

한반도 연근해 저층 트롤 조사 자료에 표본론을 적용한 개체군의 상대적 크기 추정

이효태 · 현상윤*

부경대학교 해양생물학과

Application of Sampling Theories to Data from Bottom Trawl Surveys Along the Korean Coastal Areas for Inferring the Relative Size of a Fish Population

Hyotae Lee and Saang-Yoon Hyun*

Department of Marine Biology, Pukyong National University, Busan 48513, Korea

The Korean National Institute of Fisheries Science (NIFS) has biannually (spring and fall, respectively) deployed a bottom trawl survey along the coastal areas for last decade, taking samples on a regular basis (i.e., a systematic sampling). Despite the availability of the survey data, NIFS has not yet officially reported the estimates of the groundfish population sizes as well as has not evaluated uncertainty of the estimates. The objectives of our study were to infer the relative size of a fish population, applying two different sampling techniques (namely simple and stratified sampling) with different observation units to the NIFS survey data, and to compare those two techniques in bias and precision. For demonstration purposes, we used data on Pacific cod (*Gadus macrocephalus*) collected by the 2011-2015 surveys, and the results of simple and stratified sampling showed that the point estimates and precision varied by observation unit as well as the sampling technique.

Key words: Simple sampling, Stratified sampling, Bottom trawl survey, Survey index, Uncertainty of the estimates

서론

과학조사(survey)는 체계적인 수산자원평가를 위한 필수적인 요소로, 상업적 어업과는 다르게 장기간 동안 규격화된 어구와 어선을 사용하여 환경 및 생물에 관한 정보를 수집한다. 따라서 어업에 의한 외부적 영향(시장의 변화, 어업 규제, 어업기술의 발전 등)과 무관한 자원평가가 가능하기 때문에, 세계 각국에서는 표본추출법으로 설계된 과학조사를 꾸준히 시행하고 있다. 미국해양대기국(National Oceanic and Atmospheric Administration; NOAA)의 북동부와 북서부 수산과학청은 각각 1963년과 1977년을 기점으로 매년 층화 임의 표본추출법(stratified random sampling)으로 설계한 저층트롤 과학조사를 시행하였고(NOAA, 1988; NOAA, 2011), 태평양 헬리벳 수산 위원회(International Pacific Halibut Commission; IPHC)도 1963년을 시작으로 계통 표본추출법(systematic sampling)으로 설계한 고정 주낙 과학조사(setline survey)를 지속적으로 시행하

여 개체군의 상대적 크기(relative size)를 추정하였다(IPHC, 2012). 그리고 캐나다 동부의 베드포드 해양연구소(Bedford Institute of Oceanography; BIO)도 NOAA의 수산과학청과 같은 방법으로 1970년부터 저층트롤 과학조사를 시행하여 개체군의 상대적 크기를 추정하였고, 조사 정점의 환경자료를 이용하여 생물 종의 분포 및 서식지의 특성을 평가하는 연구를 추가적으로 시행하였다(Perry and Smith, 1994). 이처럼 과학조사 자료는 개체군의 상대적 크기 추정 뿐만 아니라 조사 정점의 어획자료와 환경자료를 이용한 연구, 미성숙개체들에 관한 자료 수집을 통해 어장의 가입강도 측정과 비상업종들의 연구 등 다양한 연구를 가능하게 한다.

지속적인 과학조사는 시간의 흐름에 따라 상업적 어업과는 독립적인 자원평가를 가능하게 하지만, 자료수집을 위해 많은 비용과 시간이 소모된다. 따라서 위에 언급한 각 기관들은 소모되는 비용을 줄이고 추정치의 신뢰성을 향상시키는 효율적 과학조사 방법을 지속적으로 연구하였다. NOAA의 북동부 수산과

<https://doi.org/10.5657/KFAS.2017.0594>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Korean J Fish Aquat Sci 50(5) 594-604, October 2017

Received 26 July 2017; Revised 11 August 2017; Accepted 4 September 2017

*Corresponding author: Tel: +82. 51. 629. 5929 Fax: +82. 51. 629. 5931

E-mail address: shyun@pknu.ac.kr

학청은 조사 대상해역의 면적과 조사 기간의 상관관계를 통해 과학조사에 소요되는 기간과 비용을 줄이는 방법을 모색하였고, IPHC와 BIO는 과학조사의 설계방법을 변형하여 과학조사의 신뢰성을 높이는 연구를 진행하였다. 또한, BIO는 설계방법에 변화를 준 요인들이 추정치의 신뢰성 향상에 미치는 영향의 정도를 평가하였고, 이를 바탕으로 새로운 설계방법을 제안하였다(Smith and Gavaris, 1993).

과학조사 자료를 이용한 연구가 많은 국가에서 꾸준히 진행되는 가운데, 우리나라의 국립수산과학원도 자원조사 전용선을 이용하여 2000년대 중반부터 매년 저층트롤 과학조사를 수행하고 있다(Fig. 1). 수집된 자료를 이용하여 예망면적과 단위면적 사이의 비례식을 이용하여 자원량을 추산하였다. 본 연구에서는 먼저 단순 추출법과 층화 추출법을 적용하여 한반도 연근해에 서식하는 개체군의 상대적 크기 추정 방법과 추정치를 제시하고, 둘째, 그 추정치에 대한 불확실성을 평가하며, 셋째, 2가지 표본 추출법의 결과들(개체군의 상대적 크기 추정치, 추정치의 불확실성)을 비교하는 것이 목적이다. 국립수산과학원은 해구마다 한번씩 규칙적으로 저층 트롤 조사를 했기에, 임의(random) 추출표현을 피한다.

본 연구는 과학조사에서 어획된 어획량을 이용하여 개체군의 상대적 크기를 추정하였다. 개체군의 절대적 크기를 추정하려면 저층트롤 어구의 어획률(catchability) 정보가 있어야 하는

데, 어획률은 어종마다 다르고 이를 알기 위해서는 추가적인 자료(예, 어업의 해당 어종의 어획량 등)를 필요로 한다. 저층트롤 과학조사시 매 계절마다 300-400 어종이 어획되는데, 이렇게 많은 어종에 따라 다른 값의 어획률을 적용하는 것이 어려우므로 통상 모르는 상수로 둔다. 따라서, 본 연구에서 추정된 개체군 크기도 절대적 추정치가 아닌 상대적(최소) 추정치임을 강조한다. 미국 NOAA에서는 이 개체군의 상대적 크기 추정치를 “조사지수(survey index)”라 명명했으며, 같은 이유로 본 연구에서도 조사지수란 용어를 사용하였다.

자료 및 방법

과학조사 자료

국립수산과학원 내의 4개 연구소(동해, 남해, 서해, 본원)는 한반도 주변 해역에 각각 담당 해역을 지정하여 과학조사를 시행하였고, 이 중 2011-2015년까지 자료를 이용하여 연구를 진행하였다. 과학조사를 통해 어획 정보(어종별 개체수, 조사일자, 예망시간, 정점, 수심)와 어획된 생물들의 생물학적 정보[체장(cm), 체중(g), 성별, 성숙도]가 수집되었지만, 연령자료는 기록되어 있지 않았다. 매년 봄철(3-6월)과 가을철(9-12월)로 나누어 시행되었으며, 규격화된 어구와 조사선을 이용했고, 예망시 일정한 선속(약 3.5 kn)을 유지하여 일관성 있게 자료를 수집하였다. 봄철과 가을철 조사는 각 64-69개의 정점에서 조사를 실시하였고, 정점은 해구와 소해구를 기준으로 지정하였다. ‘해구(Grid)’는 위도와 경도를 기준으로 각 30분씩 나는 직사각형의 해역이고, ‘소해구’는 한 해구를 아홉 개의 직사각형으로 나누는 해역을 의미한다(Fig. 2). 각 해구들의 면적은 Geoscience Australia의 웹사이트에서 제공하는 위도 및 경도별 좌표간 거리 계산식(단위:km²)을 이용하여 산출하였다(Geoscience Australia, 2010). 그리고 소해구들의 면적은 각 해구의 면적의 9 (한 해구에 속한 소해구의 개수)로 나누어 계산하였다. 대부분 한 해구당 하나의 소해구를 지정하여 조사를 실시하였으나, 예외적으로 한 해구당 두 개의 소해구에서 조사를 시행한 해구도 존재했

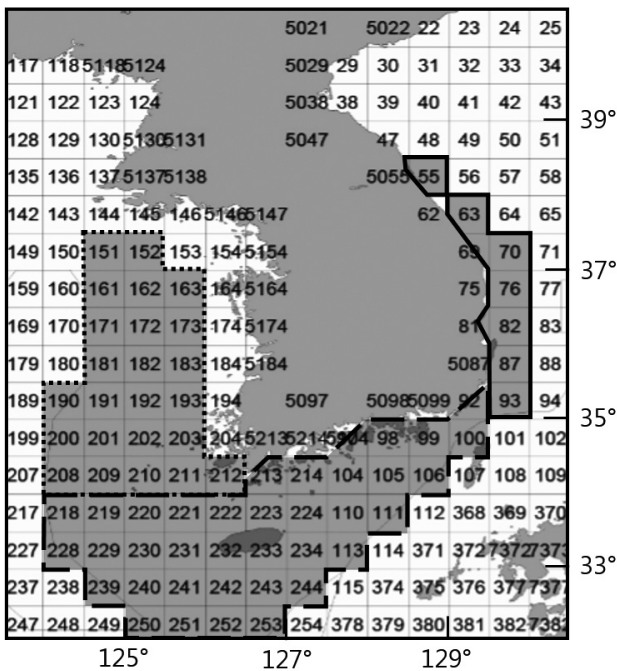


Fig. 1. Areas where the net was trawled. Survey areas were grouped into three strata. Grids encompassed by solid, dashed and dotted lines were defined as the eastern stratum ($h = 1$), the southern stratum ($h = 2$) and the western stratum ($h = 3$), respectively.

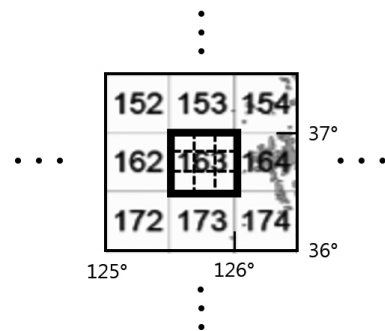


Fig. 2. A rectangle surrounded by thick line is defined as a Grid, which has nine small rectangles.

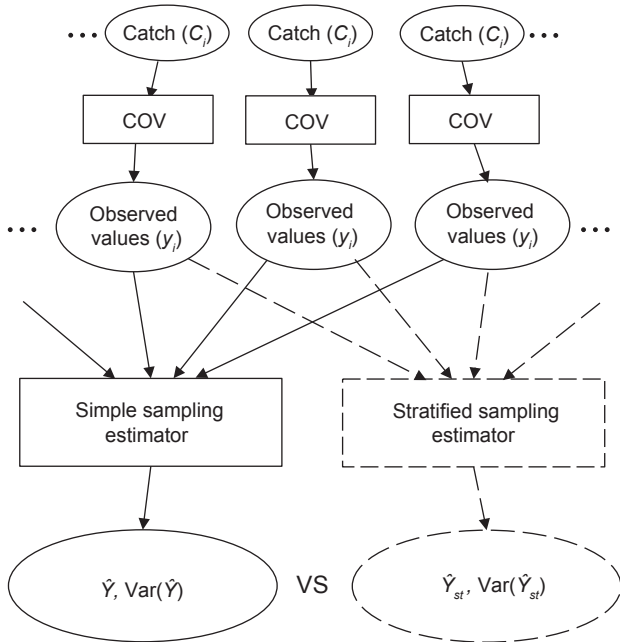


Fig. 3. Flow chart of the methodology of sampling theories. Observation unit is a small rectangle in a Grid. COV denotes the calculation of observed value with eq. 8.

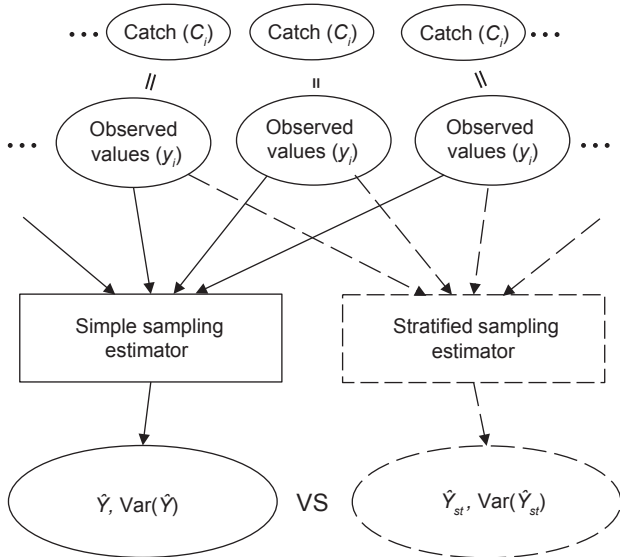


Fig. 4. Flow chart of the methodology of sampling theories. Observation unit is bottom area covered by trawl.

다. 이 경우, 두 개의 소해구에서 각각 실시한 예망시간(약 30 분)의 합이 한 소해구에서 실시한 예망시간(약 60분)과 비슷하였기 때문에, 두 개의 소해구에서 기록된 어획량과 예망면적을 합하여 한번의 조사로 간주하였다.

Table 1. Notations used in this paper

Notation	Description	
h	Stratum	East ($h = 1$) South ($h = 2$) West ($h = 3$)
N_h	Total number of possible observation units in stratum h	
n_h	Number of observation units where the net was trawled in stratum h	
N	Total number of possible observation units	$N = \sum_{h=1}^3 N_h$
n	Total number of observation units where the net was trawled	$n = \sum_{h=1}^3 n_h$
y_i	Observed value of a particular species in an observation unit i	
y_{hi}	Observed value of a particular species in an observation unit i and stratum h	
y	Sum of y_{hi} in the entire samples	$y = \sum_{h=1}^3 \sum_{i=1}^{n_h} y_{hi}$
\bar{y}_h	Mean of y_{hi} in stratum h	$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$
\bar{y}	Mean of y_{hi} in the entire samples	$\bar{y} = \frac{\sum_{h=1}^3 \sum_{i=1}^{n_h} y_{hi}}{n}$
\bar{y}_{st}	Stratified mean of y_{hi}	$\bar{y}_{st} = \frac{\sum_{h=1}^3 N_h \cdot \bar{y}_h}{N}$
Y	True population size in simple sampling	$\sum_{h=1}^3 \sum_{i=1}^{N_h} y_{hi}$
\hat{Y}	Survey index of Y	
Y_{st}	True population size in stratified sampling	
\hat{Y}_{st}	Survey index of Y_{st}	
S_h^2	The variance of observed value in stratum h	

통계적 분석

표본 추출 이론은 전체 모집단의 구성원을 일일이 조사할 수 없으므로, 표본을 수집하고 이를 이용하여 전체 모집단의 크기를 추정할 뿐만 아니라 추정치의 불확실성(e.g. 분산)을 계산하는 방법이다. 일반적인 표본추출법은 단순 ‘임의’ 추출법(simple random sampling)이나 층화 ‘임의’ 추출법이 이용되지만, 본 연구에 사용한 자료가 체계적인 방법(하나의 해구에서 한 번씩, 동일한 소해구)으로 표본 추출이 이루어졌으므로 임의적 조사가 아닌 규칙적 조사방법이다. 따라서 단순 ‘임의’ 추출법, 층화

‘임의’ 추출법이 아닌 단순 추출법, 층화 추출법임을 강조한다.

과학조사를 통해 수집된 동일한 자료에 두 가지 표본추출법(단순 추출법과 층화 추출법)을 적용하여, 개체수 및 증량에 관해 조사지수와 추정치의 분산(불확실성)을 계산 및 비교하였다. 또한, 다른 관측단위(observation unit)를 가정했을 때의 조사지수와 그 분산을 추가적으로 제시하였다(Fig. 3, 4). 위에 제시한 방법들을 통해 계산된 조사지수들 사이에는 서로 직접적인 비교가 가능하지만, 추정치의 분산은 조사지수 값이 커짐에 따라 같이 커지는 경향이 있다. 따라서 직접적인 비교가 타당하지 않아, 각 표본 추출법을 통해 계산된 표준오차에 조사지수를 나눔으로써 차이를 조정하는 변동계수(coefficient of variation)를 통해 불확실성을 비교하였다.

단순 추출법

본 연구에 사용된 표기기호는 Table 1로 정리하였다. 조사해역 전체를 하나의 묶음으로 표본을 추출하는 단순 추출법은 모집단에 대한 정보가 아주 적을 경우 유용하게 사용할 수 있는 방법으로, 모집단 조사지수(\hat{Y}) 추정 방법은 다음 과정을 통해 알 수 있다.

$$\hat{Y} = N \cdot \bar{y} \quad \dots \dots \dots (1)$$

\hat{Y} 은 식(1)로 나타낼 수 있으며,

$$\hat{Y}_{st} = N \cdot \bar{y}_{st} \quad \dots \dots \dots (4)$$

$$E(\hat{Y}) = E(N \cdot \bar{y}) = N \cdot E(\bar{y}) = N \cdot \frac{\sum_{i=1}^N Y_i}{N} = Y \quad \dots \dots \dots (2)$$

식(2)를 통해 Y 의 불편 추정치, 즉 $E(\hat{Y}) = Y$ 임을 알 수 있다 (Cochran, 1977). 추정치의 분산은 중심극한정리에 의해 다음 식을 통해 나타낼 수 있다.

$$\text{Var}(\hat{Y}) = N^2 \cdot \text{Var}(\bar{y}) = N^2 \cdot \frac{\text{Var}(y)}{n} \quad \dots \dots \dots (3)$$

참고로, 모든 추정치는 조사 대상 해역에 존재하는 특정한 하나의 종에 관한 모집단 크기를 나타내는 것으로, 모든 종의 모집단 크기의 합을 나타내는 것이 아니다.

층화 추출법

층화 추출법은 단순 추출법과 달리, 모집단에 대한 정보를 추가하여, 동질성을 가지는 해역끼리 하나의 층(stratum)으로 묶어 표본을 추출하는 방법이다. 본 연구에서는 한국해양자료센터(KODC, 2011)의 정선관측소 정점에 위도와 경도에 관한 정보를 참고하여 동해, 서해, 남해로 층을 구분하였다(Fig. 1). 층화 추출법에서 모집단의 조사지수에 관한 식은 다음과 같다.

Table 2. Illustration of application of the simple sampling technique for calculation of the survey index in number of Pacific cod (*Gadus macrocephalus*) population in the spring of 2013 under the assumption that the observation unit was a small rectangle in a Grid. The left column of the table is estimator of a metric, and the right column is its estimate

Estimator	Estimate
$y_i = \frac{A_i \times C_i}{a_i}$	$i = 208^{\text{th}}$ Grid, $y_{208} = \frac{283 \text{ km}^2 \times 3}{0.1519 \text{ km}^2} \approx 5,580 \text{ fish}$
Step1. Calculation of the observed value (y)	
$y = \sum_{i=1}^n y_i$	$y = y_{113} + y_{208} + \dots + y_{203} \approx 557,294 \text{ fish}$
Step2. Calculation of sum of the observed values (y) where the net was trawled	
$\bar{y} = \frac{y}{n}$	$\bar{y} = \frac{557,294}{68} \approx 8,196 \text{ fish}$
Step3. Calculation of mean of the observed values (\bar{y})	
$\hat{Y} = N \times \bar{y}$	$\hat{Y} \approx 612 \times 8,196 = 5,015,646 \text{ fish}$
Step4. Calculation of the survey index of population (\hat{Y})	
$\text{Var}(\hat{Y}) = N^2 \cdot \frac{\text{Var}(y)}{n}$	$\text{Var}(\hat{Y}) = 612^2 \cdot \frac{1,146,870,826}{68} \approx 6.317 \times 10^{12} \text{ fish}$
$S\hat{E}(\hat{Y}) = \sqrt{\text{Var}(\hat{Y})}$	$S\hat{E}(\hat{Y}) = \sqrt{6.317 \times 10^{12}} \approx 2,513,357 \text{ fish}$
Step5. Calculation of variance and standard error of population estimate	

만약, 각 층에서 동일한 양의 어획물이 어획되었다 하더라도 층별 관측단위의 수가 달라, 자원량이 동일하게 분포한다고 할 수 없다. 따라서 층화 추출법에서 모집단의 조사지수와 추정치의 분산 추정은 각 층별로 관측단위의 개수만큼 가중치를 부여하여, 산술평균(\bar{y}) 대신 가중평균(\bar{y}_{st})으로 계산한다. 추정치의 분산도 가중치를 부여하여 다음 식으로 나타낸다.

$$\text{Var}(\hat{Y}_{st}) = N^2 \cdot \text{Var}(\bar{y}_{st}) = \sum_{h=1}^3 [N_h(N_h - n_h) \frac{S_h^2}{n_h}] \dots (5)$$

관측단위: 소해구

단순 추출법과 층화 추출법의 관측단위를 소해구로 가정하면, 관측치(observed value)는 ‘한 소해구에 존재하는 특정 종

Table 3. Illustration of application of the stratified sampling technique for calculation of the survey index in number of Pacific cod (*Gadus macrocephalus*) population in the spring of 2013 under the assumption that the observation unit was a small rectangle in a Grid. The left column of the table is estimator of a metric, and the right column is its estimate

Estimator	Estimate
$N = \sum_{h=1}^3 N_h = N_1 + N_2 + N_3$	612 = 72 + 333 + 207
$n = \sum_{h=1}^3 n_h = n_1 + n_2 + n_3$	68 = 8 + 37 + 23

Step1. The stratification of Grid where net was trawled

$$y_{hi} = \frac{A_{hi} \times C_{hi}}{a_{hi}} \qquad y_{h=3, i=208} = \frac{283 \text{ km}^2 \times 3 \text{ fish}}{0.1519 \text{ km}^2} \approx 5,580 \text{ fish}$$

Step2. Calculation of observed value (y_{hi}) in each stratum

$$y_h = \sum_{i=1}^{n_h} y_{hi} \qquad \begin{aligned} y_{h=1} &= y_{i=55} + y_{i=63} + \dots + y_{i=93} \approx 343,193 \text{ fish} \\ y_{h=2} &= y_{i=97} + y_{i=98} + \dots + y_{i=253} \approx 108,188 \text{ fish} \\ y_{h=3} &= y_{i=151} + y_{i=152} + \dots + y_{i=212} \approx 105,913 \text{ fish} \end{aligned}$$

Step3. Calculation of sum of the observed values (y_h) in each stratum

$$\bar{y}_h = \frac{y_h}{n_h} \qquad \begin{aligned} \bar{y}_{h=1} &= \frac{343,193 \text{ fish}}{8} \approx 42,899 \text{ fish} \\ \bar{y}_{h=2} &= \frac{108,188 \text{ fish}}{37} \approx 2,924 \text{ fish} \\ \bar{y}_{h=3} &= \frac{105,913 \text{ fish}}{23} \approx 4,605 \text{ fish} \end{aligned}$$

Step4. Calculation of mean of the observed values (\bar{y}_h) in each stratum

$$\bar{y}_{st} = \frac{\sum_{h=1}^3 N_h \cdot \bar{y}_h}{N} \qquad \bar{y}_{st} = \frac{72}{612} \times 42,899 + \frac{333}{612} \times 2,924 + \frac{207}{612} \times 4,605 \approx 8,196 \text{ fish}$$

Step5. Calculation of weighted average of the observed values

$$\hat{Y}_{st} = N \cdot \bar{y}_{st} \qquad \hat{Y}_{st} \approx 612 \times 8,196 = 5,015,646 \text{ fish}$$

Step6. Calculation of the survey index (\hat{Y}_{st}) of population

$$\text{Var}(\hat{Y}_{st}) = \sum_{h=1}^3 [N_h(N_h - n_h) \frac{S_h^2}{n_h}] \qquad \begin{aligned} \text{Var}(\hat{Y}_{st}) &= \left[72 \times (72-8) \times \frac{8.838 \times 10^9}{8} + 333 \times (333-37) \times \frac{7.679 \times 10^7}{37} + 207 \right. \\ &\quad \left. \times (207-23) \times \frac{5.677 \times 10^7}{23} \right] \approx 5.232 \times 10^{12} \text{ fish} \end{aligned}$$

$$S.E(\hat{Y}_{st}) = \sqrt{\text{Var}(\hat{Y}_{st})}$$

$$S.E(\hat{Y}_{st}) = \sqrt{5.232 \times 10^{12}} \approx 2,321,507 \text{ fish}$$

Step7. Calculation of variance and standard error of population estimate

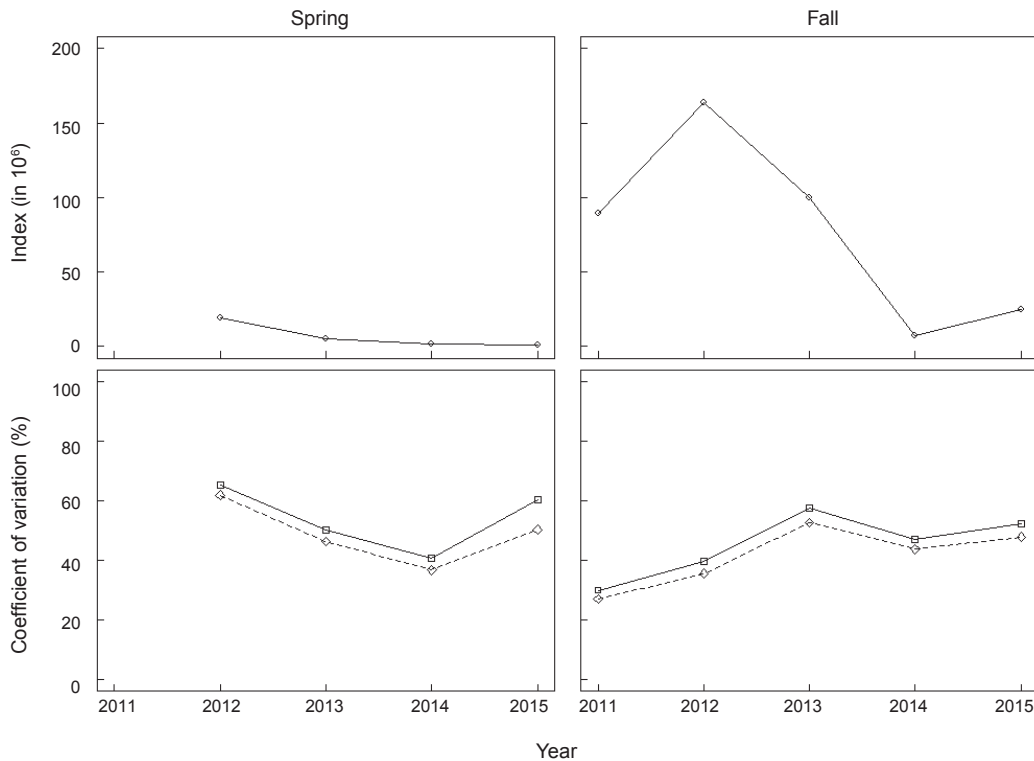


Fig. 5. The estimates of the relative sizes in number ($\times 10^6$) of cod population in years 2011-2015 under the assumption that the observation unit was a small rectangle in a Grid. The upper panels indicate estimates of the relative population size. The lower panels are the coefficient of variation. The solid line and the broken line indicate results of simple sampling, and stratified sampling respectively.

Table 4. Illustration of application of the simple sampling technique for calculation of the survey index in number of Pacific cod (*Gadus macrocephalus*) population in the spring of 2013 under the assumption that the observation unit was bottom area covered by a trawl. The left column of the table is estimator of a metric, and the right column is its estimate

Estimator	Estimate
$N = \frac{\sum_{i=1}^n T_i}{\sum_{i=1}^n a_i}$	$N = \frac{2,414 \text{ km}^2 + 2,431 \text{ km}^2 + \dots + 2,496 \text{ km}^2}{0.15 \text{ km}^2 + 0.10 \text{ km}^2 + \dots + 0.12 \text{ km}^2} = \frac{172,183 \text{ km}^2}{8.85 \text{ km}^2} \approx 19,460$
Step1. Calculation of the total number of possible tow (N)	
$y = \sum_{i=1}^n C_i$	$y = y_{113} + y_{208} + \dots + y_{203} = 261 \text{ fish}$
Step2. Calculation of sum of the observed values (y) where the net was trawled	
$\bar{y} = \frac{y}{n}$	$\bar{y} = \frac{261}{68} \approx 4 \text{ fish}$
Step3. Calculation of mean of the observed values (\bar{y})	
$\hat{Y} = N \times \bar{y}$	$\hat{Y} = 19,460 \times 4 \approx 74,693 \text{ fish}$
Step4. Calculation of the survey index of population (\hat{Y})	
$\text{Var}(\hat{Y}) = N^2 \cdot \frac{\text{Var}(y)}{n}$	$\text{Var}(\hat{Y}) = 19,460^2 \cdot \frac{246}{68} \approx 1.368 \times 10^9 \text{ fish}$
$S\hat{E}(\hat{Y}) = \sqrt{\text{Var}(\hat{Y})}$	$S\hat{E}(\hat{Y}) = \sqrt{1.368 \times 10^9} \approx 36,992 \text{ fish}$
Step5. Calculation of variance and standard error of population estimate	

의 상대적 개체군 크기'가 된다. 이를 계산하기 위해, 예망면적과 소해구면적에 관한 비례식을 적용하였고 그 식은 다음과 같다(Zhang, 2010).

$$(B \times q):A = C:a \dots\dots\dots (6)$$

B 는 단위면적 당 모집단 크기(개체수 혹은 생체량), A 는 단위면적, C 는 과학조사에서 어획된 어획량 그리고 a 는 예망면적을 뜻한다. q 는 어획률로, 모르는 상수로 두면 다음의 비례관계로 상대적 개체군 크기를 나타낼 수 있다.

Table 5. Illustration of application of the stratified sampling technique for calculation of the survey index in number of Pacific cod (*Gadus macrocephalus*) population in the spring of 2013 under the assumption that the observation unit was bottom area covered by a trawl. The left column of the table is estimator of a metric, and the right column is its estimate

Estimator	Estimate
$N_1 = \frac{\sum_{i=1}^{n_1} T_i}{\sum_{i=1}^{n_1} a_i}$	$N_1 = \frac{2,415 \text{ km}^2 + \dots + 2,512 \text{ km}^2}{0.12 \text{ km}^2 + \dots + 0.13 \text{ km}^2} \approx 17,596$
$N_h = \left\{ \begin{array}{l} N_2 = \frac{\sum_{i=1}^{n_2} T_i}{\sum_{i=1}^{n_2} a_i} \\ N_3 = \frac{\sum_{i=1}^{n_3} T_i}{\sum_{i=1}^{n_3} a_i} \end{array} \right.$	$N_2 = \frac{2,512 \text{ km}^2 + \dots + 2,602 \text{ km}^2}{0.15 \text{ km}^2 + \dots + 0.15 \text{ km}^2} \approx 20,887$
	$N_3 = \frac{2,447 \text{ km}^2 + \dots + 2,543 \text{ km}^2}{0.12 \text{ km}^2 + \dots + 0.12 \text{ km}^2} \approx 18,082$

Step1. Calculation of the number of possible tow in each stratum (N_h)

$$N = N_1 + N_2 + N_3 \qquad N = 17,596 + 20,887 + 18,082 \approx 56,564$$

Step2. Calculation of the total number of possible tow (N)

$$y_h = \sum_{i=1}^{n_h} C_{hi} \qquad \begin{array}{l} y_{h=1} = y_{i=55} + y_{i=63} + \dots + y_{i=93} = 161 \text{ fish} \\ y_{h=2} = y_{i=97} + y_{i=98} + \dots + y_{i=253} = 47 \text{ fish} \\ y_{h=3} = y_{i=151} + y_{i=152} + \dots + y_{i=212} = 53 \text{ fish} \end{array}$$

Step3. Calculation of sum of the observed values (y_h) in each stratum

$$\bar{y}_h = \frac{y_h}{n_h} \qquad \begin{array}{l} \bar{y}_{h=1} = \frac{161 \text{ fish}}{8} \approx 20 \text{ fish} \\ \bar{y}_{h=2} = \frac{47 \text{ fish}}{37} \approx 1 \text{ fish} \\ \bar{y}_{h=3} = \frac{53 \text{ fish}}{23} \approx 2 \text{ fish} \end{array}$$

Step4. Calculation of mean of the observed values (\bar{y}_h) in each stratum

$$\bar{y}_{st} = \frac{\sum_{h=1}^3 N_h \cdot \bar{y}_h}{N} \qquad \bar{y}_{st} = \frac{17,596}{56,564} \times 20 + \frac{20,887}{56,564} \times 1 + \frac{18,082}{56,564} \times 2 \approx 7 \text{ fish}$$

Step5. Calculation of weighted average of the observed values

$$\hat{Y}_{st} = N \cdot \bar{y}_{st} \qquad \hat{Y}_{st} = 56,564 \times 7 \approx 422,308 \text{ fish}$$

Step6. Calculation of the survey index (\hat{Y}_{st}) of population

$$\text{Var}(\hat{Y}_{st}) = \sum_{h=1}^3 [N_h (N_h - n_h) \frac{S_h^2}{n_h}] \qquad \text{Var}(\hat{Y}_{st}) = \left[17,596 \times (17,596 - 8) \times \frac{1,905}{8} + 20,887 \times (20,887 - 37) \times \frac{12}{37} + 18,082 \times (18,082 - 23) \times \frac{12}{23} \right]$$

$$\approx 7.401 \times 10^{10} \text{ fish}$$

$$S\hat{E}(\hat{Y}_{st}) = \sqrt{\text{Var}(\hat{Y}_{st})} \qquad S\hat{E}(\hat{Y}_{st}) = \sqrt{7.401 \times 10^{10}} \approx 272,042 \text{ fish}$$

Step7. Calculation of variance and standard error of population estimate

$$B \propto \frac{A \cdot C}{a} \dots\dots\dots (7)$$

식(2)의 B에 관측단위 개념을 추가하여, 관측치를 계산하는 식은 다음과 같다.

$$y_i = \frac{A_i \times C_i}{a_i} \dots\dots\dots (8)$$

는 관측단위의 순번을 의미하며, 관측단위는 소해구로 정의하였다. y_i 는 소해구 i 에서의 관측치, A_i 는 소해구 i 의 면적을 나타낸다. C_i 는 소해구 i 조사에서 어획된 어획량, a_i 는 소해구 i 에서의 예망면적을 의미한다. 관측치 계산의 간단한 예로, 2015년 봄철 조사 당시 63번 해구의 4번 소해구에서 5마리의 대구가 어획되었다. 소해구의 면적(A_i)은 약 270 km^2 , 예망면적(a_i)은 0.155 km^2 이므로 63번 해구의 4번 소해구에 존재하는 대구의 관측치는 $[(270 \text{ km}^2 \times 5\text{마리}) / 0.155 \text{ km}^2] \approx 8,714\text{마리}$ 로 추정할 수 있었다. 관측단위를 소해구로 가정하였을 때, 단순 추출법과 층화 추출법을 이용한 조사지수 추정과정의 예시로, 개체수에 관한 2013년 봄철 대구 모집단의 조사지수(\hat{Y} ; \hat{Y}_{st})와 추정치의 표준오차[$S\hat{E}(\hat{Y})$; $S\hat{E}(\hat{Y}_{st})$] 계산과정 및 결과를 Table 2와

Table 3으로 정리하였다.

관측단위: 예망면적

표본 추출법의 관측단위를 다르게 가정할 경우, 조사지수와 N의 계산이 달라진다. 관측단위를 예망면적으로 가정한다면, 관측치(y_i)는 소해구와 예망면적 사이의 비례식을 적용하지 않은 ‘한 소해구 i 의 조사에서 어획된 특정 종의 어획량(C_i)’이 된다. n 은 전체 조사대상 해역에서 예망을 시행한 횟수(Smith and Lundy, 2006)를 나타낸다. N 은 조사 가능한 전체 예망 횟수로, 다음 식을 통해 계산할 수 있다.

$$N = \frac{\sum_{i=1}^n T_i}{\sum_{i=1}^n a_i} \dots\dots\dots (9)$$

T_i 는 소해구 i 가 속해 있는 해구(Grid)의 면적을 의미하며, 이후 나머지 계산과정은 관측단위를 소해구로 지정하였을 때의 단순 추출법과 동일하다. 층화 추출법에서 y_{hi} 는 ‘한 층 h 내에 소해구 i 의 조사에서 어획된 특정 종의 어획량’을 의미하고 n_h 는 하나의 층내에서 예망을 시행한 횟수를 나타낸다. N_h 는 한 층내에서의 조사 가능한 예망수를 의미하며, 다음 식으로 계산이 가능하다.

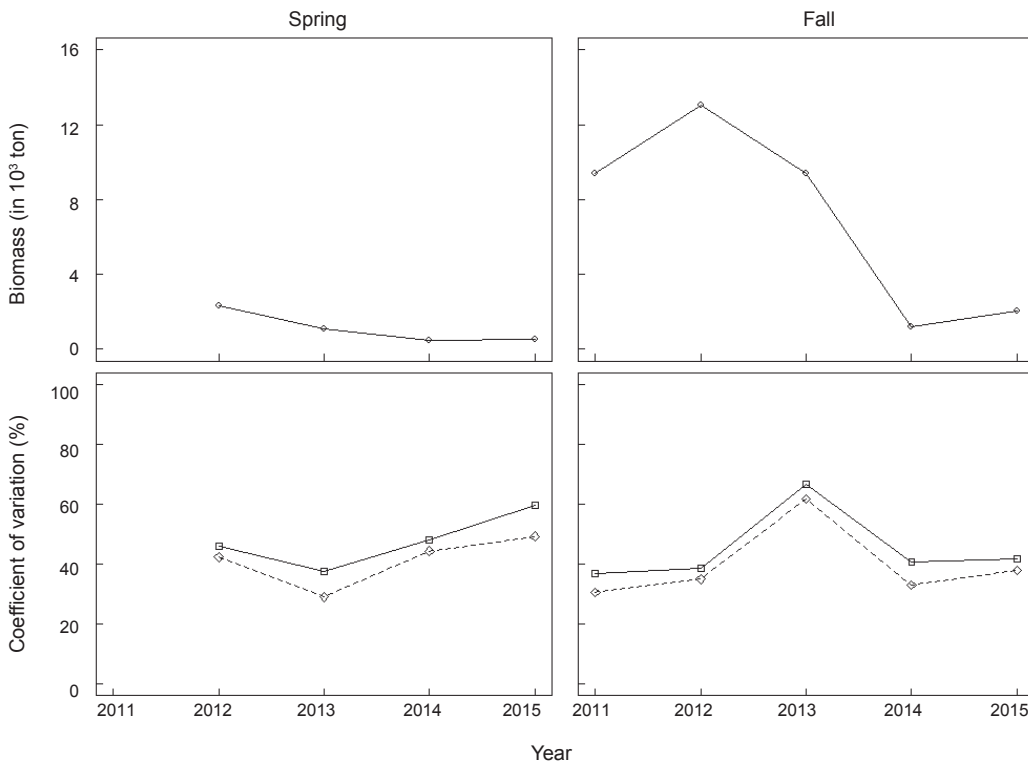


Fig. 6. The estimates of the relative sizes (in 10³ metric tons) of cod population in years 2011-2015 under the assumption that the observation unit was a small rectangle in a Grid. The upper panels indicate estimates of the relative population size. The lower panels are the coefficient of variation. The solid line and the broken line indicate results of simple sampling, and stratified sampling respectively.

$$N_h = \begin{cases} N_1 = \frac{\sum_{i=1}^{n_1} T_i}{\sum_{i=1}^{n_1} a_i} \\ N_2 = \frac{\sum_{i=1}^{n_2} T_i}{\sum_{i=1}^{n_2} a_i} \\ N_3 = \frac{\sum_{i=1}^{n_3} T_i}{\sum_{i=1}^{n_3} a_i} \end{cases} \dots\dots\dots (10)$$

이후 나머지 계산과정은 관측단위를 소해구로 지정하였을 때의 층화 추출법과 동일하다. 관측단위를 예망면적으로 가정하였을 때, 단순 추출법과 층화 추출법을 이용한 조사지수 및 추정치 분산 추정과정의 예시로, 개체수에 관해 2013년 봄철 대구 모집단의 조사지수($\hat{Y}; \hat{Y}_{st}$)와 추정치의 표준오차($\widehat{SE}(\hat{Y}); \widehat{SE}(\hat{Y}_{st})$) 계산과정 및 결과를 Table 4와 Table 5로 정리하였다.

결과 및 고찰

본 연구에 이용된 방법론의 예시를 보이기 위해, 과학조사에서 어획된 종들 중에서 대구(*Gadus macrocephalus*)만을 대상으로 결과를 제시하였다. 2011년 봄철의 경우, 예망면적을 계

산하기 위해 필요한 전개판 간격 정보의 누락으로, 결과에서 제외하였다. 관측단위를 소해구로 가정한 경우, 모든 층에서 해구당 관측단위 개수가 동일하여 모집단 크기의 조사지수는 단순 추출법과 층화 추출법에서 동일하게 나타났다(Fig. 5, 6). 그리고 변동계수는 모든 결과에서 단순추출법이 더 높은 값을 나타냈기 때문에, 정밀도(precision) 측면에서 층화 추출법이 더 우수한 추정치를 보였다. 이는 층화 추출법이 표본을 추출하고자 하는 모집단에 관한 정보를 이용하여, 동질한 특성을 가지는 범위로 묶는 층화(stratification) 과정 이후 표본을 추출하기 때문에 단순 추출법보다 정밀도의 측면에서 우수한 추정치를 제시한다는 일반적 사실(Scheaffer et al., 1971; Lohr, 1998)과 일치하였다.

하지만, 관측단위를 예망면적으로 가정한 경우, 각 해구마다 관측단위의 수가 달라 단순 추출법과 층화 추출법의 점추정치가 다르게 나타났으며, 대부분 층화 추출법에서 높은 추정치를 나타냈다(Fig. 7, 8). 변동계수는 층화 추출법의 결과에서 더 높은 변동계수를 나타내기도 하여, 기존의 이론과 다른 결과를 보였다. 이는 이미 수집된 자료에 표본이론을 적용하여, 각각의 표본이론에 이용된 가정이 달라 나타나는 결과라 생각된다. 본 연구의 결과들은 같은 자료를 이용하더라도, 과학조사의 사전 설

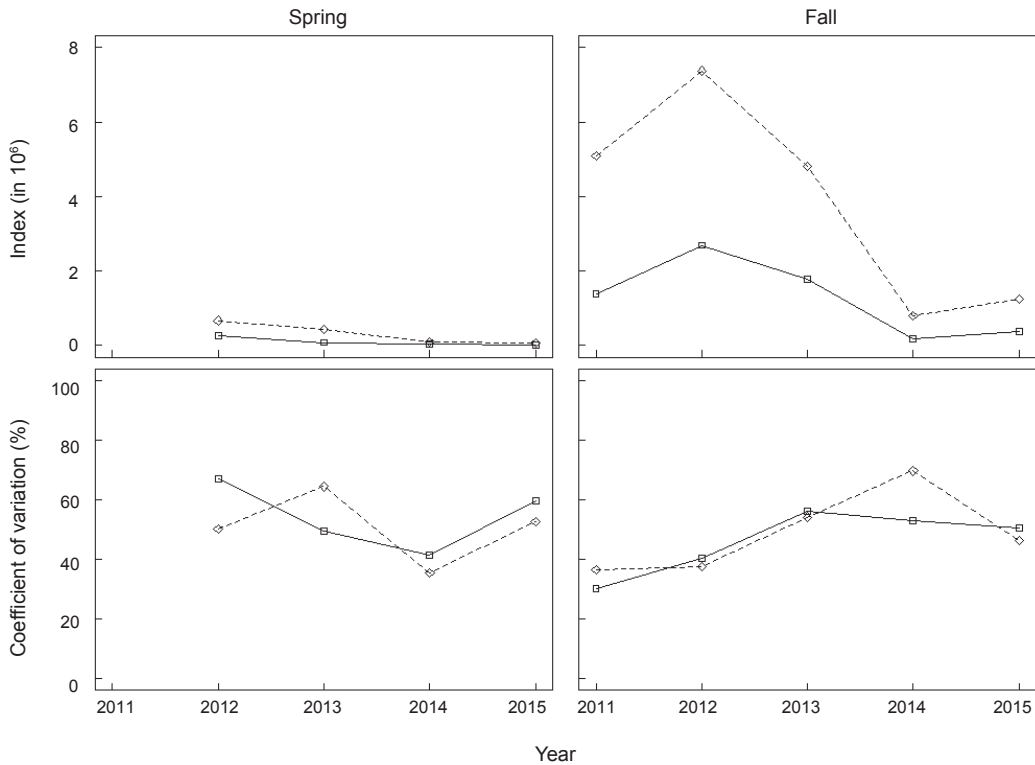


Fig. 7. The estimates of the relative sizes in number ($\times 10^6$) of cod population in years 2011-2015 under the assumption that the observation unit was bottom area covered by a trawl. The upper panels indicate estimates of the relative population size. The lower panels are the coefficient of variation. The solid line and the broken line indicate results of simple sampling, and stratified sampling respectively.

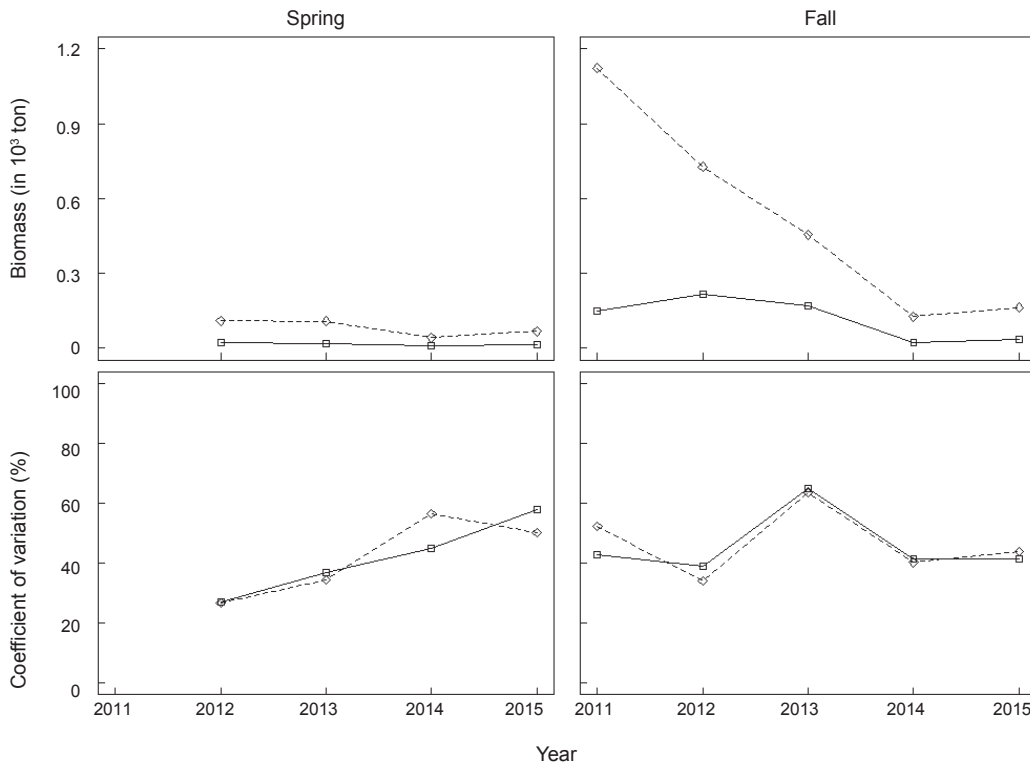


Fig. 8. The estimates of the relative sizes (in 10³ metric tons) of cod population in years 2011-2015 under the assumption that the observation unit was bottom area covered by a trawl. The upper panels indicate estimates of the relative population size. The lower panels are the coefficient of variation. The solid line and the broken line indicate results of simple sampling, and stratified sampling respectively.

계방법의 가정 따라 전혀 다른 조사지수와 그 분산을 보였다. 따라서, 보다 정확한 자원평가를 위해 특정 표본이론과 그에 맞는 적절한 관측단위를 설정하여 사전 설계된 과학조사가 요구된다는 것을 시사한다.

관측단위를 소해구로 가정할 경우, 관측치 계산을 위한 비례식에 ‘한 소해구 전체 면적에 생물이 균일하게 분포한다’는 가정이 뒤따른다. 하지만 대부분의 생물들은 무리를 형성하여 생활하거나 생존에 적절한 환경에 모여 군집을 형성하기 때문에 생물학적으로 위 가정은 적절하다 할 수 없다. 또한 층화 추출법을 적용하여 조사지수와 추정치의 분산 계산 과정에서 층화의 동질성 또는 이질성을 가늠할 수 있는 추가적인 환경 자료를 적용하지 않고 단순히 위도와 경도에 관한 정보를 적용하여 임의적으로 층군(strata)을 가정하였기 때문에, 층화에 의한 정밀도의 증가를 적절히 반영한다 할 수 없다. 본 연구는 위에 제시된 약점과 보완점들이 존재하지만, 한반도 해역에서 이루어진 과학조사 자료를 이용하여 개체군의 상대적 크기를 추정하고 추정치의 불확실성을 평가하는 첫 번째 연구라는 점에서 의의를 가진다.

앞으로 진행될 과학조사를 새롭게 설계한다면 다양한 환경 정보를 반영할 수 있는 층화 임의 표본 추출법으로 설계할 것을 제안한다. 이미 층화 임의 추출법은 많은 과학조사의 설계에 이용

되는 방법으로, 환경 또는 모집단에 관한 정확한 정보를 많이 반영할수록 추정치의 정밀도가 높아져 효과적인 과학조사를 진행할 수 있을 것이다. 추가적으로, 관측단위를 예망면적으로 설계하는 것을 제안한다. 본 논문에서 제시한 조사지수와 그 분산들은 여러 가정들(2가지 표본론과 2가지 관측단위)의 적용에 의한 결과이다. 따라서 어떤 결과가 더 적절한지 판단할 수 없지만, 관측단위의 측면에서 본다면 ‘한 소해구 전체 면적에 생물이 균일하게 분포한다’는 부적절한 가정을 필요로 하지 않는 관측단위, 즉 관측치를 ‘예망시 어획된 특정종의 어획량’으로 가정하는 것이 더 현실적이라 사료된다.

사 사

본 논문의 연구는 국립수산과학원 수산과학연구사업의 자료 지원과, 부경대학교 자율창의학술연구비(2016년)의 재정지원을 받고 수행되었습니다. 저층트롤 조사자료를 제공해주신 국립수산과학원에 깊은 감사를 드립니다. 논문의 검토 및 수정에 도움을 주신 김동주 선생님, Victoria University of Wellington에 재학 중인 김규한 선생님, 부경대학교 자원생물학과 최소희 학생에게 감사드립니다. 또한 국외의 저층트롤 과학조사 방법 조사에 도움을 주신, International Pacific Halibut Commis-

sion 소속의 Dr. Ian Stewart께도 사의를 표합니다.

Zhang CI. 2010. Marine fisheries resource ecology. Pukyong National University, Busan, Korea, 368-371.

References

- Cochran WG. 1977. Sampling techniques. John Wiley & Sons, New York, NY, U.S.A., 18-95.
- Geoscience Australia. 2010. Geodetic Techniques. Retrieved from <http://www.ga.gov.au/scientific-topics/positioning-navigation/geodesy/geodetic-techniques/calculation-methods> on Aug 26, 2017.
- IPHC (International Pacific Halibut Commission). 2012. IPHC Setline charters 1963 through 2003. Technical Report 58, 12-69.
- KODC (Korea Oceanographic Data Center). 2011. NIFS (National Institute of Fisheries Science) Serial Oceanographic Observation. Retrieved from http://kodc.nifs.go.kr/kodc/eng_eng_soo_summary.kodc on Oct 11, 2016.
- Lohr SL. 1998. Sampling: Design and analysis. Julet M, ed. Brooks/Cole, Boston, MA, U.S.A., 73-82.
- NOAA (National Oceanic and Atmospheric Administration). 1988. An evaluation of the bottom trawl survey program of the northeast fisheries center. NOAA Technical Memorandum NMFS-F/NEC 52, 7-33.
- NOAA (National Oceanic and Atmospheric Administration). 2011. The 2003 to 2008 U.S. west coast bottom trawl surveys of groundfish resources off Washington, Oregon, and California: Estimates of distribution, abundance, length, and age composition. NOAA Technical Memorandum NMFS-NWFSC 114, 3-10.
- NOAA (National Oceanic and Atmospheric Administration). 2012. Assessment or data updates of 13 northeast groundfish stocks through 2010. Northeast Fish Sci Cent Ref Doc 12-06, 185-186.
- Perry RI and Smith SJ. 1994. Identifying habitat associations of marine fishes using survey data: An application to the north-west Atlantic. *Can J Fish Aquat Sci* 51, 589-602. <https://doi.org/10.1139/f94-061>.
- Scheaffer RL, Mendenhall W, Ott RL and Gerow KG. 1971. Elementary survey sampling, Julet M, ed. Brooks/Cole, Boston, MA, U.S.A., 114-133.
- Smith SJ and Gavaris S. 1993. Improving the precision of abundance estimates of eastern Scotian shelf Atlantic cod from bottom trawl surveys. *North Am J of Fish Management* 13, 35-47. [http://dx.doi.org/10.1577/1548-8675\(1993\)013<0035:ITPOAE>2.3.CO;2](http://dx.doi.org/10.1577/1548-8675(1993)013<0035:ITPOAE>2.3.CO;2).
- Smith SJ and Hubley B. 2014. Impact of survey design changes on stock assessment advice: Sea scallops. *ICES J Mar Sci* 71, 320-327. <https://doi.org/10.1093/icesjms/fst115>.
- Smith SJ and Lundy MJ. 2006. Improving the precision of design-based scallop drag surveys using adaptive allocation methods. *Can J Fish Aquat Sci* 63, 1639-1646. <https://doi.org/10.1139/f06-063>.