because there is likely to be less distortion of the Hessian for  $w$  and  $r$  values close to one and parameter estimation is likely to be more efficient.

Surplus production during each year of the model can be computed approximately from biomass and catch estimates (Jacobson et al., 2002):

$$P_t = B_{t+1} - B_t + C_t$$

In future versions of the CASA model, surplus production will be more calculated more accurately by projecting the population at the beginning of the year forward one year assuming only natural mortality.

### *Weight at length<sup>13</sup>*

The assumed body weight for size bins except the last is calculated using user-specified length-weight parameters and the middle of the size group. Different length-weight parameters are used for the population and for the commercial fishery. Mean body weight in the last size bin is read from the input file and can vary from year to year. Typically, mean weight in the last size bin for the population would be computed based on survey length composition data for large individuals and the population length –weight relationship. Mean weight in the last size bin for the fishery would be computed in the same manner based on fishery size composition data.

In principle, these calculations could be carried out in the model itself because all of the required information is available. In practice, it seems better to do the calculations externally and supply them to the model as inputs because of decisions that typically have to be made about smoothing the estimates and years with missing data.

### *Population summary variables*

Total abundance at the beginning of the year is the sum of abundance at length  $N_{y,L}$  at the beginning of the year. Average annual abundance for a particular length group is:

---

<sup>12</sup> Standard beta distributions used to describe recruit size distributions and in priors are often constrained to be unimodal in the CASA model. Beta distributions  $B(w, r)$  with mean  $\mu = w/w + r$  and variance

$\sigma^2 = wr / [(w+r)^2 (w+r+1)]$  are unimodal when  $w > 1$  and  $r > 1$ . See

[http://en.wikipedia.org/wiki/Beta\\_distribution](http://en.wikipedia.org/wiki/Beta_distribution) for more information.

<sup>13</sup> Model input data include a scalar that is used to convert the units for length-weight parameters (e.g. grams) to the units of the biomass estimates and landings data (e.g. mt).

$$\bar{N}_{y,L} = N_{y,L} \frac{1 - e^{-Z_{y,L}}}{Z_{y,L}}$$

The current implementation of the assessment model assumes different weight-at-length relationships for the stock and the fishery. Average stock biomass is computed using the population weight at length information.

Total stock biomass is:

$$B_y = \sum_{L=1}^{n_L} N_{y,L} w_L$$

where  $w_L$  is weight at length for the population on January 1. Total catch weight is:

$$W_y = \sum_{L=1}^{n_L} C_{y,L} w'_L$$

where  $w'_L$  is weight at length in the fishery.

$F_y$  estimates for two years are comparable only when the fishery selectivity in the model was the same in both years. A simpler exploitation index is calculated for use when fishery selectivity changes over time:

$$U_y = \frac{C_y}{\sum_{j=x}^{n_L} N_{y,L}}$$

where  $x$  is a user-specified length bin (usually at or below the first bin that is fully selected during all fishery selectivity periods).  $U_y$  exploitation indices from years with different selectivity patterns may be relatively comparable if  $x$  is chosen carefully.

Spawner abundance in each year is ( $T_y$ ) is computed:

$$T_y = \sum_{L=1}^{n_L} N_{y,L} e^{-\tau Z_{y,L}} g_L$$

Where  $0 \leq \tau \leq 1$  is the fraction of the year elapsed before spawning occurs (supplied by the user). Maturity at length ( $g_L$ ) is from an ascending logistic curve:

$$g_L = \frac{1}{1 + e^{a-bL}}$$

with parameters  $a$  and  $b$  supplied by the user. Spawner biomass is computed using the population length-weight values.

Egg production ( $S_y$ ) in each year is computed:

$$S_y = \sum_{L=1}^{n_L} N_{y,L} e^{-\tau Z_{y,L}} g_L x_L$$

where:

$$x_L = cL^v$$

Where the fecundity parameters ( $c$  and  $v$ ) for fecundity are supplied by the user. Fecundity parameters per se include no adjustments for maturity or survival. They should represent reproductive output for a spawner of given size.

### Fishery and survey selectivity

The current implementation of CASA includes six options for calculating fishery and survey selectivity patterns. Fishery selectivity may differ among “fishery periods” defined by

the user. Selectivity patterns that depend on length are calculated using lengths at the mid-point of each bin ( $\ell$ ). After initial calculations (described below), selectivity curves are rescaled to a maximum value of one.

Option 1 is a flat with  $s_L=1$  for all length bins. Option 2 is an ascending logistic curve:

$$s_{y,\ell} = \frac{1}{1 + e^{A_Y - B_Y \ell}}$$

Option 3 is an ascending logistic curve with a minimum asymptotic minimum size for small size bins on the left.

$$s_{y,\ell} = \left( \frac{1}{1 + e^{A_Y - B_Y \ell}} \right) (1 - D_Y) + D_Y$$

Option 4 is a descending logistic curve:

$$s_{y,\ell} = 1 - \frac{1}{1 + e^{A_Y - B_Y \ell}}$$

Option 5 is a descending logistic curve with a minimum asymptotic minimum size for large size bins on the right:

$$s_{y,\ell} = \left( 1 - \frac{1}{1 + e^{A_Y - B_Y \ell}} \right) (1 - D_Y) + D_Y$$

Option 6 is a double logistic curve used to represent “domed-shape” selectivity patterns with highest selectivity on intermediate size groups:

$$s_{y,\ell} = \left( \frac{1}{1 + e^{A_Y - B_Y \ell}} \right) \left( 1 - \frac{1}{1 + e^{D_Y - G_Y \ell}} \right)$$

The coefficients for selectivity curves  $A_Y$ ,  $B_Y$ ,  $D_Y$  and  $G_Y$  carry subscripts for time because they may vary between fishery selectivity periods defined by the user. All options are parameterized so that the coefficients  $A_Y$ ,  $B_Y$ ,  $D_Y$  and  $G_Y$  are positive. Under options 3 and 5,  $D_Y$  is a proportion that must lie between 0 and 1.

Depending on the option, estimable selectivity parameters may include  $\alpha$ ,  $\beta$ ,  $\delta$  and  $\gamma$ . For options 2, 4 and 6,  $A_Y = e^{\alpha_Y}$ ,  $B_Y = e^{\beta_Y}$ ,  $D_Y = e^{\delta_Y}$  and  $G_Y = e^{\gamma_Y}$ . Options 3 and 5 use the same conventions for  $A_Y$  and  $B_Y$ , however, the coefficient  $D_Y$  is a proportion estimated as a logit-transformed parameter (i.e.  $\delta_Y = \ln[D_Y/(1-D_Y)]$ ) so that:

$$D_Y = \frac{e^{\delta_Y}}{1 + e^{\delta_Y}}$$

The user can choose, independently of all other parameters, to either estimate each fishery selectivity parameter or to keep it at its initial value. Under Option 2, for example, the user can estimate the intercept  $\alpha_Y$ , while keep the slope  $\beta_Y$  at its initial value.

### *Per recruit modeling*

The per recruit model in CASA uses the same population model as in other model calculations under conditions identical to the last year in the model. It is a standard length-based approach except that discard and incidental mortality are accommodated in all calculations. In per recruit calculations, fishing mortality rates and associated yield estimates are understood to include landings and discard mortality, but to exclude incidental mortality. Thus, landings per recruit  $L$  are:

$$L = \frac{C}{(1 + \Delta)}$$

where  $C$  is total catch (yield) per recruit and  $\Delta$  is the ratio of discards  $D$  to landings in the last year of the model. Discards per recruit are calculated:

$$D = \Delta L$$

Losses due to incidental mortality ( $G$ ) are calculated:

$$G = \frac{I(1 - e^{-Z})B}{Z}$$

$$= IK$$

where  $I = Fu$  is the incidental mortality rate,  $u$  is a user-specified multiplier (see above) and  $B$  is stock biomass per recruit. Note that  $C = FK$  so that  $K = C/F$ . Then,

$$G = \frac{FuC}{F}$$

$$G = uC$$

The model will estimate a wide variety ( $F_{\%SBR}$ ,  $F_{max}$  and  $F_{0.1}$ ) of per recruit model reference points as parameters. For example,

$$F_{\%SBR} = e^{\theta_j}$$

where  $F_{\%SBR}$  is the fishing mortality reference point that provides a user specified percentage of maximum SBR.  $\theta_j$  is the model parameter for the  $j^{th}$  reference point.

A complete per recruit output table is generated in all model runs that can be used for evaluating the shape of YPR and SBR curves, including the existence of particular reference points.

Per recruit reference points are time consuming to estimate and it is usually better to estimate them after other more important population dynamics parameters are estimated. Phase of estimation can be controlled individually for %SBR,  $F_{MAX}$  and  $F_{0.1}$  so that per recruit calculations can be delayed as long as possible. If the phase is set to zero or a negative integer, then the reference point will not be estimated. As described below, estimation of  $F_{max}$  always entails an additional phase of estimation. For example, if the phase specified for  $F_{max}$  is 2, then the parameter will be estimated initially in phase 2 and finalized the last phase (phase  $\geq 3$ ). This is done so that the estimate from phase 2 can be used as an initial value in a slightly different goodness of fit calculation during the latter phase.

Per recruit reference points should have no effect on other model estimates. Residuals (calculated – target) for %SBR,  $F_{0.1}$  and  $F_{max}$  reference points should always be very close to zero. Problems may arise, however, if reference points (particularly  $F_{max}$ ) fall on the upper bound for fishing mortality. In such cases, the model will warn the user and advise that the offending reference points should not be estimated. *It is good practice to run CASA with reference point calculations turned on and then off to see if biomass and fishing mortality estimates change.*

The user specifies the number of estimates required and the target %SBR level for each. For example, the target levels for four %SBR reference points might be 0.2, 0.3, 0.4 and 0.5 to estimate  $F_{20\%}$ ,  $F_{30\%}$ ,  $F_{40\%}$  and  $F_{50\%}$ . The user has the option of estimating  $F_{max}$  and/or  $F_{0.1}$  as model parameters also but it is not necessary to supply target values.

## Tuning and goodness of fit

There are two steps in calculating the negative log likelihood (NLL) used to measure how well the model fits each type of data. The first step is to calculate the predicted values for data. The second step is to calculate the NLL of the data given the predicted value. The overall goodness of fit measure for the model is the weighted sum of NLL values for each type of data and each constraint:

$$\Lambda = \sum \lambda_j L_j$$

where  $\lambda_j$  is a weighting factor for data set  $j$  (usually  $\lambda_j=1$ , see below), and  $L_j$  is the NLL for the data set. The NLL for a particular data is itself usually a weighted sum:

$$L_j = \sum_{i=1}^{n_j} \psi_{j,i} L_{j,i}$$

where  $n_j$  is the number of observations,  $\psi_{j,i}$  is an observation-specific weight (usually  $\psi_{j,i}=1$ , see below), and  $L_{j,i}$  is the NLL for a single observation.

Maximum likelihood approaches reduce the need to specify *ad-hoc* weighting factors ( $\lambda$  and  $\phi$ ) for data sets or single observations, because weights can often be taken from the data (e.g. using CVs routinely calculated for bottom trawl survey abundance indices) or estimated internally along with other parameters. In addition, robust maximum likelihood approaches (see below) may be preferable to simply down-weighting an observation or data set. However, despite subjectivity and theoretical arguments against use of *ad-hoc* weights, it is often useful in practical work to manipulate weighting factors, if only for sensitivity analysis or to turn an observation off entirely. Observation specific weighting factors are available for most types of data in the CASA model.

### Missing data

Availability of data is an important consideration in deciding how to structure a stock assessment model. The possibility of obtaining reliable estimates will depend on the availability of sufficient data. However, NLL calculations and the general structure of the CASA model are such that missing data can usually be accommodated automatically. With the exception of catch data (which must be supplied for each year, even if catch was zero), the model calculates that NLL for each datum that is available. No NLL calculations are made for data that are not available and missing data do not generally hinder model calculations.

### Likelihood kernels

Log likelihood calculations in the current implementation of the CASA model use log likelihood “kernels” or “concentrated likelihoods” that omit constants. The constants can be omitted because they do not affect slope of the NLL surface, final point estimates for parameters or asymptotic variance estimates. For data with normally distributed measurement errors, the complete NLL for one observation is:

$$L = \ln(\sigma) + \ln(\sqrt{2\pi}) + 0.5 \left( \frac{x-u}{\sigma} \right)^2$$

The constant  $\ln(\sqrt{2\pi})$  can always be omitted. If the standard deviation is known or assumed known, then  $\ln(\sigma)$  can be omitted as well because it is a constant that does not affect derivatives. In such cases, the concentrated NLL is:

$$L = 0.5 \left( \frac{x - \mu}{\sigma} \right)^2$$

If there are  $N$  observations with possible different variances (known or assumed known) and possibly different expected values:

$$L = 0.5 \sum_{i=1}^N \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2$$

If the standard deviation for a normally distributed quantity is not known and is estimated (implicitly or explicitly) by the model, then one of two equivalent calculations is used. Both approaches assume that all observations have the same variance and standard deviation. The first approach is used when all observations have the same weight in the NLL:

$$L = 0.5N \ln \left[ \sum_{i=1}^N (x_i - u)^2 \right]$$

The second approach is equivalent but used when the weights for each observation ( $w_i$ ) may differ:

$$L = \sum_{i=1}^N w_i \left[ \ln(\sigma) + 0.5 \left( \frac{x_i - u}{\sigma} \right)^2 \right]$$

In the latter case, the maximum likelihood estimator:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x})^2}{N}}$$

(where  $\hat{x}$  is the average or predicted value from the model) is used explicitly for  $\sigma$ . The maximum likelihood estimator is biased by  $N/(N-d_f)$  where  $d_f$  is degrees of freedom for the model. The bias may be significant for small sample sizes, which are common in stock assessment modeling, but  $d_f$  is usually unknown.

If data  $x$  have lognormal measurement errors, then  $\ln(x)$  is normal and  $L$  is calculated as above. In some cases it is necessary to correct for bias in converting arithmetic scale means to log scale means (and *vice-versa*) because  $\bar{x} = e^{\bar{\chi} + \sigma^2/2}$  where  $\chi = \ln(x)$ . It is often convenient to convert arithmetic scale CVs for lognormal variables to log scale standard deviations using  $\sigma = \sqrt{\ln(1 + CV^2)}$ .

For data with multinomial measurement errors, the likelihood kernel is:

$$L = n \sum_{i=1}^n p_i \ln(\theta_i) - K$$

where  $n$  is the known or assumed number of observations (the “effective” sample size),  $p_i$  is the proportion of observations in bin  $i$ , and  $\theta_i$  is the model’s estimate of the probability of an observation in the bin. For surveys,  $\theta_i$  is adjusted for mortality up to the date of the survey and for growth up to the mid-point of the month in which the survey occurs. For fisheries,  $\theta_i$  accommodates all of the mortality during the current year and is adjusted for growth during January 1 to mid-July. The constant  $K$  is used for convenience to make  $L$  easier to interpret. It measures the lowest value of  $L$  that could be achieved if the data fit matched the model’s expectations exactly:

$$K = n \sum_{i=1}^n p_i \ln(p_i)$$

For data  $x$  that have measurement errors with expected values of zero from a gamma distribution:

$$L = (\gamma - 1) \ln\left(\frac{x}{\beta}\right) - \frac{x}{\beta} - \ln(\beta)$$

where  $\beta > 0$  and  $\gamma > 0$  are gamma distribution parameters in the model. For data that lie between zero and one with measurement errors from a beta distribution:

$$L = (p - 1) \ln(x) + (q - 1) \ln(1 - x)$$

where  $p > 0$  and  $q > 0$  are parameters in the model.

In CASA model calculations, distributions are usually described in terms of the mean and CV. Normal, gamma and beta distribution parameters can be calculated mean and CV by the method of moments.<sup>14</sup> Means, CV's and distributional parameters may, depending on the situation, be estimated in the model or specified by the user.

The NLL for a datum  $x$  from gamma distribution is:

$$L = (1 - k) * \ln(x) + \frac{x}{\theta} + \ln[\Gamma(k)] + k \ln(\theta)$$

where  $k$  is the shape parameter and  $\theta$  is the scale parameter. The last two terms on the right are constants and can be omitted if  $k$  and  $\theta$  are not estimated. Under these circumstances,

$$L = (1 - k) * \ln(x) + \frac{x}{\theta}$$

## Robust methods

Goodness of fit for survey data may be calculated using a “robust” maximum likelihood method instead of the standard method that assumes lognormal measurement errors. The robust method may be useful when survey data are noisy or include outliers.

Robust likelihood calculations in CASA assume that measurement errors are from a Student's  $t$  distribution with user-specified degrees of freedom  $d_f$ . Degrees of freedom are specified independently for each observation so that robust calculations can be carried out for as many (or as few) cases as required. The  $t$  distribution is similar to the normal distribution for  $d_f \geq 30$ . As  $d_f$  is reduced, the tails of the  $t$  distribution become fatter so that outliers have higher probability and less effect on model estimates. If  $d_f = 0$ , then measurement errors are assumed in the model to be normally distributed.

The first step in robust NLL calculations is to standardize the measurement error residual  $t = (x - \bar{x})/\sigma$  based on the mean and standard deviation. Then:

---

<sup>14</sup> Parameters for standard beta distributions  $B(w,r)$  with mean  $\mu = w/(w+r)$  and variance

$\sigma^2 = wr/[(w+r)^2(w+r+1)]$  are calculated from user-specified means and variances by the method of moments. In particular,  $w = \mu[\mu(1-\mu)/\sigma^2 - 1]$  and  $r = (1-\mu)[\mu(1-\mu)/\sigma^2 - 1]$ . Not all combinations of  $\mu$  and  $\sigma^2$  are feasible. In general, a beta distribution exists for combinations of  $\mu$  and  $\sigma^2$  if  $0 < \mu < 1$  and  $0 < \sigma^2 < \mu(1-\mu)$ . Thus, for a user-specified mean  $\mu$  between zero and one, the largest feasible variance is  $\sigma^2 < \mu(1-\mu)$ . These conditions are used in the model to check user-specified values for  $\mu$  and  $\sigma^2$ . See [http://en.wikipedia.org/wiki/Beta\\_distribution](http://en.wikipedia.org/wiki/Beta_distribution) for more information.

$$L = \ln\left(1 + \frac{t^2}{d_f}\right) \left(1 - \frac{1-d_f}{2}\right) - \frac{\ln(d_f)}{2}$$

### *Catch weight data*

Catch data (landings plus discards) are assumed to have normally distributed measurement errors with a user specified CV. The standard deviation for catch weight in a particular year is  $\sigma_y = \kappa \hat{C}_y$ , where “^” indicates that the variable is a model estimate and errors in catch are assumed to be normally distributed. The standardized residual used in computing NLL for a single catch observation and in making residual plots is  $r_y = (C_y - \hat{C}_y) / \sigma_y$ .

### **Specification of landings, discards, catch**

Landings, discard and catch data are in units of weight and are for a single or “composite” fishery in the current version of the CASA model. The estimated fishery selectivity is assumed to apply to the discards so that, in effect, the length composition of catch, landings and discards are the same.

Discards are from external estimates ( $d_t$ ) supplied by the user. If  $d_t \geq 0$ , then the data are used as the ratio of discard to landed catch so that:

$$D_t = L_t \Delta_t$$

where  $\Delta_t = D_t / L_t$  is the ratio of discard and landings (a.k.a.  $d/K$  ratios) for each year. If  $d_t < 0$  then the data are treated as discard in units of weight:

$$D_t = \text{abs}(d_t).$$

In either case, total catch is the sum of discards and landed catch ( $C_t = L_t + D_t$ ). It is possible to use discards in weight  $d_t < 0$  for some years and discard as proportions  $d_t > 0$  for other years in the same model run.

If catches are estimated (see below) so that the estimated catch  $\hat{C}_t$  does not necessarily equal observed landings plus discard, then estimated landings are computed:

$$\hat{L}_t = \frac{\hat{C}_t}{1 + \Delta_t}$$

Estimated discards are:

$$\hat{D}_t = \Delta_t \hat{L}_t.$$

Note that  $\hat{C}_t = \hat{L}_t + \hat{D}_t$  as would be expected.

### *Fishery length composition data*

Data describing numbers or relative numbers of individuals at length in catch data (fishery catch-at-length) are modeled as multinomial proportions  $c_{y,L}$ :

$$c_{y,L} = \frac{C_{y,L}}{\sum_{j=1}^{n_L} C_{y,j}}$$

The NLL for the observed proportions in each year is computed based on the kernel for the multinomial distribution, the model’s estimate of proportional catch-at-length ( $\hat{c}_y$ ) and an

estimate of effective sample size  ${}^cN_y$  supplied by the user. Care is required in specifying effective sample sizes, because catch-at-length data typically carry substantially less information than would be expected based on the number of individuals measured. Typical conventions make  ${}^cN_y \leq 200$  (Fournier and Archibald, 1982) or set  ${}^cN_y$  equal to the number of trips or tows sampled (Pennington et al., 2002). Effective sample sizes are sometimes chosen based on goodness of fits in preliminary model runs (Methot, 2000; Butler et al., 2003).

Standardized residuals are not used in computing NLL fishery length composition data. However, approximate standardized residuals  $r_y = (c_{y,L} - \hat{c}_{y,L}) / \sigma_{y,L}$  with standard deviations  $\sigma_{y,L} = \sqrt{\hat{c}_{y,L}(1 - \hat{c}_{y,L}) / {}^cN_y}$  based on the theoretical variance for proportions are computed for use in making residual plots.

### *Survey index data*

In CASA model calculations, “survey indices” are data from any source that reflect relative proportional changes in an underlying population state variable. In the current version, surveys may measure stock abundance at a particular point in time (e.g. when a survey was carried out), stock biomass at a particular point in time, or numbers of animals that dies of natural mortality during a user-specified period. For example, the first option is useful for bottom trawl surveys that record numbers of individuals, the second option is useful for bottom trawl surveys that record total weight, and the third option is useful for survey data that track trends in numbers of animals that died due to natural mortality (e.g. survey data for sea scallop “clappers”). Survey data that measure trends in numbers dead due to natural mortality can be useful in modeling time trends in natural mortality. In principle, the model will estimate model natural mortality and other parameters so that predicted numbers dead and the index data match in either relative or absolute terms.

In the current implementation of the CASA model, survey indices are assumed to be linear indices of abundance or biomass so that changes in the index (apart from measurement error) are assumed due to proportional changes in the population. Nonlinear commercial catch rate data are handled separately (see below). Survey index and fishery length composition data are handled separately from trend data (see below). Survey data may or may not have corresponding length composition information.

In general, survey index data give one number that summarizes some aspect of the population over a wide range of length bins. Selectivity parameters measure the relative contribution of each length bin to the index. Options and procedures for estimating survey selectivity patterns are the same as for fishery selectivity patterns, but survey selectivity patterns are not allowed to change over time.

NLL calculations for survey indices use predicted values calculated:

$$\hat{I}_{k,y} = q_k A_{k,y}$$

where  $q_k$  is a scaling factor for survey index  $k$ , and  $A_{k,y}$  is stock available to the survey. The scaling factor is computed using the maximum likelihood estimator:

$$q_k = e^{\frac{\sum_{i=1}^{N_k} \left[ \ln \left( \frac{I_{k,i}}{A_{k,i}} \right) / \sigma_{k,i}^2 \right]}{\sum_{j=1}^{N_k} \left( 1 / \sigma_{k,j}^2 \right)}}$$

where  $N_v$  and  $\sigma_{k,j}^2$  is the log scale variance corresponding to the assumed CV for the survey observation.<sup>15</sup>

Available stock for surveys measuring trends in abundance or biomass is calculated:

$$A_{k,y} = \sum_{L=1}^{n_L} s_{k,L} N_{y,L} e^{-Z_{y,L} \tau_{k,y}}$$

where  $s_{k,L}$  is size-specific selectivity of the survey,  $\tau_{k,y} = J_{k,y} / 365$ ,  $J_{k,y}$  is the Julian date of the survey in year  $y$ , and  $e^{-Z_{y,L} \tau_{k,y}}$  is a correction for mortality prior to the survey. Available biomass is calculated in the same way except that body weights  $w_L$  are included in the product on the right hand side.

Available stock for indices that track numbers dead by natural mortality is:

$$A_{k,y} = \sum_{L=1}^{n_k} s_{k,L} \tilde{M}_{y,L} \bar{N}_{y,L}$$

where  $\bar{N}_{y,L}$  is average abundance during the user-specified period of availability and  $\tilde{M}_{y,L}$  is the instantaneous rate of natural mortality for the period of availability. Average abundance during the period of availability is:

$$\bar{N}_{y,L} = \frac{\tilde{N}_{y,L} (1 - e^{-\tilde{Z}_{y,L}})}{\tilde{Z}_{y,L}}$$

where  $\tilde{N}_{y,L} = N_{y,L} e^{-Z\Delta}$  is abundance at elapsed time of year  $\Delta = \tau_{k,y} - \nu_k$ ,  $\nu_k = j_k / 365$ , and  $j_k$  is the user-specified duration in days for the period of availability. The instantaneous rates for total  $\tilde{Z}_{y,L} = Z_{y,L} (\tau_{k,y} - \nu_k)$  and natural  $\tilde{M}_{y,L} = M_{y,L} (\tau_{k,y} - \nu_k)$  mortality are also adjusted to correspond to the period of availability. In using this approach, the user should be aware that the length based selectivity estimated by the model for the dead animal survey ( $s_{k,L}$ ) is conditional on the assumed pattern of length-specific natural mortality ( $\bar{u}$ ) which was specified as data in the input file.

NLL calculations for survey index data assume that log scale measurement errors are either normally distributed (default approach) or from a  $t$  distribution (robust estimation approach). In either case, log scale measurement errors are assumed to have mean zero and log scale standard errors either estimated internally by the model or calculated from the arithmetic CVs supplied with the survey data.

---

<sup>15</sup> Scaling factors in previous versions were calculated  $q_s = e^{\varpi_s}$  where  $\varpi_s$  is an estimable and survey-specific parameter. However, prior distributions were shown to have a strong effect on the parameters such that the relationship  $N=qA$  did not hold. The approach in the current model avoids this problem.

The standardized residual used in computing NLL for one survey index observation is  $r_{k,y} = \ln(I_{k,y}/\hat{I}_{k,y})/\sigma_{k,y}$  where  $I_{k,y}$  is the observation. The standard deviations  $\sigma_{k,y}$  will vary among surveys and years if CVs are used to specify the variance of measurement errors. Otherwise a single standard deviation is estimated internally for the survey as a whole.

#### *Survey length composition data*

Length bins for fishery and survey length composition data are flexible and the flexibility affects goodness of fit calculations in ways that may be important to consider in some applications. The user specifies the starting size (bottom of first bin) and number of bins used for each type of fishery and survey length composition. The input data for each length composition record identifies the first/last length bins to be used and whether they are plus groups that should include all smaller/larger length groups in the data and population model when calculating goodness of fit. Goodness of fit calculations are carried out over the range of lengths specified by the user. Thus length data in the input file may contain large or small size bins that are ignored in goodness of fit calculations. As described above, the starting size and bin size for the population model are specified separately. In the ideal and simplest case, the minimum size and same length bins are used for the population and for all length data. However, as described below, length specifications in data and the population model may differ.

For example, the implicit definitions of plus groups in the model and data may differ. If the first bin used for length data is a plus group, then the first bin will contain the sum of length data from the corresponding and smaller bins of the original length composition record. However, the first bin in the population model is never a plus group. Thus, predicted values for a plus group will contain the sum of the corresponding and smaller bins in the population. The observed and predicted values will not be perfectly comparable if the starting sizes for the data and population model differ. Similarly, if the last bin in the length data is a plus group, it will contain original length composition data for the corresponding and all larger bins. Predicted values for a plus group in the population will be the sum for the corresponding bin and all larger size groups in the population, implicitly including sizes  $> L_{\infty}$ . The two definitions of the plus group will differ and goodness of fit calculation may be impaired if the original length composition data does not include all of the large individuals in samples.

In the current version of the CASA model, the size of length composition bins must be  $\geq L_{bin}$  in the population model (this constraint will be removed in later versions). Ideally, the size of data length bins is the same or a multiple of the size of length bins in the population. However, this is not required and the model will prorate the predicted population composition for each bin into adjacent data bins when calculating goodness of fit. With a 30-34 mm population bin and 22-31 and 32-41 mm population bins, for example, the predicted proportion in the population bin would be prorated so that 2/5 was assigned to the first data bin and 3/5 was assigned to the second data bin. This proration approach is problematic when it is used to prorate the plus group in the population model into two data bins because it assumes that abundance is uniform over lengths within the population group. The distribution of lengths in a real population might be far from uniform between the assumed upper and lower bounds of the plus group.

The first bin in each length composition data record must be  $\geq L_{min}$  which is the smallest size group in the population model. If the last data bin is a plus group, then the *lower* bound of the last data bin must be  $\leq$  the upper bound of the last population bin. Otherwise, if the last data bin is not a plus group, the *upper* bound of the last data bin must be  $\leq$  the upper bound of the

population bin.

NLL calculations for survey length composition data are similar to calculations for fishery length composition data. Surveys index data may measure trends in stock abundance or biomass but survey length composition data are always for numbers (not weight) of individuals in each length group. Survey length composition data represent a sample from the true stock which is modified by survey selectivity, sampling errors and, if applicable, errors in recording length data. For example, with errors in length measurements, individuals belonging to length bin  $j$ , are mistakenly assigned to adjacent length bins  $j-2, j-1, j+1$  or  $j+2$  with some specified probability. Well-tested methods for dealing with errors in length data can be applied if some information about the distribution of the errors is available (e.g. Methot 2000).

Prior to any other calculations, observed survey length composition data are converted to multinomial proportions:

$$i_{k,y,L} = \frac{n_{k,y,L}}{\sum_{j=L_{k,y}^{first}}^{L_{k,y}^{last}} n_{k,y,j}}$$

where  $n_{k,y,j}$  is an original datum and  $i_{k,y,L}$  is the corresponding proportion. As described above, the user specifies the first  $L_{k,y}^{first}$  and last  $L_{k,y}^{last}$  length groups to be used in calculating goodness of fit for each length composition and specifies whether the largest and smallest groups should be treated as “plus” groups that contain all smaller or larger individuals.

Using notation for goodness of fit survey index data (see above), predicted length compositions for surveys that track abundance or biomass are calculated:

$$A_{k,y,L} = \frac{s_{k,L} N_{y,L} e^{-Z_{y,j} \tau_{k,y}}}{\sum_{L=L_{k,y}^{first}}^{L_{k,y}^{last}} s_{k,j} N_{y,j} e^{-Z_{y,j} \tau_{k,y}}}$$

Predicted length compositions for surveys that track numbers of individuals killed by natural mortality are calculated:

$$A_{k,y} = \frac{s_{k,L} \tilde{M}_{y,L} \bar{N}_{y,L}}{\sum_{L=L_{k,y}^{first}}^{L_{k,y}^{last}} s_{k,L} \tilde{M}_{y,L} \bar{N}_{y,L}}$$

Considering the possibility of structured measurement errors, the expected length composition  $\bar{A}'_{k,y}$  for survey catches is:

$$\bar{A}'_{k,y} = \bar{A}_{k,y} E_k$$

where  $E_k$  is an error matrix that simulates errors in collecting length data by mapping true length bins in the model to observed length bins in the data.

The error matrix  $E_k$  has  $n_L$  rows (one for each true length bin) and  $n_L$  columns (one for each possible observed length bin). For example, row  $k$  and column  $j$  of the error matrix gives the conditional probability  $P(k|j)$  of being assigned to bin  $k$ , given that an individual actually belongs to bin  $j$ . More generally, column  $j$  gives the probabilities that an individual actually belonging to length bin  $j$  will be recorded as being in length bins  $j-2, j-1, j, j+1, j+2$  and so on. The columns of  $E_k$  add to one to account for all possible outcomes in assigning individuals to observed length bins.  $E_k$  is the identity matrix if there are no structured measurement errors.

In CASA, the probabilities in the error matrix are computed from a normal distribution with mean zero and  $CV = e^{\pi_k}$ , where  $\pi_k$  is an estimable parameter. The normal distribution is truncated to cover a user-specified number of observed bins (e.g. 3 bins on either side of the true length bin).

The NLL for observed proportions at length in each survey and year is computed with the kernel for a multinomial distribution, the model's estimate of proportional survey catch-at-length ( $\hat{i}_{k,y,L}$ ) and THE effective sample size  ${}^L N_y$  supplied by the user. Standardized residuals for residual plots are computed as for fishery length composition data.

#### *Effective sample size for length composition data*

Effective sample sizes that are specified by the user are used in goodness of fit calculations for survey and fishery length composition data. A post-hoc estimate of effective sample size can be calculated based on goodness of fit in a model run (Methot 1989). Consider the variance of residuals for a single set of length composition data with  $N$  bins used in calculations. The variance of the sum based on the multinomial distribution is:

$$\sigma^2 = \sum_{j=1}^N \left[ \frac{\hat{p}_j (1 - \hat{p}_j)}{\varphi} \right]$$

where  $\varphi$  is the effective sample size for the multinomial and  $\bar{p}_j$  is the predicted proportion in the  $j^{\text{th}}$  bin from the model run. Solve for  $\varphi$  to get:

$$\varphi = \frac{\sum_{j=1}^N [\hat{p}_j (1 - \hat{p}_j)]}{\sigma^2}$$

The variance of the sum of residuals can also be calculated:

$$\sigma^2 = \sum_{j=1}^N (p_j - \hat{p}_j)^2$$

This formula is approximate because it ignores the traditional correction for bias. Substitute the third expression into the second to get:

$$\varphi = \frac{\sum_{j=1}^N [\hat{p}_j (1 - \hat{p}_j)]}{\sum_{k=1}^N (p_j - \hat{p}_j)^2}$$

which can be calculated based on model outputs. The assumed and effective sample sizes will be similar in a reasonable model when the assumed sample sizes are approximately correct. Effective sample size calculations can be used iteratively to manually adjust input values to

reasonable levels (Methot 1989).

#### *Variance constraints on dev parameters*

Variability in dev parameters (e.g. for natural mortality, recruitment or fishing mortality) can be limited using variance constraints that assume the deviations are either independent or that they are autocorrelated and follow a random walk. When a variance constraint for independent deviations is activated, the model calculates the NLL for each log scale residual  $\gamma_y / \sigma_\gamma$ , where  $\gamma_y$  is a dev parameter and  $\sigma$  is a log-scale standard deviation. If the user supplies a positive value for the arithmetic scale CV, then the NLL is calculated assuming the variance is known. Otherwise, the user-supplied CV is ignored and the NLL is calculated with the standard deviation estimated internally. Calculations for autocorrelated deviations are the same except that the residuals are  $(\gamma_y - \gamma_{y-1}) / \sigma_\gamma$  and the number of residuals is one less than the number of dev parameters.

#### *LPUE data*

Commercial landings per unit of fishing effort (LPUE) data are modeled in the current implementation of the CASA model as a linear function of average biomass available to the fishery, and as a nonlinear function of average available abundance. The nonlinear relationship with abundance is meant to reflect limitations in “shucking” capacity for sea scallops.<sup>16</sup> Briefly, tows with large numbers of scallops require more time to sort and shuck and therefore reduce LPUE from fishing trips when abundance is high. The effect is exaggerated when the catch is composed of relatively small individuals. In other words, at any given level of stock biomass, LPUE is reduced as the number of individuals in the catch increases or, equivalently, as the mean size of individuals in the catch is reduced.

Average available abundance in LPUE calculations is:

$${}^a\bar{N}_y = \sum_{L=1}^{n_L} s_{y,L} \bar{N}_{y,L}$$

and average available biomass is:

$${}^a\bar{B}_y = \sum_{L=1}^{n_L} s_{y,L} w_L^f \bar{N}_{y,L}$$

where the weights at length  $w_L^f$  are for the fishery rather than the population. Predicted values for LPUE data are calculated:

$$\hat{L}_y = \frac{{}^a\bar{B}_y \eta}{\sqrt{\phi^2 + {}^a\bar{N}_y^2}}$$

Measurement errors in LPUE data are assumed normally distributed with standard deviations  $\sigma_y = CV_y \hat{L}_y$ . Standardized residuals are  $r_y = (L_y - \hat{L}_y) / \sigma_y$ .

---

<sup>16</sup> D. Hart, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA, pers. comm.

### *Per recruit (SBR and YPR) reference points<sup>17</sup>*

The user specifies a target %SBR value for each reference point that is estimated. Goodness of fit is calculated as the sum of squared differences between the target %SBR and %SBR calculated based on the reference point parameter. Except in pathological situations, it is always possible to estimate %SBR reference point parameters so that the target and calculated %SBR levels match exactly. Reference point parameters should have no effect on other model estimates and the residual (calculated – target %SBR) should always be very close to zero.

Goodness of fit for  $F_{0.1}$  estimates is calculated in a manner similar to %SBR reference points. Goodness of fit is calculated as the squared difference between the slope of the yield curve at the estimate and one-tenth of the slope at the origin. Slopes are computed numerically using central differences if possible or one-sided (right hand) differences if necessary.

$F_{max}$  is estimated differently in preliminary and final phases. In preliminary phases, goodness of fit for  $F_{max}$  is calculated as  $(1/Y)^2$ , where  $Y$  is yield per recruit at the current estimate of  $F_{max}$ . In other words, yield per recruit is maximized by finding the parameter estimate that minimizes its inverse. This preliminary approach is very robust and will find  $F_{max}$  if it exists. However, it involves a non-zero residual  $(1/Y)$  that interferes with calculation of variances and might affect other model estimates. In final phases, goodness of fit for  $F_{max}$  is calculated as  $(d^2)$  where  $d$  is the slope of the yield per recruit curve at  $F_{max}$ . The two approaches give the same estimates of  $F_{MAX}$  but the goodness of fit approach used in the final phases has a residual of zero (so that other model estimates are not affected) and gives more reasonable variance estimates. The latter goodness of fit calculation is not used during initial phases because the estimates of  $F_{MAX}$  tend to “drift down” the right hand side of the yield curve in the direction of decreasing slope. Thus, the goodness of fit calculation used in final phases works well only when the initial estimate of  $F_{MAX}$  is very close to the best estimate.

Per recruit reference points should have little or no effect on other model estimates. Problems may arise, however, if reference points (particularly  $F_{max}$ ) fall on the upper bound for fishing mortality. In such cases, the model will warn the user and advise that the offending reference points should not be estimated. *It is good practice to run CASA with and without reference point calculations to ensure that reference points do not affect other model estimates including abundance, recruitments and fishing mortality rates.*

### *Growth data*

Growth data in CASA consist of records giving initial length, length after one year of growth, and number of corresponding observations. Growth data may be used to help estimate growth parameters that determine the growth matrix  $P$ . The first step is to convert the data for each starting length to proportions:

$$P(b, a) = \frac{n(b, a)}{\sum_{j=n_L-b+1}^{n_L} n(j, a)}$$

where  $n(b, a)$  is the number of individuals starting at size  $a$  that grew to size  $b$  after one year. The NLL is computed assuming that observed proportions  $p(a|b)$  at each starting size are a sample from a multinomial distribution with probabilities given by the corresponding column in the

---

<sup>17</sup> This approach is not currently estimated because of performance problems. The user can, however, estimate per recruit reference point from a detailed table written in the main output file (nc.rep). However, variances are not available in the table.

models estimated growth matrix  $P$ . The user must specify an effective sample size  $^P N_j$  based, for example, on the number of observations in each bin or the number of individuals contributing data to each bin. Observations outside bin ranges specified by the user are ignored. Standardized residuals for plotting are computed based on the variance for proportions.

### *Survey gear efficiency data*

Survey gear efficiency for towed trawls and dredges is the probability of capture for individuals anywhere in the water column or sediments along the path swept by the trawl. Ideally, the area surveyed and the distribution of the stock coincides so that:

$$I_{k,y} = q_k B_{k,y}$$

$$q_k = \frac{a_k e_k u_k}{A}$$

$$e_k = \frac{A q_k}{a_k u_k}$$

$$K_t = \frac{A}{a_k u_k}$$

$$e_k = K_t q_t$$

Where  $I_{k,y}$  is a survey observation in units equivalent to biomass (or numerical) density (e.g. kg per standard tow),  $B_{k,y}$  is the biomass (or abundance) available to the survey,  $A$  is the area of the stock,  $a_k$  is the area swept during one tow,  $0 < e_k \leq 1$  is efficiency of the survey gear, and  $u_k$  is a constant that adjusts for different units.

Efficiency estimates from studies outside the CASA model may be used as prior information in CASA. The user supplies the mean and CV for the prior estimate of efficiency, along with estimates of  $A_k$ ,  $a_k$  and  $u_k$ . At each iteration if the model, the gear efficiency implied by the current estimate of  $q_k$  is computed. The model then calculates the NLL of the implied efficiency estimate assuming it was sampled from a unimodal beta distribution with the user-specified mean and CV.

If efficiency estimates are used as prior information (if the likelihood weight  $\lambda > 0$ ), then it is very important to make sure that units and values for the survey data ( $I$ ), biomass or abundance ( $B$ ), stock area ( $A$ ), area per tow ( $a$ ), and adjustments for units ( $u$ ) are correct (see Example 1). The units for biomass are generally the same as the units for catch data. In some cases, incorrect specifications will lead to implied efficiency estimates that are  $\leq 0$  or  $\geq 1$  which have zero probability based on a standard beta distribution used in the prior. The program will terminate if  $e \leq 0$ . If  $e \geq 1$  during an iteration, then  $e$  is set to a value slightly less than one and a penalty is added to the objective function. In some cases, incorrect specifications will generate a cryptic error that may have a substantial impact on estimates.

Implied efficiency estimates are useful as a model diagnostic even if very little prior information is available because some model fits may imply unrealistic levels of implied efficiency. The trick is to down weight the prior information (e.g.  $\lambda = 1e^{-6}$ ) so that the implied efficiency estimate has very little effect on model results as long as  $0 < e < 1$ . Depending on the situation, model runs with  $e$  near a bound indicate that estimates may be implausible. In

addition, it may be useful to use a beta distribution for the prior that is nearly a uniform distribution by specifying a prior mean of 0.5 and variance slightly less than  $1/12=0.083333$ .

Care should be taken in using prior information from field studies designed to estimate survey gear efficiency. Field studies usually estimate efficiency with respect to individuals on the same ground (e.g. by sampling the same grounds exhaustively or with two types of gear). It seems reasonable to use an independent efficiency estimate and the corresponding survey index to estimate abundance in the area surveyed. However, stock assessment models are usually applied to the entire stock, which is probably distributed over a larger area than the area covered by the survey. Thus the simple abundance calculation based on efficiency and the survey index will be biased low for the stock as a whole. In effect, efficiency estimates from field studies tend to be biased high as estimates of efficiency relative to the entire stock.

#### *Maximum fishing mortality rate*

Stock assessment models occasionally estimate absurdly high fishing mortality rates because abundance estimates are too small. The NLL component used to prevent this potential problem is:

$$L = \lambda \sum_{t=0}^N (d_t^2 + q^2)$$

where:

$$d_t = \begin{cases} Ft - \Phi & \text{if } Ft > \Phi \\ 0 & \text{otherwise} \end{cases}$$

and

$$q_t = \begin{cases} \ln(Ft/\Phi) & \text{if } Ft > \Phi \\ 0 & \text{otherwise} \end{cases}$$

with the user-specified threshold value  $\Phi$  set larger than the largest value of  $F_t$  that might possibly be expected (e.g.  $\Phi=3$ ). The weighting factor  $\lambda$  is normally set to a large value (e.g. 1000).

## References

- J.L. Butler, L.D. Jacobson, J.T. Barnes, and H.G. Moser. 2003. Biology and population dynamics of cowcod (*Sebastes levis*) in the southern California Bight. *Fish. Bull.* 101: 260-280.
- Fournier, D., and Archibald, CP. 1982. General theory for analyzing catch at age data. *Can. J. Fish. Aquat. Sci.* 39: 1195-1207.
- Jacobson, L.D., Cadrin, S.X., and J.R. Weinberg. 2002. Tools for estimating surplus production and  $F_{MSY}$  in any stock assessment model. *N. Am. J. Fish. Mgmt.* 22: 326-338.
- Methot, R. D. 2000. Technical description of the stock synthesis assessment program. NOAA Tech. Memo. NMFS-NWFSC-43: 1-46.
- Pennington, M., Burmeister, L-M., and Hjellvik, V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fish. Bull.* 100: 74-80.
- Press, W.H., Flannery, B.P., S.A. Teukolsky, and W.T. Vetterling. 1990. Numerical recipes. Cambridge Univ. Press, NY.
- Sullivan, P.J., Lai, H.L., and Gallucci, V.F. 1990. A catch-at-length analysis that incorporates a stochastic model of growth. *Can. J. Fish. Aquat. Sci.* 47: 184-198.

## Appendix B10. Forecasting methodology (SAMS model)

Dvora Hart, Northeast Fisheries Science Center, Woods Hole, MA.

The model presented here is a version of the SAMS (Scallop Area Management Simulator) model used to project sea scallop abundance and landings as an aid to managers since 1999. Subareas were chosen to coincide with current management. In particular, Georges Bank was divided into four open areas (two portions of the South Channel, Northern Edge and Peak, and Southeast Part), the three access portions of the groundfish closures, and the three no access portions of these areas. The Mid-Atlantic was subdivided into seven areas: Virginia Beach, the Delmarva, Elephant Trunk Closed Area and Hudson Canyon South Rotational Areas, New York Bight, Inshore New York Bight, and Long Island.

### Methods

The model tracks population vectors  $\mathbf{p}(i,t) = (p_1, p_2, \dots, p_n)$ , where  $p_j(i,t)$  represents the density of scallops in the  $j$ th size class in area  $i$  at time  $t$ . The model uses a difference equation approach, where time is partitioned into discrete time steps  $t_1, t_2, \dots$ , with a time step of length  $\Delta t = t_{k+1} - t_k$ . The landings vector  $\mathbf{h}(i,t_k)$  represents the catch at each size class in the  $i$ th region and  $k$ th time step. It is calculated as:

$$h(i,t_k) = [I - \exp(\Delta t H(i,t_k))] p(i,t_k),$$

where  $I$  is the identity matrix and  $H$  is a diagonal matrix whose  $j$ th diagonal entry  $h_{jj}$  is given by:

$$h_{jj} = 1/(1 + \exp(s_0 - s_1 * s))$$

where  $s$  is the shell height of the mid-point of the size-class.

The landings  $L(i,t_k)$  for the  $i$ th region and  $k$ th time step are calculated using the dot product of landings vector  $\mathbf{h}(i,t_k)$  with the vector  $\mathbf{m}(i)$  representing the vector of meat weights at shell height for the  $i$ th region:

$$L(i,t_k) = A_i \mathbf{h}(i,t_k) \bullet \mathbf{m}(i)$$

where  $e_i$  represents the dredge efficiency in the  $i$ th region.

Even in the areas not under special area management, fishing mortalities tend not to be spatially uniform due to the sessile nature of sea scallops (Hart 2001). Fishing mortalities in open areas were determined by a simple “fleet dynamics model” that estimates fishing mortalities in open areas based on area-specific catch rates, and so that the overall DAS or open-area  $F$  matches the target. Based on these ideas, the fishing mortality  $F_i$  in the  $i$ th region is modeled as:

$$F_i = k * f_i * L_i$$

where  $L_i$  is the estimated LPUE (landings per day charged) in the  $i$ th region,  $f_i$  is an area-specific adjustment factor to take into account preferences for certain fishing grounds (due to lower costs, shorter steam times, ease of fishing, habitual preferences, etc.), and  $k$  is a constant adjusted so that the total DAS or fishing mortality meets its target. For these simulations,  $f_i = 1$  for all areas.

Scallops of shell height less than a minimum size  $s_d$  are assumed to be discarded, and suffer a discard mortality rate of  $d$ , taken here, as in the rest of the assessment to be 20%. There is also evidence that some scallops not actually landed may suffer mortality due to incidental damage from the dredge. Let  $F_L$  be the landed fishing mortality rate and  $F_I$  be the rate of incidental mortality on scallops not caught. For Georges Bank, which is a mix of sandy and hard bottom, we used  $F_I = 0.2F_L$ . For the Mid-Atlantic (almost all sand), we used  $F_I = 0.1F_L$ . Incidental mortality for a given shell height bin was then calculated using equations (4.3) and (4.4) of the main document.

Growth in each subarea was specified by a growth transition matrix  $G$ , based on area-specific growth increment data from 2001-2012. Recruitment was modeled stochastically, and was assumed to be log-normal in each subarea. The mean, variance and covariance of the recruitment in a subarea was set to be equal to that observed in the historical time-series between 1979-2013. New recruits enter the first size bin at each time step at a rate  $r_i$  depending on the subarea  $i$ , and stochastically on the year. These simulations assume that recruitment is a stationary process, i.e., no stock-recruitment relationship is assumed. This may underestimate recruitment in the Mid-Atlantic if the recent strong recruitment there are due to a stock-recruit relationship.

The population dynamics of the scallops in the present model can be summarized in the equation:

$$p(i, t_{k+1}) = \rho_i + G \exp(-M\Delta t H) p(i, t_k),$$

where  $\rho_i$  is a random variable representing recruitment in the  $i$ th area. The model was run with 10 time steps per year. The population and harvest vectors are converted into biomass by using the shell-height meat-weight relationship:

$$W = \exp[a + b \ln(s)],$$

where  $W$  is the meat weight of a scallop of shell height  $s$ . These relationships are subarea-specific; see Appendix B3 for details. For calculating biomass, the shell height of a size class was taken as its midpoint.

Commercial landing rates (LPUE, landed meat weight per day) were estimated using an empirical function based on the observed relationship between annual landing rates, expressed as number caught per day (NLPUE) and survey exploitable numbers per tow. At low biomass levels, NLPUE increases roughly linearly with survey abundance. However, at high abundance levels, the catch rate of the gear will exceed that which can be shucked by a seven-man crew. This is similar to the situation in predator/prey theory, where a predator's consumption rate is limited by the time required to handle and consume its prey (Holling 1959). The original Holling Type-II predator-

prey model assumes that handling and foraging occur sequentially. It predicts that the per-capita predation rate  $R$  will be a function of prey abundance  $N$  according to a Monod functional response:

$$R = \frac{\alpha N}{\beta + N},$$

where  $\alpha$  and  $\beta$  are constants. In the scallop fishery, however, some handling (shucking) can occur while foraging (fishing), though at a reduced rate because the captain and one or two crew members need to break off shucking to steer the vessel during towing and to handle the gear during haulback.

The fact that a considerable amount of handling can occur at the same time as foraging means that the functional response of a scallop vessel will saturate quicker than predicted by the above equation. To account for this, a modified Holling Type-II model was used, so that the landings (in numbers of scallops) per unit effort (DAS)  $L$  (the predation rate, i.e., NLPUE) will depend on scallop (prey) exploitable numbers  $N$  according to the formula:

$$L = \frac{\alpha N}{\sqrt{\beta^2 + N^2}}.$$

The parameters  $\alpha$  and  $\beta$  to this model were fit to the observed fleet-wide LPUE vs. exploitable biomass relationship during the years 1994-2012 (previous years were not used because of the change from port interviews to logbook reporting). The number of scallops that can be shucked should be nearly independent of size provided that the scallops being shucked are smaller than about a 20 count. The time to shuck a large scallop will go up modestly with size. To model this, if the mean meat weight of the scallops caught,  $g$ , in an area is more than 20 g, the parameters  $\alpha$  and  $\beta$  in the above equation are reduced by a factor  $\sqrt{20/g}$ . This means, for example, that a crew could shuck fewer 10 count scallops per hour than 20 count scallops in terms of numbers, but more in terms of weight.

An estimate of the fishing mortality imposed in an area by a single DAS of fishing in that area can be obtained from the formula  $F_{\text{DAS}} = L_a/N_a$ , where  $L_a$  is the NLPUE in that area obtained as above, and  $N_a$  is the exploitable abundance (expressed as absolute numbers of scallops) in that area. This allows for conversion between units of DAS and fishing mortality.

Initial conditions for the population vector  $\mathbf{p}(i,t)$  were estimated using the 2013 surveys, with the overall estimates scaled to match the 2013 biomass as estimated by CASA. The 2013 initial conditions were varied depending on the survey standard errors in each subarea, and scaled so that the initial standard error in biomass was about 15,000 mt, a figure that the working group considered a fair measure of the true uncertainty in the initial estimates.