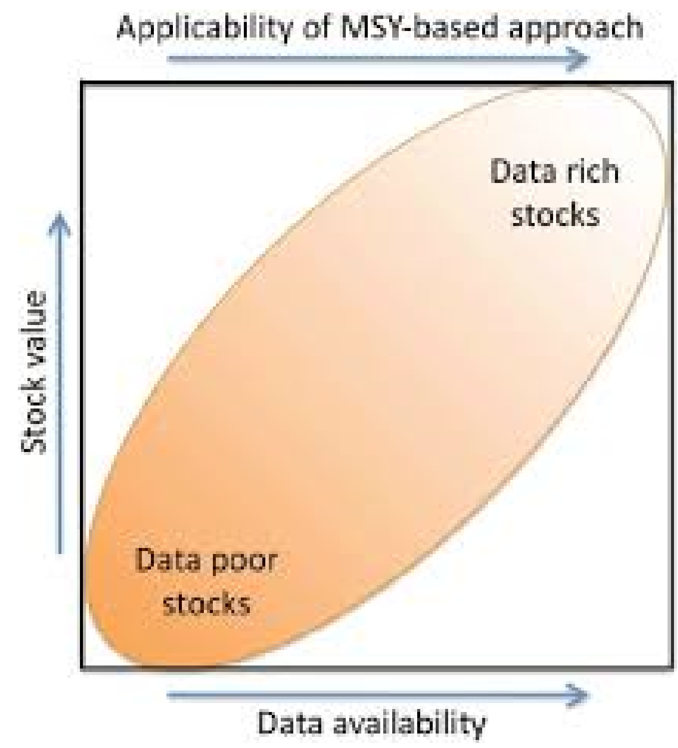
Marine Institute

Lecture 2: Fisheries Dependent and Independent (Surveys) Indices of Stock Size



Noel Cadigan

CFER

Centre for Fisheries Ecosystems Research

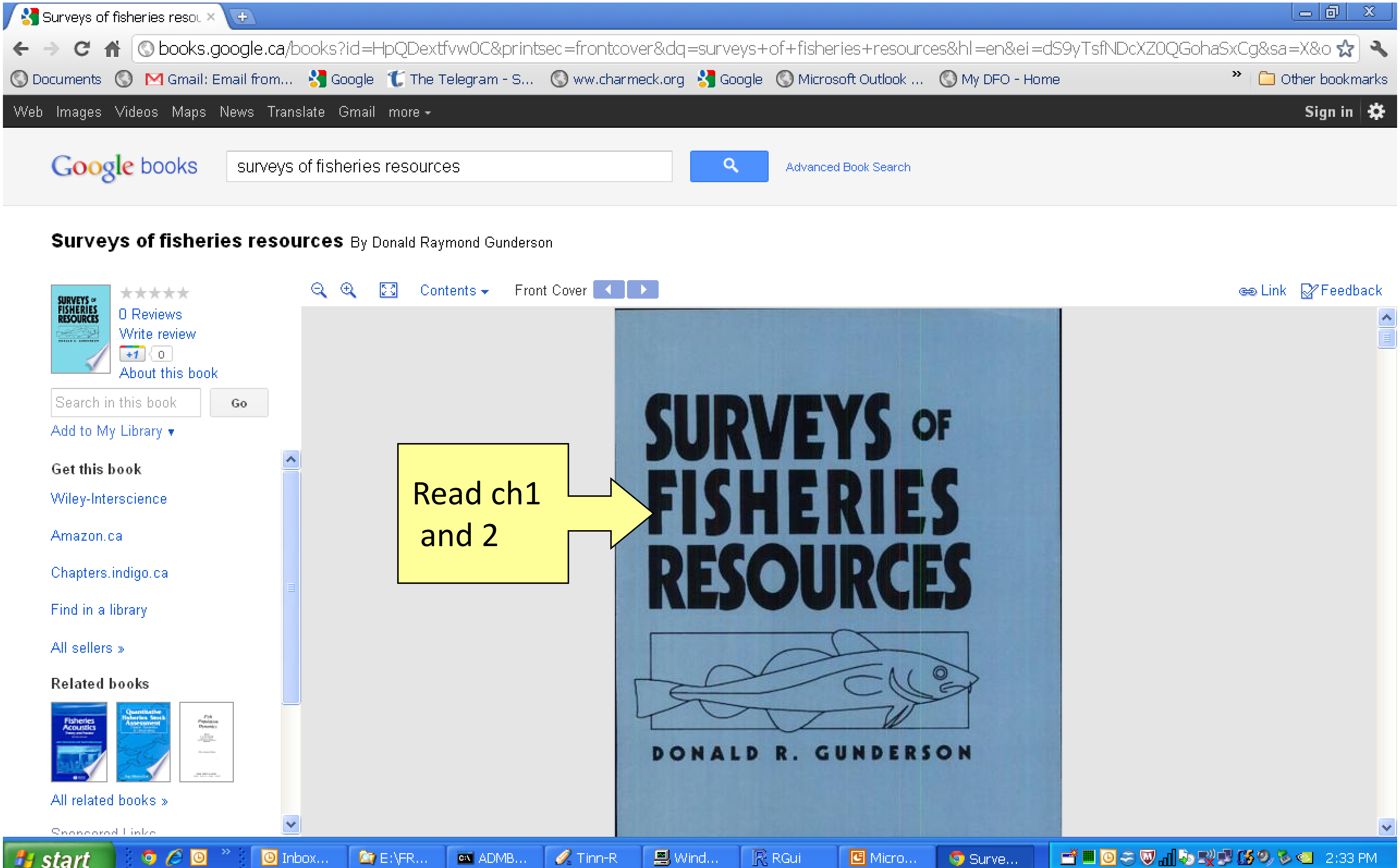
Marine Institute

Memorial University

F6004 Lecture 2 Outline

- Data from fisheries
  - catch per unit effort (CPUE),
  - interpreting fishery catch data (relation to stock abundance and distribution)
  - measures from individual fish (weights, lengths, age determination, condition factor)
- Data from surveys
  - Survey types (trawl, acoustic, fishery gear surveys, questionnaires to fishermen, visual counts)
  - Survey design (random, stratification, systematic, transect methods)
  - Survey indices and relation to stock abundance and distribution
  - Measures from individual fish

Readings: 1) For general reading on fisheries data, read Hilborn and Walters Chapter 5. 2) A general treatment of Fisheries surveys is found in “Surveys of Fisheries Resources” by D.R. Gunderson. 3) Kimura and Somerton (2006). 4) Doubleday (1981) for Northwest Atlantic bottom trawl surveys


Surveys of fisheries resou
books.google.ca/books?id=HpQDextfvw0C&printsec=frontcover&dq=surveys+of+fisheries+resources&hl=en&ei=dS9yTsfNDcXZ0QGohaSxCg&sa=X&o
Documents
Gmail: Email from...
Google
The Telegram - S...
ww.charmeck.org
Google
Microsoft Outlook ...
My DFO - Home
Other bookmarks
Web Images Videos Maps News Translate Gmail more
Sign in
Google books
surveys of fisheries resources
Advanced Book Search
Surveys of fisheries resources By Donald Raymond Gunderson
0 Reviews
Write review
About this book
Search in this book
Go
Add to My Library
Get this book
Wiley-Interscience
Amazon.ca
Chapters.indigo.ca
Find in a library
All sellers »
Related books
All related books »
Contents
Front Cover
Link
Feedback
Read ch1 and 2
SURVEYS OF FISHERIES RESOURCES
DONALD R. GUNDERSON
start
Inbox...
E:\FR...
ADMB...
Tinn-R
Wind...
RGui
Micro...
Surve...
2:33 PM

Reviews in Fisheries Science, 14:245–283, 2006

ISSN: 1064-1262 print
DOI: 10.1080/10641260600621761

Taylor & Francis

Review of Statistical Aspects of Survey Sampling for Marine Fisheries

DANIEL K. KIMURA AND DAVID A. SOMERTON

National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle, Washington, USA

Fishery surveys are an essential ingredient of modern fisheries stock assessment. To understand this, one must understand that the survey time series are the essential anchor that makes modern fishery stock assessment modeling possible. Without fishery surveys, fisheries stock assessment scientists would have great difficulty modeling absolute abundance, and therefore agencies would have difficulty setting levels of allowable catch. In this review, we explore the basic assumptions that need to be fulfilled for valid surveys to be accomplished. Although these assumptions are generally well understood, the complexity of survey sampling gear and the complexity of target animal behavior makes them difficult to fulfill in practice and can easily lead to problems when interpreting survey results from even the most carefully designed survey. In reviewing the literature surrounding fishery surveys, their sampling design, modeling, and methods of analysis, it became clear that fisheries scientists have long been preoccupied with coping with the intense variability found in fish catches. This variability is found within fishing hauls, between hauls, between area, time and depth strata. Coping with this variability, which is due to animal behavior, habitat variation, and the nature of fishing gear, will be a constant theme that ties together our review.

Keywords trawl surveys, hydroacoustic surveys, abundance indices, sampling design, data analysis, spatial sampling

Commercial Fishery Catch data

- Fishery catch statistics are commonly used in stock assessment.
- Described these in last class
- I will describe more today and in future lectures
- catch statistics are collected in a variety of ways.

Catch and Effort

– Effort is commonly defined as hours fished for the entire fleet and season

– Conceptually, effort should indicate the “swept-area” of the fishery

– How much ground was covered to produce the catch.

– Let E = effort, and C = catch

– CPUE = C/E can be a useful index of stock size.

Catch and Effort

– An index (*I*) is a measurement that is proportional to stock size;

– That is, $I = q x N$

– If *q* is the same from year to year then we can discern trends in stock size from trends in *I*.

Catch and Effort

– A CPUE index should be standardized so that it represents the same region, time of year, fleet composition, etc

– Otherwise you can get biases

– Methods to standardize CPUE series can be quite complex

Simple example: no standardization

		Year					
		1	2	3	4	5	6
Vessel1	Catch	800	720	640	560	480	400
	Effort	200	180	160	140	120	100
	CPUE	4	4	4	4	4	4
Vessel2	Catch	400	480	560	640	720	800
	Effort	200	240	280	320	360	400
	CPUE	2	2	2	2	2	2
Total	Catch	1200	1200	1200	1200	1200	1200
	Effort	400	420	440	460	480	500
	CPUE	3.0	2.9	2.7	2.6	2.5	2.4

Total: A 20% reduction since start
But no change in CPUE for each vessel!

Simple example

		Year					
		1	2	3	4	5	6
Vessel1	Catch	400	360	320	280	240	200
	Effort	200	180	160	140	120	100
	CPUE	2	2	2	2	2	2
Vessel2	Catch	400	480	560	640	720	800
	Effort	200	240	280	320	360	400
	CPUE	2	2	2	2	2	2
Total	Catch	800	840	880	920	960	1000
	Effort	400	420	440	460	480	500
	CPUE	2	2	2	2	2	2

Standardize vessel 1 to vessel 2

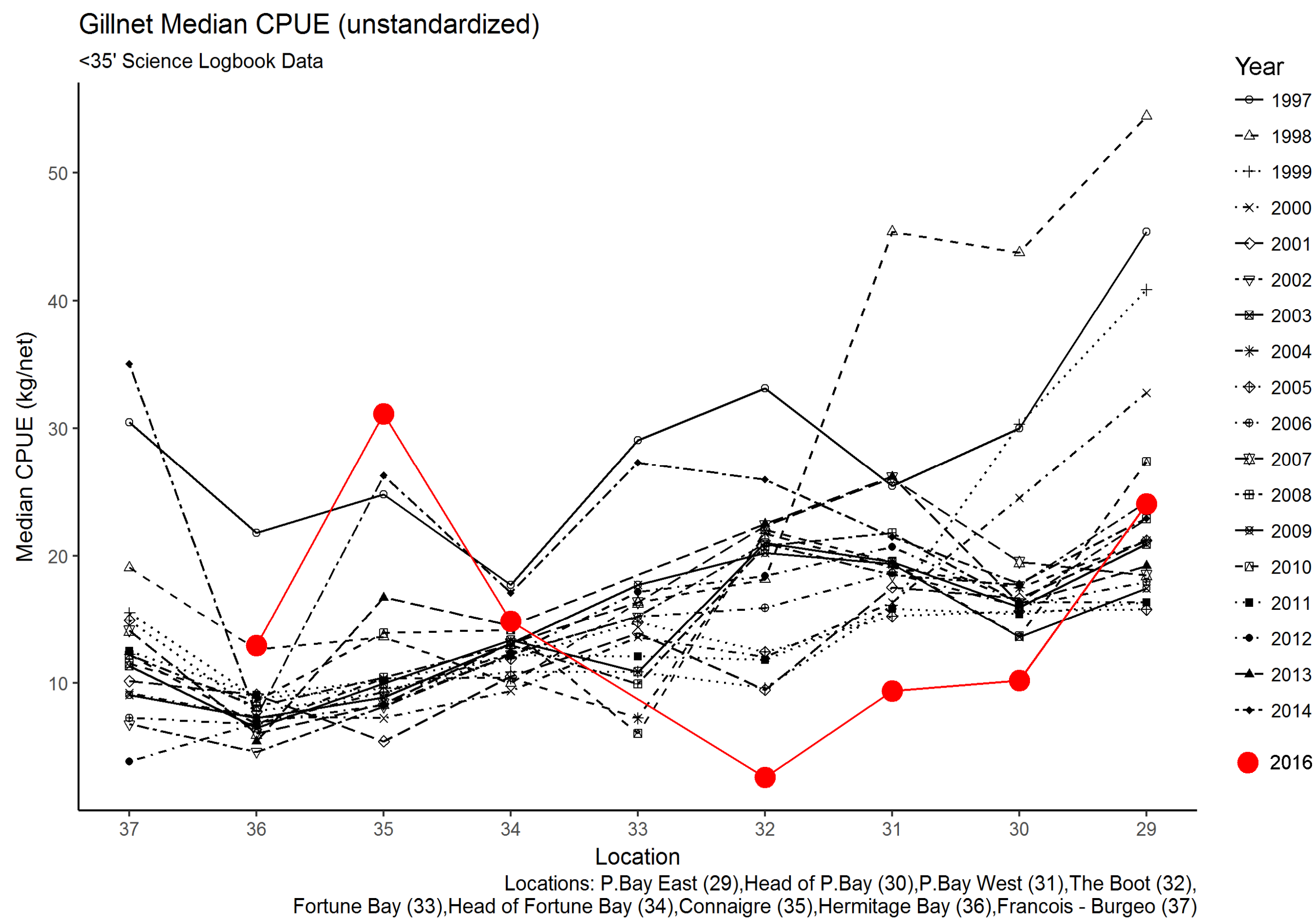
Example: 3Ps cod Commercial Logbooks <35' Vessels

- A logbook for Canadian vessels <35' was initiated in 1997 to provide information for stock assessment.
- For 3Ps cod, there are approximately 171,000 records in the database from 1997-2016. 1 Record=1 Logbook page.
- Return rates of this logbook have been low (≈50% of total <35' fleet catch) in recent years – and return rate has decreased with time.
- CPUE computed as catch weight per net (gillnet) or catch weight per 1000 hooks (linetrawl).

Unstandardized Gillnet CPUE

Values for 2017 in red.

Average except at head / western P. Bay.

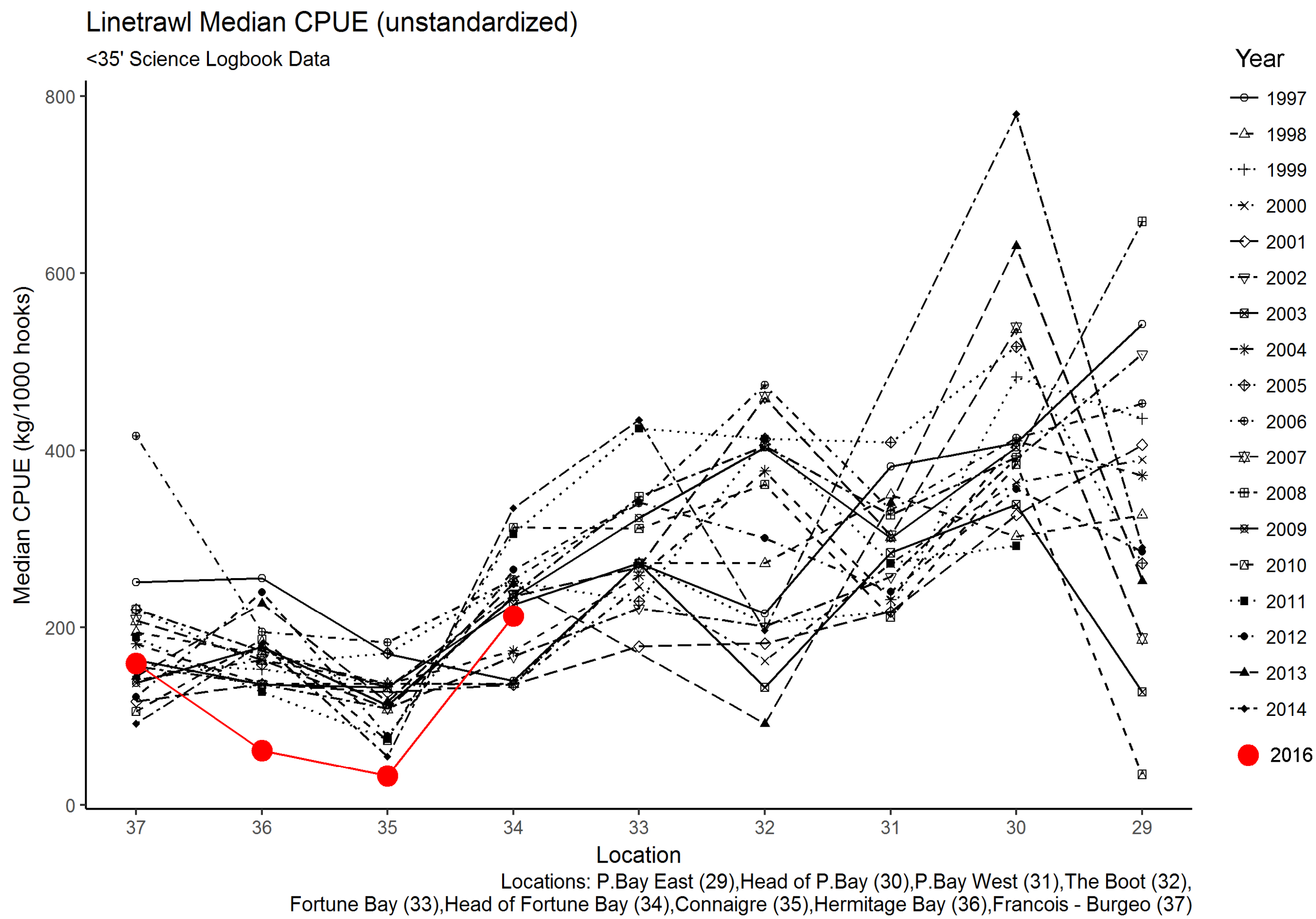


Location

# CFV_{2014}	2	14	3	14	2	10	14	22	15
# CFV_{2016}	0	7	3	6	0	2	6	4	3

Unstandardized CPUE

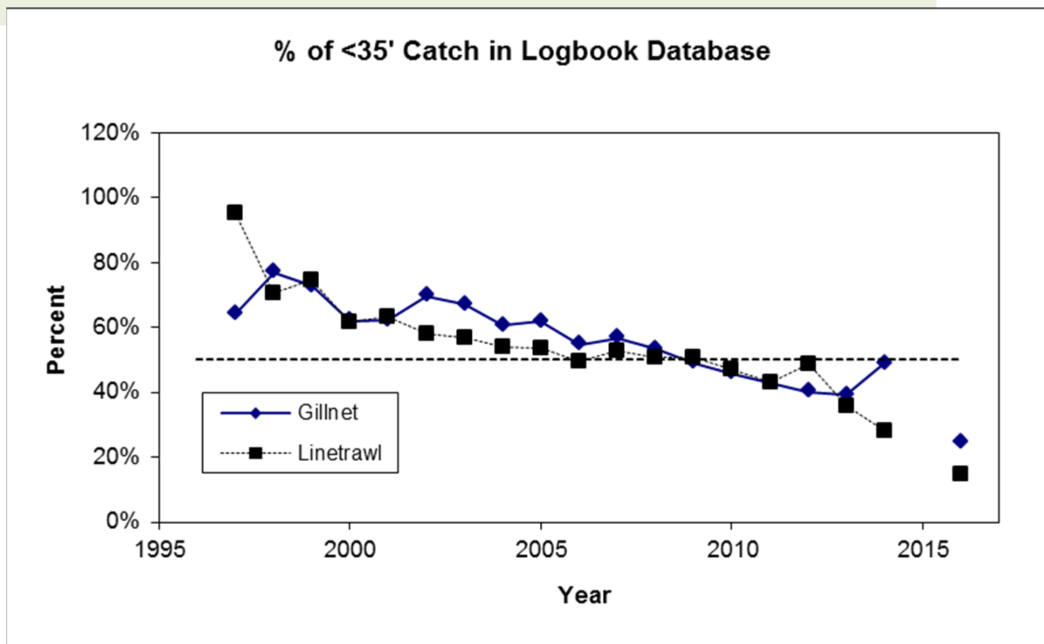
Values for 2017 in red. Highly variable across areas



	Location								
# CFV_{2014}	6	9	1	7	2	2	0	5	3
# CFV_{2016}	1	5	4	3	0	0	0	0	0

Commercial CPUE (<35' Vessels)

- Prior to modelling, data are screened out if:
  - Amount of Gear unavailable.
  - Statistical Area unavailable. This is the main reason for excluding data from analysis
  - the proportion of logbooks reporting something other than areas 29-37 has increased over time.

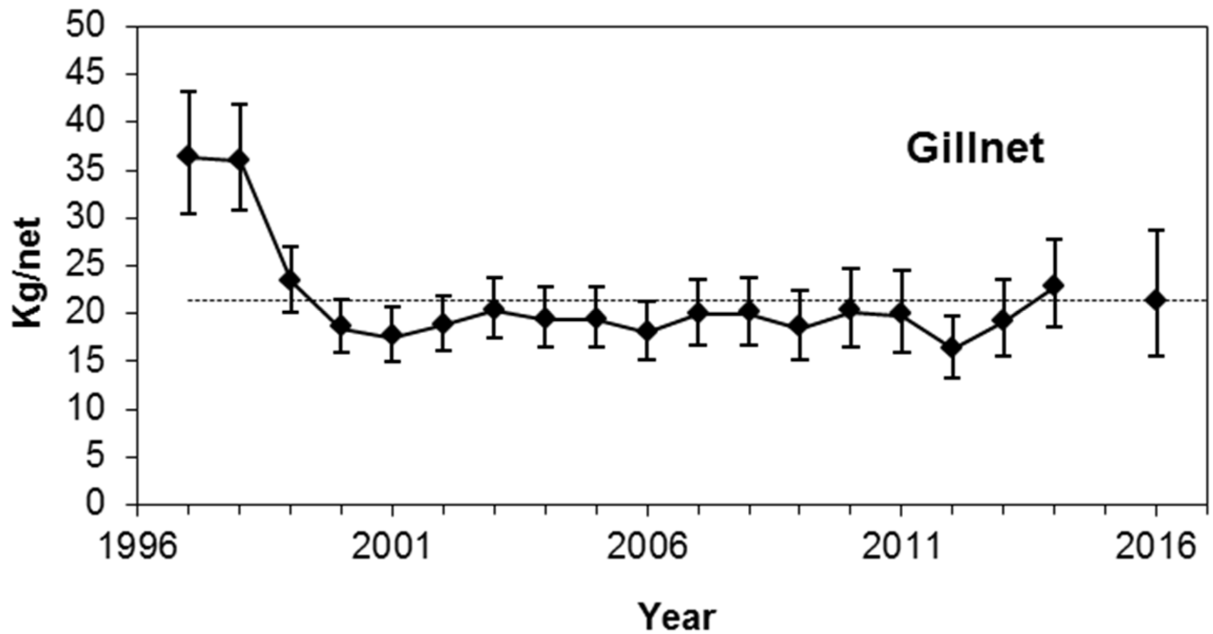


Commercial CPUE (<35' Vessels)

- Standardization: Generalized Linear Model
  - Gamma distribution with a log link
  - estimate year and month within location (=LFA) effects
  - No intercept.
  - Effort (amount of gear) used as offset variable.

C/E = effects
Log(C/E) = log(effects) = log_effects
=> log(C) = log_effects + log(E)

Commercial GN CPUE (<35' Vessels)



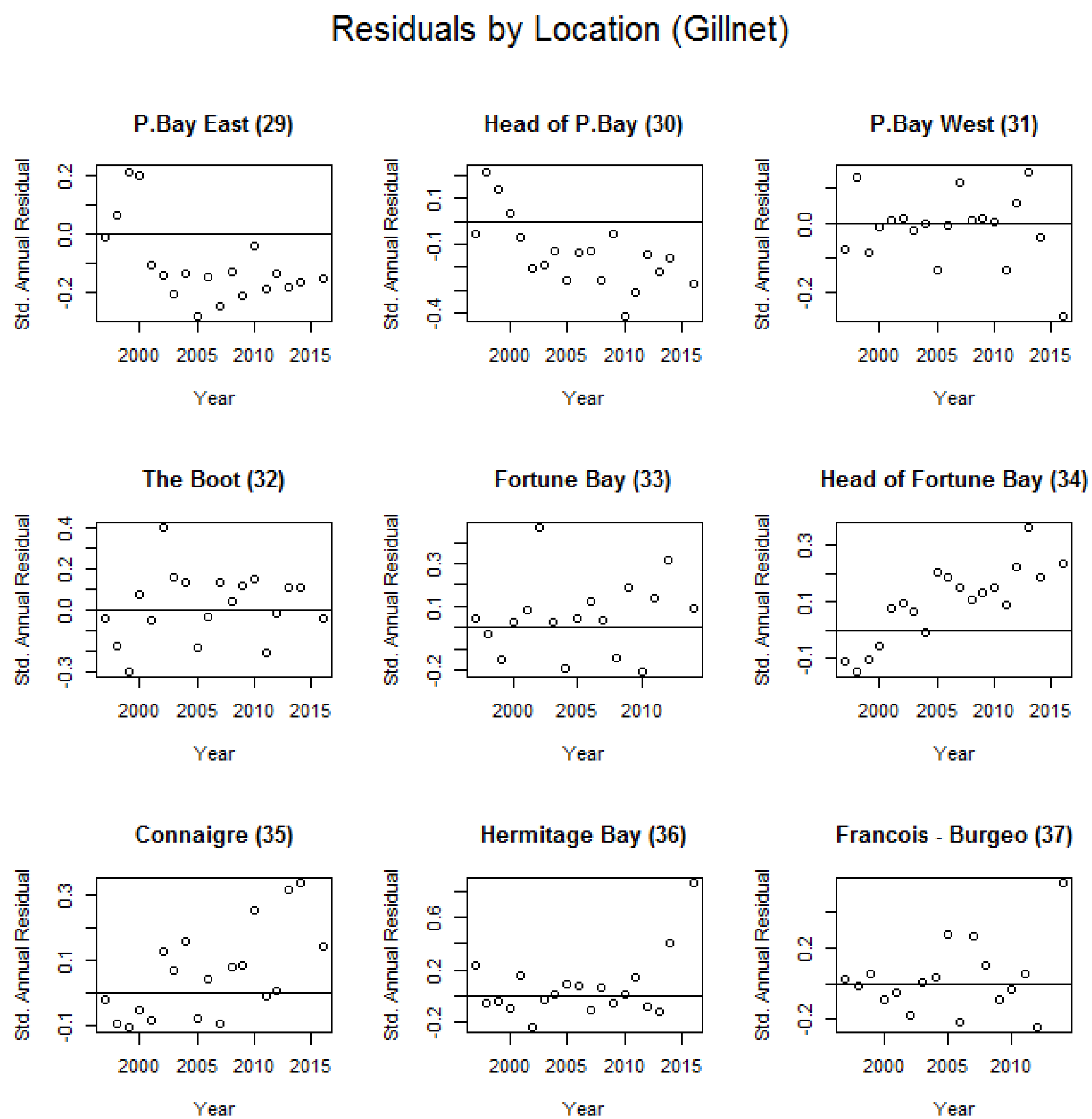
Remarkably steady after initial two years post-moratorium. Lack of change as various year-classes have passed through the fishery is not understood. Most recent value is near the time-series mean.

Residuals (E)

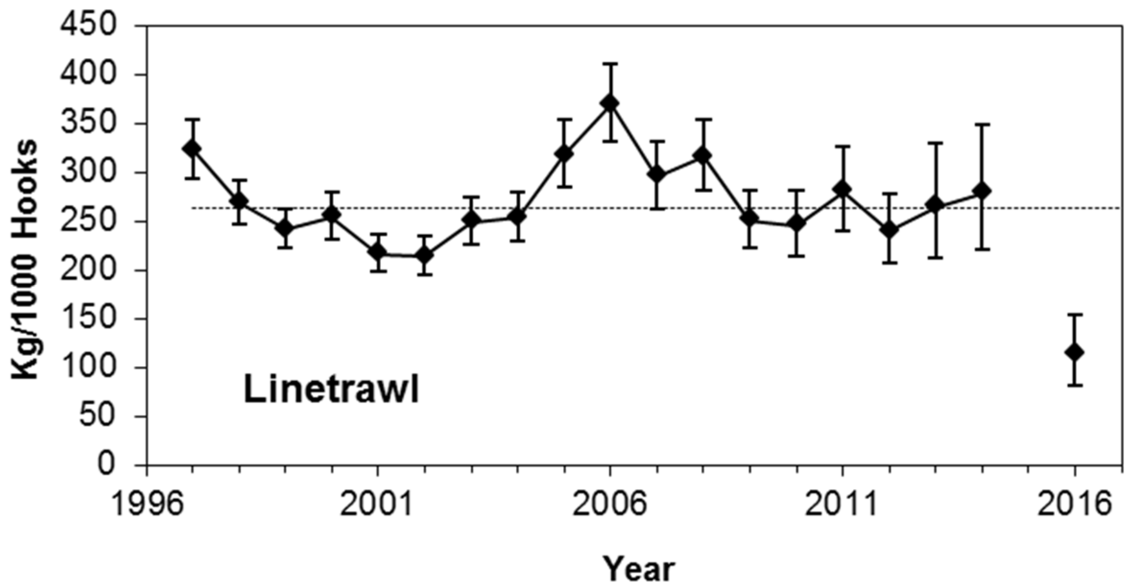
- Computed a standardized annual residual:

E={sum(obs)-sum(exp)}/ Sqrt[phi*sum(exp^2)]

- Different trends over time in different parts of the stock area.

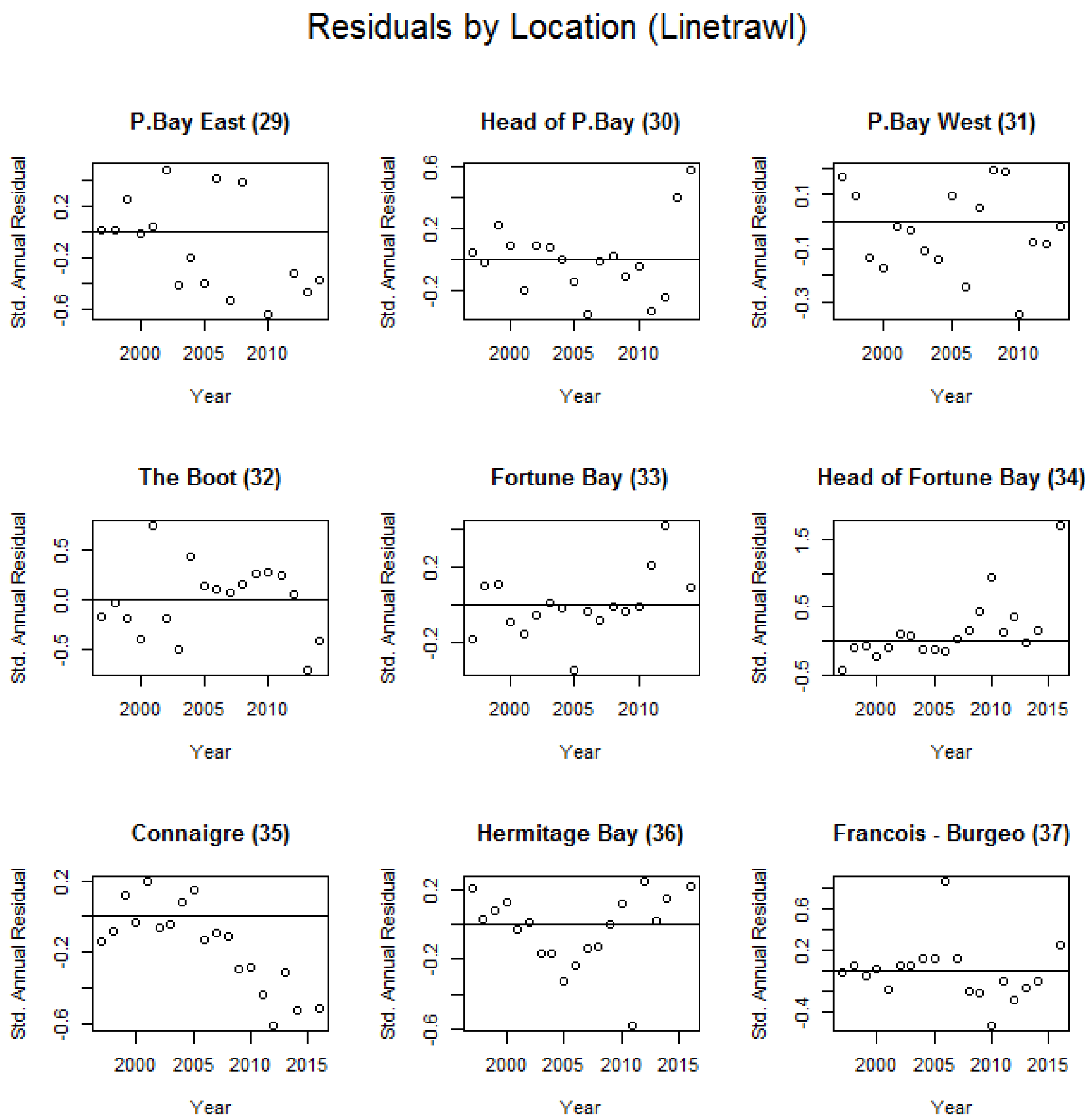


3Ps cod Commercial LT CPUE (<35' Vessels)



More variation over time; current value is lowest in the time-series, but based on low returns / poor coverage.

Residuals



- Computed a standardized annual residual:

E={sum(obs)-sum(exp)}/

Sqrt[phi*sum(exp^2)]

- Different trends over time in different parts of

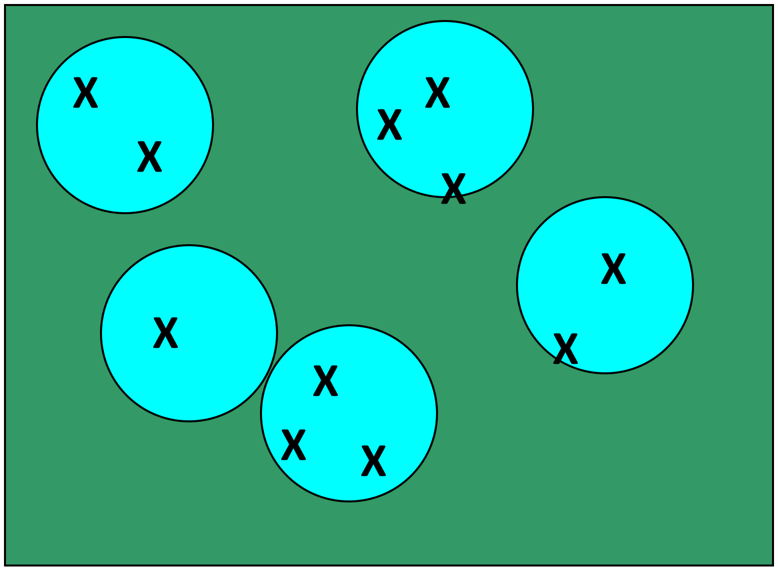
Catch and Effort

– Even when properly standardized, a CPUE index may not reflect stock trends because of fleet aggregation on the hot spots.

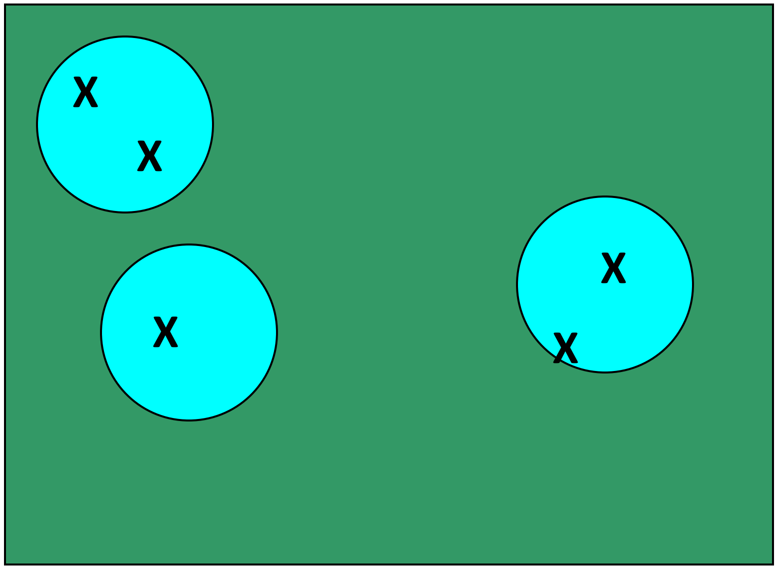
– CPUE often reflects stock trends in major aggregations rather than trends in the stock as a whole.

Catch and Effort

Year 1



Year 2



In this example the stock declines by 40% in year 2, but CPUE does not change

A Complex Analysis – you don't need to read this

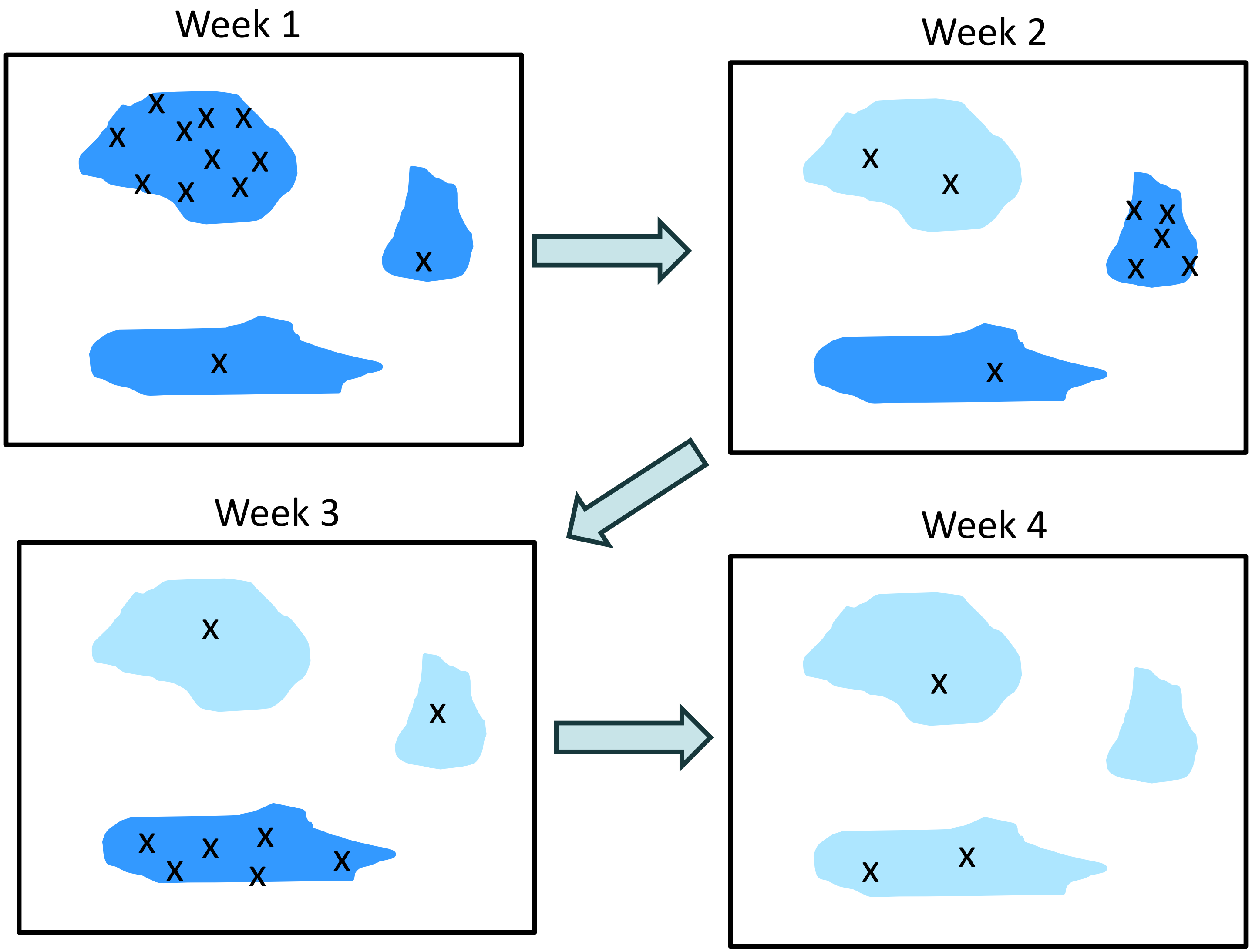
NRC Research Press

ARTICLE

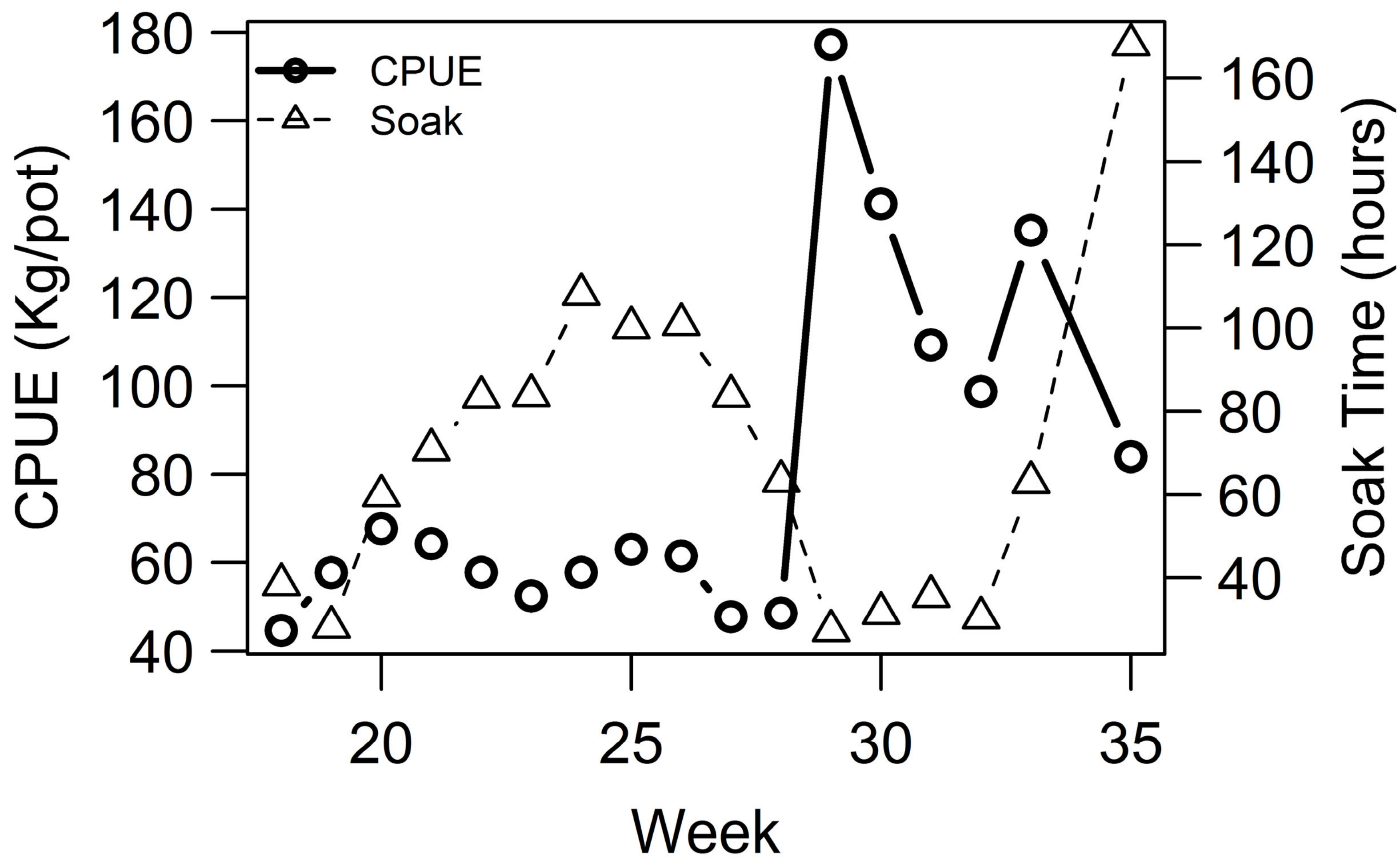
A spatiotemporal model for snow crab (*Chionoecetes opilio*) stock size in the southern Gulf of St. Lawrence[1]

Noel G. Cadigan, Elmer Wade, and Anders Nielsen

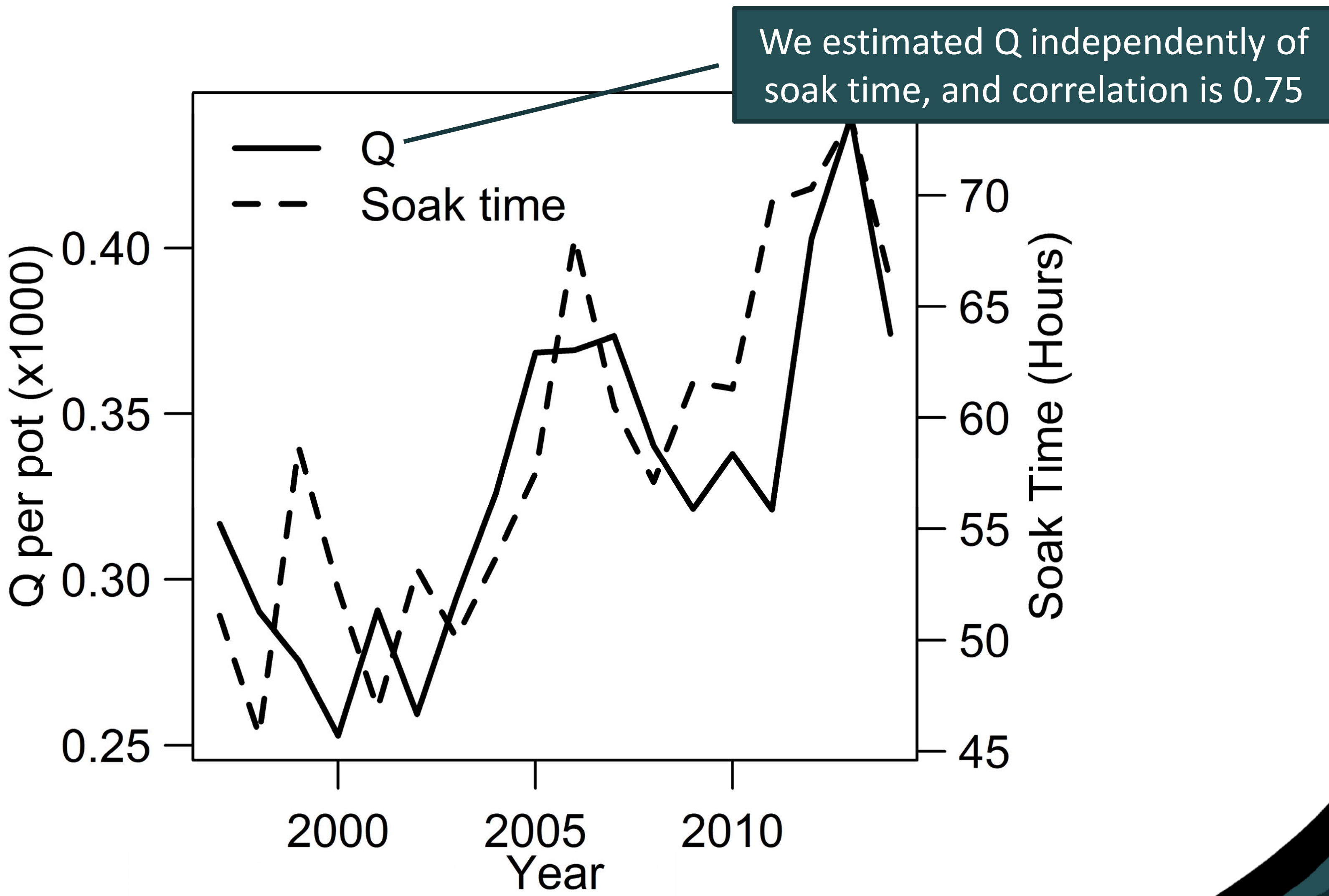
Snow Crab CPUE Problem: Targeting



Snow Crab CPUE Problem: change in fleet behavior



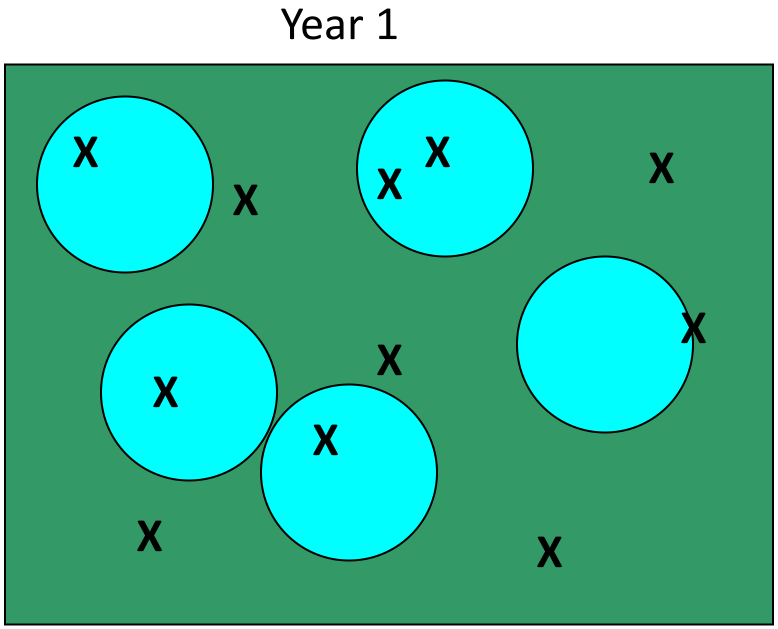
Annual Changes in CPUE Catchability

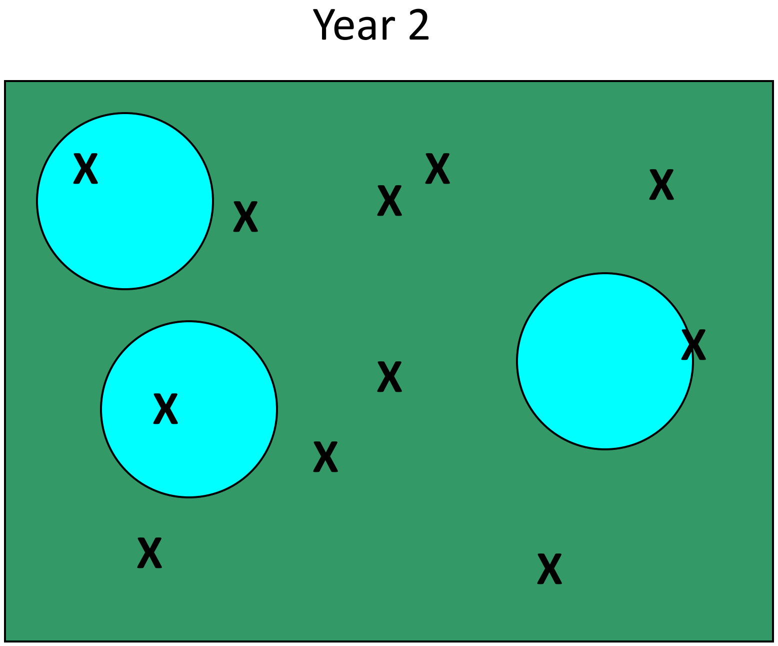


Surveys

- Directed stock surveys do not usually target high-density aggregations of fish and they give a better indication of stock trends
- Good reference:
- "Surveys of Fisheries Resources" by D.R. Gunderson.

Survey





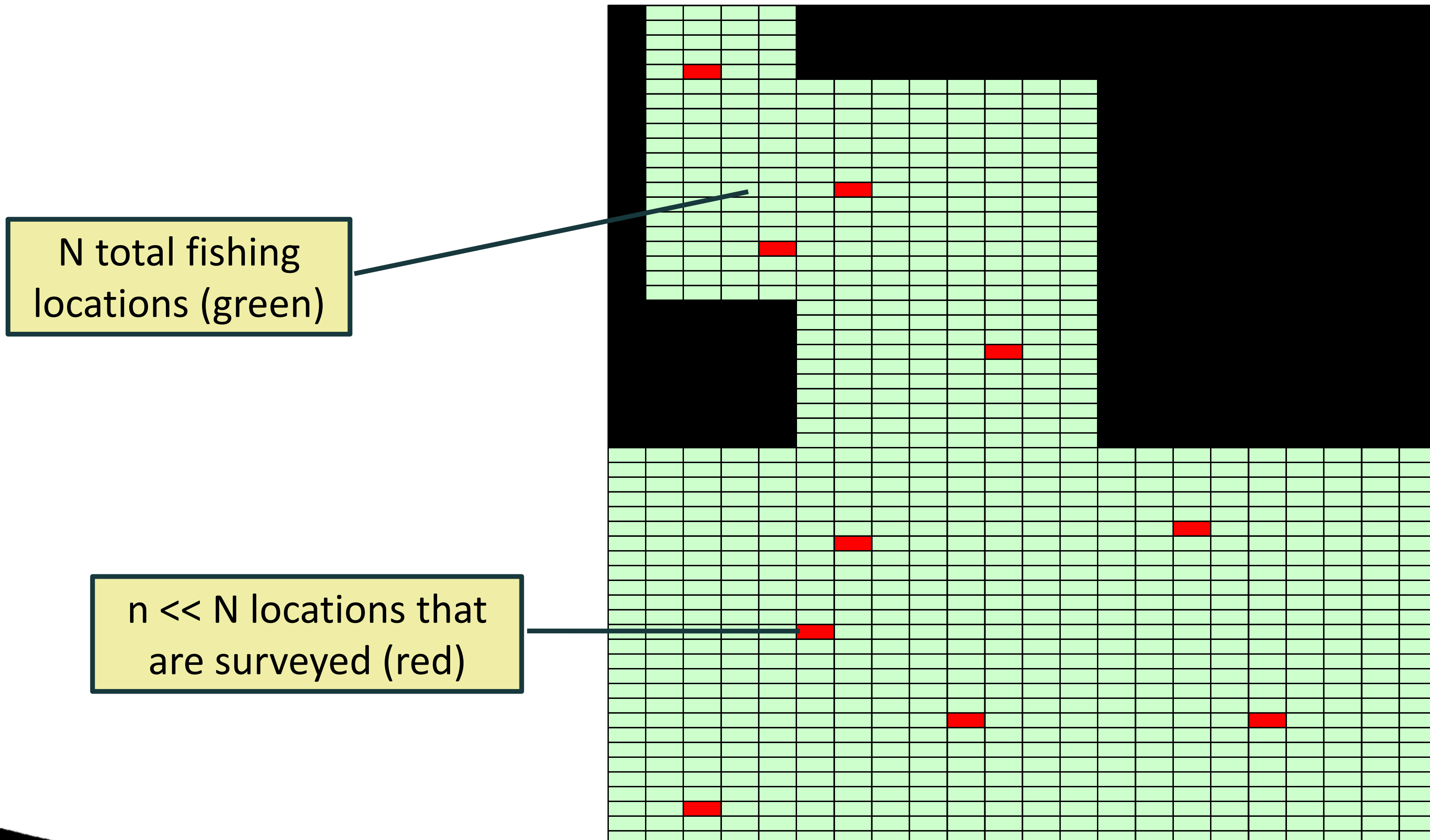
In year 1 the survey detected 4 clusters, but only 3 in year 2

Surveys

- A fishery-independent survey is important for successful stock management
- But it does not guarantee successful stock management
- There are many types of surveys: trawl, acoustic, fishery gear surveys, questionnaires to fishermen, visual counts via shipboard and air
- And many designs: random, stratified, systematic, transect methods

Survey Designs – random

– Involves choosing *n* fishing locations in the survey completely at random from a total of *N* possible fishing locations (a population of locations).



Survey Designs – stratified-random

– Assume there is some natural way to divide the population of N fishing locations into H strata, each with N_h fishing locations, h = 1,…,H.

– $N_1 + N_2 + N_3 + \ldots + N_H = N$.

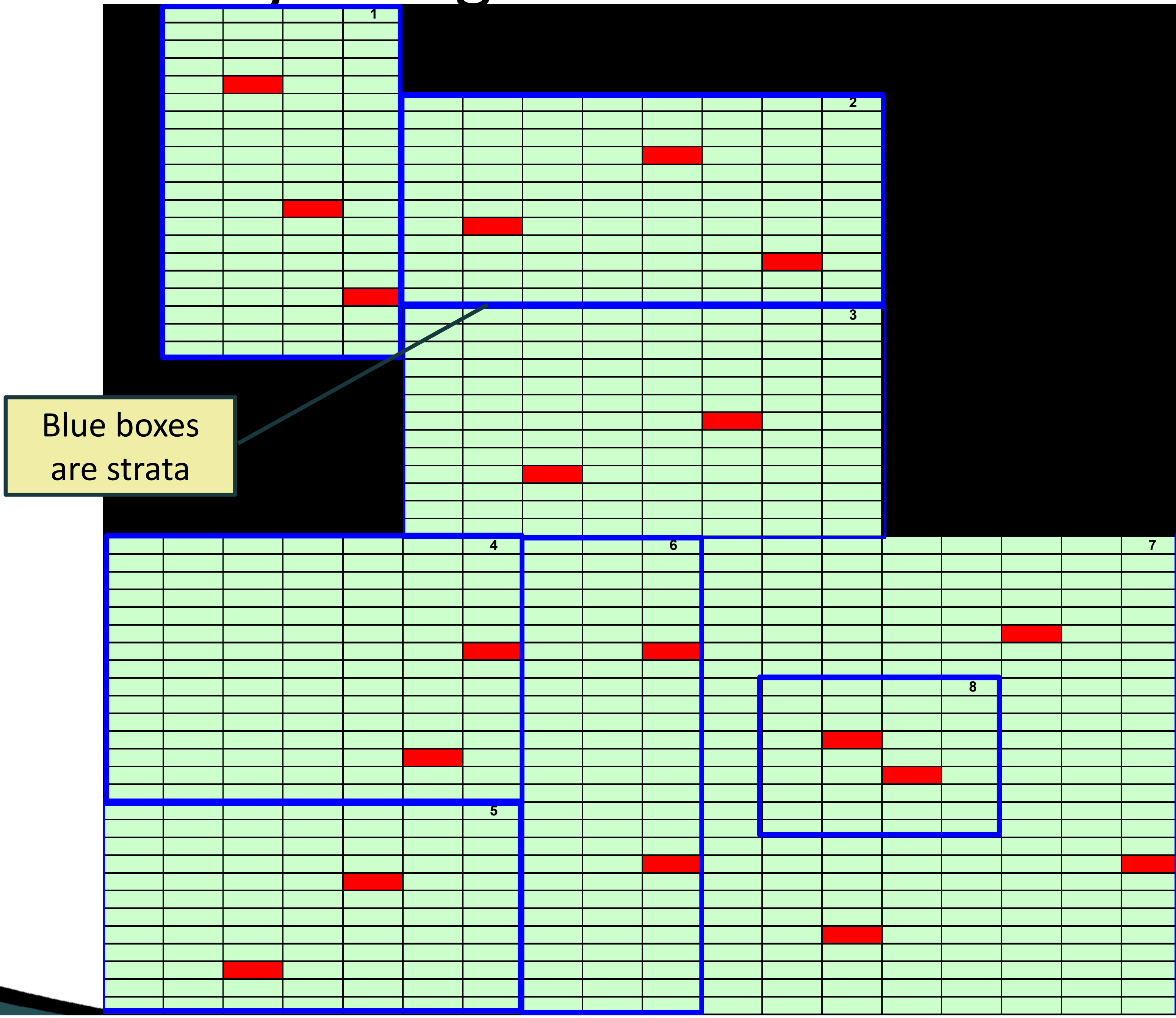
– A convenient way to express the sum is $\sum_{h=1}^{H} N_h = N$

Often in fishery surveys, strata can be formed for depth intervals, banks, etc.

Survey Designs – stratified-random

- Stratification is effective when the variation in catch rates in a strata is much lower than the variation in catch rates between strata.
- In a stratified random survey design n_h locations are chosen completely at random from the N_h possible locations in stratum h.
- The total survey sample size is $\Sigma_h n_h = n$.
- Spatial strata ensures sampling is spread out.

Survey Designs – stratified-random



Survey Designs – systematic

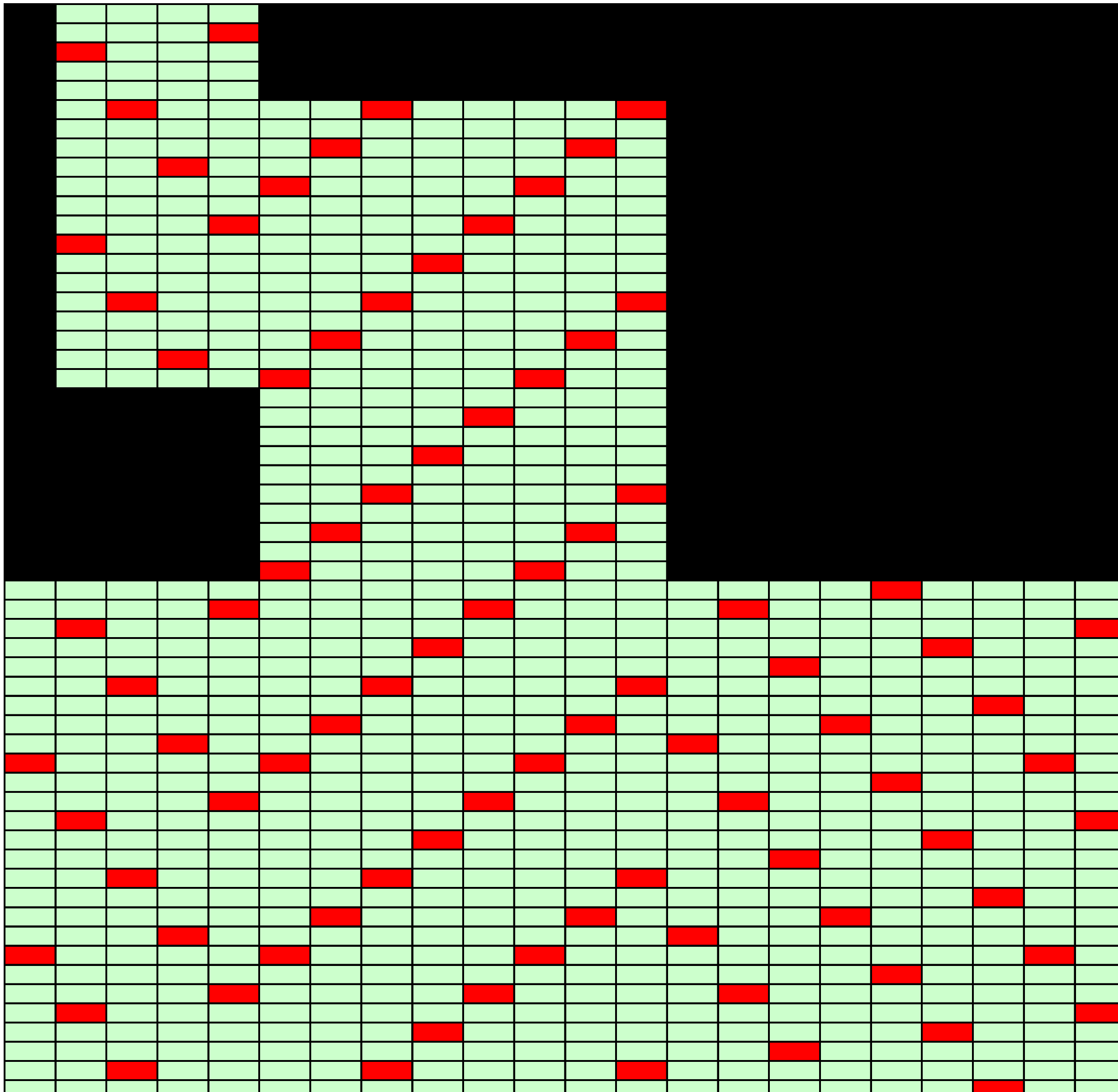
– The population of all possible sampling sites is ordered somehow, *1,….,N*.

– A systematic sample involves dividing the sites into bins (or strata) and selecting the same ordered unit from each bin.

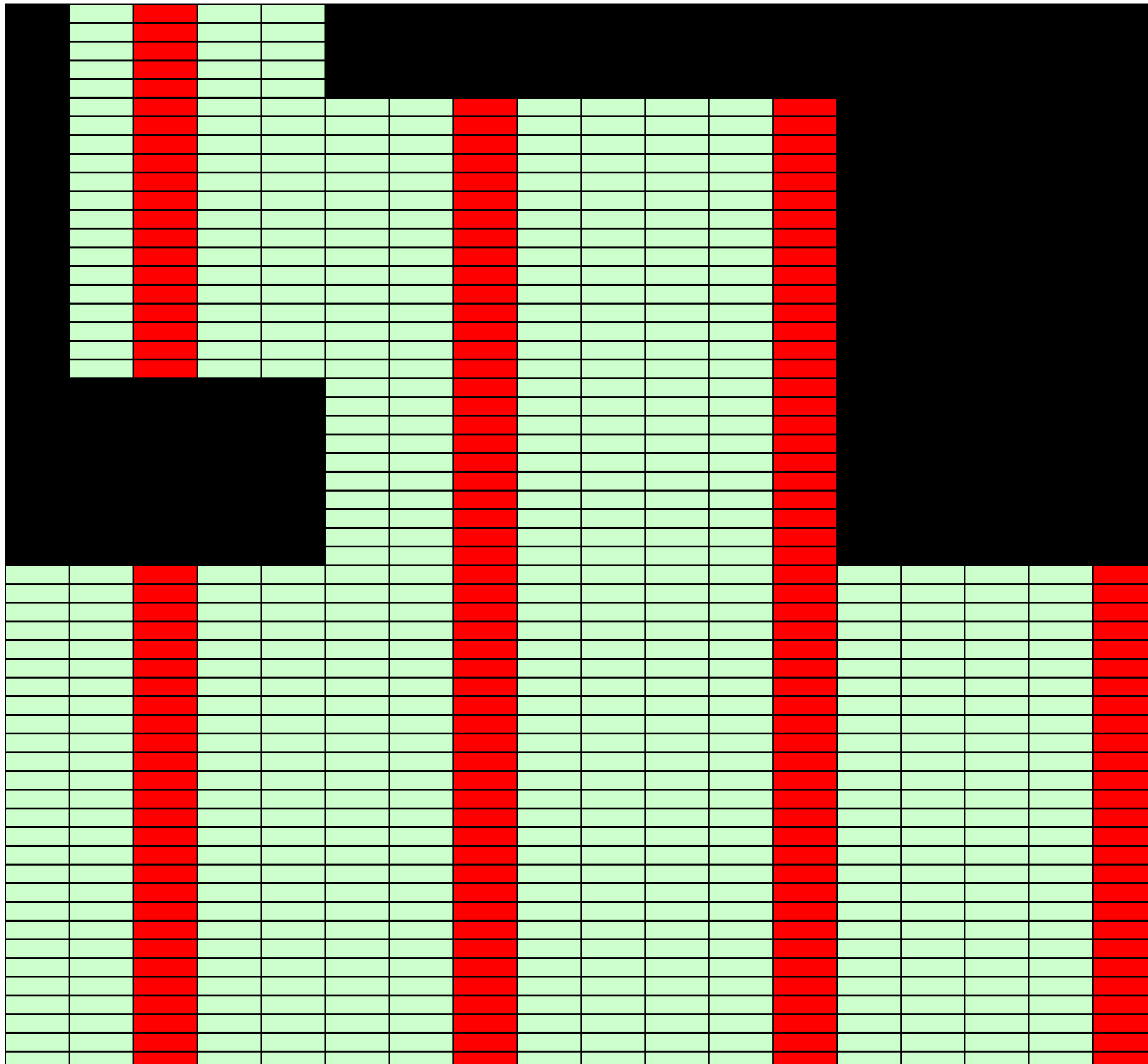
– Eg. Say we divide the *N* units into groups of 10, and decide to sample the 5^{th} unit in each bin (i.e. 5, 15, 25,….)

– It is stratified sampling, but not random

Survey Designs – systematic



Survey Designs – line transect



Survey Designs – others

– An important feature of these designs is that where you fish is determined before the survey.

– There are many other designs

– Some are adaptive, in which additional samples are chosen based on the results of the survey.

– Eg. 2-phase stratified random sampling

– Need to be careful because adaptive sampling can lead to biased estimates

Surveys

- Trawling and fishery gear sampling could be used with any of these designs, although stratified random sampling and systematic are common choices.
- Acoustic and visual (ship or air) sampling could also be used with any of these designs, although transect sampling is common.
- In fact, we could use these designs with questionnaires to fishermen.

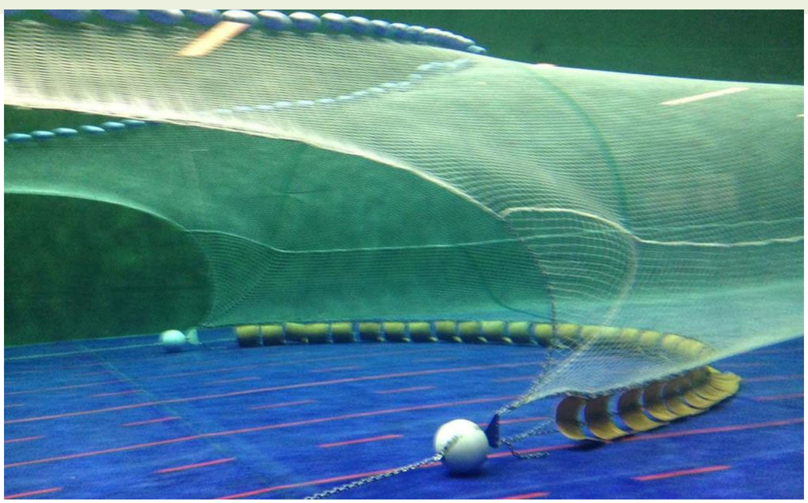
Stratified random bottom trawl

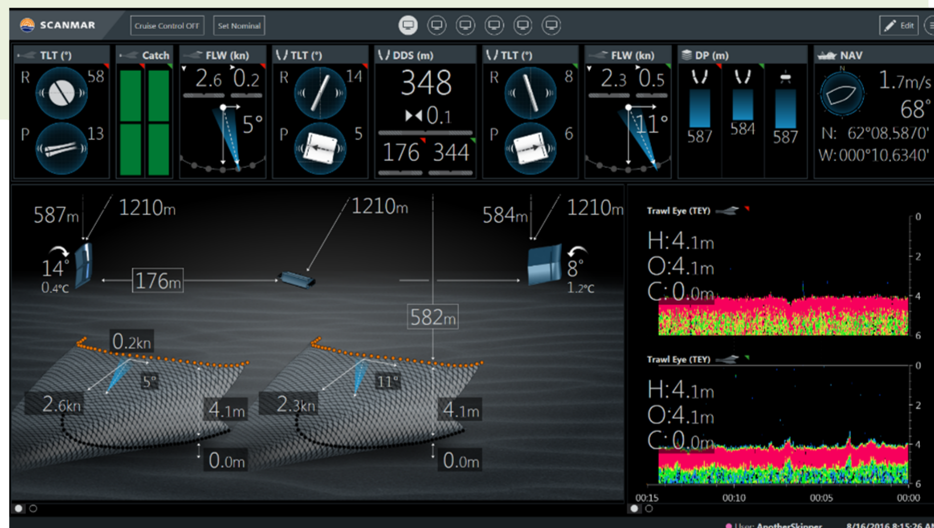
- Involves towing a standard net at a fixed speed and for a fixed duration.
- The intent is to ensure the swept-area is the same for each sample (i.e. tow).
- Measure the catch of many species, and take a variety of biological measurements.
- The sampling unit is very small relative to the stock area – there are 000's of units that could be sampled.

DFO Survey Trawl

– Campelen 1800 survey trawl with a 12.7 mm liner in the lower 7 m of the codend, and a standard tow of 15 minutes on bottom at 3.0 knots

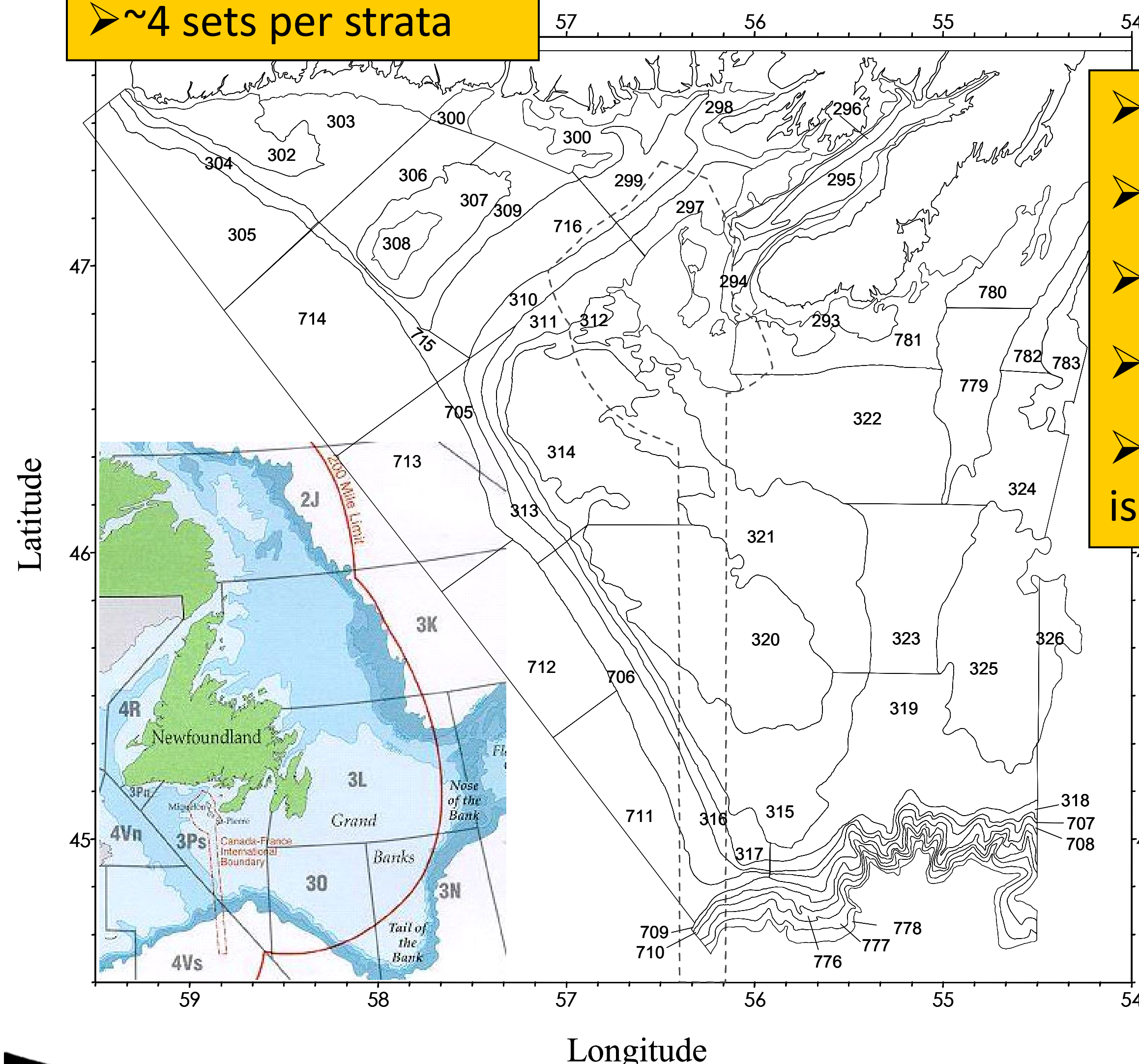
– Use net monitoring systems to determine bottom contact and gear configuration (doors and/or wingspread)





3Ps survey

- 44 strata
- ~4 sets per strata



- 3Ps survey area is **67 000** Km^2,
- larger than NS (**55 000** Km^2)
- smaller than NB (**73 000** Km^2)
- or NL (**406 000** Km^2)
- About 60% of Newfoundland island (**114 000** Km^2).

3Ps survey

– Stratified random sampling, with proportional allocation (n_h approx. proportional to N_h)

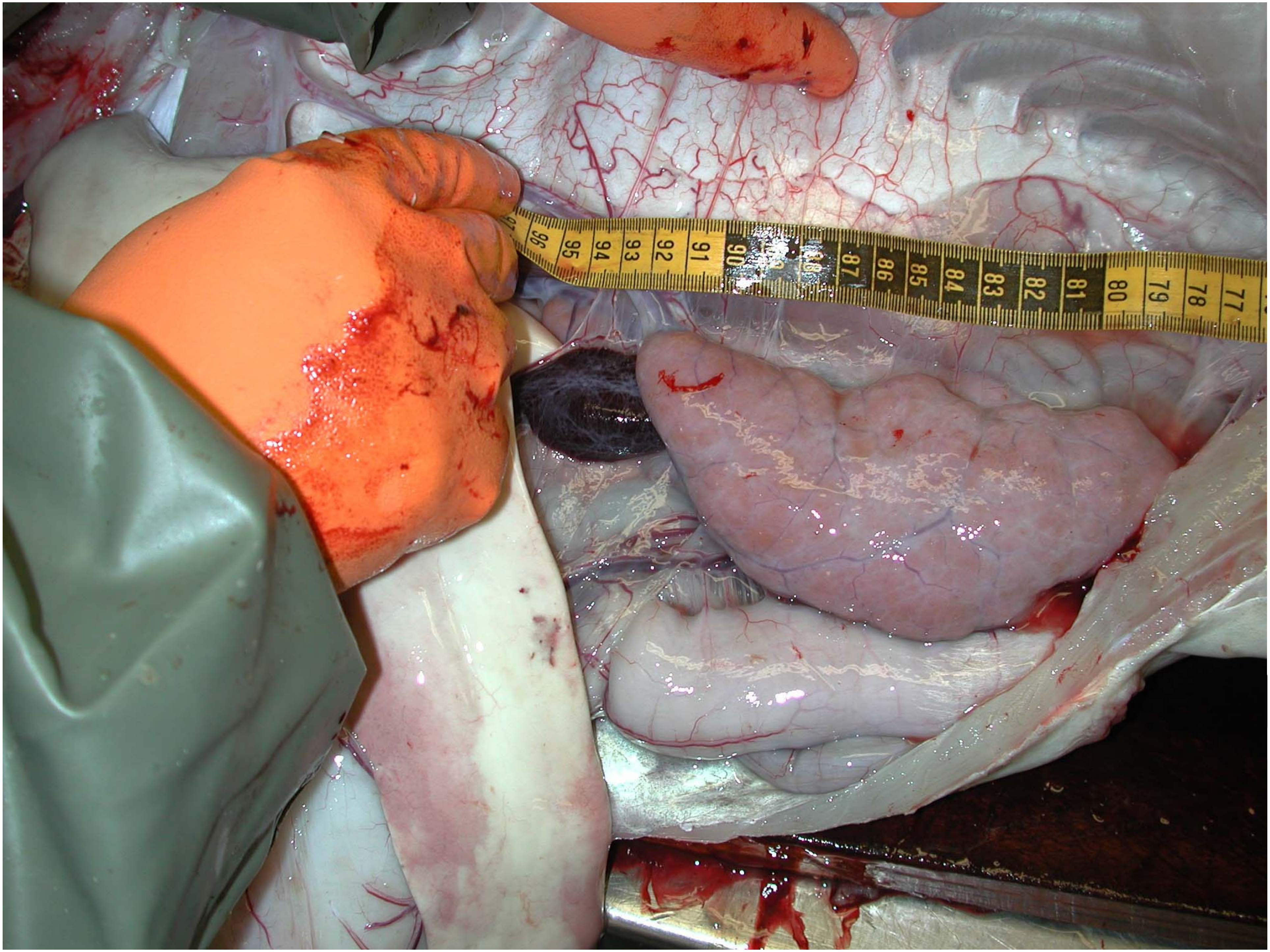
– A minimum of two sets per strata.

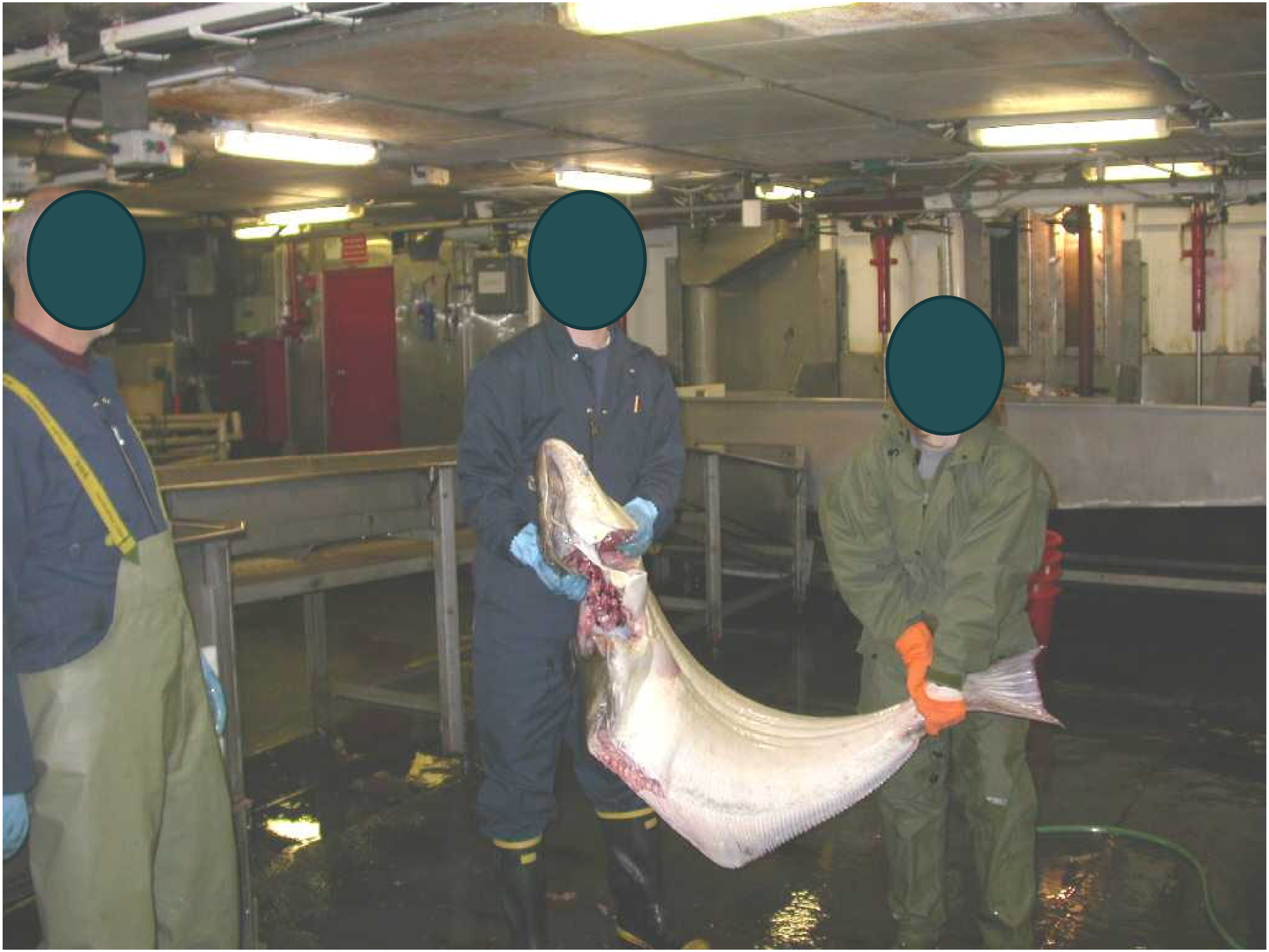
– The number of cod caught per tow are standardized to a 15 minute tow, at 3 knots. This has usually been the practise since 1996.

– Approximately 50 species counted, and sampled for length, age, sex, maturity, and condition status.

Some photos from a survey about 15 years ago


FARNELL


96
95
94
93
92
91
90
87
86
85
84
83
82
81
80
79
78
77




EXIT


EXIT
RANGER
TOUGH


NOTE
DAVE'S BAG TO BE CODED IN FFS AS
8018
Dave

DFO 3Ps Research Vessel trawl survey

- Survey area divided into areas of same depth (strata).
- Several fishing sets with positions assigned randomly within each stratum. Currently 178 total sets allocated across 45 strata.
- Catch counted & weighed. Area trawled scaled up to whole stratum area.
- Estimates from all strata combined to get total abundance & biomass indices.

DFO 3Ps Research Vessel trawl survey

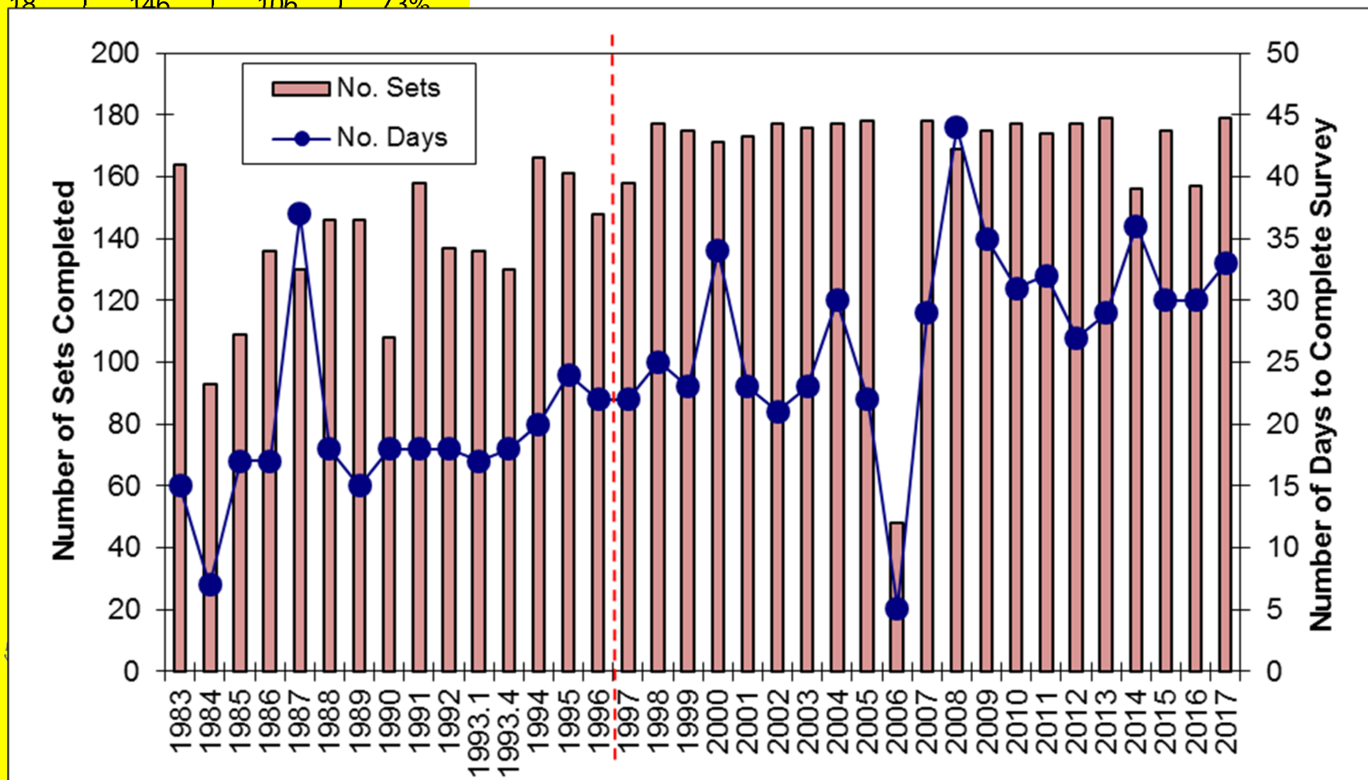
- Biological Sampling provides information on maturity, growth, condition, etc.
- Campelen trawl, 15 min tows. Surveyed in April/May (since 1993).
- Survey not completed during Spring of 2006.

Survey Statistics

Year	Start Date	End Date	No. Days	No. Sets	No. Sets w/ Cod	% w/ cod
1983	23-Apr-83	8-May-83	15	164	117	71%
1984	10-Apr-84	17-Apr-84	7	93	59	63%
1985	8-Mar-85	25-Mar-85	17	109	78	72%
1986	6-Mar-86	23-Mar-86	17	136	88	65%
1987	13-Feb-87	22-Mar-87	37	130	95	73%
1988	27-Jan-88	14-Feb-88	18	146	106	73%
1989	1-Feb-89	16-Feb-89	15	146	90	62%
1990	1-Feb-90	19-Feb-90	18	108	66	61%
1991	2-Feb-91	20-Feb-91	18	158	104	66%
1992	6-Feb-92	24-Feb-92	18	137	63	46%
1993.1	6-Feb-93	23-Feb-93	17	136	52	38%
1993.4	2-Apr-93	20-Apr-93	18	130	63	48%
1994	6-Apr-94	26-Apr-94	20	166	73	44%
1995	04-Apr-95	28-Apr-95	24	161	65	40%
1996	10-Apr-96	01-May-96	22	148	105	71%
1997	02-Apr-97	23-Apr-97	22	158	104	66%
1998	10-Apr-98	05-May-98	25	177	113	64%
1999	13-Apr-99	06-May-99	23	175	128	73%
2000	08-Apr-00	11-May-00	34	171	136	80%
2001	07-Apr-01	29-Apr-01	23	173	134	77%
2002	05-Apr-02	27-Apr-02	21	177	117	66%
2003	05-Apr-03	02-May-03	23	176	117	66%
2004	11-Apr-04	11-May-04	30	177	107	60%
2005	17-Apr-05	09-May-05	22	178	134	75%
2006	*13-Apr-06*	*18-Apr-06*	*5.1*	*48*	*43*	
2007	04-Apr-07	02-May-07	29	178	135	76%
2008	10-Apr-08	23-May-08	44	169	115	68%
2009	08-Apr-09	13-May-09	35	175	137	78%
2010	08-Apr-10	08-May-10	31	177	132	75%
2011	07-Apr-11	08-May-11	32	174	131	75%
2012	31-Mar-12	26-Apr-12	27	177	137	77%
2013	26-Mar-13	23-Apr-13	29	179	133	74%
2014	05-Apr-14	10-May-14	36	156	105	67%
2015	11-Apr-15	10-May-15	30	175	116	66%
2016	02-Apr-16	01-May-16	30	157	110	70%
2017	06-Apr-17	08-May-17	33	179	121	68%

Survey Statistics

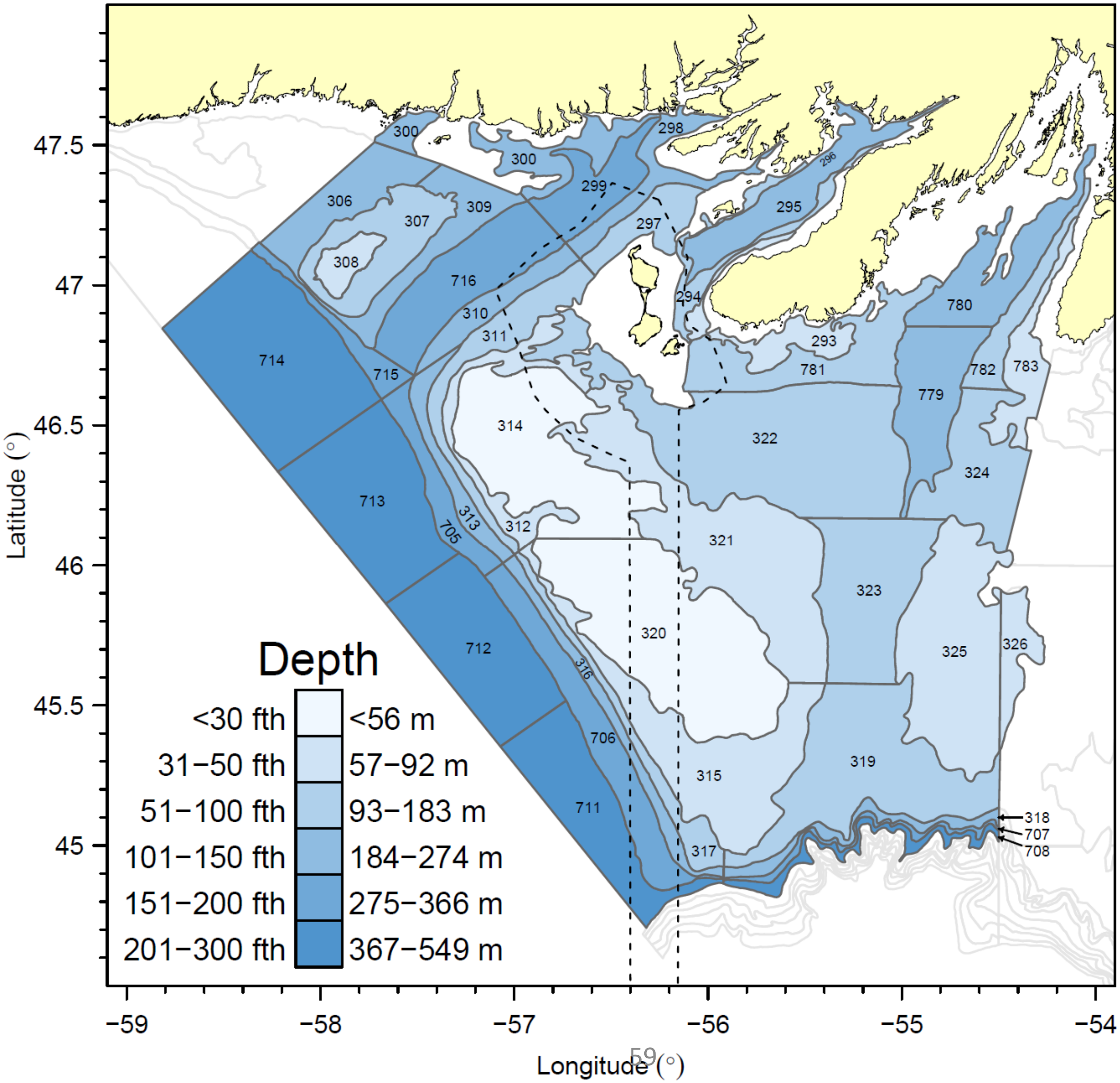
Year	Start Date	End Date	No. Days	No. Sets	No. Sets w/ Cod	% w/ cod
1983	23-Apr-83	8-May-83	15	164	117	71%
1984	10-Apr-84	17-Apr-84	7	93	59	63%
1985	8-Mar-85	25-Mar-85	17	109	78	72%
1986	6-Mar-86	23-Mar-86	17	136	88	65%
1987	13-Feb-87	22-Mar-87	37	130	95	73%
1988	27-Jan-88	14-Feb-88	18	146	106	73%
1989	1-Feb-89	16-Feb-89				
1990	1-Feb-90	19-Feb-90				
1991	2-Feb-91	20-Feb-91				
1992	6-Feb-92	24-Feb-92				
1993.1	6-Feb-93	23-Feb-93				
1993.4	2-Apr-93	20-Apr-93				
1994	6-Apr-94	26-Apr-94				
1995	04-Apr-95	28-Apr-95				
1996	10-Apr-96	01-May-96				
1997	02-Apr-97	23-Apr-97				
1998	10-Apr-98	05-May-98				
1999	13-Apr-99	06-May-99				
2000	08-Apr-00	11-May-00				
2001	07-Apr-01	29-Apr-01				
2002	05-Apr-02	27-Apr-02				
2003	05-Apr-03	02-May-03				
2004	11-Apr-04	11-May-04				
2005	17-Apr-05	09-May-05				
2006	*13-Apr-06*	*18-Apr-06*				
2007	04-Apr-07	02-May-07				
2008	10-Apr-08	23-May-08				
2009	08-Apr-09	13-May-09				
2010	08-Apr-10	08-May-10	31	177	132	73%
2011	07-Apr-11	08-May-11	32	174	131	75%
2012	31-Mar-12	26-Apr-12	27	177	137	77%
2013	26-Mar-13	23-Apr-13	29	179	133	74%
2014	05-Apr-14	10-May-14	36	156	105	67%
2015	11-Apr-15	10-May-15	30	175	116	66%
2016	02-Apr-16	01-May-16	30	157	110	70%
2017	06-Apr-17	08-May-17	33	179	121	68%



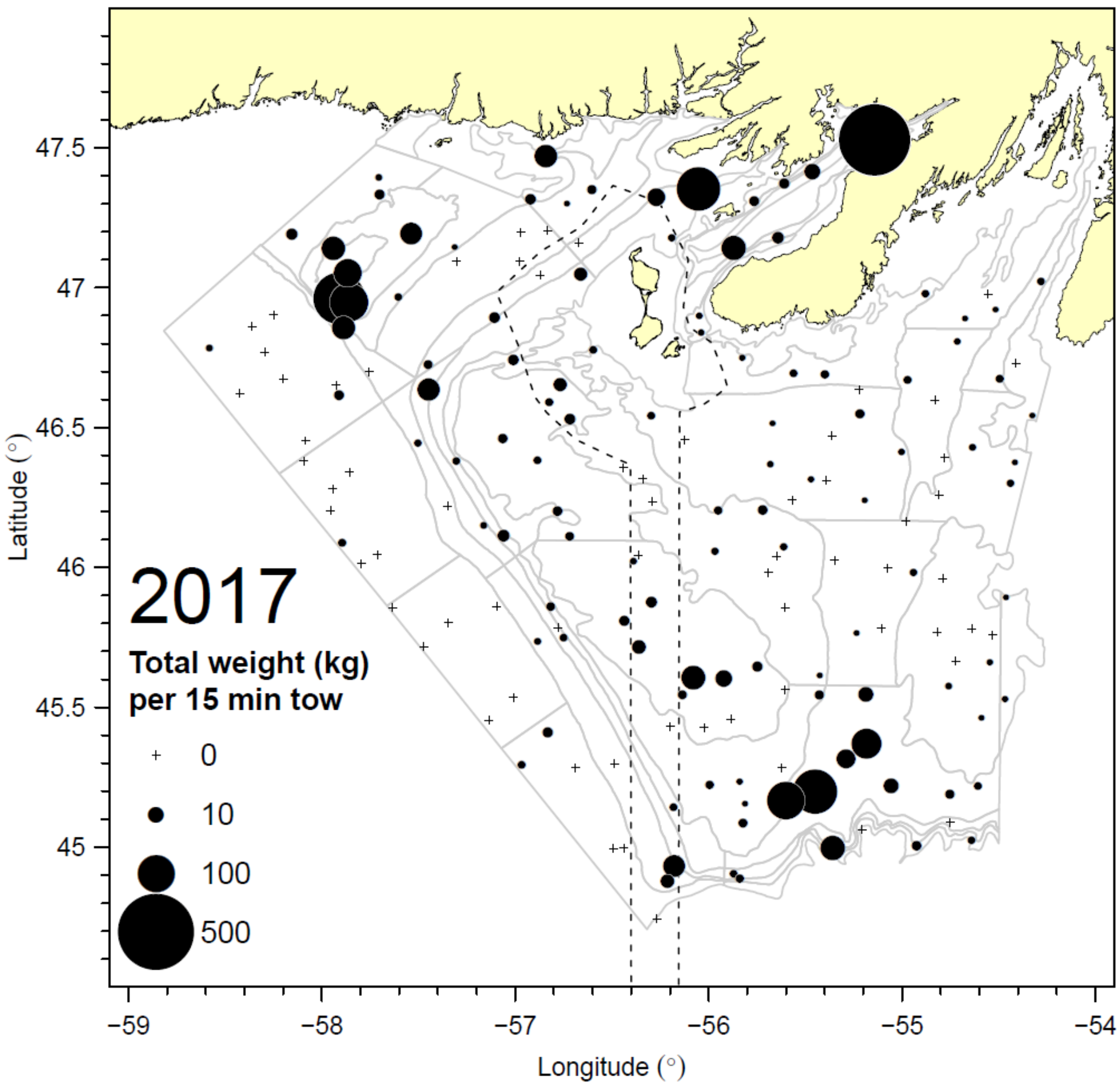
<u>2017 survey</u>

179/178 sets compl

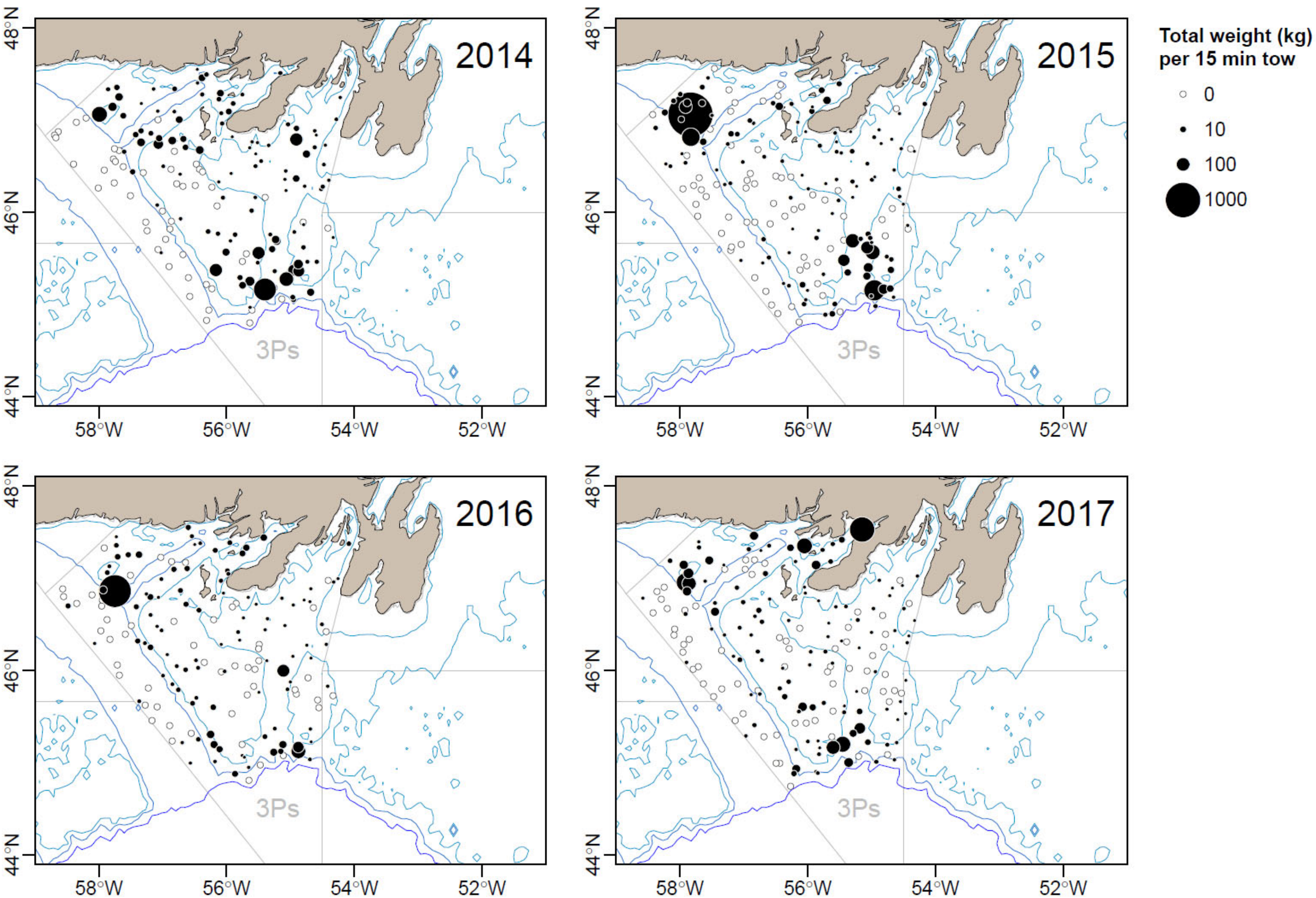
3Ps Stratification



2017 Survey Sets



RV Survey – Distribution
Weight per Standard Tow
*Bubble size proportional to weight caught.



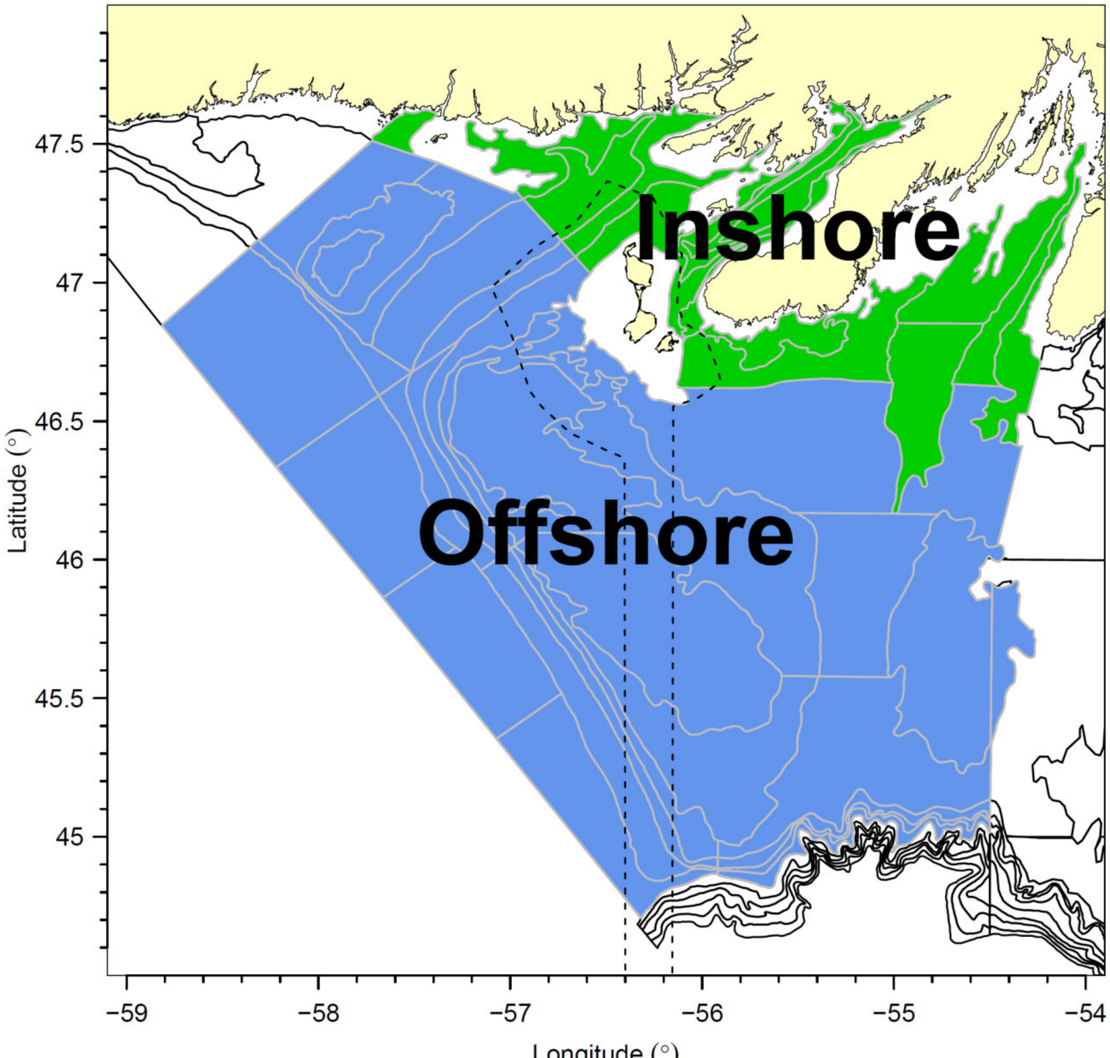
DFO research vessel trawl survey

New strata added 1994 and 1997 – 12% increase in surveyed area.

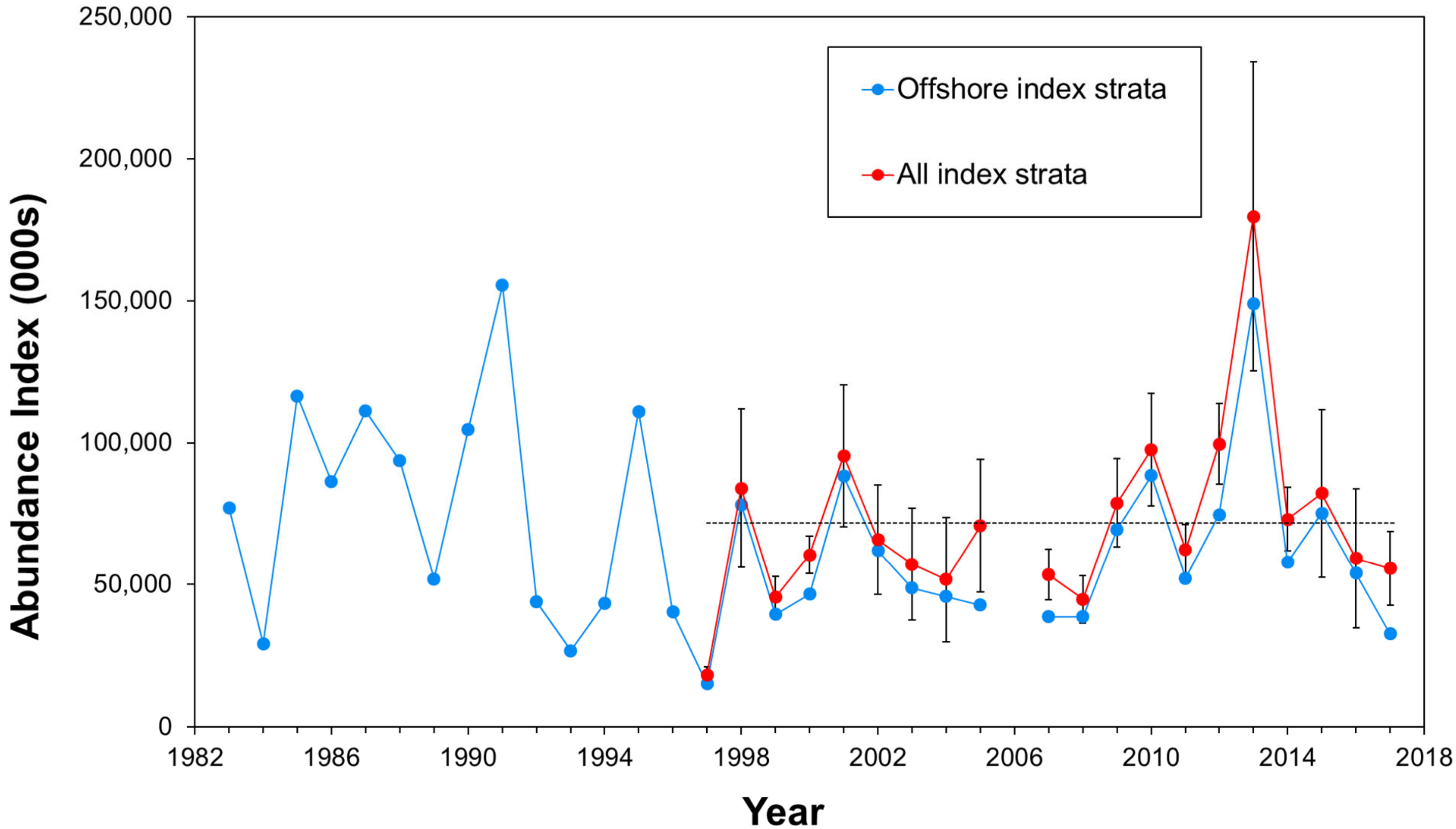
- Strata 779-783 added in 1994
- Strata 293-300 added in 1997
- Two time series:

  Inshore/Offshore (1997-2017)

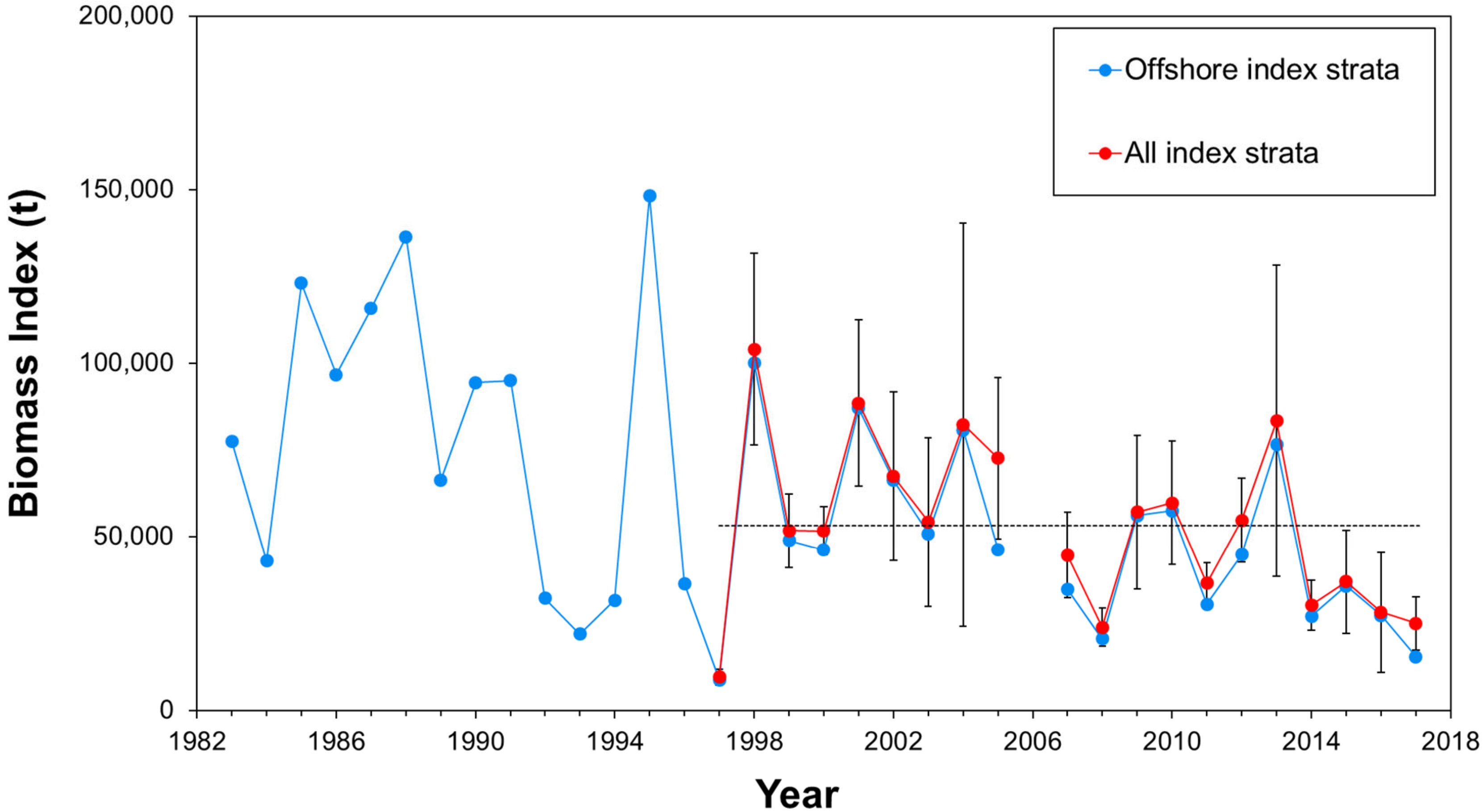
  Offshore (1983-2017)



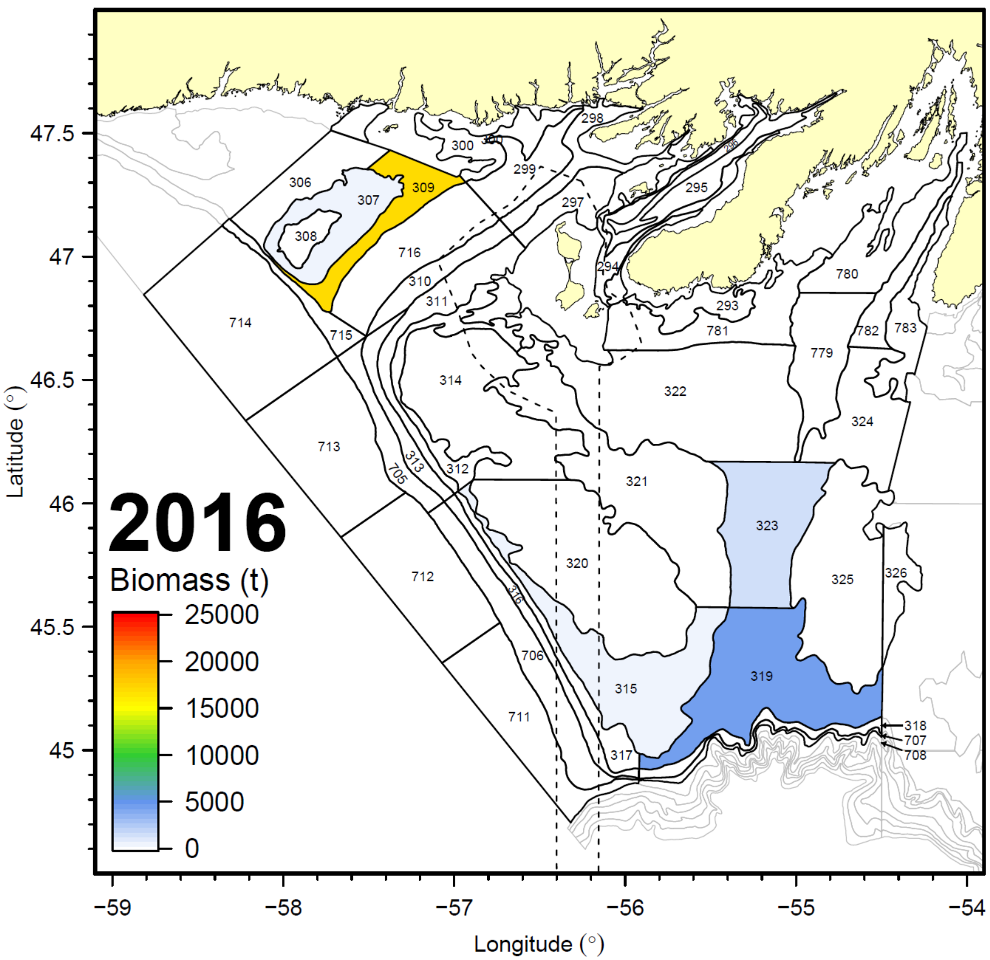
RV Survey
Abundance Index



RV Survey
Biomass Index



Concentration of Survey Index



Index includes 45 strata, yet much of abundance and biomass estimates typically found in just a few strata.

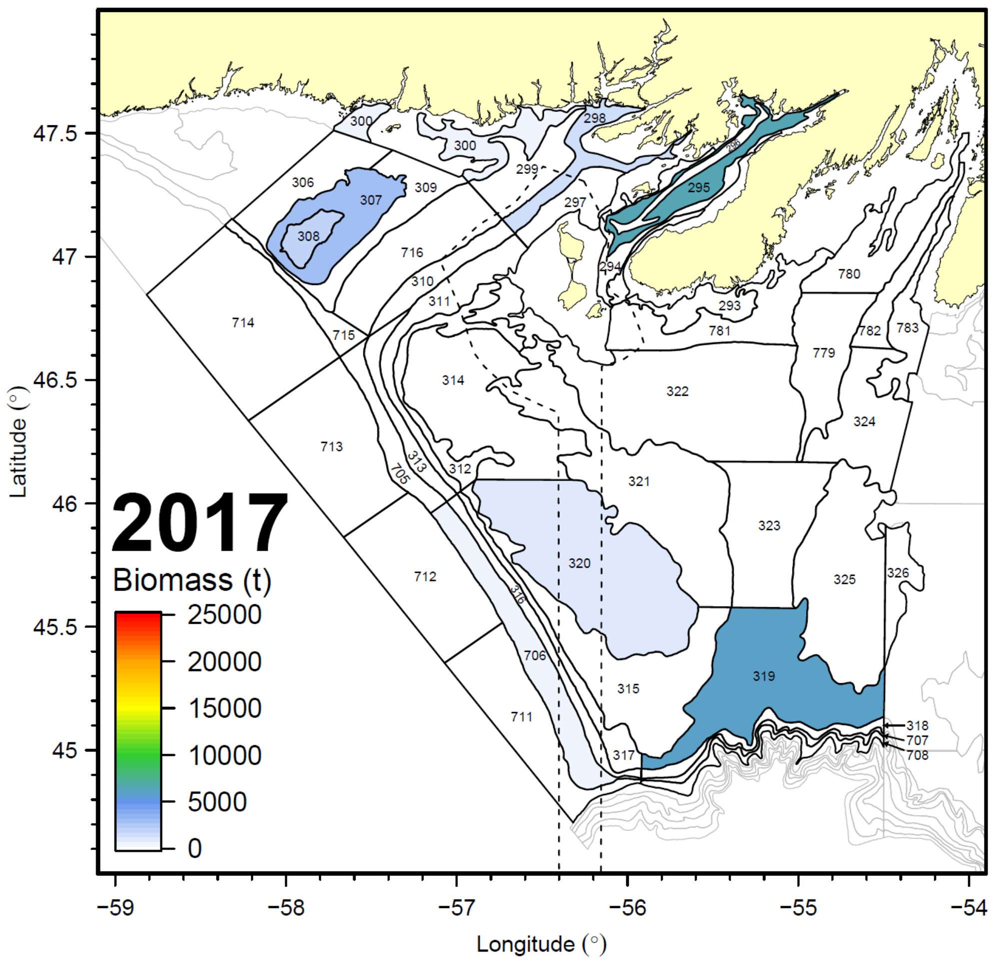
Typically, Stratum 319 most dominant. More than 15% of annual abundance and/or biomass in all but 3 years over 1997-2017.

<u>2016 survey</u>

Stratum 319:
23% (14 mil) of abundance index
16% (4.5 kt) of biomass index.

Stratum 309:
38% (22.5 mil) of abundance index
60% ((17 kt) of biomass index.

Concentration of Survey Index

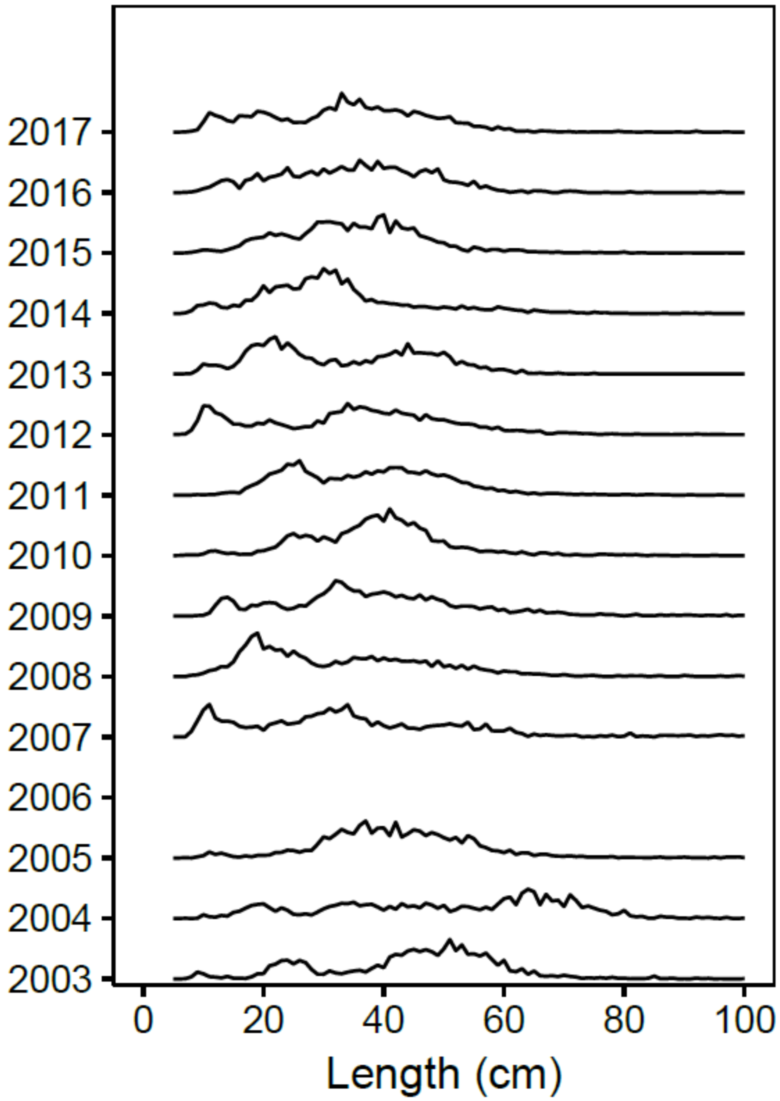


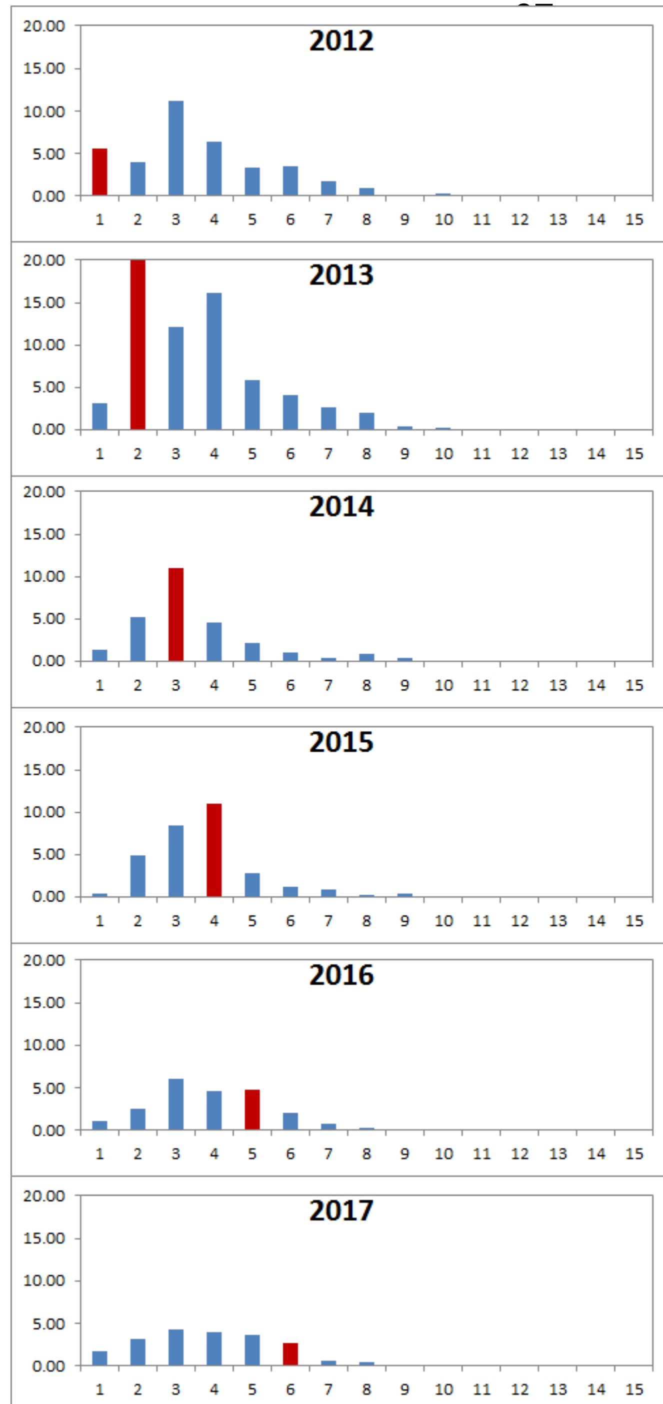
<u>2017 survey</u>

Stratum 319:
21% (11.6 mil) of abundance index
25% (6.2 kt) of biomass index.

Stratum 295:
24% (13.5 mil) of abundance index
27% ((6.7 kt) of biomass index.

Survey MNPT at Age



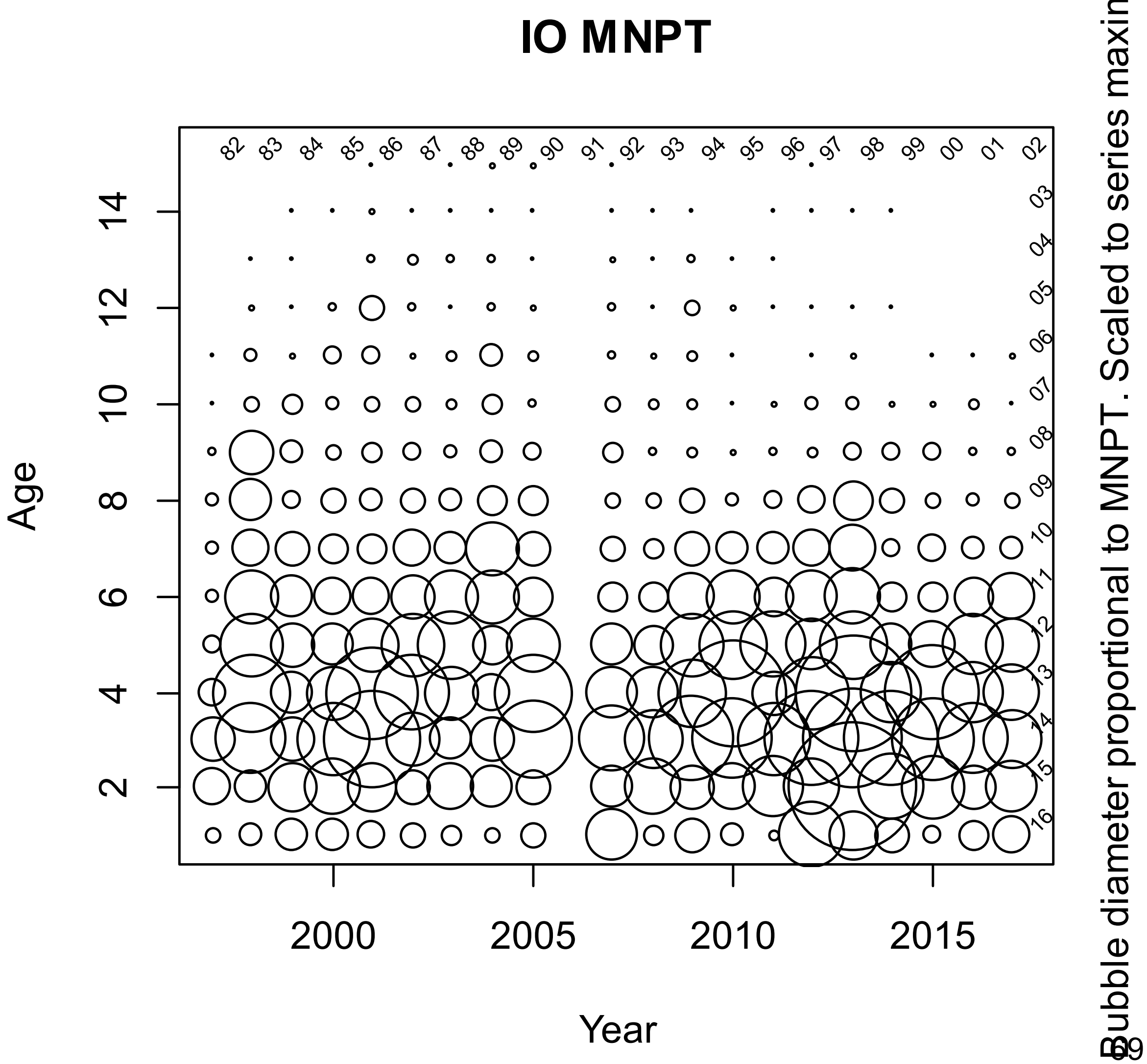


“Bubble” Plots

- Several ways in which to produce bubble plots, all of which provide different perspective of catch or survey age-disaggregated information.

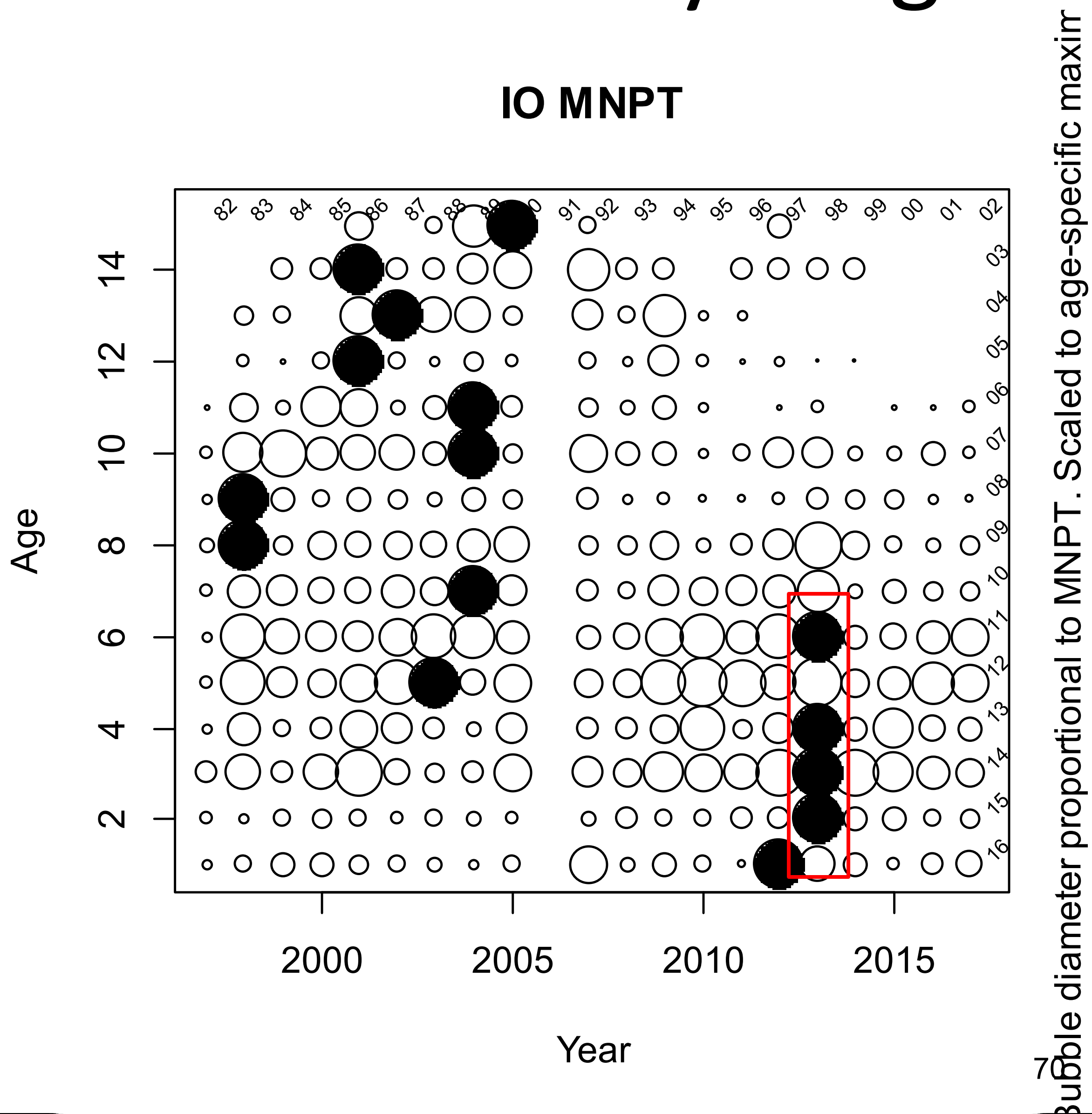
- Will consider:
  - Raw bubbles, area proportional to diameter
  - Age-standardized bubbles
  - standardized proportion at age per year plots (“spay”).

RV Survey – Age Comps



- 1997-2017
- Majority of fish caught in surveys are 2-6 years old. Much reduced catches of older fish.

RV Survey – Age Comps



- Scaled independently within each age. Filled circle indicates each age-specific maximum.
- Indication of year-effects.

"Spay" Plots (learn how to do these)

- Compute annual proportions at age.

$$P_{a,y} = \frac{I_{a,y}}{\sum_a I_{a,y}}$$

- Compute the mean and standard deviation of proportion at each age.

$$\overline{P_a} = \frac{\sum_y P_{a,y}}{N_y}$$

$$s_a = \sqrt{\frac{\sum_y \left(P_{a,y} - \overline{P_a}\right)^2}{Y-1}}$$

- Compute standardized proportion at age.

$$P_{a,y}{}^{s} = \frac{P_{a,y} - \overline{P_a}}{s_a}$$

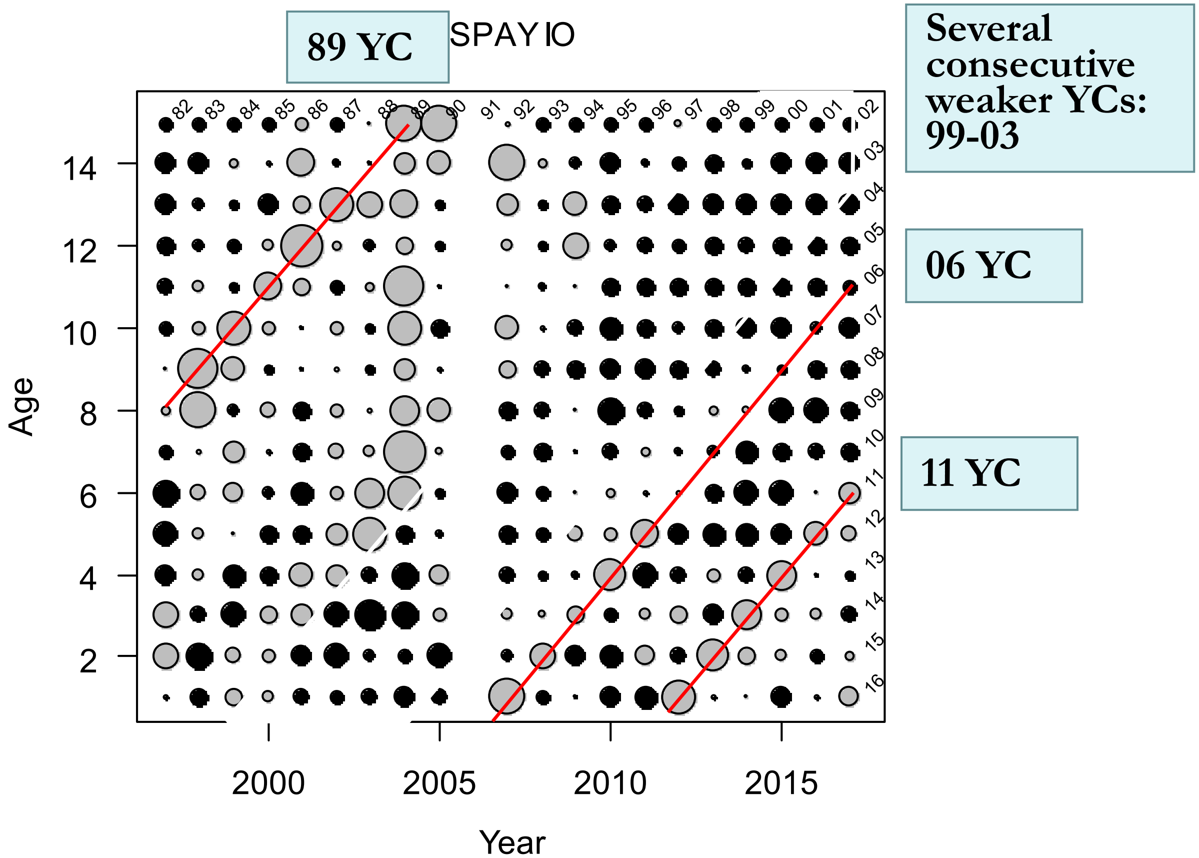
“Spay” Plots

- What can be inferred from these plots:
  - Contribution of each age group to that year’s index.
  - How the contribution of age *a* compares to all other age *a* data points.
  - How well a given cohort is “tracked” by the index.

- What cannot be interpreted from these plots:
  - Relative strength of cohorts.

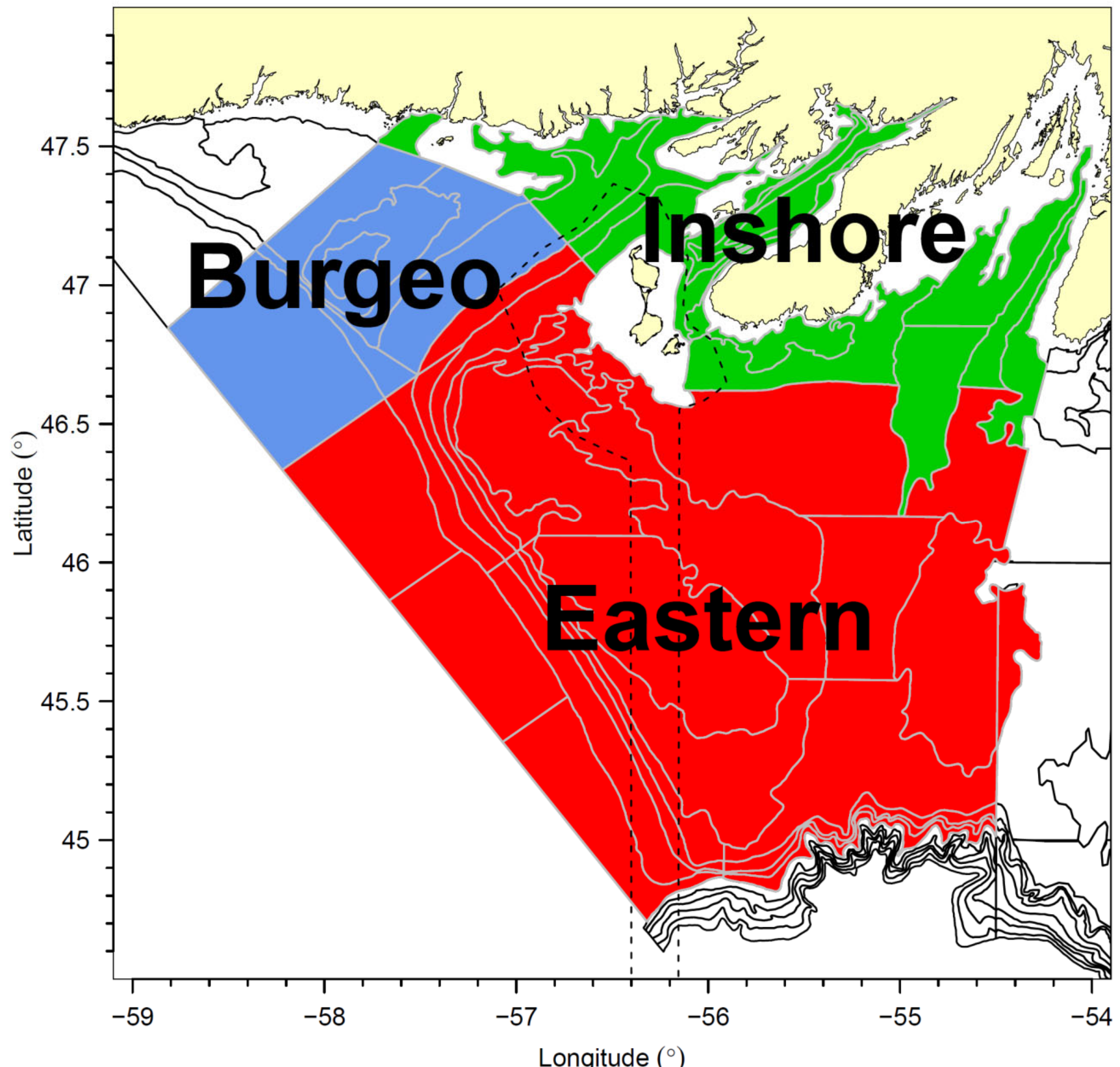
RV Survey – Age Comps

Grey – above average; black below average. Small dot is average

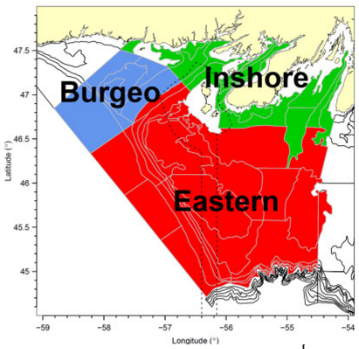


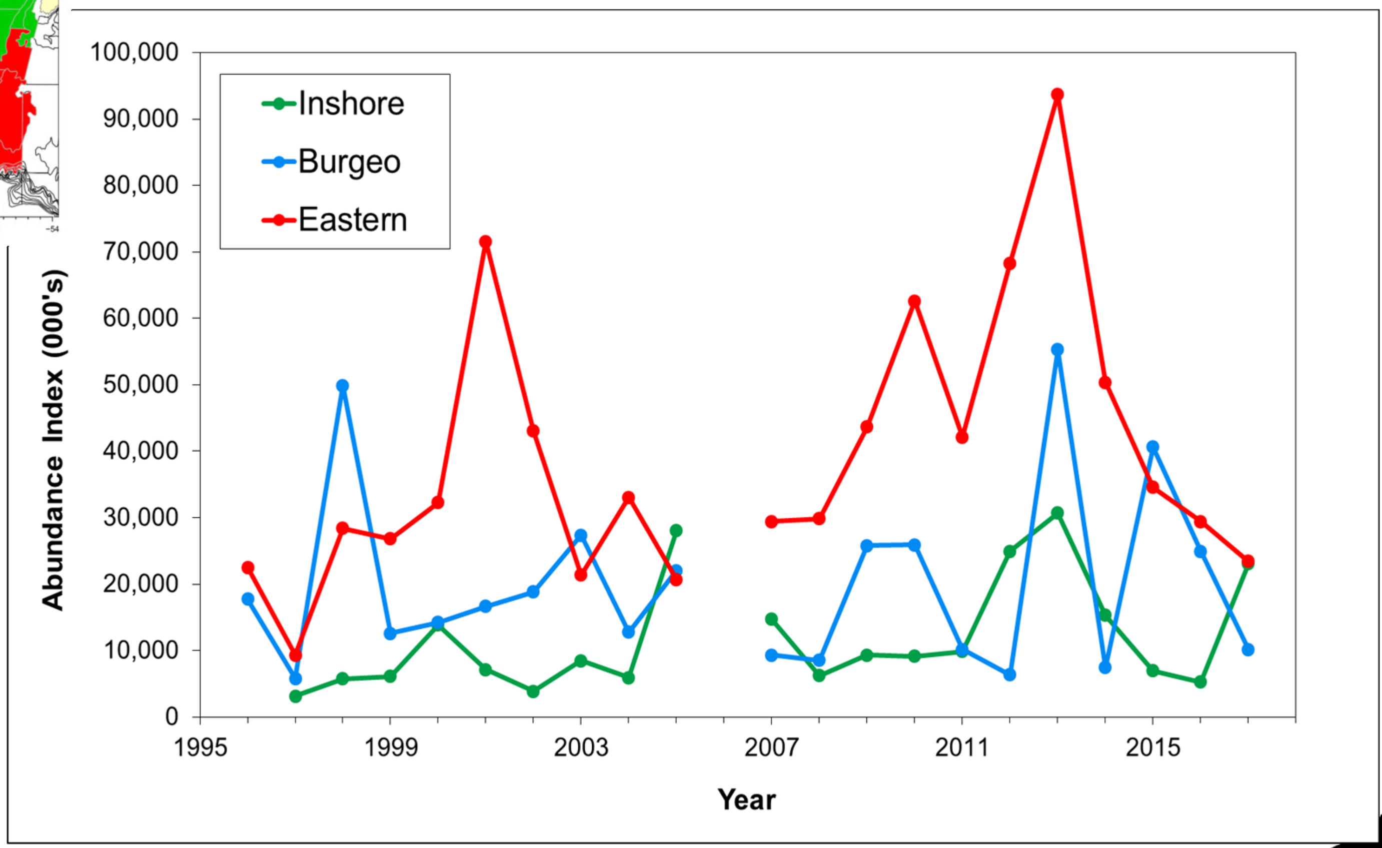
- Some cohorts track well (e.g. 89 YC).
- Indication of consecutive weaker YCs
- Low proportion of older fish in recent years
- 2011 YC tracking well so far

'Split Indices'

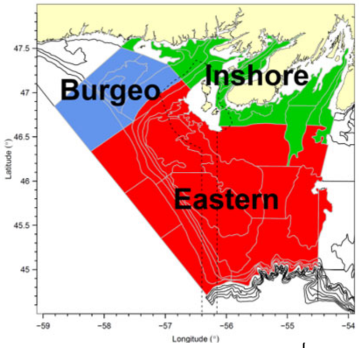


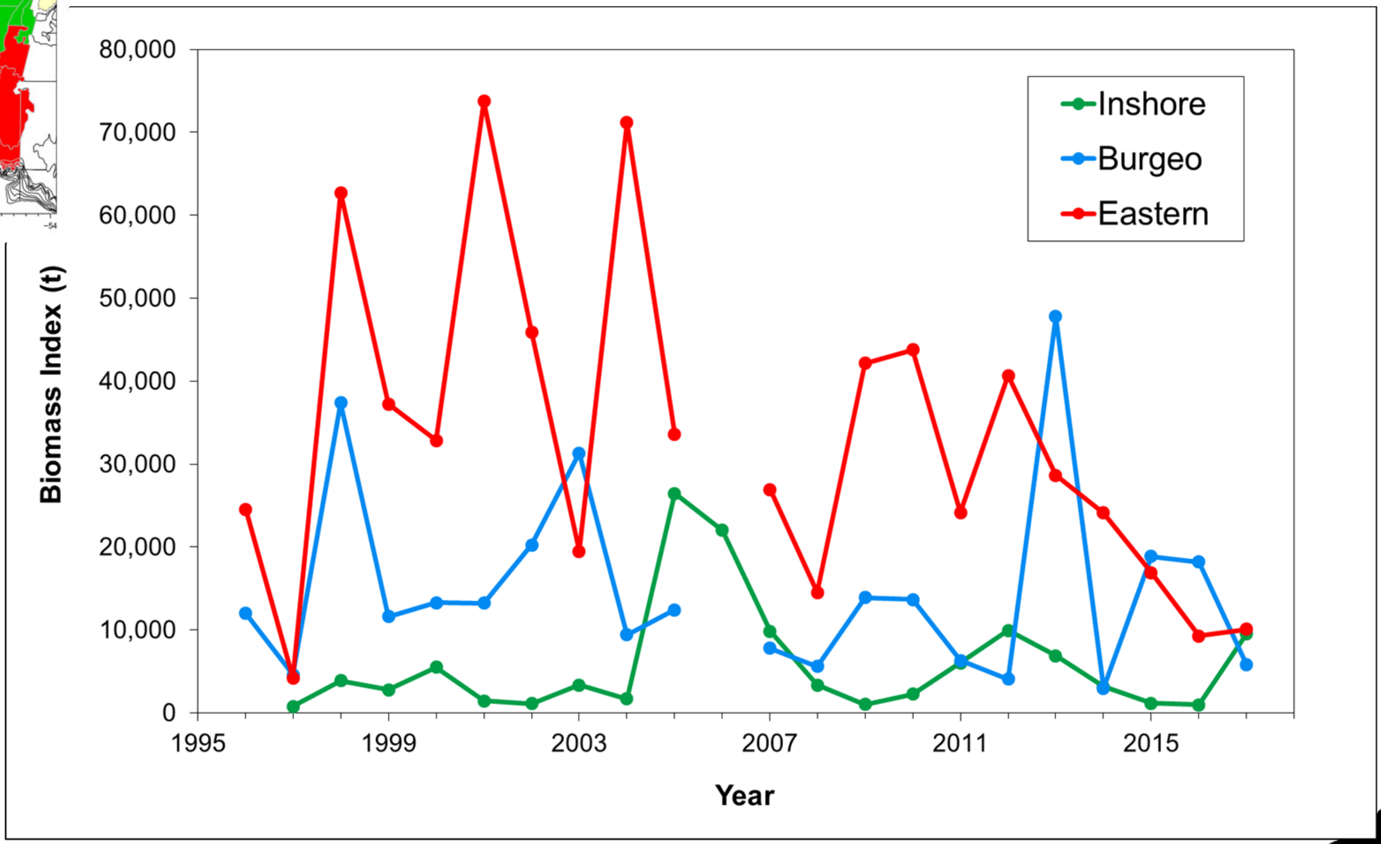
'Split Indices' - Abundance



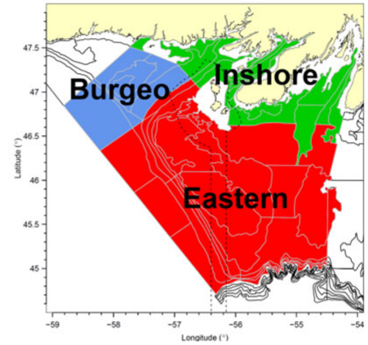


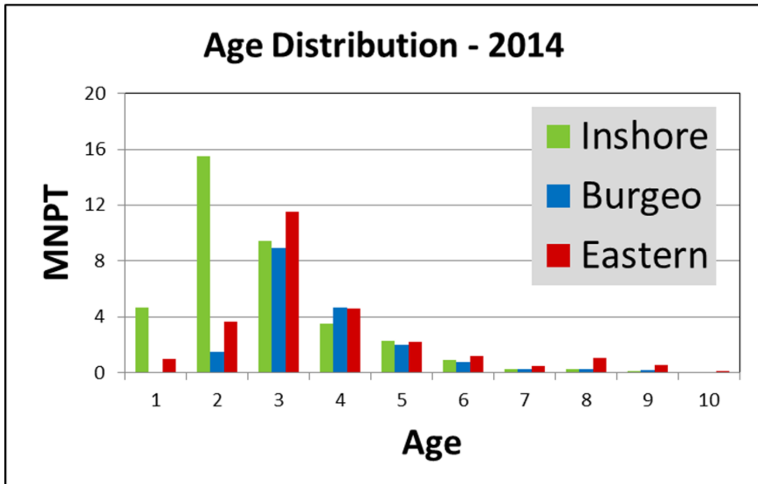
‘Split Indices’ - Biomass

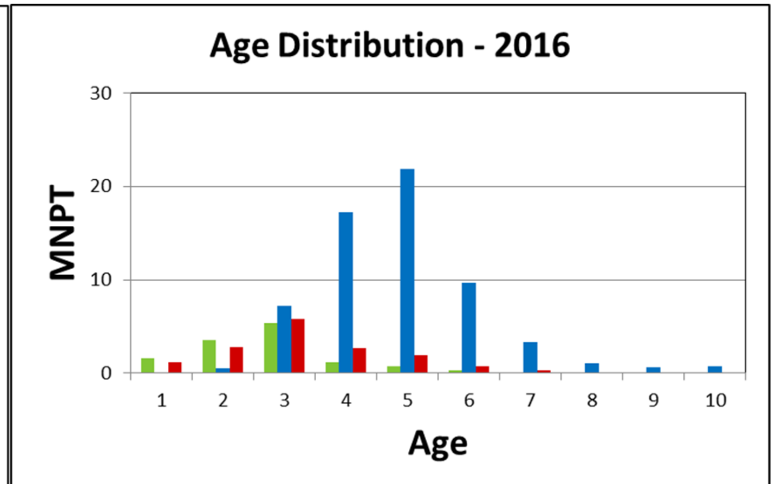


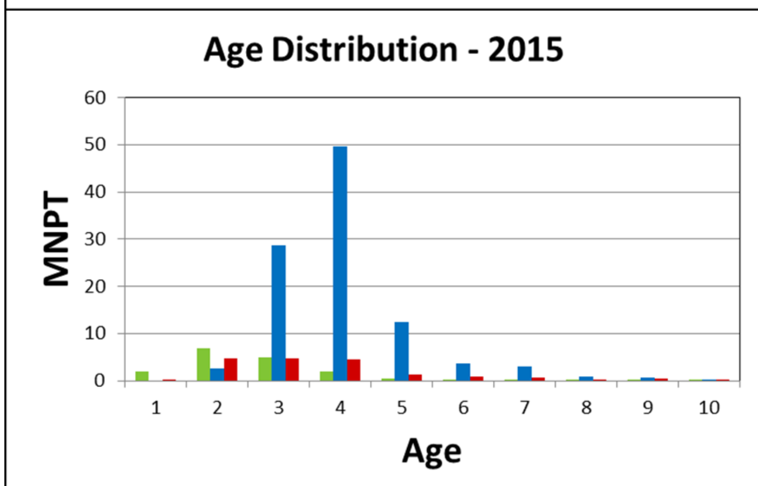


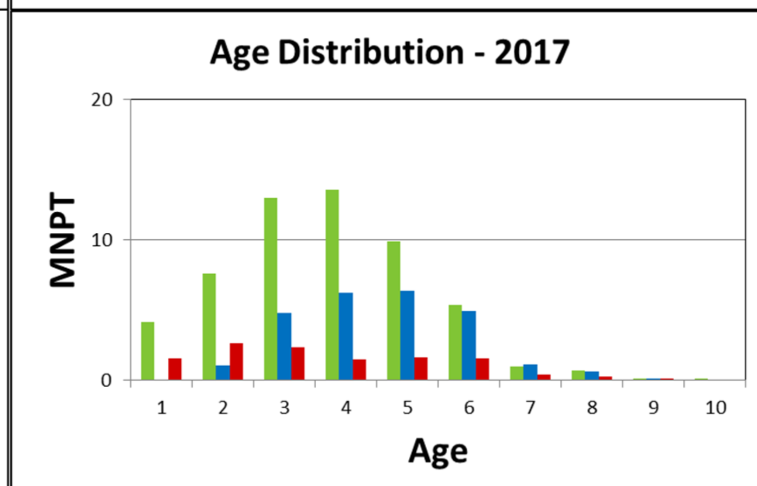
'Split Indices'
Age Distribution



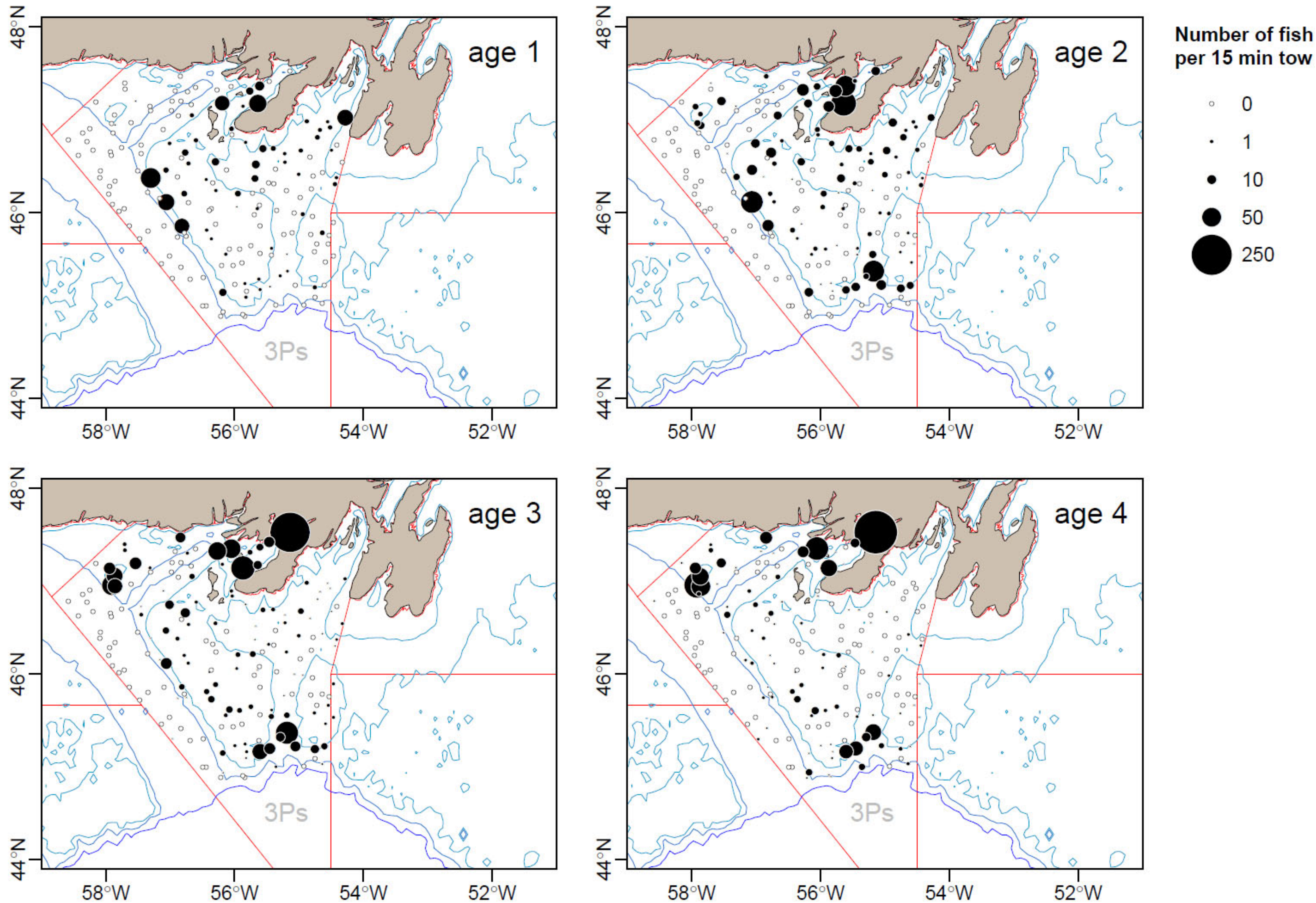




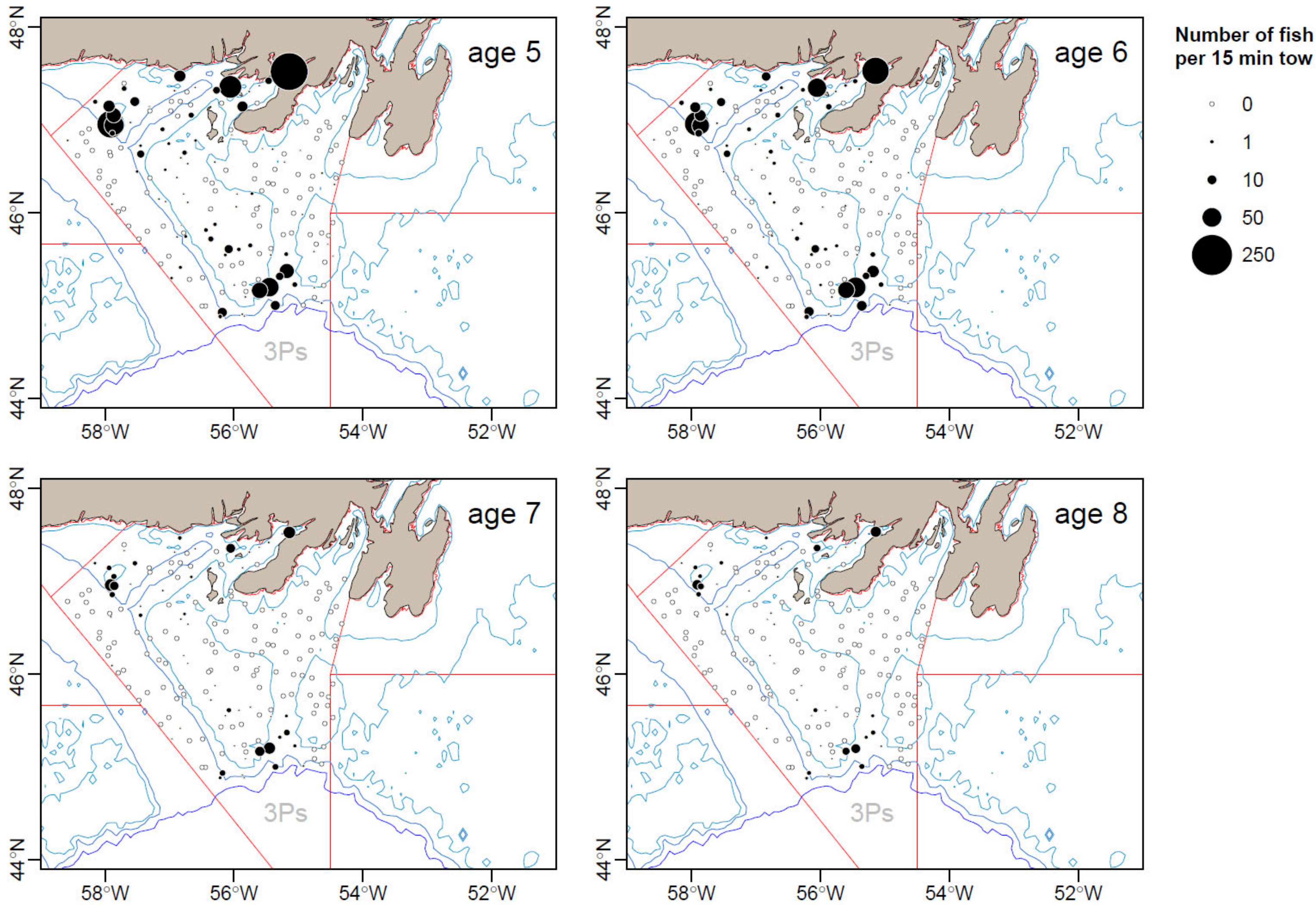




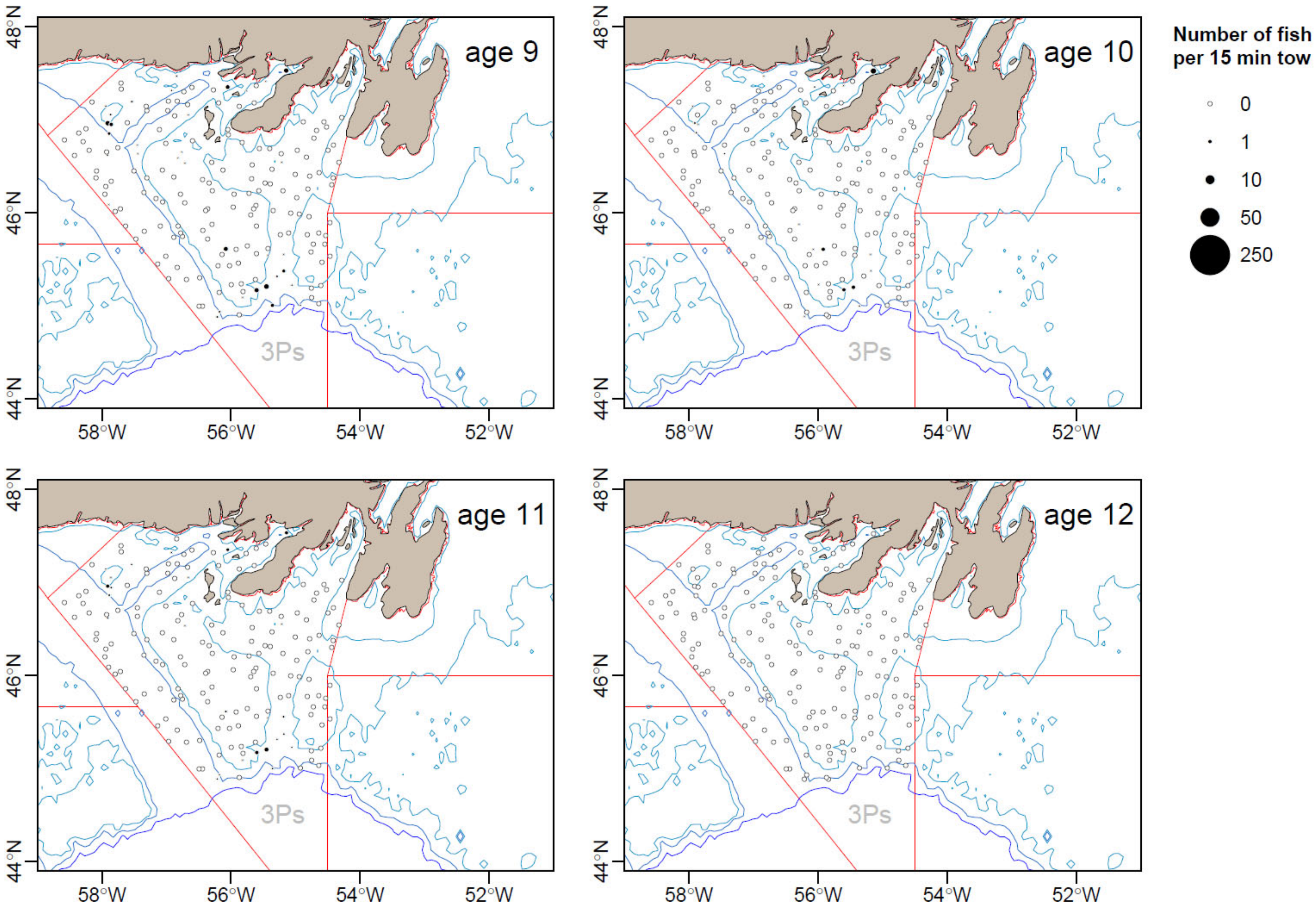
RV Survey – Distribution
Ages 1-4 number/tow 2017



RV Survey – Distribution
Ages 5-8 number/tow 2017



RV Survey – Distribution
Ages 9-12 number/tow 2017

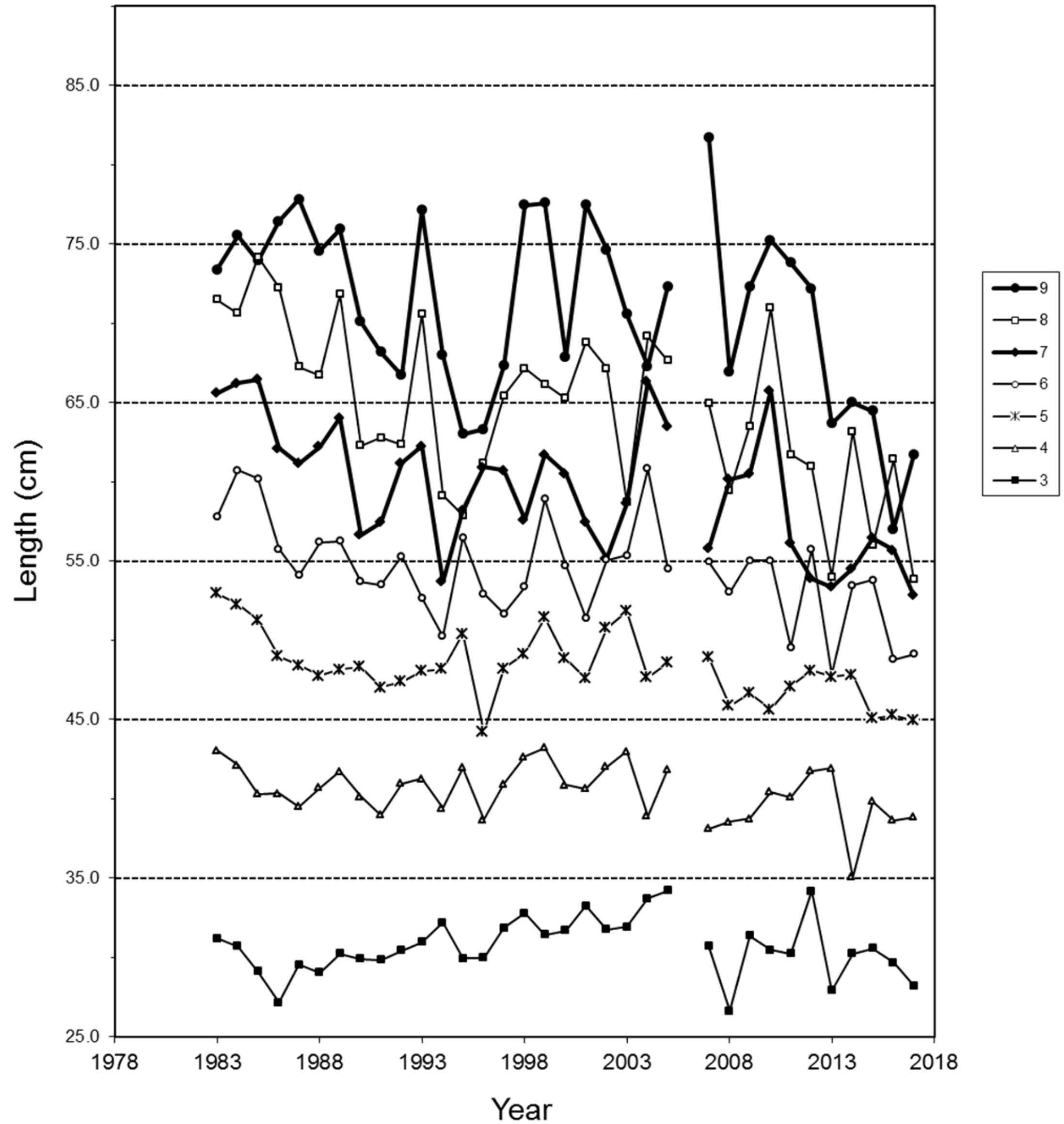


Biological Data

- Biological sampling during research vessel surveys provides data on cod:
  - Ages
  - Growth (with respect to both Length & Weight)
  - Condition
  - Maturity

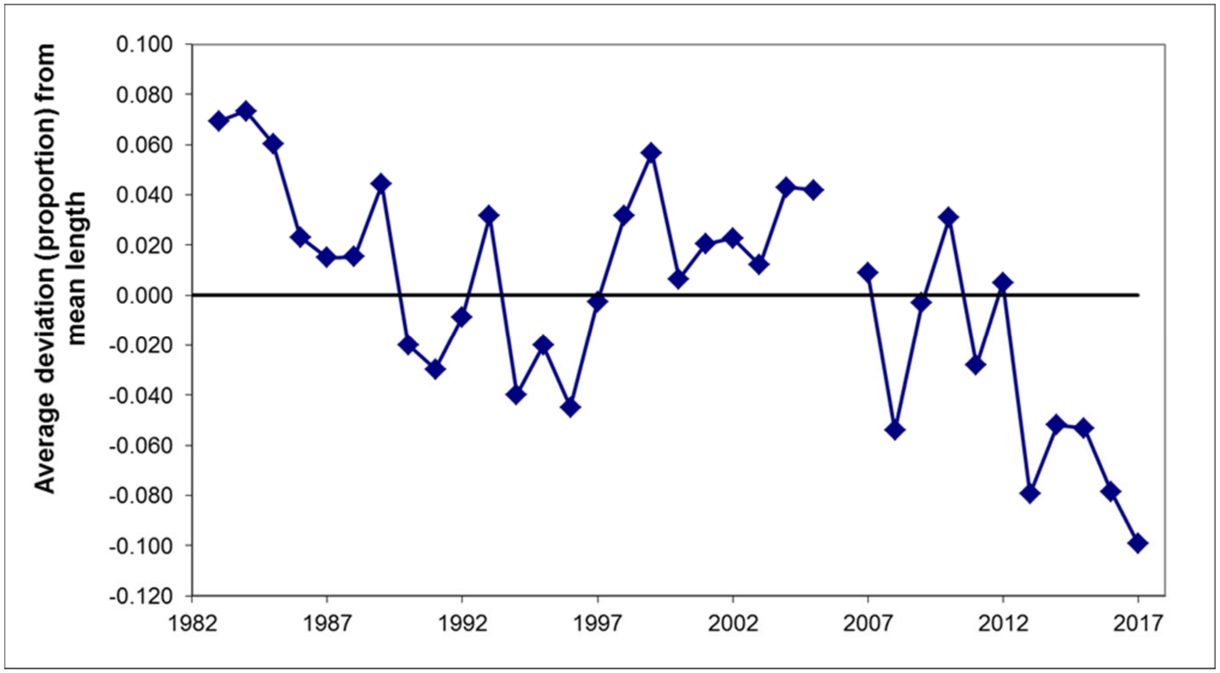
Biological Data

Mean Length at Age



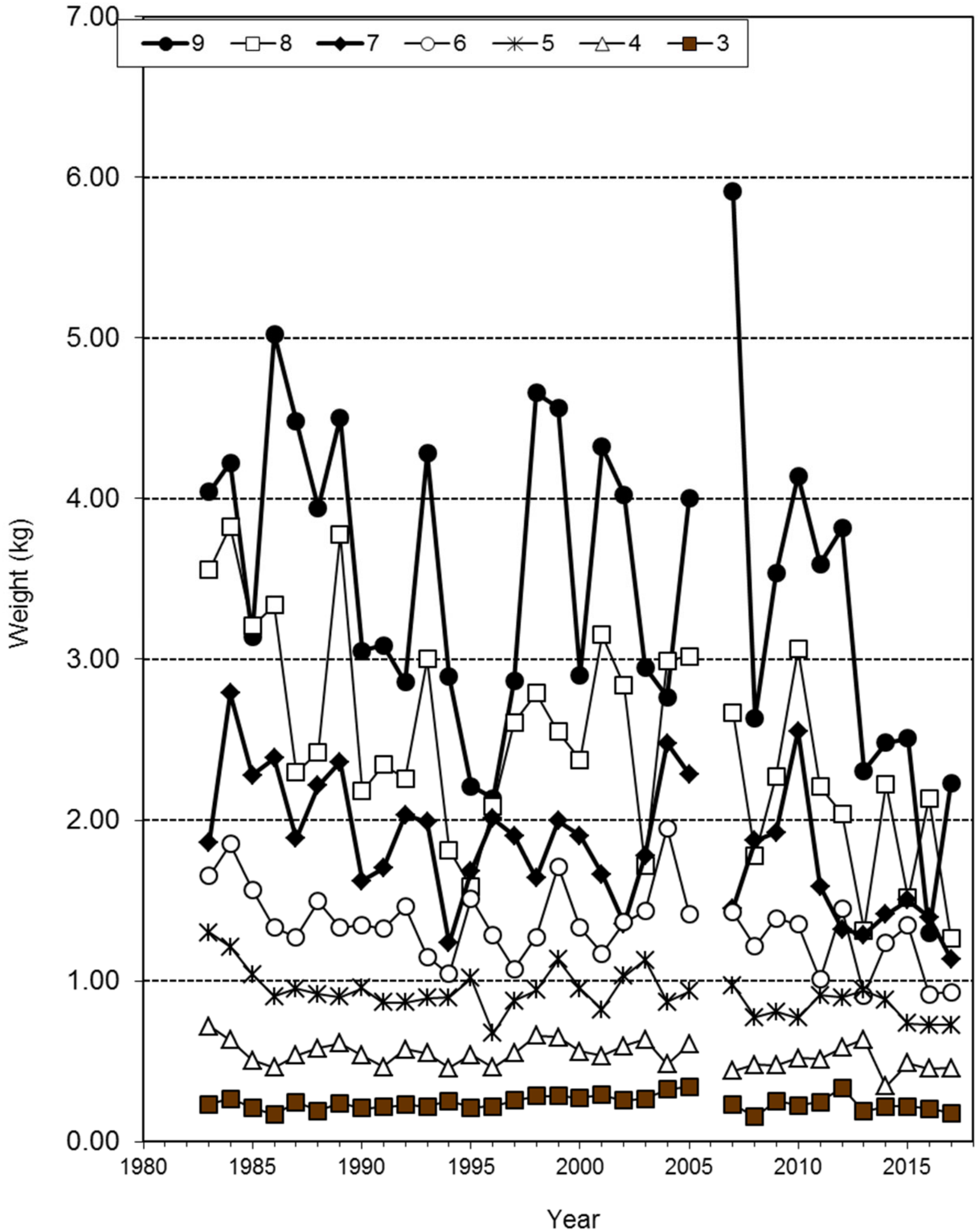
Biological Data

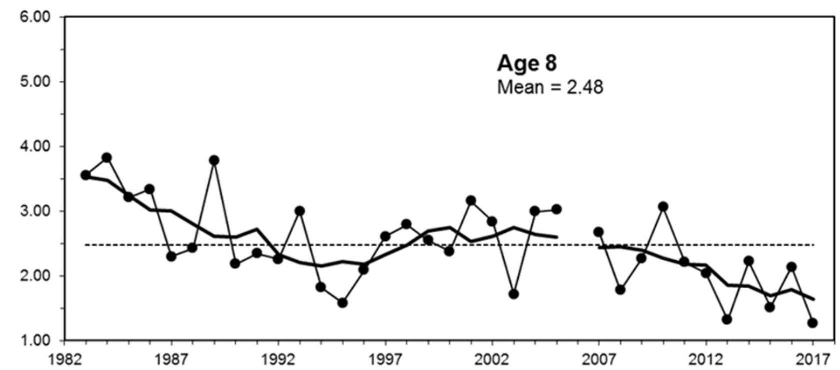
Mean Length at Age



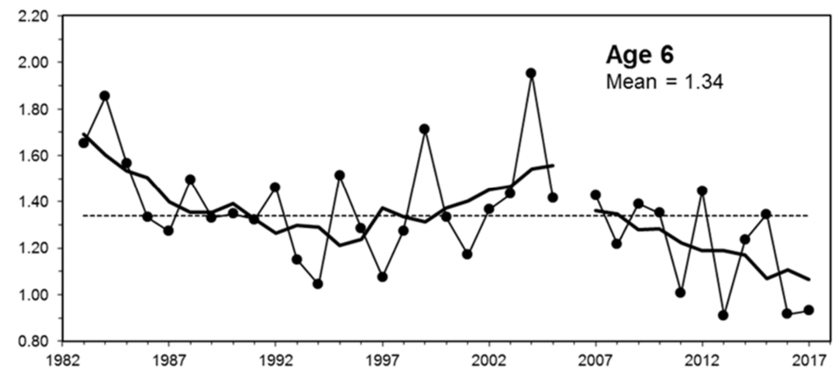
Biological Data

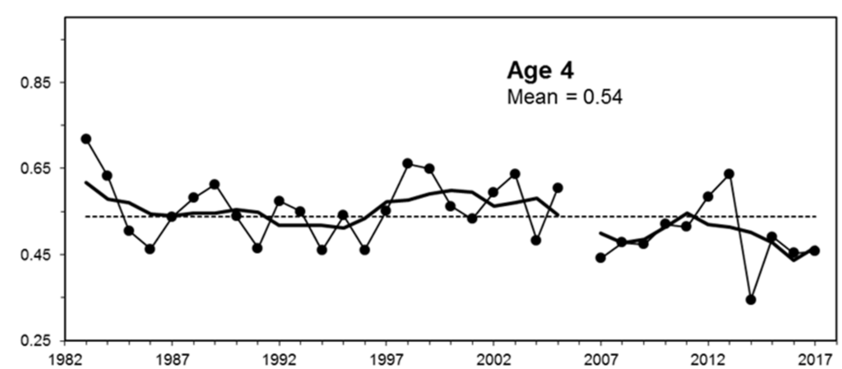
Mean Weight at Age





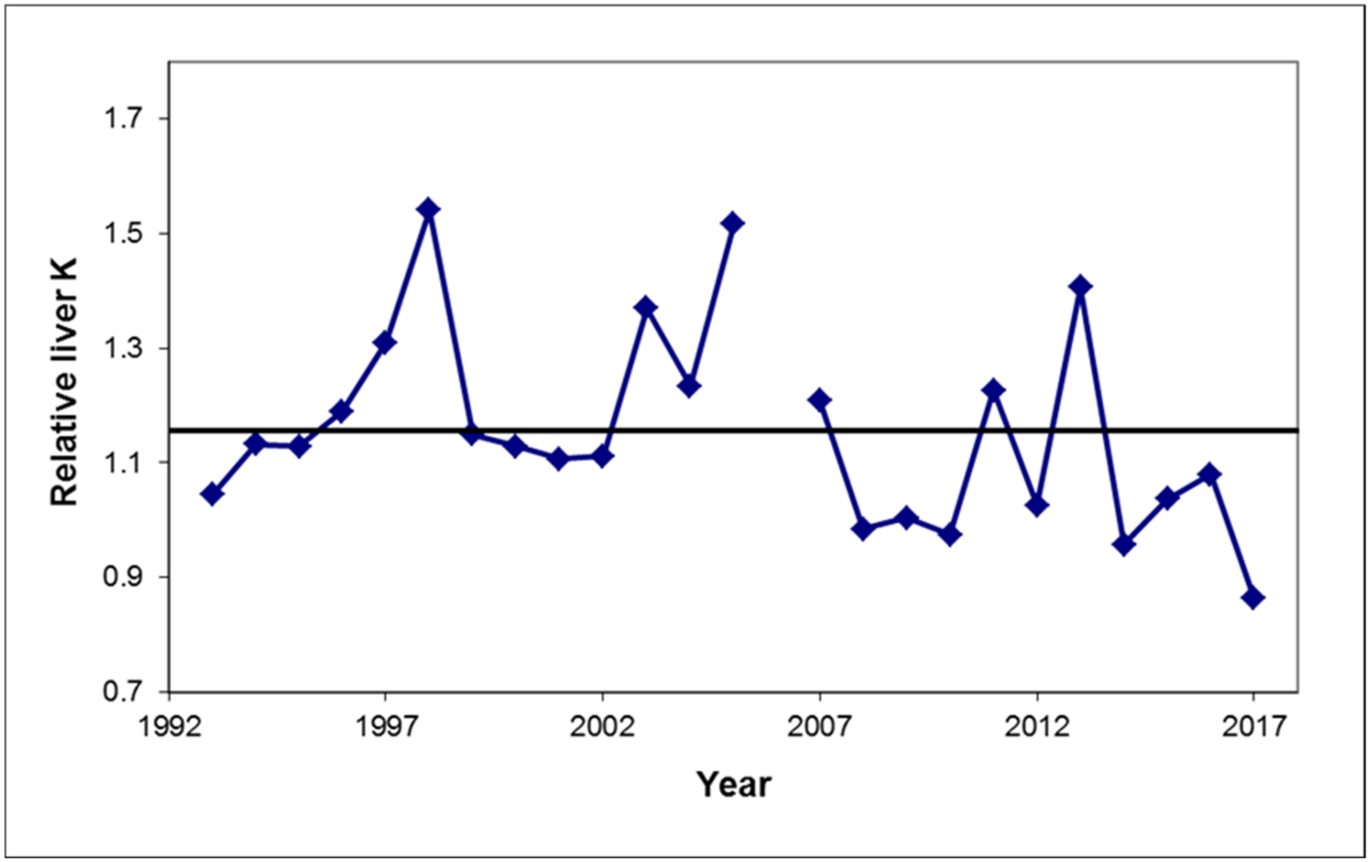
downward trend in weight-at-age since about 2007 for ages 6 and 8, with the trend starting a few years later for age 4.





RV Survey – Biological Data

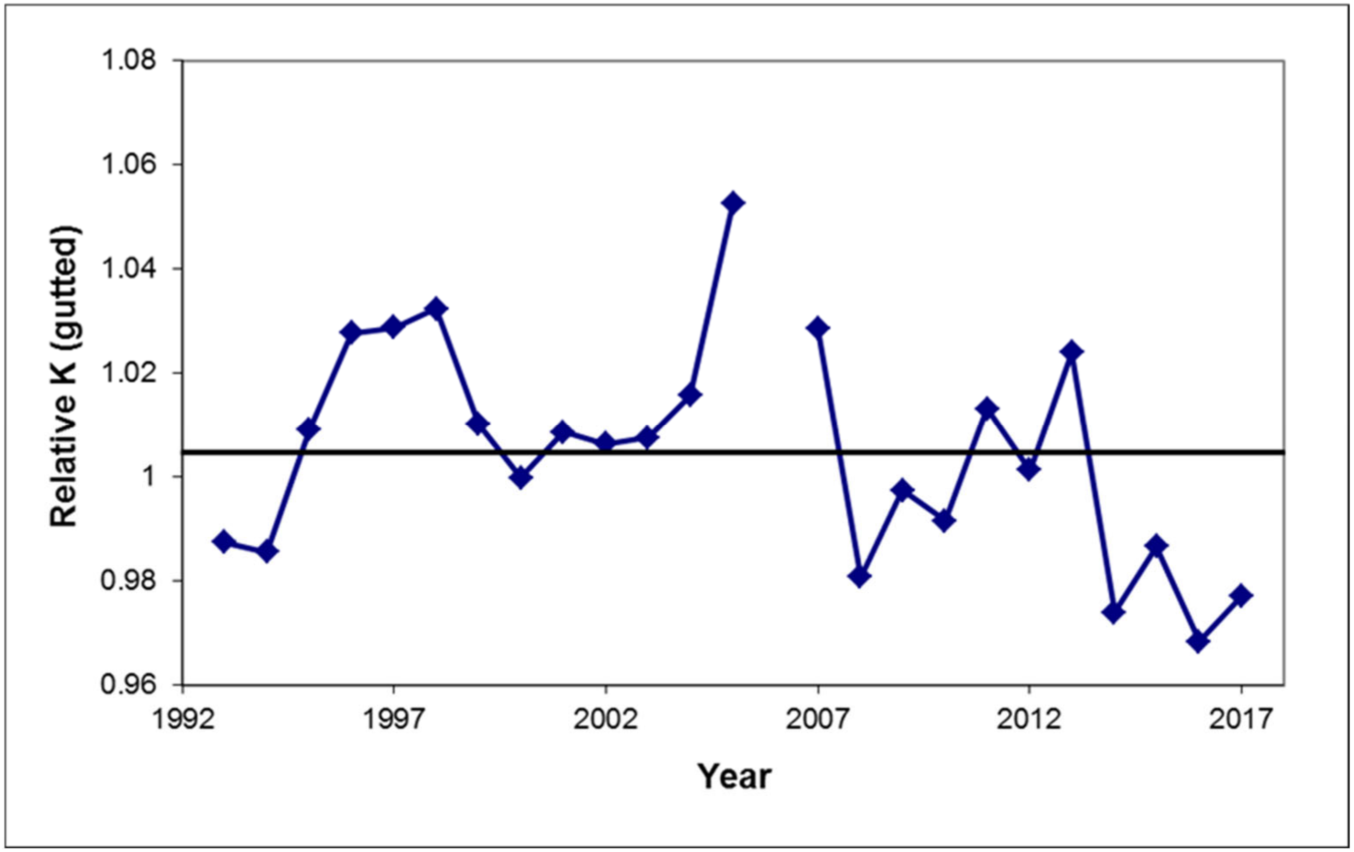
Cod Condition (Liver Index)



ries.

RV Survey – Biological Data

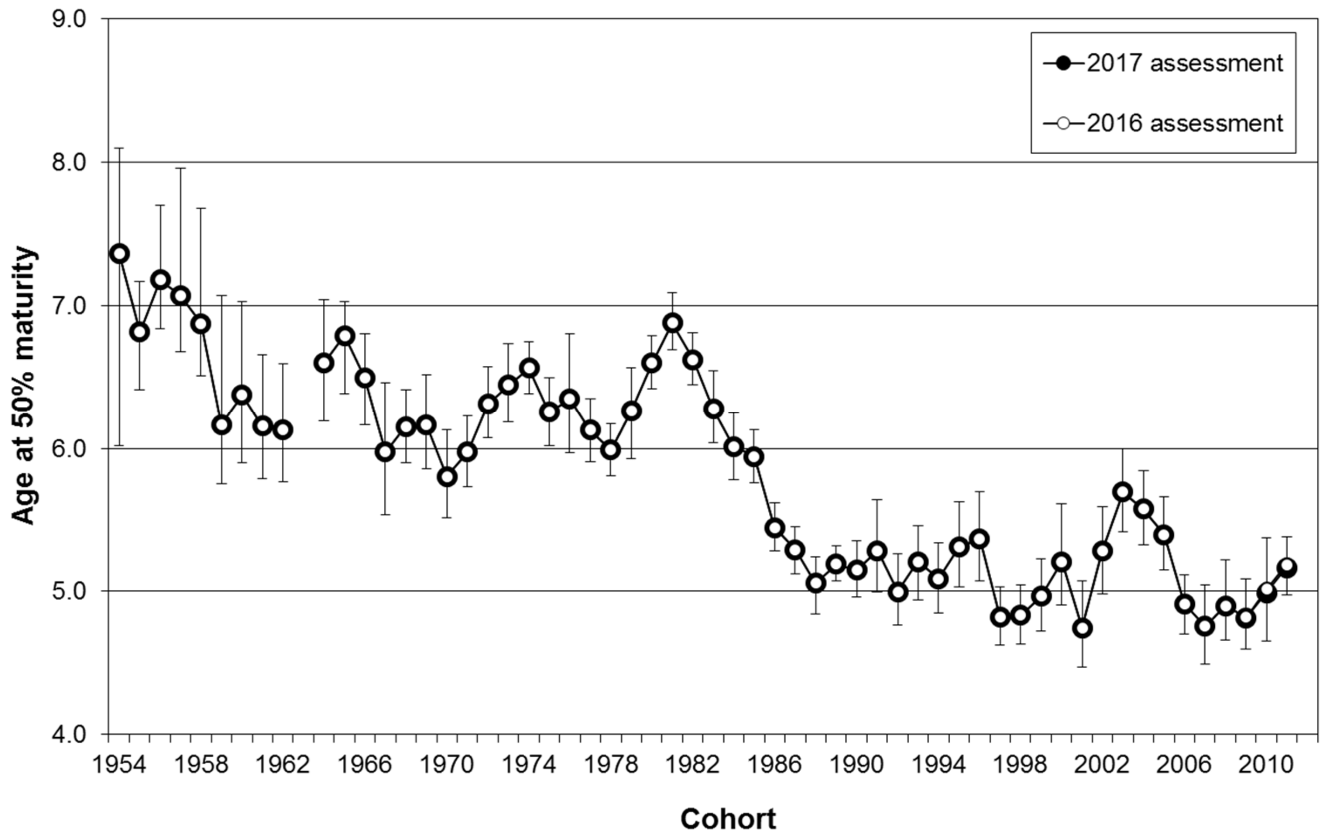
Cod Condition (Gutted)



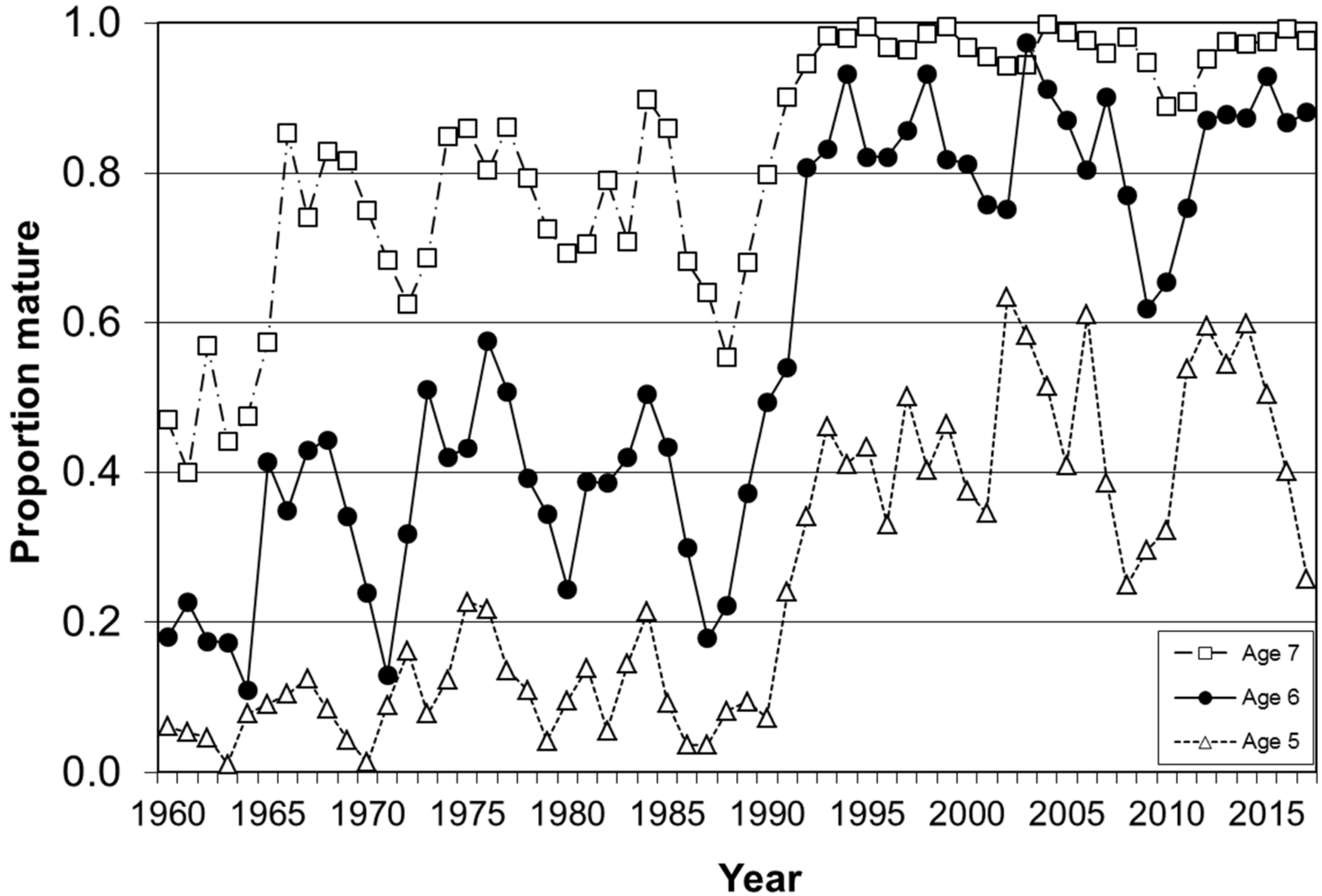
ries.

RV Survey – Biological Data

Maturity



RV Survey – Biological Data
Maturity



3Ps Survey Problems

– Coverage

– Movement

– Measurement error

– Large catches

– Change in survey protocols (e.g. vessels)

Survey problems and model-based solutions for indices of cod stock size off the south coast of Newfoundland

Nan Zheng and Noel Cadigan

Draft Paper, but more to be done yet!!

Abstract: Survey indices for cod in NAFO Subdivision 3Ps cod have high between-year variability caused by a number of factors, including occasional very large catches or annual variations in the timing of migrations that could affect the amount of fish in the survey area. Other survey problems are due to changes in the area surveyed, which was expanded in 1994 and 1997 with the addition of strata further inshore, and changes in the timing of the survey in 1993 from winter to spring. There is also incomplete coverage in some years, particularly in 2006 where less than half of the survey was completed. These changes add uncertainty and possible bias to survey indices of cod stock size. We develop a spatiotemporal survey model that addresses these issues. The model consists of four components, for 1) number caught per tow, 2) length composition of a catch, 3) maturity-at-length, and 4) weight-at-length. We combine model results to estimate catch-per-tow survey averages of: 1) number-at-length, 2) total number of mature fish, 3) total biomass, and 4) total mature biomass. Our results indicate that in 2018 both survey number per tow and mature number per tow were above the average; however, biomass and SSB per tow were below average. The difference in trends between abundance and biomass is mostly related to the decline in survey catch rates for larger sized cod.

Change in survey protocols

Executive Summary

In preparation for the pending transition from existing research platforms to new offshore fisheries research vessels (OFSVs), regional Fisheries & Oceans Canada staff, Coast Guard crew and invited subject matter experts participated in a workshop to discuss the study design, resource requirements and operational constraints of comparative trawl studies on both Atlantic and Pacific coasts.

….

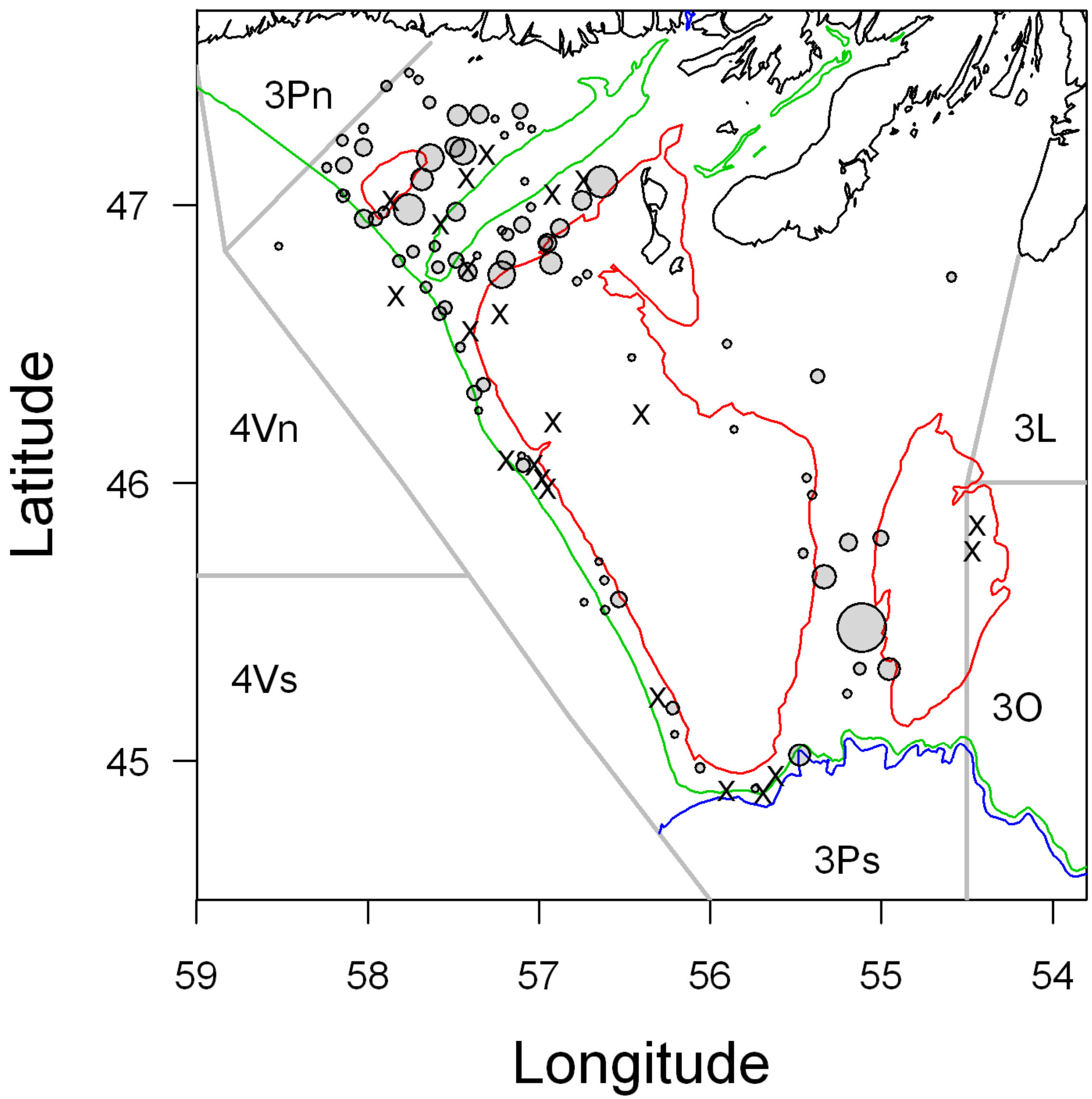
By the end of the workshop, a set of 15 principles had been drafted to guide comparative fishing studies on both coasts. It was also recognized that further work would be required by both DFO Science and Coast Guard and that a working group should be formed to oversee the coordination of these efforts.

Proceedings of the National Comparative Trawl Workshop, November 28-30, 2017, Nanaimo, BC

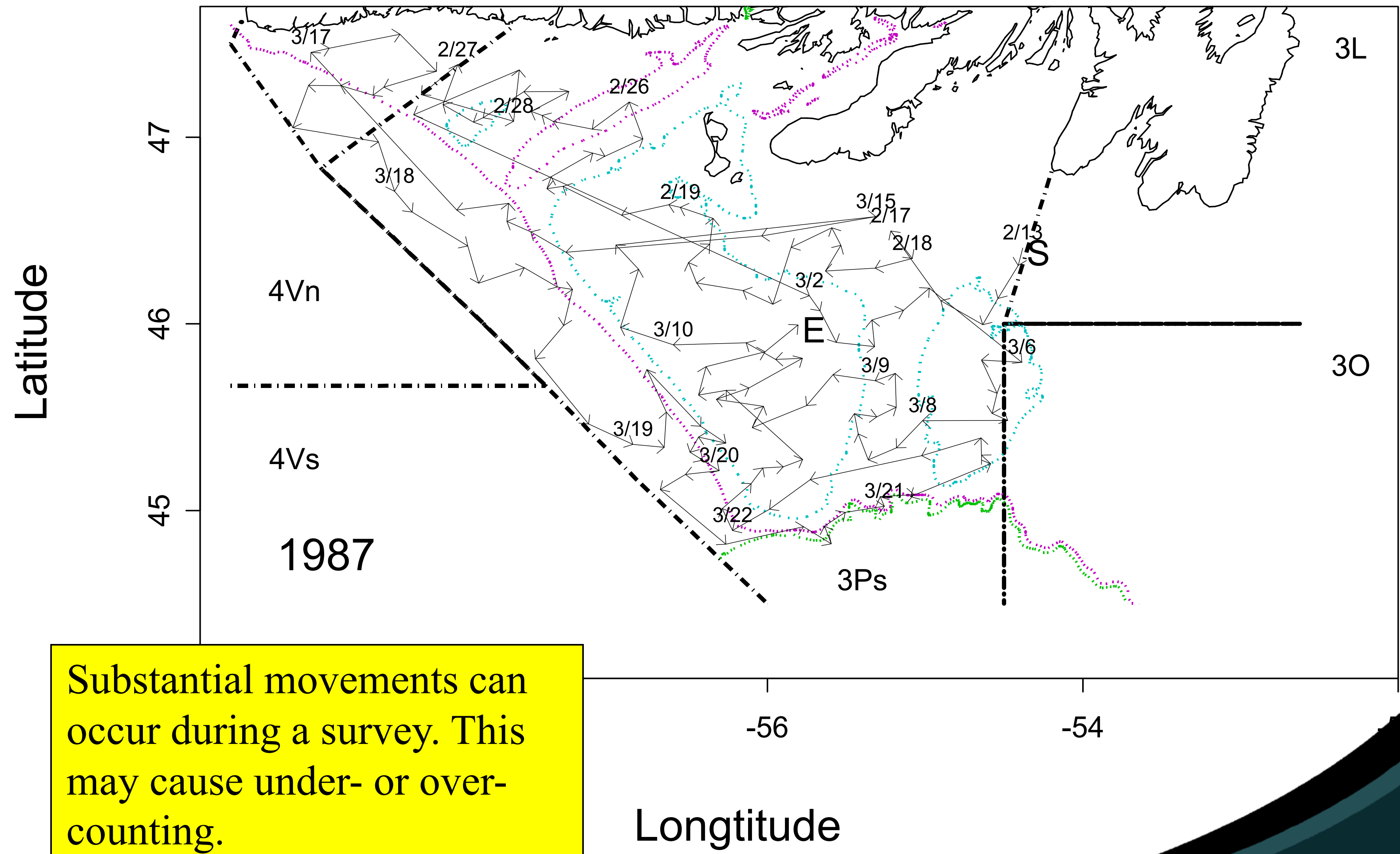
M.E. Thiess, H. Benoit, D.S. Clark, K. Fong, L.G.S. Mello, F. Mowbray, P. Pepin, N. Cadigan, T. Miller, D. Thirkell, and L. Wheeland

Fisheries and Oceans Canada
Science Branch, Pacific Region
Pacific Biological Station
Nanaimo, BC
V9T 6N7

3Ps survey catches of cod, 1978

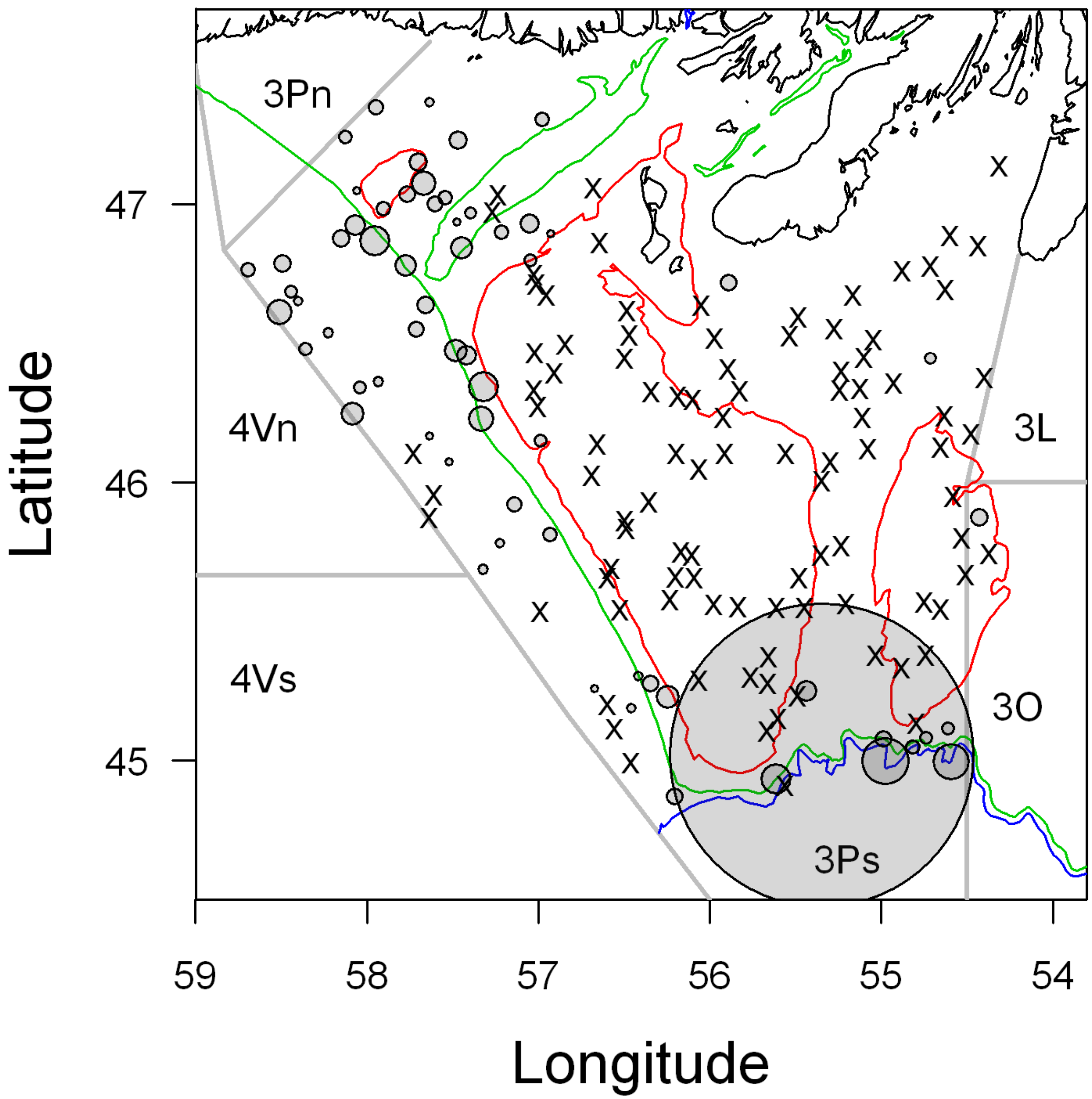


Vessel Trajection

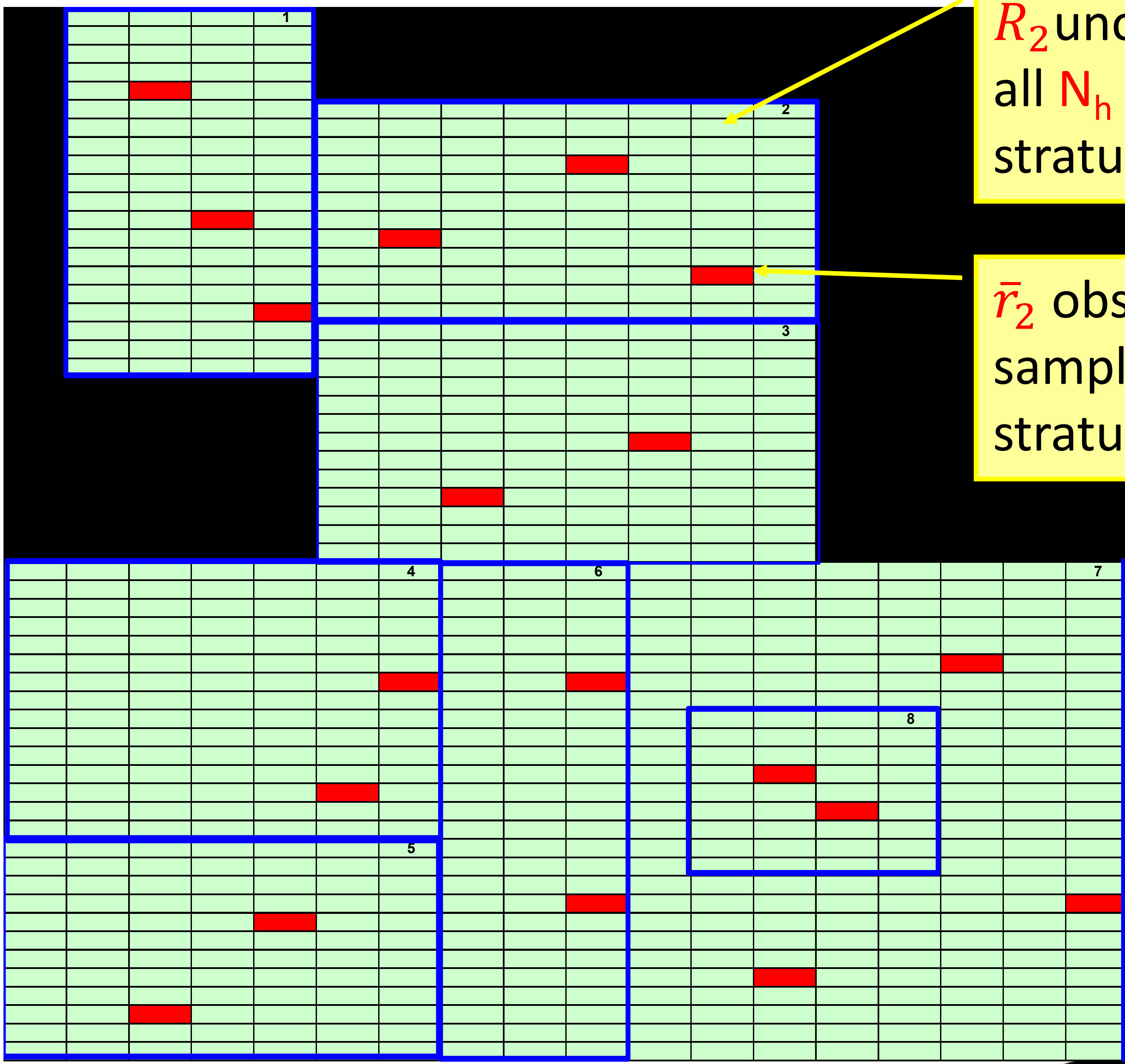


Substantial movements can occur during a survey. This may cause under- or over-counting.

3Ps survey catches of cod, 1995



Stratified-Random Sample (SRS) Index



$\bar{R}_2$ unobserved mean at all $N_h = 96$ green sites in stratum $h = 2$

$\bar{r}_2$ observed mean at 3 sampled (red) sites in stratum 2 ($n_2 = 3$)

Survey indices of abundance

– Survey index is an estimate of the average catch (R_+/N) at all (i.e. $N = \sum_{h=1}^{H} N_h$) tow stations in the survey region

– $$R_+ = \sum_{i=1}^{N} R_i = \sum_{h=1}^{H} \sum_{i=1}^{N_h} R_{hi}$$

NOTE:
$$\sum_{i=1}^{N_h} R_{hi} = N_h \bar{R}_h$$

– $$\bar{R} = \frac{R_+}{N} = \frac{\sum_{h=1}^{H} \sum_{i=1}^{N_h} R_{hi}}{N} = \frac{\sum_{h=1}^{H} N_h \bar{R}_h}{\sum_{h=1}^{H} N_h}$$

– A common (and unbiased) estimator of $\bar{R}$ is

– $$\hat{\bar{R}} = \frac{\sum_{h=1}^{H} N_h \bar{r}_h}{N} = \sum_{h=1}^{H} \frac{N_h}{N} \bar{r}_h$$

– Where $\bar{r}_h$ is the sample mean in stratum h.

Survey indices of abundance

- Survey index is an estimate of the average catch (R_+/N) at <u>all</u> (i.e. $N = \sum_{h=1}^{H} N_h$) tow stations in the survey region

- $$R_+ = \sum_{i=1}^{N} R_i = \sum_{h=1}^{H} \sum_{i=1}^{N_h} R_{hi}$$

NOTE:
$$\sum_{i=1}^{N_h} R_{hi} = N_h \bar{R}_h$$

- $$\bar{R} = \frac{R_+}{N} = \frac{\sum_{h=1}^{H} \sum_{i=1}^{N_h} R_{hi}}{N} = \frac{\sum_{h=1}^{H} N_h \bar{R}_h}{\sum_{h=1}^{H} N_h}$$

- A common (and unbiased) estimator of $\bar{R}$ is

- $$\hat{\bar{R}} = \frac{\sum_{h=1}^{H} N_h \bar{r}_h}{\sum_{h=1}^{H} N_h} = \sum_{h=1}^{H} \frac{N_h}{N} \bar{r}_h$$

an estimate of

- Where $\bar{r}_h$ is the <u>sample</u> mean in stratum h.

Survey indices of abundance

– Note that $\bar{r}_h = n_h^{-1} \sum_{i=1}^{n_h} R_{hi}$ is an unbiased estimator of $\bar{R}_h = N_h^{-1} \sum_{i=1}^{N_h} R_{hi}$

– and $N_h \bar{r}_h$ is an unbiased estimator for the total of stratum h, $R_{h+} = \sum_{i=1}^{N_h} R_{hi}$.

– Hence $\hat{\bar{R}}$ is an unbiased estimator for $\bar{R}$

Survey indices of abundance

– If the number of sample units in a stratum is proportional to the stratum area (i.e. A_h), and

– The total area of all strata is $\mathrm{A} = \sum_{h=1}^{H} A_h$, then

– $N_h = cA_h$, $N = cA$, and $\frac{N_h}{N} = \frac{A_h}{A}$

– Let the fraction of total area in stratum h be $W_h = \frac{A_h}{A}$

– The stratified estimator of $\bar{R}$ can also br written as

– $$\hat{\bar{R}} = \sum_{h=1}^{H} W_h \bar{r}_h \text{ where } \sum_{h=1}^{H} W_h = 1$$

– Sometimes we say $\hat{\bar{R}}$ is an area-weighted average, or an area-weighted average of stratum averages

Variance

– The estimate depends on what sites were randomly chosen.

– If we chose a different set of sites then we would get a different estimate

– Statisticians have figured out a way to measure how different

Variance

– Imagine repeating the survey thousands of times

– And computing thousands of $\hat{\hat{R}}$'s

– Could then plot a histogram of these estimates

– And we could compute their variance

– Which is the mean squared difference from the average of the thousands of estimates

$$B^{-1}\sum_{b=1}^{B=xxxx}\left\{\hat{\hat{R}}_b - ave(\hat{\hat{R}}_1, \dots, \hat{\hat{R}}_B)\right\}^2$$

Variance of the stratified mean

– Statisticians have figured out that this variance will be

$$Var\left(\hat{\bar{R}}\right) = \sum_{h=1}^{H} \left(\frac{N_h}{N} \right)^2 \frac{1 - f_h}{n_h} \frac{\sum_{i=1}^{N_h} \left(R_{hi} - \bar{R}_h \right)^2}{N_h - 1}$$

where $f_h = n_h / N_h$ is the sampling fraction

– You don't have to repeat the survey to know what the variance will be if you could repeat the survey thousands of times.

This is amazing!

Variance of the stratified mean

– unbiased estimator for $\mathrm{Var}(\hat{\bar{R}})$ is

– $\widehat{\mathrm{Var}}\left(\hat{\bar{R}}\right) = \mathrm{var}\left(\hat{\bar{R}}\right) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \frac{1-f_h}{n_h} \frac{\sum_{i=1}^{n_h}(R_{hi}-\bar{r}_h)^2}{n_h-1}$

Variance of the stratified mean

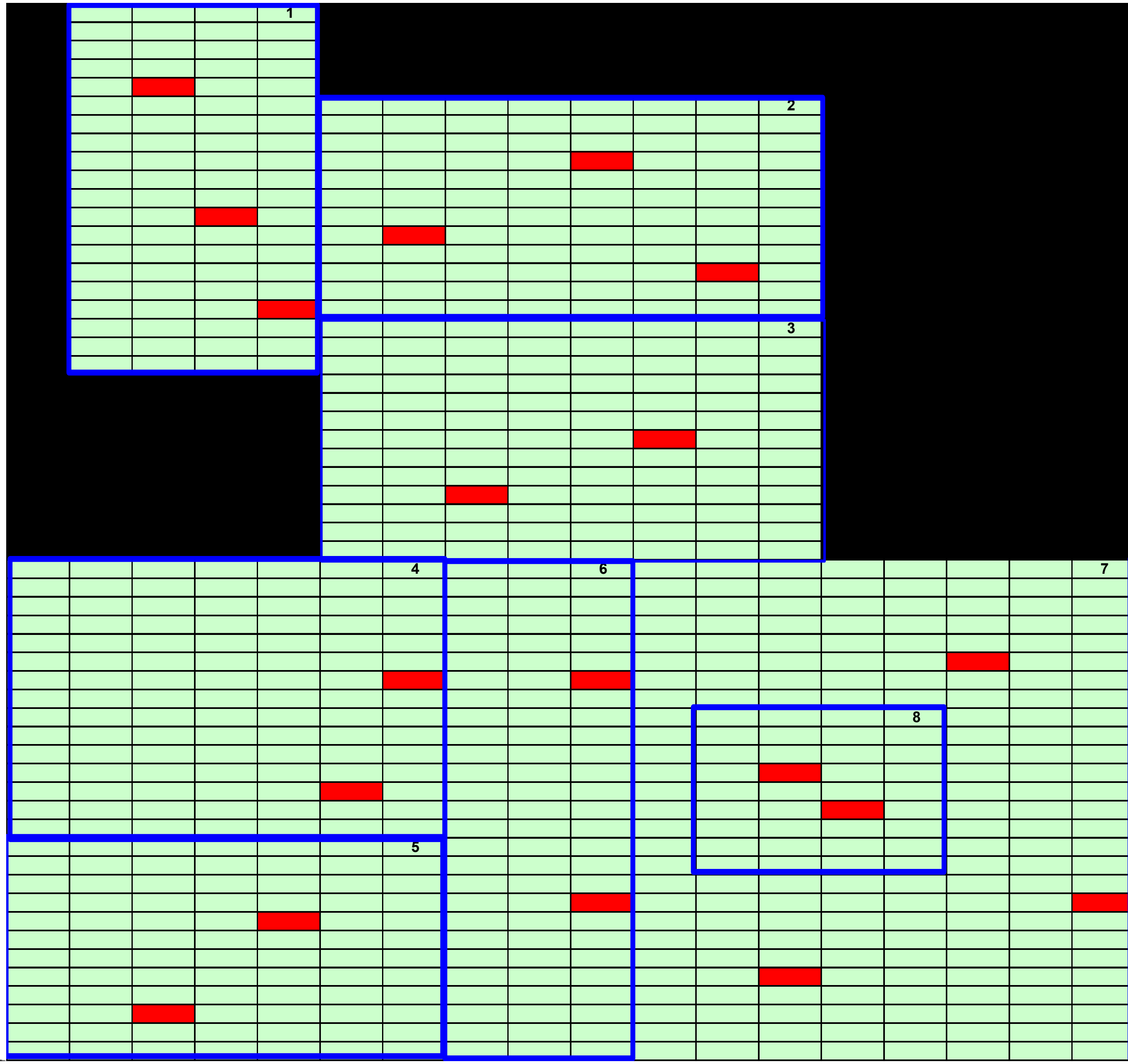
– unbiased estimator for $\mathrm{Var}(\hat{\bar{R}})$ is

– $\widehat{\mathrm{Var}}\left(\hat{\bar{R}}\right) = \mathrm{var}\left(\hat{\bar{R}}\right) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \frac{1-f_h}{n_h} \frac{\sum_{i=1}^{n_h}(R_{hi}-\bar{r}_h)^2}{n_h-1}$

An estimate of $\mathrm{Var}\left(\hat{\bar{R}}\right)$

$$Var\left(\hat{\bar{R}}\right) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \frac{1-f_h}{n_h} \frac{\sum_{i=1}^{N_h}\left(R_{hi}-\bar{R}_h\right)^2}{N_h-1}$$

example



example

stratum	area	catch
1	10	0
1	10	1
1	10	2
2	13	2
2	13	0
2	13	0
3	14	5
3	14	7
4	13	10
4	13	22
5	12	35
5	12	5
6	11	50
6	11	10
7	17	0
7	17	18
7	17	47
8	5	100
8	5	34

Area is proportional to N_h

example

```
> ## compute the mean for each stratum
> stratum.means <- tapply(data$catch,data$stratum,mean)
>
> ## display the results
> stratum.means
        1          2          3          4          5          6          7          8
 1.0000000  0.6666667  6.0000000 16.0000000 20.0000000 30.0000000 21.6666667
67.0000000
>
> ## get the areas fr each stratum
> stratum.areas <- tapply(data$area,data$stratum,unique)
>
> ## display the results
> stratum.areas
 1  2  3  4  5  6  7  8
10 13 14 13 12 11 17  5
>
```

example

```
> ## compute the stratum area-weighted average
> index.mean = sum(stratum.areas*stratum.means)/sum(stratum.areas)
>
> ## display the result
> index.mean
[1] 16.67368
>
> ## compare with the unweighted average
> mean(data$catch)
[1] 18.31579
>
```

In General ….

- Consider a *probability-based survey sampling design* in which each member (i) of the population has a <u>pre-determined</u> and nonzero probability $\pi_i > 0$ of being selected for sampling.
- Let I_i be an indicator variable that is one if member i is selected for sampling and is zero otherwise
- $\pi_i = E(I_i)$
- In a fixed-size sampling design in which the sample size is n, then $\sum_{i=1}^{N} I_i = n$
- Therefore $\sum_{i=1}^{N} \pi_i = n$

Horvitz-Thompson(HT) estimator

- An unbiased HT estimator of the population total $R = \sum_{i=1}^{N} R_i$ is

$$\hat{R}_{HT} = \sum_{i=1}^{n} \frac{R_i}{\pi_i}$$

- Proof: We can write

$$\hat{R}_{HT} = \sum_{i=1}^{n} \frac{R_i}{\pi_i} = \sum_{i=1}^{N} I_i \frac{R_i}{\pi_i}$$

- and $\mathrm{E}\left(\hat{R}_{HT}\right) = \sum_{i=1}^{N} E(I_i) \frac{R_i}{\pi_i} = \sum_{i=1}^{N} R_i$

HT Variance – End Lecture 2

- Recall $Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \, Cov(X_i, X_j)$
- Hence

$$Var(\hat{R}_{HT}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{R_i}{\pi_i} \frac{R_j}{\pi_j} Cov(I_i, I_j)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{R_i}{\pi_i} \frac{R_j}{\pi_j} (\pi_{ij} - \pi_i \pi_j)$$

- An unbiased estimator of $Var(\hat{R}_{HT})$ is
- $\widehat{Var}\,(\hat{R}_{HT}) = var(\hat{R}_{HT}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{R_i}{\pi_i} \frac{R_j}{\pi_j} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_{ij}}$