

개체군 체장자료를 이용한 연령조성 추정

김규한 · 현상윤^{1*} · 서영일²

웰링턴 빅토리아대학교 수학 통계학부, ¹부경대학교 자원생물학과, ²국립수산과학원 연근해자원과

Inference of Age Compositions in a Sample of Fish from Fish Length Data

Kyuhan Kim, Saang-Yoon Hyun^{1*} and Young Il Seo²

School of Mathematics and Statistics, Victoria University of Wellington, Wellington 6140, New Zealand

¹Department of Marine Biology, Pukyong National University, Busan 48513, Korea

²Coastal Water Fisheries Resources Research Division, National Institute of Fisheries Science, Busan 46083, Korea

Fish ages are critical information in fish stock assessments because they are required for age-structure models such as virtual population analysis and stochastic catch-at-age models, whose outputs include recruitment strengths, a spawning stock size (abundance or biomass), and the projection of a fish population size in future. However, most countries other than the developed countries have not identified ages of fish caught by fisheries or surveys in a consistent manner for a long time (e.g., >20 years). Instead, data about fish body sizes (e.g., lengths) have been well available because of ease of measurement. To infer age compositions of fish in a target group using fish length data, we intended to improve the length frequency analysis (LFA), which Schnute and Fournier had introduced in 1980. Our study was different in two ways from the Schnute and Fournier's method. First we calculated not only point estimates of age compositions but also the uncertainty in those estimates. Second, we modified LFA based on the von Bertalanffy growth model (ν B-based model) to allow both individual-to-individual and cohort-to-cohort variability in estimates of parameters in the ν B-based model. For illustration, we used data about lengths of Korean mackerel *Scomber japonicas* caught by purse-seine fisheries from 2000-2016.

Key words: Automatic differentiation model builder, age compositions, length frequency distributions, *Scomber japonicus*

서론

미국, 캐나다 등 여러 수산 선진국가들에서는 효율적인 자원 관리를 위해 다양한 수산자원평가모델을 개발하였으며, 실제로 이러한 모델을 이용하여 자원관리의 주요 지표[e.g., 개체수, 사망률 (또는 생산율), 최대지속어획량 (maximum sustainable yield), 어획률, 가입강도 등]를 추정하고 있다. 뿐만 아니라, 추정된 지표의 불확실성(uncertainty)까지 계산하여 관리자가 효과적인 수산자원관리 정책을 수립 할 수 있는 과학적 정보를 제공하고 있다. 이러한 수산자원평가모델 중 대표적인 것이 연령구조 확률모델(stochastic age-structured models)인데 연령별로 개체수를 추정하고 미래의 개체수도 예보할 수 있는 장

점이 있다(Fournier and Archibald, 1982; Methot and Wetzel, 2013). 그러나 우리나라의 경우 어류의 연령자료가 부재한 상황이며, 현존하는 어류 연령자료의 대부분도 단기간에 국한되어 있거나 일관성이 결여되어 있어 어업 및 과학조사(survey) 자료를 활용한 연령구조모델의 적용이 상당히 제한적인 상황이다. 하지만 이러한 연령자료의 부재 속에서도, 대부분의 어업 자료에 어류의 체장과 체중이 장기간에 걸쳐 기록되어 있다는 점은 두 생물학적 자료를 기반으로 한 연령구조모델의 활용 가능성을 제시하였다.

어류 개체군의 체장빈도분포를 이용하여 어류의 성장 모수와 연령 조성을 추정하는 체장빈도분석법(Length Frequency Analysis; LFA)은 이미 Schnute and Fournier (1980)가 소

<https://doi.org/10.5657/KFAS.2018.0079>



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 51(1) 79-90, February 2018

Received 2 January 2018; Revised 22 January 2018; Accepted 23 January 2018

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

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

개하였고, 현재 이용되고 있는 수산자원평가모델 중 하나인 MULTIFAN-CL 소프트웨어의 한 모듈(module)로 실행화되고 있다. 하지만, Schnute and Fournier (1980)의 연령조성 추정치는 점 추정치에만 국한되어 있을 뿐만 아니라 어류의 연급군(cohort)간 성장 차이를 고려하지 못하였고 MULTIFAN-CL의 LFA는 독립된 2개 이상의 체장빈도수 자료를 활용하여 하나의 von Bertalanffy 성장 모수를 추정하기 때문에 어류 개체군의 시기별 또는 연급군간 성장 차이를 나타내지 못하는 단점을 지니고 있다.

따라서 본 연구는 크게 두가지 목적에 기반을 두고 있다. 첫째로는 어류의 연급군간 성장 차이를 반영하기 위해 기존 Schnute and Fournier의 LFA (Schnute and Fournier, 1980)를 개선하는 것이며, 둘째로는, 개체군 어획물의 체장자료만 이용하여 그 어획물의 시기별 성장과 연령조성치를 추정하고 그 추정치들의 불확실성까지 계량화하는 것이다. 실례로서, 국립수산과학원에서 2000년부터 2016년까지 수집한 고등어(*Scomber japonicus*)의 체장 자료를 이용하였다. 일반적으로 성장이 빠르고 수명이 짧은 어종은 연령별 체장의 편차가 작기 때문에 체장빈도수 분포의 최빈값의 수(the number of modes)가 뚜렷하게 나타나므로 최빈값의 개수가 체장빈도분포에 숨겨진 연령수(the number of age classes)를 나타낸다(Pauly and Morgan, 1987). 그러나 성장이 비교적 느리거나 수명이 긴 어종은 최빈값의 수가 뚜렷하게 나타나지 않기 때문에 그 개수가 연령수보다 적게 나타나며 이러한 경우 자료의 분포 형태를 바탕으로 연령조성을 추정하는 것은 더욱 어렵다. 특히, 본 연구에서 적용한 한국 연근해 고등어의 체장분포는 후자의 경우에 해당되므로 이 연구의 실용성을 증명하였다.

재료 및 방법

체장 조성 자료

국립수산과학원에서 2000년부터 2016년까지 수집한 고등어의 월별 체장 조성 자료를 본 연구에 이용하였다. 우리나라 해역에서 대형선망으로 어획된 고등어를 샘플하여 가랑이 체장(fork length: FL)이 측정되었으며, 개체별 체장의 치수는 누락되어 있으나 계급 구간(class interval)이 1 cm인 도수 분포표(frequency table)에 계급별 빈도수가 기록되어있었다. 17년동안 수집된 체장 자료의 총 빈도수가 241,677개에 이를 정도로 자료의 표본 크기(sample size)가 방대하여 체장빈도분석을 위한 자료로 적합하였다.

2015년 5월을 제외한 모든 월별 체장빈도수 자료가 기록되어 있으나 월별 분석을 수행하기에는 표본 크기가 충분하지 않았다. 표본 크기의 충분 여부는 히스토그램을 관찰하여 판단하였으며, 월별 자료의 경우 특정 계급의 상대빈도수가 이웃한 계급에 비해 극도로 높거나 낮게 나타나는 경우가 빈번하게 나타났다. 따라서 충분한 표본 크기를 확보하기 위해 연도별 자

료를 상반기 (1월-6월)와 하반기 (7월-12월)로 나누어 분석하였으며, 총 34개의 데이터 세트를 체장빈도분석에 적용하였다 (Fig. 1-3).

체장빈도분석법(Length frequency analysis)

본 연구는 두가지 점에서 Schnute and Fournier (1980) 방법과 다르다. 우선 수치적 추정방법으로 Automatic Differentiation Model Builder (ADMB) 스크립트 소프트웨어(Fournier et al., 2012)를 사용하여 모수들의 점 추정치뿐만 아니라 그 추정치의 불확실성도 계량화하였다. 두번째로는 기존 Schnute and Fournier의 방법인 von Bertalanffy (vB)모델에 기반을 두고 점 추정치를 계산한 결과와 연급군간 성장 차이를 고려하기 위해 vB 모델에 업데이트 않고 계산한 결과를 함께 제시하였다.

von Bertalanffy기반 체장빈도분석법 (vB-based)

Schnute and Fournier 방법(Schnute and Fournier, 1980) (이하 Schnute and Fournier의 LFA)에서 연령별 평균 체장(μ_a)은 von Bertalanffy 성장식(von Bertalanffy, 1938)을 다음과 같이 재모수화(reparametrize)한다(Ford, 1933; Walford, 1946; Ricker, 1975).

$$\mu_a = H(L - I) \frac{1 - k^{a-1}}{1 - k^{M-1}}, \text{ where } a=1, \dots, M \dots \dots (1)$$

여기서 a 는 연령, I 은 최소 연령 그룹의 평균 체장, L 은 최대 연령 그룹의 평균 체장, k 는 Brody coefficient (e.g., $K = -\log(k)$); 여기서 K 는 von Bertalanffy 성장식의 성장 모수), M 은 최대연령을 의미한다.

연령이 a 인 같은 종의 어류 집단이 서로 다른 크기의 체장을 가진다는 사실은 자명하며, 대부분의 어종에서 이러한 연령별 체장의 표준편차(σ_a)는 연령이 커짐에 따라 증가한다. 이를 연령에 관한 선형식으로 나타내면 다음과 같다(Schnute and Fournier, 1980).

$$\sigma_a = s + (S - s) \frac{a - 1}{M - 1}, \text{ where } a=1, \dots, M \dots \dots (2)$$

여기서, s 는 최소 연령 그룹의 체장의 표준편차, S 는 최대 연령 그룹의 체장의 표준편차를 나타내는데, 어류의 연령별 체장의 분포가 정규분포를 따른다는 가정을 하면 eq. (1)과 (2)에서 유도한 연령별 평균 체장(μ_a)과 표준편차(σ_a)를 대입하여 연령이 a 인 어류의 체장이 j 번째 계급(j th length interval)에 속할 확률(q_{aj})을 계산할 수 있다(Schnute and Fournier, 1980).

$$q_{aj} = \frac{1}{\sigma_a \sqrt{2\pi}} \int_{x_j - w/2}^{x_j + w/2} \exp\left[-\frac{1}{2} \left(\frac{x - \mu_a}{\sigma_a}\right)^2\right] dx \dots \dots (3)$$

여기서 x 는 체장을 의미하며, x_j 는 j 번째 계급(체장)의 중간

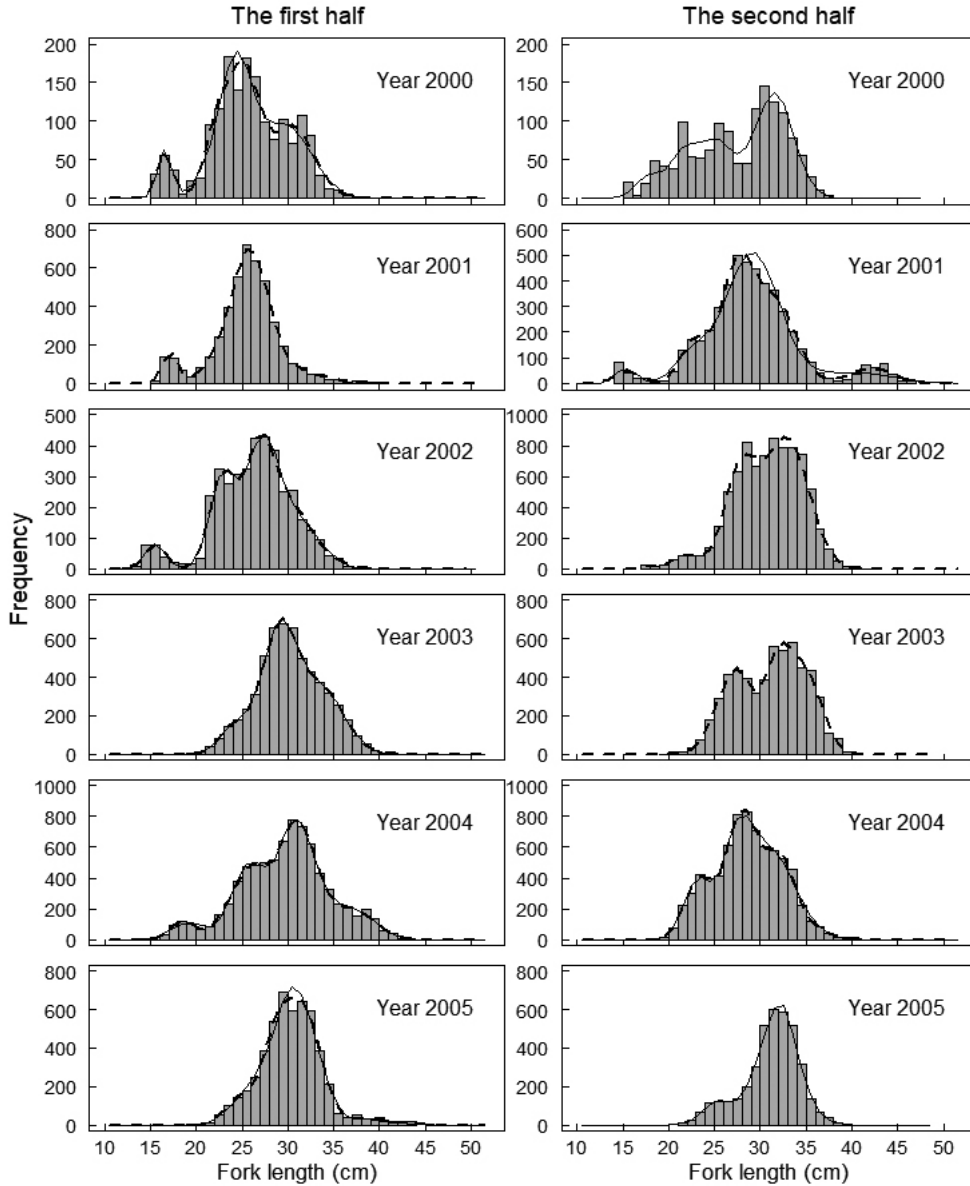


Fig. 1. Length frequency distributions from 2000-2005. Bars represent actual length frequencies, and lines are expected frequencies made by the vB-based method (solid line) and the vB-free method (broken line).

값, w 는 각 계급의 구간 넓이를 나타내는데, 체장 자료의 총 빈도수(f)에 q_{aj} 와 각 연령 그룹의 조성 비율(P_a)을 곱하면 j 번째 계급의 기대 빈도수(expected frequency; f_j)를 구할 수 있다.

$$\hat{f}_j = f \sum_{a=1}^M (P_a q_{aj}) \quad \dots \dots \dots (4)$$

Eq. (1)-(4)의 모수들(e.g., l, L, k, s, S, P_a through P_{M+1})을 추정하기 위해 separation statistic (A)를 목적 함수(objective func-

tion)로 이용하였다(Macdonald and Pitcher, 1979; Schnute and Fournier, 1980).

$$A = 2 \sum_{j=1}^N f_j^{(1)} \log\left(\frac{f_j}{\hat{f}_j}\right) \quad \dots \dots \dots (5)$$

여기서 $f_j^{(1)}$ 는 j 번째 계급의 관측 빈도수(observed frequency), $\sum_{j=1}^N f_j^{(1)}$ 은 1이상의 빈도수를 가지는 계급의 빈도수 합(e.g., $\sum_{j=1}^{(m)}$

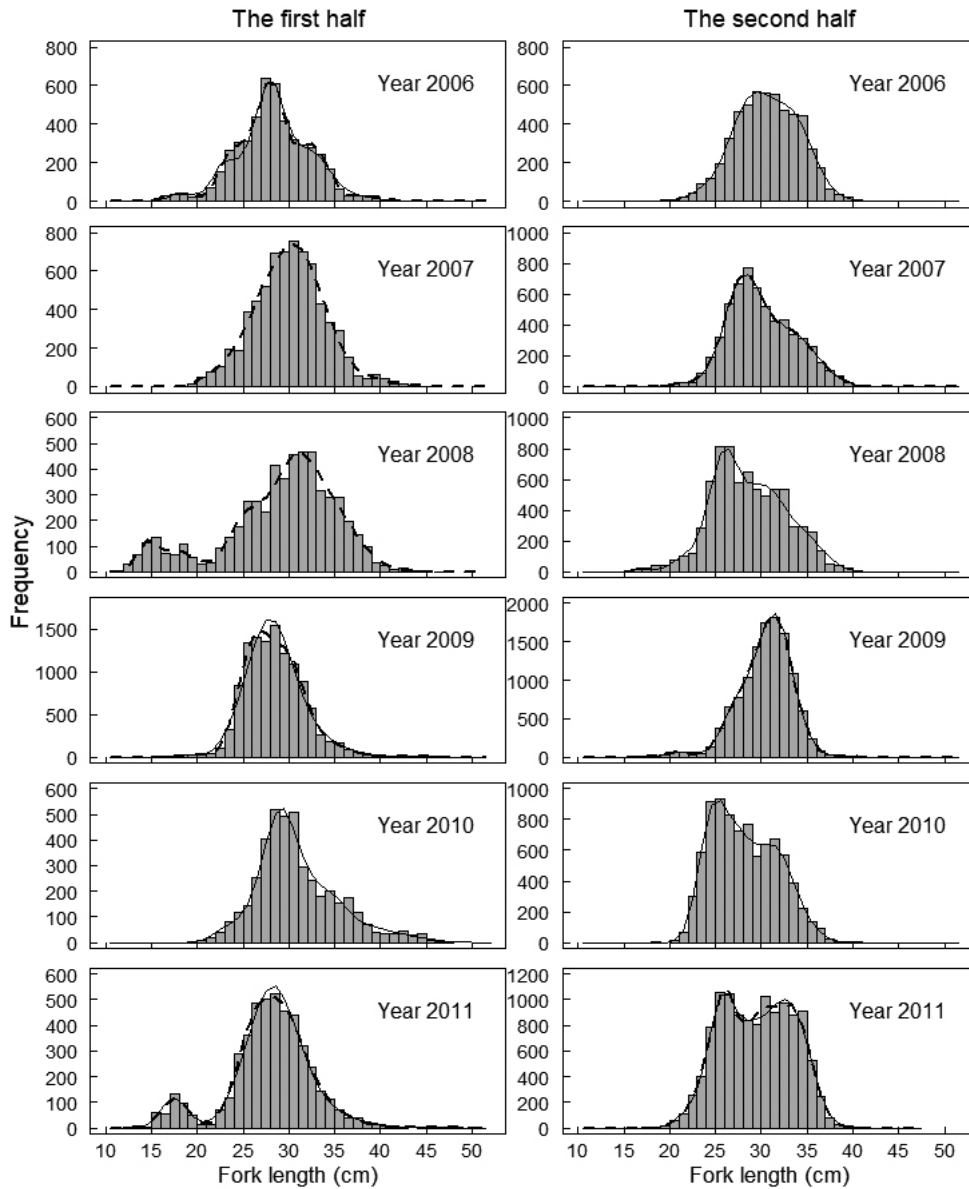


Fig. 2. Length frequency distributions from 2006-2011. Bars represent actual length frequencies, and lines are expected frequencies made by the vB-based method (solid line) and the vB-free method (broken line).

은 $f_j \geq m$ 의 합), N 은 총 계급의 수를 나타내며 A 를 최소화 하는 \hat{f}_j 를 구하여 모수들의 추정치를 계산 할 수 있다.

von Bertalanffy에 얽매이지 않는 체장빈도분석법 (vB-free)

von Bertalanffy 성장모델의 기본 가정 중의 하나가, '한 개체 (an individual)의 순간 성장률(instantaneous growth rate)은 체장에 비례하여 같은 속도로 감소한다'라는 것을 되새긴다면 (von Bertalanffy 1938), von Bertalanffy 성장모델을 여러 개의

연급군에 적용할 때, 상당한 오차가 수반되는 것에 동의할 것이다. 즉, von Bertalanffy 성장모델을 기반으로 유도된 Schnute and Fournier의 LFA도 여러 연급군을 포함하고 있는 어체장 표본에 적용할 때 상당한 오차가 있을 수 있다. 같은 이유로, 본 연구에서 이용한 고등어의 상반기 및 하반기별 체장 자료가 수 개의 연급군으로부터 수집되었다는 점을 상기하면, 연령별 평균 체장이 von Bertalanffy 성장식의 점근적(asymptotic) 패턴을 따르지 않을 가능성이 있다. 즉, 연급군간 성장률의 차이로 인해 수개의 연급군이 섞여있는 체장 조성 자료를 분석할 경우

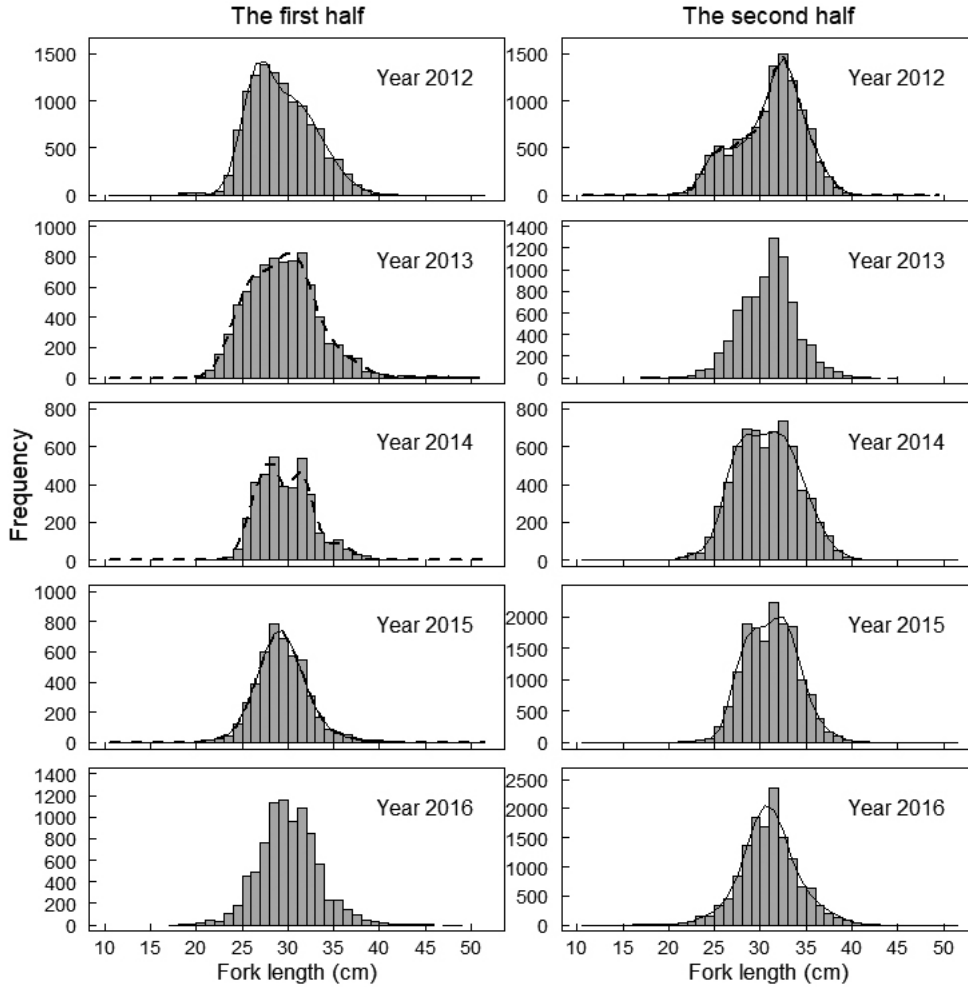


Fig. 3. Length frequency distributions from 2012-2016. Bars represent actual length frequencies, and lines are expected frequencies made by the vB-based method (solid line) and the vB-free method (broken line).

연령이 a 인 그룹과 $a+1$ 인 그룹의 평균 체장의 차이가 $a+2$ 그룹과 $a+3$ 그룹의 평균 체장의 차이보다 작은 경우가 발생 할 수 있다는 것이다.

연급군별 분석이 불가능한 상황에서, 앞서 언급한 연령별 평균 체장의 불규칙성을 반영하기 위해 eq. (1)의 μ_a 를 유도된 모수(derived parameter)가 아닌 자유 모수(free parameter)로 이용하였다. 따라서, 기존의 von Bertalanffy식을 이용한 Schnute and Fournier의 LFA (이하 vB-based) 방법은 eq. (1)의 3개의 모수(i.e., l , L , and k)를 추정하여 전체 연령별 평균 체장(μ_a)을 계산하였기 때문에 전체 연령 그룹의 수(M)와 관계없이 추정해야 할 모수의 수가 3개로 한정되어 있으나, μ_a 를 자유 모수로 이용한(이하 vB-free)방법은 M 이 4 이상인 경우 추정해야 할 모수의 수가 vB-based보다 1개씩 증가하게 된다(e.g., vB-based에서 l , L , k ; vB-free에서 $\mu_1, \mu_2, \mu_3, \mu_4$).

모수의 추정(Inference of parameters)

Eq. (5)의 A 를 목적함수(objective function)로 취급하여 A 의 최소값을 계산하기 위해, ADMB 소프트웨어를 이용하였다. ADMB를 이용하는 과정을 간단히 기술하면, A 를 모수들에 관하여 수치적으로 미분하여 각 모수들의 추정치를 계산하고, 또한 A 를 2차 미분한 헤시안 행렬(Hessian matrix)을 바탕으로 모수 추정치의 분산-공분산(variance-covariance) 행렬을 구한다. ADMB 코드(TPL file)와 관련된 자료(DAT file)는 Hyun et al. (2017)의 보고서에 첨부되어 있다.

모델 적합도 검증 및 비교

일반적으로, 추정할 모수의 갯수가 증가할수록 A 값은 줄어들게 된다. 즉, A 값을 모델의 적합도 지수로 이용할 경우 모수의 수가 더 많은 vB-free방법이 더 적합한 모델로 평가받을 가능성

이 높다. 그러나, 모델의 적합도를 검증할때는 모델의 간결성 (parsimony) 측면을 함께 고려해야하며, Schnute and Fournier의 LFA에서는 χ^2 테스트를 통해 이러한 모델 적합도 검증을 수행하였다.

Schnute and Fournier (1980)는 χ^2 통계량 (B_m)을 계산하기 위해 아래 식을 이용하였다.

$$B_m = \sum_{j=1}^N \frac{{}^{(m)}(f_j - \hat{f}_j)^2}{\hat{f}_j} + \frac{(\hat{g}_m - g_m)^2}{f - \hat{g}_m}, \text{ where } m=2, \dots, 5 \dots \dots (6)$$

여기서, \hat{g}_m 과 g_m 은 각각 $\sum_{j=1}^N \hat{f}_j$ 과 $\sum_{j=1}^N {}^{(m)}f_j$ 을 나타내며, 각 χ^2 통계량의 자유도 (E_m)는 다음의 식을 이용하여 계산하였다.

$$E_m = N_m - R \dots \dots \dots (7)$$

여기서, N_m 은 $f_j \geq m$ 인 계급의 수, R 은 모델에서 추정되어야 할 자유 모수의 총 개수이다. 본 연구에서는 eq. (6)과 (7)을 통해 계산된 χ^2 의 P 값을 바탕으로 vB-based와 vB-free의 모델 적합도를 비교하였다.

결 과

체장빈도분석법(Length frequency analysis)

Choi (2003)가 추체골을 연령 형질로 이용하여 추정된 von Bertalanffy 성장식, $L_t = 51.67 \cdot [1 - \exp\{-0.299 \cdot (t + 0.428)\}]$ 을

바탕으로 적절한 연령 그룹의 수를 파악하였다. 한국 연근해에 서식하는 고등어의 수명은 6세로 알려져 있으며, Choi (2003)의 성장식에 따르면 1-6세의 평균 체장은 각각 17.96 cm, 26.67 cm, 33.13 cm, 37.92 cm, 41.47 cm, 44.11 cm 이다. 본 연구에서 이용한 고등어의 체장 조성 자료에서 5세의 평균 체장이 속하는 계급(41-42 cm)보다 큰 계급의 빈도수는 전체 년도에 걸쳐 583개로 나타났다. 이는 전체 총 빈도수(241,677개)에서 0.2%를 차지하는 비율로 42 cm 보다 큰 개체의 출현 빈도가 상당히 낮다는 것을 알 수 있다. 물론, 6세 그룹의 체장의 표준편차가 가장 크다는 것을 고려하였을 때[eq. (2) 참고], 42 cm 보다 작은 개체들 중 일부는 5세가 아닌 6세일 가능성을 배제할 수 없다. 그럼에도 불구하고, 6세 그룹에서 가장 높은 빈도수(최빈값)를 보여야 할 44-45 cm 계급(또는 인접한 계급) 역시 그 출현 빈도수가 상당히 낮거나 존재하지 않았기 때문에, 본 연구에서는 고등어의 총 연령 그룹의 수를 6개가 아닌 5개로 설정하였다. 그러나 5세 그룹에 일부 포함될 6세 또는 7세이상 개체들을 고려하여 마지막 연령 그룹인 5세는 ‘age-5’가 아닌 ‘age-5+’로 고려하였다.

von Bertalanffy기반 체장빈도분석법(vB-based)

총 34개의 반기별 체장 자료 중 25개의 자료에서 vB-based의 모수가 성공적으로 추정되었으며 각 모수의 추정치와 표준오차는 Table 2에 정리하였다[모수 추정치간의 상관관계는 Hyun et al. (2017) 보고서를 참고]. 나머지 9개 자료의 모수 추정 실패 요인을 살펴보면 2000-I, 2013-I, 2014-I (숫자는 년도, I와 II는

Table 1. Notations used in this paper

Notation	Description
a	age.
M	The number of age classes.
μ_a	The mean of the lengths of age- a fish.
l	The mean of the lengths of fish at the first age class.
L	The mean of the lengths of fish at the last age class.
s	The standard deviation of the lengths at the first age class.
S	The standard deviation of the lengths at the last age class.
K	Fraction parameter in the von Bertalanffy growth model: e.g., $\mu_a = L_\infty \cdot (1 - \exp(-K \cdot (a - t_0)))$.
k	Brody coefficient; $k = \exp(-K)$, where K is the above fraction parameter in the von Bertalanffy growth model.
P_a	Proportion of age- a fish in a total of age classes: i.e., $\sum_{a=1}^M P_a = 1$.
A	The objective function used in the numerical optimization (eq. 5).
f_j	The observed frequency (fish count) that belongs to the j th length interval. And, $f = \sum_{j=1}^N f_j$, where N is the last length interval.
B_m	χ^2 statistic in eq. 6.
E_m	Degrees of freedom in the χ^2 statistic in eq. 6.

Table 2. Estimates of parameters and their respective standard errors under the vB-based method. Values inside parentheses are the standard errors of parameter point estimates. A indicates the value of the objective function (Eq. 5). I and II represent the first half and the second half of each year. Because $P_{5+} = 1 - \sum_{a=1}^4 P_{a^+}$, we don't show the estimate of P_{5+} . See Table 1 for the other notations

		parameters									A
		l	L	s	S	k	P_1	P_2	P_3	P_4	
2000	I	20.90 (0.20)	37.1 (1.20)	1.16 (0.12)	3.88 (0.16)	0.9 (0.08)	0.01 (0.00)	0.06 (0.02)	0.81 (0.04)	0.11 (0.01)	82.92
	II	17.79 (0.21)	31.6 (0.15)	1.38 (0.10)	2.07 (0.07)	0.92 (0.02)	0.07 (0.01)	0.17 (0.01)	0.22 (0.01)	0.03 (0.04)	119.57
2001	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	15.34 (0.07)	39.66 (0.24)	1.04 (0.05)	4.71 (0.09)	0.88 (0.01)	0.03 (0.00)	0.11 (0.01)	0.75 (0.01)	0.00 (0.00)	394.18
2002	I	15.51 (0.06)	30.94 (0.18)	1.08 (0.03)	2.46 (0.08)	0.58 (0.01)	0.06 (0.00)	0.27 (0.01)	0.43 (0.01)	0.00 (0.00)	107.72
	II	21.57 (0.11)	33.44 (0.08)	1.98 (0.06)	2.16 (0.04)	0.46 (0.01)	0.04 (0.00)	0.42 (0.01)	0.00 (0.00)	0.10 (0.01)	89.27
2003	I	24.28 (0.15)	34.54 (0.43)	1.65 (0.07)	2.63 (0.26)	0.59 (0.02)	0.11 (0.01)	0.54 (0.02)	0.00 (0.00)	0.30 (0.13)	17.67
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2004	I	19.16 (0.10)	38.02 (0.16)	1.82 (0.05)	2.00 (0.09)	0.78 (0.01)	0.06 (0.00)	0.29 (0.01)	0.47 (0.01)	0.09 (0.01)	93.82
	II	23.08 (0.06)	32.03 (0.22)	1.42 (0.03)	2.71 (0.09)	0.54 (0.01)	0.18 (0.01)	0.38 (0.02)	0.00 (0.00)	0.28 (0.06)	59.53
2005	I	25.97 (0.18)	39.13 (0.28)	2.15 (0.07)	2.60 (0.18)	0.75 (0.02)	0.17 (0.01)	0.80 (0.01)	0.00 (0.00)	0.00 (0.00)	99.19
	II	25.26 (0.01)	34.87 (0.40)	1.60 (0.06)	1.77 (0.14)	0.61 (0.04)	0.13 (0.01)	0.17 (0.03)	0.62 (0.03)	0.00 (0.00)	20.88
2006	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	24.55 (0.42)	34.40 (0.56)	2.17 (0.17)	2.23 (0.15)	0.63 (0.04)	0.06 (0.01)	0.50 (0.03)	0.00 (0.00)	0.39 (0.13)	39.49
2007	I	19.33 (0.57)	37.44 (0.41)	2.58 (0.15)	3.25 (0.12)	0.77 (0.04)	0.00 (0.00)	0.20 (0.00)	0.75 (0.01)	0.00 (0.00)	109.41
	II	13.96 (0.94)	37.71 (0.49)	0.84 (0.26)	3.30 (0.22)	0.83 (0.04)	0.00 (0.00)	0.01 (0.00)	0.60 (0.02)	0.39 (0.02)	35.00
2008	I	16.16 (0.07)	35.25 (0.39)	2.11 (0.05)	2.93 (0.16)	0.55 (0.02)	0.12 (0.00)	0.21 (0.01)	0.41 (0.03)	0.08 (0.05)	178.03
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2009	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	20.63 (0.08)	35.71 (0.15)	1.41 (0.04)	2.49 (0.07)	0.63 (0.01)	0.02 (0.00)	0.21 (0.01)	0.76 (0.01)	0.00 (0.00)	73.72
2010	I	23.80 (0.16)	43.12 (0.38)	1.78 (0.07)	2.69 (0.18)	0.94 (0.02)	0.06 (0.01)	0.61 (0.01)	0.23 (0.01)	0.06 (0.01)	128.76
	II	24.50 (0.08)	32.62 (0.30)	1.38 (0.03)	2.43 (0.08)	0.82 (0.04)	0.31 (0.02)	0.28 (0.01)	0.00 (0.00)	0.31 (0.05)	69.36

Table 2. Continued

		parameters									A
		<i>l</i>	<i>L</i>	<i>s</i>	<i>S</i>	<i>k</i>	P_1	P_2	P_3	P_4	
2011	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	22.17 (0.12)	34.05 (0.11)	1.42 (0.05)	1.72 (0.05)	0.9 (0.02)	0.03 (0.00)	0.31 (0.01)	0.19 (0.01)	0.21 (0.01)	101.43
2012	I	19.67 (0.09)	32.4 (0.26)	0.86 (0.05)	3.8 (0.09)	0.5 (0.02)	0.01 (0.00)	0.37 (0.01)	0.21 (0.03)	0.36 (0.03)	167.73
	II	12.85 (2.21)	32.48 (0.04)	1.20 (0.11)	2.48 (0.03)	0.71 (0.06)	0.00 (0.00)	0.00 (0.00)	0.21 (0.01)	0.00 (0.00)	103.14
2013	I	15.00 (0.00)	33.03 (0.18)	0.61 (0.12)	4.84 (0.21)	0.46 (0.01)	0.00 (0.00)	0.20 (0.02)	0.59 (0.02)	0.13 (0.03)	107.31
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2014	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	22.83 (0.24)	32.25 (0.51)	1.26 (0.10)	2.87 (0.07)	0.54 (0.04)	0.01 (0.00)	0.28 (0.04)	0.00 (0.00)	0.14 (0.24)	58.42
2015	I	21.85 (0.29)	35.32 (1.60)	0.99 (0.20)	6.45 (0.60)	0.50 (0.08)	0.01 (0.00)	0.88 (0.02)	0.11 (0.02)	0.00 (0.00)	59.65
	II	23.00 (0.17)	36.69 (0.23)	1.19 (0.05)	2.59 (0.11)	0.68 (0.02)	0.01 (0.00)	0.36 (0.01)	0.55 (0.01)	0.06 (0.01)	224.72
2016	I	20.9 (0.20)	37.1 (1.20)	1.16 (0.12)	3.88 (0.16)	0.9 (0.08)	0.01 (0.00)	0.06 (0.02)	0.81 (0.04)	0.11 (0.01)	168.05
	II	18.64 (0.15)	38.93 (0.28)	1.56 (0.07)	2.85 (0.09)	0.83 (0.01)	0.01 (0.00)	0.06 (0.00)	0.79 (0.01)	0.15 (0.01)	232.19

각각 상반기와 하반기를 의미)에서 k 값이 1.0 이상으로 추정되었고, 2002-II, 2003-I, 2007-I, 2013-II에서는 s 가 S 보다 크게 추정되었으며, 2008-I와 2016-I는 최적화(optimization)에 실패하였다. Eq. (1)의 k 가 $K = -\log(k)$ 과 같이 von Bertalanffy 성장식의 성장 모수 k 를 재모수화 한 것이므로 k 의 값이 항상 0과 1 사이에 위치하여야 한다는 것은 자명하다. 따라서 k 가 1.0 이상으로 추정되는 경우는 각 모수의 추정치가 계산되었다 하더라도 실패한 것으로 간주하였다. $\hat{s} > \hat{S}$ 경우는 보다 명확하다. Eq. (2)가 연령이 증가함에 따라 체장의 표준편차가 함께 증가한다는 가정을 내포하고 있다는 것을 상기하면, s 가 S 보다 항상 작게 추정되어야 한다는 것을 쉽게 알 수 있다. 모수 추정치간의 상관관계는 자료에 따라 큰 편차를 보였으나, k 와 L 은 거의 모든 자료에서 높은 양의 상관관계를 보였다.

von Bertalanffy에 얽매이지 않는 체장빈도분석법 (vB-free)

총 16개의 자료에서 vB-based 방법으로 추정되지 않은 자료가 vB-free에 의해 추정되거나, 반대로 vB-based 방법으로 추정된 자료가 vB-free로는 추정되지 않는 결과를 보였다(Fig. 1-3). 총 34개의 반기별 체장 자료 중 22개의 자료에서 vB-free의 모

수가 성공적으로 추정되었으며, 각 모수의 추정치와 표준오차는 Table 3에 정리하였다[모수 추정치간의 상관관계는 Hyun et al. (2017) 보고서를 참고]. 나머지 12개 자료의 모수 추정 실패 요인을 살펴보면 2000-II, 2005-II, 2006-II, 2008-II, 2013-II, 2014-II에서 s 가 S 보다 크게 추정되었고, 2010-I, 2016-II는 3세의 평균 체장이 4세보다 크게 추정되었으며, 2010-II, 2012-I, 2015-II, 2016-I에서는 최적화에 실패하였다.

모수추정치간의 상관관계는 vB-based의 결과처럼 모델에 이용된 자료에 따라 큰 편차를 보였으나, 2세와 3세, 3세와 4세 그룹의 평균 체장은 거의 모든 자료에서 큰 양의 상관관계를 보였다.

모델 적합도 검증 및 비교

vB-based와 vB-free의 추정 결과를 비교·검증하기 위해 두 모델에서 모두 모수 추정이 가능했던 15개 자료 결과의 P 값을 비교하였다(Table 4). 전체적으로 M 값이 상당히 작기 때문에 두 모델의 적합도 차이가 크다고 볼 수는 없으나, vB-free가 vB-based에 비해 2개의 모수가 더 많음에도 불구하고, 총 15개의 자료 중 13개의 자료에서 더 높은 적합도를 보였다.

Table 3. Continued

		Parameters											A
		μ_1	μ_2	μ_3	μ_4	μ_5	s	S	P_1	P_2	P_3	P_4	
2011	I	17.56 (0.05)	26.08 (0.11)	29.58 (0.18)	34.43 (0.39)	44.14 (0.88)	1.42 (0.04)	3.18 (0.18)	0.09 (0.00)	0.32 (0.03)	0.50 (0.02)	0.09 (0.01)	71.01
	II	20.47 (0.59)	22.76 (0.39)	25.89 (0.07)	29.95 (0.15)	33.67 (0.10)	1.21 (0.16)	1.79 (0.04)	0.00 (0.00)	0.04 (0.01)	0.34 (0.01)	0.28 (0.01)	80.23
2012	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	21.15 (0.22)	24.87 (0.06)	28.08 (0.10)	32.12 (0.06)	34.99 (0.13)	1.01 (0.06)	1.81 (0.05)	0.00 (0.00)	0.12 (0.00)	0.17 (0.00)	0.49 (0.01)	45.46
2013	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2014	I	26.06 (0.13)	28.07 (0.11)	31.45 (0.06)	35.66 (0.10)	41.09 (0.51)	1.09 (0.05)	1.48 (0.07)	0.13 (0.02)	0.38 (0.02)	0.40 (0.01)	0.08 (0.00)	81.44
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2015	I	22.00 (0.36)	26.73 (0.46)	29.33 (0.26)	34.38 (0.72)	41.15 (2.55)	1.00 (0.21)	3.57 (0.40)	0.01 (0.00)	0.08 (0.08)	0.84 (0.08)	0.07 (0.01)	57.45
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2016	I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

고찰

체장빈도분석법(Length frequency analysis)

vB-based와 vB-free 모두 Schnute and Fournier LFA의 기본 전제 조건인 ‘연령별 체장의 분포가 정규분포(a normal probability distribution)를 따른다’는 것을 전제로 하고 있지만, 이러한 가정은 어구선택성 등 표본 추출 편향(sampling bias)가 가미된 체장 빈도 자료에는 부적합하며 이는 수치적 최적화(numerical optimization) 방법의 실패와도 연결된다. 즉, Schnute and Fournier의 LFA의 방법은 체장 빈도 자료가 M 개의 정규분포로 설명될 수 없을 경우 추정이 되지 않는데, 특히 본 연구에서 이용된 고등어의 체장 자료와 같이 과학조사(survey)가 아닌 상업어업(commercial fisheries)에서 수집된 자료는 무작위 추출(random selection)을 통한 자료 수집이 이루어지지 않았기 때문에 이러한 정규분포 가정을 충족하지 않을 가능성은 더욱 높아진다. 고등어 자료의 반기별 체장 분포를 살펴보면 1세 그룹으로 추정되는 20 cm 미만의 개체들의 출현이 다소 불규칙적인 것을 알 수 있다(Fig. 1-3). 자료의 수집 위치와 어구 선택성 등을 알 수 없는 상황에서 1세 그룹의 불규칙적 출현 원인을 규명할 수는 없으나, 고등어의 연령별 공간적 분포의 차이와 고등어의 어획 금지 체장으로 인한 어구 선택성의 변화 등을 주된 요인으로 추측할 수 있다. 따라서, 수집된 자료가 고등어의 연령별 체장 분포를 충분히 대표하고 있지 않을 가능성이 높으며, 선

망이라는 어구의 특성과 고등어의 어획 금지 체장의 설정은 특히 1세 또는 2세 그룹의 체장 분포를 정규 분포가 아닌 다소 왜곡된 형태를 가지도록 하였을 것이다.

M 개의 정규분포(연령 그룹별 체장 분포)가 특정 비율로 합쳐져서 하나의 혼합된 분포를 이루며, 이러한 기본 원리를 이용하여 어류의 체장 조성 자료를 분석하는 것이 Schnute and Fournier LFA의 기본 틀이다. 즉, M 개의 정규분포가 서로 겹치는 비율이 높을수록 모수 추정치가 큰 오차를 수반하게 되는데, 분석에 이용된 체장 자료의 분포가 단봉형(unimodal)에 가깝고 빈도(frequency)를 가지는 계급의 수가 작을수록 이러한 문제점에 큰 영향을 받는 것으로 보인다. 본 연구에서 이용한 반기별 고등어의 체장 분포 대부분은 이러한 단봉형의 형태를 보였는데, vB-based와 vB-free 모두 모수 추정치간의 상관관계가 대체적으로 크게 나타난 것이 이러한 사실을 뒷받침 한다(Hyun et al., 2017). 상반되는 예로서, 2000-I, 2002-I, 2011-I의 체장빈도수 분포의 경우, 다봉형(multi-modal)을 보였고 vB-based와 vB-free에서 모두 모수 추정이 잘 이루어졌다. 특히, 두 방법 모두 P_1 (1세 개체수의 비율)의 추정치와 다른 모수 추정치와의 상관관계가 거의 없는 것으로 나타났다.

Schnute and Fournier의 LFA에서 eq. (1)은 M 개의 정규분포의 평균, eq. (2)은 M 개의 정규분포의 표준편차를 의미한다. eq. (1)로 인한 제약(i.e., 어류의 연령별 평균 체장이 vB 모델을 따르기 때문에 μ_a 가 l, L, k 에 의해 결정된다)은 μ_a 를 자유 모수로 이용한 vB-free를 통해 해결할 수 있었으나, 표준편차의 경

Table 4. Comparison of vB-based (“von” below) and vB-free (“free” below) in χ^2 test, based on data from 15 years, with which these two methods commonly resulted in the successful estimation of parameters. To help readers to quickly compare in P-values of the χ^2 test between “von” and “free” methods, we used bold font for the larger P-value. I and II represent the first half and the second half of each year. See Table 1 for the other notations

		B_2	B_3	B_4	B_5	E_2	E_3	E_4	E_5	P-value of the χ^2 test			
										$E_m = 2$	$E_m = 3$	$E_m = 4$	$E_m = 5$
2000-I	von	78.91	77.73	77.73	77.73	14	13	13	13	0.05×10 ⁻⁹	0.03×10 ⁻⁹	0.03×10 ⁻⁹	0.03×10 ⁻⁹
	free	63.44	61.04	61.04	61.04	12	11	11	11	5.30×10⁻⁹	5.94×10⁻⁹	5.94×10⁻⁹	5.94×10⁻⁹
2001-II	von	351.45	351.45	351.45	351.45	24	24	24	24	6.36×10 ⁻⁶⁰	6.36×10 ⁻⁶⁰	6.36×10 ⁻⁶⁰	6.36×10 ⁻⁶⁰
	free	135.56	135.56	135.56	135.56	22	22	22	22	2.41×10⁻¹⁸	2.41×10⁻¹⁸	2.41×10⁻¹⁸	2.41×10⁻¹⁸
2002-I	von	105.91	104.51	104.14	104.14	18	17	16	16	0.02×10 ⁻¹²	0.01×10 ⁻¹²	5.76×10 ⁻¹⁵	5.76×10 ⁻¹⁵
	free	96.54	95.46	95.14	95.14	16	15	14	14	0.15×10⁻¹²	0.09×10⁻¹²	0.04×10⁻¹²	0.04×10⁻¹²
2002-II	von	97.48	97.48	97.48	97.48	15	15	15	15	0.04×10 ⁻¹²	0.04×10 ⁻¹²	0.04×10 ⁻¹²	0.04×10 ⁻¹²
	free	42.70	42.70	42.70	42.70	13	13	13	13	0.05×10⁻³	0.05×10⁻³	0.05×10⁻³	0.05×10⁻³
2003-I	von	19.25	19.25	19.25	16.76	14	14	14	13	0.16	0.16	0.16	0.21
	free	18.33	18.33	18.33	16.06	12	12	12	11	0.11	0.11	0.11	0.14
2004-I	von	77.71	77.36	77.36	77.36	19	18	18	18	4.60×10 ⁻⁹	2.47×10 ⁻⁹	2.47×10 ⁻⁹	2.47×10 ⁻⁹
	free	47.96	42.93	42.93	42.93	17	16	16	16	0.09×10⁻³	0.29×10⁻³	0.29×10⁻³	0.29×10⁻³
2004-II	von	60.92	60.92	60.92	57.73	14	14	14	12	0.08×10 ⁻⁶	0.08×10 ⁻⁶	0.08×10 ⁻⁶	0.06×10 ⁻⁶
	free	22.05	22.05	22.05	20.37	12	12	12	10	0.04	0.04	0.04	0.03
2005-I	von	93.33	93.33	93.16	93.16	15	15	14	14	0.24×10 ⁻¹²	0.24×10 ⁻¹²	0.10×10 ⁻¹²	0.10×10 ⁻¹²
	free	65.64	65.64	65.35	65.35	13	13	12	12	5.05×10⁻⁹	5.05×10⁻⁹	2.35×10⁻⁹	2.35×10⁻⁹
2007-I	von	102.64	101.30	100.78	100.78	19	17	16	16	0.18×10 ⁻¹²	0.05×10 ⁻¹²	0.03×10 ⁻¹²	0.03×10 ⁻¹²
	free	89.83	88.57	88.57	88.57	17	15	14	14	6.57×10⁻¹²	1.83×10⁻¹²	0.71×10⁻¹²	0.71×10⁻¹²
2007-II	von	30.41	30.41	27.68	27.68	15	15	13	13	0.01	0.01	0.01	0.01
	free	26.97	26.97	22.22	22.22	13	13	11	11	0.01	0.01	0.02	0.02
2008-I	von	166.37	166.37	159.79	158.57	25	25	24	23	0.08×10 ⁻²¹	0.08×10 ⁻²¹	0.49×10 ⁻²¹	0.31×10 ⁻²¹
	free	121.79	121.79	119.69	118.65	23	23	22	21	1.98×10⁻¹⁵	1.98×10⁻¹⁵	1.99×10⁻¹⁵	1.26×10⁻¹⁵
2009-II	von	154.75	151.16	84.85	84.85	19	18	16	16	0.03×10 ⁻²¹	0.04×10 ⁻²¹	0.02×10 ⁻⁹	0.02×10 ⁻⁹
	free	69.53	67.95	41.17	41.17	17	16	14	14	0.03×10⁻⁶	2.27×10⁻³	0.17×10⁻³	0.17×10⁻³
2011-II	von	91.90	89.23	89.23	82.02	15	13	13	12	0.44×10 ⁻¹²	0.20×10 ⁻¹²	0.20×10 ⁻¹²	1.70×10 ⁻¹²
	free	66.82	65.06	65.06	59.89	13	11	11	10	3.08×10⁻⁹	1.05×10⁻⁹	1.05×10⁻⁹	3.80×10⁻⁹
2012-II	von	100.75	99.52	99.52	97.71	15	14	14	12	9.40×10 ⁻¹⁵	5.87×10 ⁻¹⁵	5.87×10 ⁻¹⁵	1.56×10 ⁻¹⁵
	free	51.25	49.95	49.95	39.97	13	12	12	10	1.82×10⁻⁶	1.42×10⁻⁶	1.42×10⁻⁶	17.2×10⁻⁶
2015-I	von	53.58	52.82	52.32	52.32	14	13	12	12	1.52×10⁻⁶	0.97×10⁻⁶	0.54×10⁻⁶	0.54×10⁻⁶
	free	52.67	52.08	51.69	51.69	12	11	10	10	0.47×10 ⁻⁶	0.26×10 ⁻⁶	0.13×10 ⁻⁶	0.13×10 ⁻⁶

우 과모수화(over-parameterization)를 염려하여 vB-based와 vB-free 모두 eq. (2)를 이용하였다. 따라서, eq. (2)의 가정이 충족되지 않을 경우 모수가 추정 되었다 하더라도 실패 한 것으로 간주 하였는데 vB-based와 vB-free 모두 s 가 S 보다 크게 추정되는 경우가 모수 추정 실패 원인의 약 50%를 차지하였다 (vB-based에서 4/9; vB-free에서 6/12).

어류의 연령이 증가함에 따라 연령별 체장의 표준편차가 증가한다는 것은 타당한 가정이다. 그러나 어구선택성과 공간적 분

포의 차이로 인해 표본추출 편이가 발생하였다면 연령별 체장의 표준편차가 eq. (2)의 가정을 따르지 않을 가능성이 있다. 실제로, 대부분의 고등어 체장 분포 자료에서 특정 계급의 빈도수가 상당히 높게 나타나는 현상이 관찰되었다. 각 계급의 구간 넓이가 1 cm라는 점을 고려하면 체장 선택성의 정도가 상당히 크다는 것을 알 수 있는데, 이로 인해 그 해당 계급을 주로 포함하는 연령 그룹의 체장의 표준편차가 과소 추정될 수 있는 것이다. 고등어의 공간적 분포의 경우, 샘플의 어획 장소(즉, 공간적

분포)에 대한 정보가 누락되어 있으므로 그 영향의 유무를 단정할 수는 없으나, Choi (2003)에 따르면 고등어의 어장 분포가 시기별로 큰 차이를 보인다. 즉, 어구 선택성과 더불어 고등어의 생활사에 따른 회유 패턴이 수집된 자료의 체장 분포에 영향을 끼쳤을 가능성 역시 배제할 수 없다.

모델 적합도 검증 및 비교

χ^2 테스트를 통해 vB-free의 방법이 vB-based에 비해 더 우수하다는 결과를 얻었으나 그 차이는 미미하였다(Table 4). 또한, vB-free방법이 vB-based방법에 비해 2개 더 많은 모수를 갖고 있기에, 그만큼 모수 추정도 쉽지 않다. 실제로 vB-free가 vB-based에 비해 모수 추정을 실패한 경우가 더 많았으며 (vB-based에서 25/34; vB-free에서 22/34), 각 모수의 적절한 초기값을 지정하는 것에도 더 많은 어려움이 있었다. 그러나 두 모델의 연령 조성 추정치를 생물학적 관점에서 비교하였을 때 vB-free가 vB-based보다 현실적인 추정치를 도출하였다(Hyun et al., 2017). 예를 들면, 2002-II와 2003-I의 vB-based 연령 조성 추정치는 3세가 거의 출현하지 않는데, 2세와 4세 개체들이 확실히 존재함에도 3세 개체들이 출현하지 않는다는 것은 생물학적 관점에서 이해하기 어려운 현상이다.

일반적으로 vB-free가 vB-based에 비해 모수 추정의 어려움은 있으나, 고등어와 같이 수명이 짧은 어종을 대상으로 할 경우 이러한 문제가 큰 단점으로 부각되지는 않는다. 또한, vB-free가 vB-based 보다 적합도면에서(미미한 차이일지라도) 우수할 뿐만 아니라 생태학적 관점에서 더욱 타당한 결과치를 보였기 때문에 수명이 짧은 어종에는 vB-free를 이용하여 체장 조성 자료를 분석해야 할 것이다. 다만, 어종에 따라 성장 패턴과 형태가 다르기 때문에 본 연구 결과를 바탕으로 체장 빈도 분석법의 적용 방식을 일반화 하기는 어렵다. 따라서, 이러한 일반화를 위해서는 어종의 형태, 성장 패턴, 수명 등을 바탕으로 대표 어종을 선별한 다음 vB-based와 vB-free 방법을 비교하는 후속 연구가 수행되어야 할 것이다.

사 사

이 논문은 2018년도 국립수산물과학원 연근해어업 자원평가 및 관리연구 (R2018024)의 지원으로 수행되었습니다.

References

Choi Y. 2003. Stock assessment and management implications of chub mackerel, *Scomber japonicus* in Korean waters. Ph.D. Thesis, Pukyong National University, Busan, Korea.
 Ford E. 1933. An account of the herring investigations conducted at Plymouth during the years from 1924 to 1933. J Mar Biol Assoc UK 19, 305-384. <http://dx.doi.org/10.1017/S0025315400055910>.
 Fournier D and Archibald CP. 1982. A general theory for analyzing

catch at age data. Can J Fish Aquat Sci 39, 1195-1207. <http://dx.doi.org/10.1139/f82-157>.
 Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder MN, Nielsen A and Sibert J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim Methods Software 27, 233-249. <http://dx.doi.org/10.1080/10556788.2011.597854>.
 Hyun S-Y, Kim K, Lee H, Choi S, Kang M, Gim J and Jung YR. 2017. A study on fish stock assessments and the applicaiton package under the Korean situation. National Institute of Fisheries Science, Busan, Korea, 91.
 Macdonald P and Pitcher T. 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. J Fish Res Board Can 36, 987-1001. <http://dx.doi.org/10.1139/f79-137>.
 Method RD and Wetzel CR. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish Res 142, 86-99. <http://dx.doi.org/10.1016/j.fishres.2012.10.012>.
 Pauly D and Morgan G. 1987. Length-based methods in fisheries research. In: Proceedings of the International Conference on the Theory and Application of Length-Based Methods for Stock Assesment. Pauly D and Morgan G, eds. The International Center for Living Aquatic Resources Management, Manila, Philippines, 137-146.
 Ricker WE. 1975. Computation and interpretation of biological statistics of fish populations. B Fish Res Board Can 191, 1-382.
 Schnute J and Fournier D. 1980. A new approach to length-frequency analysis: growth structure. Can J Fish Aquat Sci 37, 1337-1351. <http://dx.doi.org/10.1139/f80-172>.
 von Bertalanffy L. 1938. A quantitative theory of organic growth (inquiries on growth laws. II). Hum Biol 10, 181-213.
 Walford LA. 1946. A new graphic method of describing the growth of animals. Biol Bull 90, 141-147.