이학석사 학위논문

한국 고등어(*Scomber japonicus*) 자원 평가를 위한 Stock Synthesis 소프트웨어의 체장 기반 모델 적용

2021년 8월

부경대학교 대학원

해양생물학과

이 승 준

이학석사 학위논문

한국 고등어(*Scomber japonicus*) 자원 평가를 위한 Stock Synthesis 소프트웨어의 체장 기반 모델 적용

지도교수 현 상 윤

이 논문을 이학석사 학위논문으로 제출함.

2021 년 8 월

부경대학교 대학원

해양생물학과

이 승 준

이승준의 이학석사 학위논문을 인준함.

2021 년 8 월 27 일

위 원 장 이학박사 오 철 웅 (인)

위 원 이학박사 현 상 윤 (인)

위 원 이학박사 강 희 중 (인)

목차

그림목차

표목차

Application of a length-based module of Stock Synthesis software for assessment of Korea chub mackerel (*Scomber japonicus*)

Seung Joon Lee

Department of Marine Biology, The Graduate School,
Pukyong National University

Abstract

Stock Synthesis (SS) is one of the fishery stock assessment software packages designated as the standard package of the US NOAA Fisheries. SS started to be developed in 1983 by Methot, and is implemented in Auto-Differentiation Model Builder (ADMB), which is a numerical optimization software language for estimating multiple parameter in a non-linear model. Although SS has been applied in a wide variety of fish assessments, there is no case of applying SS to a fish stock assessment in Korea. The main objective of this study was to apply the size-structured model in SS for assessment of Korea chub mackerel (*Scomber japonicus*). In addition, I performed the model validation through a simulation study and retrospective error analysis. The data used in this study were body sizes from 2000-2019, and fishery total catch (MT), and catch-per-unit-effort (CPUE in MT/haul) from 1975-2019. To perform a retrospective error analysis, I considered two cases. Case 1 (C1) was assumed to use all available data, and Case 2 (C2) was assumed to use data from 2000 to 2019, the period when all data were simultaneously available. The results of estimates in two cases showed significant differences in the estimates of growth parameters and annual recruitment deviations. While the model appeared to fit the data better in C1 than in C2, the simulation study showed

that the model was more stable in C2 in terms of numerical convergence than in C1. The model did not represent a retrospective pattern, and the Mohn's ρ values were also found to have negligible retrospective errors. The estimated annual stock spawning biomass in both cases showed a similar trend, and was found to increase gradually over time.

I. 서론

Stock Synthesis (SS)는 현재, 미국 해양대기국(National Oceanic and Atmospheric Administration; NOAA) 수산청의 표준 자원 평가 소프트웨어 중 하나이며, 연령 및 체장 자료를 기반으로 통합 분석(Maunder and Punt, 2013) 접근법을 고려한 연령구조 모델이다. Methot은 통합 분석 접근법의 시초인 Fournier and Archibald (1982)의 많은 특징들을 채택하여 1980년대 초(Methot, 1986, 1989)에 주도적으로 SS를 구현하기 시작했으며, 현재에도 꾸준히 업데이트가 진행되고 있다(Methot et al., 2013). SS는 성장모델, 산란-가입 모델, 어구 선택성 모델 등을 고려하고 있으며, 대상 개체군의 생물학적 특징 및 가용한 자료를 바탕으로 다양한 모델 옵션(예: 산란-가입 모델의 경우 Ricker model, Beverton-Holt model, Hockey stick model, etc.)을 선택할 수 있도록 구성되어 있다. 한편, SS는 ADMB (Automatic Differentiation Model Builder, Fournier et al., 2012) 언어로 구현되어 있다. 따라서, 다소 복잡한 비선형 통계 모델의 모수를 추정하는 경우에도 수치 최적화가 효과적으로 수행된다는 특징이 있다(Fournier et al., 2012). 또한, 결과를 효율적으로 진단할 수 있도록 R 소프트웨어 패키지 'r4ss'가 개발되어 있어 연구를 효율적으로 수행할 수 있다.

오늘날 개체군의 자원량 크기를 추정하기 위해서는 일반적으로 연령

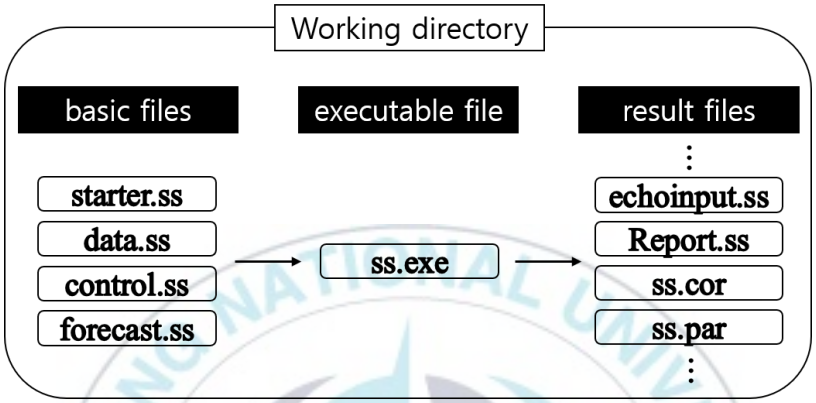
구조 모델을 적용한다. 연령 구조 모델은 연령 자료가 필수적으로 요구되지만, 연령 자료가 제한적인 경우 대안으로 체장-기반 모델을 적용할 수 있다. 최근 국외에서는 SS의 체장-기반 모델을 활용한 연구들(Szuwalski, 2016, Sharma et al., 2014)이 꾸준히 수행되고 있다. 반면, 국내의 경우에는 체장-기반 모델을 적용하여 자원 평가를 수행한 연구들(Gim, 2019, Park, 2021)이 수행되고 있지만, 현재까지 국내에서 SS를 적용한 연구 사례는 아직 없는 것으로 알려져 있다. 따라서, 본 논문은 두 가지 목적을 가지고 수행되었다. 첫 번째로는, SS의 구성과 구현 방법에 대해 소개함으로써 앞으로 SS를 활용하여 국내 어종에 대한 자원 평가를 수행하고자 하는 연구자들에게 보다 쉽게 SS를 적용할 수 있도록 하는 것이다. 두번째로는, 한국 고등어(*Scomber japonicus*)의 체장 조성 자료를 기반으로 한 SS의 체장-기반 모델을 적용하여 한국 고등어 자원을 평가하는 것이다. 이외에도, 본 저자가 구성한 SS 모델 프레임워크의 성능을 평가하기 위해 모델 검증을 수행했으며, 이를 위해 시뮬레이션 연구와 레트로스펙티브 오차 분석(retrospective error analysis)을 수행했다. SS를 활용하고자 하는 독자 및 연구원들의 이해를 돕기 위해 한국 고등어를 대상으로 실제 연구에 사용된 SS 기본 파일을 부록에 제시했다.

II. 재료 및 방법

II. 1. SS의 구성 및 구현

본 논문은 한국 고등어의 체장 조성 자료를 기반으로 한 SS의 체장-기반 모델을 적용했다. NOAA 수산청에서 관리하고 있는 SS 공식 홈페이지(https://vlab.ncep.noaa.gov/web/stock-synthesis)에서 초보자들에게 제공하고 있는 네 가지 예제 파일(starter.ss, data.ss, control.ss, forecast.ss)을 토대로, 주어진 자료 및 한국 고등어의 특성에 맞게 이를 수정했다. 제공되는 예제 파일은 연령 및 체장 자료를 기반으로 한 연령 구조 모델에 대한 예제로 구성되어 있기 때문에, 이를 체장-기반 모델로 고려하도록 파일들을 수정하는 과정을 거쳤다. SS의 전반적인 구현 과정의 모식도를 Figure 1에 제시했다.

Figure 1. SS의 구성 요소 및 실행 과정 모식도. 작업 폴더 안에 네 가지 기본적인 파일(starter.ss, data.ss, control.ss, forecast.ss)이 실행 파일(ss.exe)에 의해 성공적으로 실행되면 여러 가지 결과 파일(echoinput.ss, Report.ss, ss.cor, ss.par, etc.)이 생성 된다.



기본적으로 필요한 네 가지 파일 중 (1). starter.ss 파일은 data.ss와 control.ss 파일의 이름을 입력하여 실행파일(ss.exe)이 파일을 순서대로 읽을 수 있도록 하고, 출력할 결과를 자세하게, 또는 간소화하게 제시할지를 선택하는 기능을 가진다. 시뮬레이션 연구와 같이 다소 시간이 오래 걸리는 경우에는 결과를 간소하게 제시하여 시간을 단축하였고, 결과를 검토하기 위한 경우에는 결과를 자세히 제시했다. (2). data.ss 파일은 어업(fishery) 또는 조사(survey)로부터 얻은 시계열 자료(예: CPUE 자료, 어획 노력량 자료, 체장-조성 자료, 연령-조성 자료 및 체중-조성 자료, tag-recapture 자료 등), 자료의 단위 및 수집 시기와 기간에 대한 정보를 다루며, 각 코호트의 시간 단위, 산란 시기와 최대 연령과 같은 생물학적 정보, 그리고

각 자료에 가정할 확률 분포 및 분포의 분산에 대한 정보를 다룬다. data.ss 파일과 관련된 내용은 **II. 2. 자료, II. 3. 가정** 부분에서 자세하게 다뤘다. (3). control.ss 파일은 주로 모수 추정 및 모델 선택과 관련이 있다. 각 모델 (예: 성장 모델, 산란−가입 모델, 어구 선택성 모델 등)의 추정할 모수들의 하한(lower bound), 상한(upper bound), 그리고 초깃값(initial guess value)에 대한 정보를 입력하며, 모수를 자유 모수로써 추정할지, 추정하지 않고 입력 값(input value)으로 고려할지에 대한 내용을 다룬다. SS는 모델별로 고려할 수 있는 다양하 모델 옵션을 제시하고 있기 때문에, 가용한 자료와 한국 고등어의 특징 등을 반영하여 적절한 모델 옵션을 선택했다. control.ss 파일과 관련된 내용은 **II. 3. 다. 모델 가정** 부분에서 자세하게 다뤘다. forecast.ss 파일은 자료 이후의 기간 즉, 미래 예보(projection)와 관련 되어 있다. 수산 자원 관리 및 평가의 기준이 되는 지표인 참조점 (reference point), 예를 들어 MSY, F_{MSY}, B_{MSY}, SPR 등과 관련된 옵션을 설정하며, 미래 예보와 관련된 기준점(benchmark)에 대한 옵션을 설정한다. 본 논문에서는, 미래 예보를 고려하지 않았기 때문에 SS에서 제공되는 예제 파일에서 일부분만을 수정했다. 본 연구에서 사용된 네 가지 파일은 **부록 3**에 제시하였다.

실행 파일(ss.exe)은 네 가지 파일을 순서대로 읽은 후 결과 파일을 생성하는 역할을 수행하며, 실행 파일의 구성을 파악하기 위해 SS 홈페이지에서 제공하는 ss.tpl (templet) 파일을 참고했다. ss.tpl 파일은 크게 자료에

대한 부분(DATA_SECTION), 모수에 대한 부분(PARAMETER_SECTION), 그리고 실제 모델 계산(PROCEDURE_SECTION)에 대한 부분으로 나누어져 있기 때문에, SS 개발팀이 자료를 어떻게 다루고 모델을 구성하여 결과를 제시하는지 세부적으로 파악할 수 있다. 실행 파일을 실행하기 위한 방법으로는 ADMB 홈페이지(https://www.admb-project.org/)에서 제공하는 ADMB 명령 프롬프트를 활용하는 방법, 작업 공간 내에서 PowerShell 창을 열어 실행하는 방법, 마지막으로는 R software 환경에서 실행 파일을 실행하는 방법이 있다. 세 가지 방법 모두 네 가지 파일을 읽은 후 결과 파일을 생성한다는 공통점이 있지만, 본 논문에서는 이 세 가지 방법 중 R software 환경에서 SS를 구현하는 방법을 적용했다. R software 환경에서 SS를 구현할 경우, R software 패키지인 'r4ss'를 활용할 수 있으며, SS를 구현할 때 생성되는 결과들 중에서 중요한 결과들 만을 간추려 제시해주기 때문에 추정치를 신속하게 진단할 수 있는 장점이 있다.

기본적인 네 가지 파일을 구성하고 이를 실행 파일로 구현하면 작업 공간에 결과 파일들이 생성되며, 대표적으로는 echoinput.ss, Report.ss, ss.cor. ss.par 파일이 있다. echoinput.ss 파일은 자료들이 제대로 입력되었는지를 검토하고, SS를 구현하는 과정에서 발생하는 오류를 찾는 기능을 한다. starter.ss 파일부터 forecast.ss 파일까지 입력한 자료와 모델 옵션들을 순서대로 제시해주기 때문에 오류가 발생한 경우 이 지점을 정확히 파악하고 쉽게 해결할 수 있다. Report.ss 파일은 모수 추정의 전반적인 결과, 모

수 추정치로부터 계산된 다양한 유도 값등을 제시하기 때문에 각 모수의 추정치들이 적절하게 추정되었는지를 검토할 수 있다. ss.cor 파일은 모수 추정치와 표준 오차, 추정치 간의 상관계수를 제시하며, ss.par 파일은 자유 모수로써 추정한 모수의 수, 목적함수 값, 최대 기울기 및 모수 추정치를 제시한다.

II. 2. 자료

한국 고등어의 자원평가에 사용된 자료는 고등어를 잡는 모든 어업으로부터 얻어진 연도별 총 어획량 자료(total catch in MT, 1975-2019)가 있으며, 대형 선망(large purse-seine) 어업으로부터 얻어진 단위노력당 어획량 자료(catch-per-unit-effort; CPUE in MT/haul, 1975-2019)와 체장-빈도 자료(length-frequency data, 2000-2019)가 있다. 연도별 총 어획량 자료는 국가 통계 포털(Korean Statistical Information Service; KOSIS)에서 공개적으로 제공하고 있는 자료를 사용하였으며, 단위노력당 어획량 자료와 체장-빈도 자료는 국립수산과학원(National Institute of Fisheries Science; NIFS)에서 제공받은 것을 사용했다. 어획량 자료(Figure 2) 및 단위노력당 어획량 자료(Figure 3)는 연 단위로 구성되어 있으며, 체장-빈도 자료(Figure 7)는 연도별로 10 cm부터 52 cm까지 1 cm의 체장 계급 간격을 적용하여 각 체장 계급 구간에서 관측되는 고등어의 체장

자료[가랑이 체장(fork length in cm)]를 이산형 자료로 고려하였다. 이후, 각 체장 계급에서 관측된 체장-빈도 자료를 체장 계급별 비율 자료(체장 조성 자료)로 취급하였다.

Figure 2. 본 연구에서 사용된 연도별 총 어획량 자료(total catch in MT).

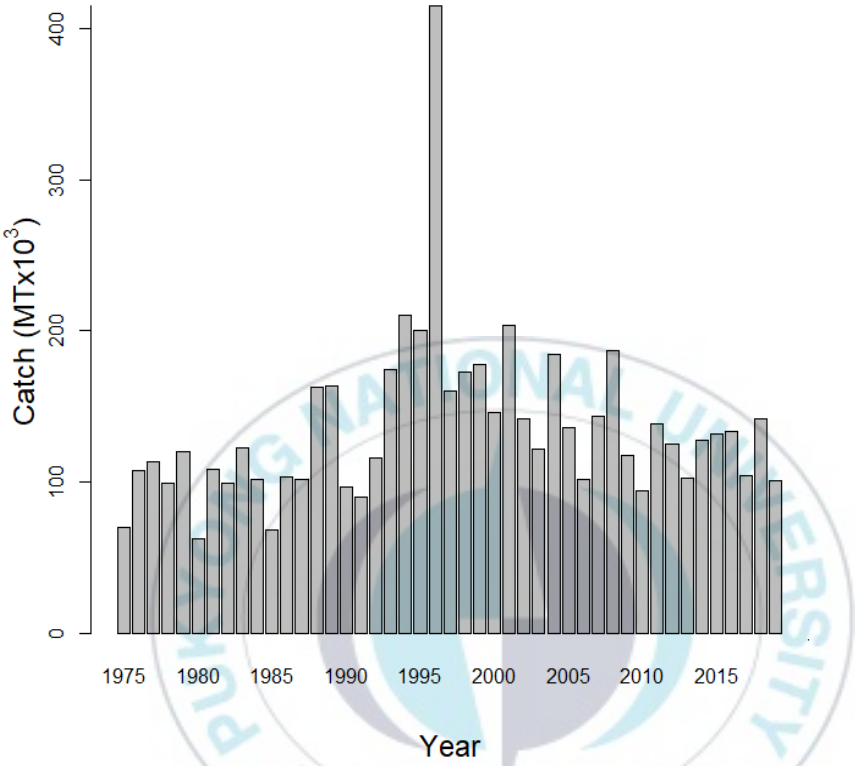
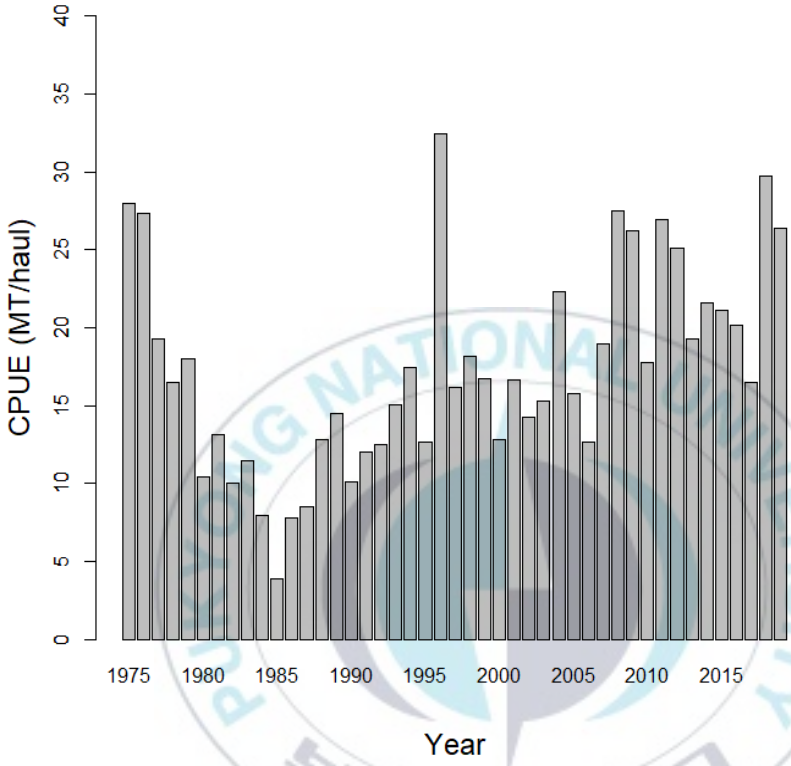


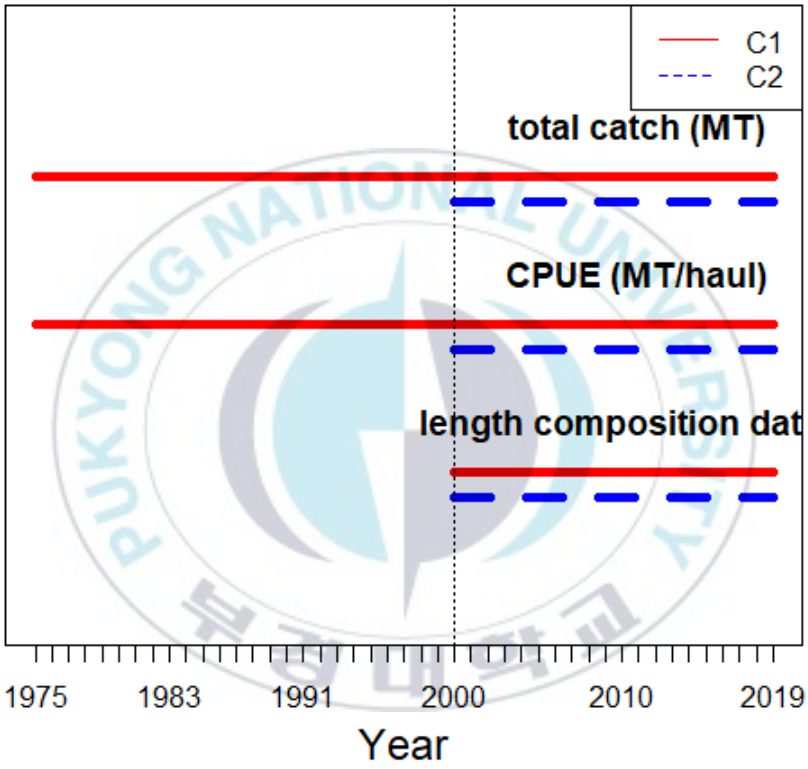
Figure 3. 본 연구에서 사용된 연도별 단위노력당 어획량 자료(CPUE in MT/haul).



II. 3. 모델 가정

대형 선망 어업은 한국 고등어 총 어획량의 약 90%를 차지하고 있기 때문에, 대형 선망 어업으로부터 얻은 단위노력당 어획량 및 체장 조성 자료는 한국 고등어를 어획하는 총 어업의 특성을 대표한다고 가정하였다. 한편, 본 논문에서는 자료의 사용 기간에 따라 두 가지 경우를 가정했다. 자료를 모두 사용하는 경우를 case 1(C1)이라고 가정했고, 레트로스펙티브 오차 분석을 수행하기 세 가지 자료가 공통으로 존재하는 기간 즉, 2000년도부터 2019년도까지의 자료를 사용하는 경우를 case 2(C2)로 가정했다(Figure 4).

Figure 4. 두 경우(C1, C2)에서 사용된 시계열 자료의 범위. C1과 C2는 공통적으로, 한국 고등어(*Scomber japonicus*)의 연도별 총 어획량(total catch in MT) 자료와 단위노력당 어획량(CPUE in MT/haul) 자료, 그리고 체장 조성 자료(length composition data)를 2000년도부터 2019년도까지 동일한 자료를 사용하였으며, C1의 경우, 총 어획량 자료와 단위노력당 어획량 자료를 1975년도부터 사용하였다.



한국 고등어의 각 코호트의 시간 단계는 1 년을 단위로 하였으며, 자원평가의 용이성을 위해서 고등어의 명목상 연령은 매년 1 월 1 일에 1 씩 증가하며, 동일에 일괄적으로 산란을 한다고 가정했다. 고등어의 최대 연령은 선행 연구로 알려진 6 세로 가정했으며(NFRDI, 2010), 이를

+그룹으로 고려했으며, 이를 6+세로 표기했다. SS 는 알부터 가입까지의 개체군 생활사 기간을 명시적으로 모델링하고 있지 않기 때문에, 산란이 발생하는 순간에서의 연령을 0 세로 고려했고(Methot et al., 2013), 이때를 가입(recruitment)으로 정의했다. 따라서, 본 논문에서는 한국 고등어의 가상의 연령을 0-(6+)세로 총 7 개의 연령 계급을 구성했다. 고등어의 성비는 2013 년 1 월부터 2017 년 12 월까지 한국 연근해에서 대형 선망 어업으로 어획된 암·수 총 7,659 개체에 대한 선행 연구(Kim et al., 2020)에 따라서 암컷과 수컷을 60:40 으로 매년 동일하게 적용했으며, 이 비율은 이후, 각 연도 초기(beginning of the year)의 산란가능한 생물량(Stock Spawning Biomass; SSB in MT)을 계산하기 위해서 가정되었다.

SS 의 특징 중 하나는, 수많은 모델 옵션을 사용자가 상황에 맞게 선택하여 하나의 모델 프레임워크를 구성할 수 있다는 것이다. 따라서, 이어지는 모델의 수식 및 모델 가정들은 SS 에서 제시하는 것들이며, 본 저자는 이를 참고하여 수식 및 표기법을 제시하였다(Appendix A in Methot et al., 2013).

II. 3. 다. (1). 사망

순간 자연사망률은 개체군이 어획과 관련되지 않고 질병, 먹이 경쟁, 노화 등의 요인으로 제거되는 것과 관련된다. 순간 자연사망률은 M 으로

표기되며, 일반적으로 추정이 어려운 모수로 알려져 있다(Vetter, 1988). 본 연구에서는 M을 선행 연구(Gim et al., 2020)를 참고하여 0.38 $year^{-1}$을 입력 값으로 지정하였고, 연도별, 연령별로 동일하게 즉, 상수로 가정하였다. 순간 어획사망률은 개체군이 어획에 의해 제거되는 것과 관련된다. 연도별 순간 어획사망률은 F_y로 표기하며, 이를 SS 에서 제시하고 있는 hybrid 방법을 적용해 계산했다(Methot et al., 2013). 즉, F_y를 자유 모수로서 추정하지 않고, 근사적 방법을 적용하여 유도한다.

II. 3. 다. (2). 체장-체중 관계식

체장-체중 관계식은 각 체장 계급의 중앙값(mid-point)에 대응하는 고등어의 무게를 유도하기 위해서 사용되었다. 본 논문에서 사용된 자료들과는 별개로, SS 소프트웨어의 외부 환경에서 2005 년부터 2017 년까지 대형선망 어업으로부터 얻은 총 15,648 개의 체장-체중 자료를 사용하여 모수 α, β 를 추정했다. 모수 추정은 ADMB 소프트웨어로 수행했으며, 모수 추정치 $\hat{\alpha}, \hat{\beta}$ ($\hat{\alpha}$=2.8×10^{-06}, $\hat{\beta}$=3.4428)은 SS 의 체장-체중 관계식[식 (1)]의 모수 α, β에 입력 값으로 지정했다. 식 (1):

$$W_l = \alpha\,(L_l')^{\beta}; \qquad (1)$$

l 은 각 체장 계급 구간(0-1 cm, 1–2 cm ···, 51-52 cm)의 최젓값을 나타내며, L'_l은 각 체장 계급 구간의 중앙값(0.5 cm, 1.5 cm, ···, 52.5 cm)을 나타낸다. W_l는 각 체장 계급 구간의 중앙값 L'_l에 대응하는 고등어의 무게(kg)를 나타낸다. 모수 α는 중량 계수(weight coefficient)를 의미하고, β는 중량지수(weight exponent)를 뜻한다.

II. 3. 다. (3). 체장−성숙률 관계식 및 산란력

체장−성숙률(maturity) 관계식은 로지스틱 모델(logistic model)을 가정했으며, ADMB 소프트웨어 환경에서 모수를 추정했다. Kim et al., 2020 의 Fig. 8 에서 체장−성숙률 자료를 스캔 및 이를 사용하여 두 모수 ω_1, ω_2를 추정했다. 이후, 추정치 $\hat{\omega}_1$, $\hat{\omega}_2$ ($\hat{\omega}_1$=-0.82, $\hat{\omega}_2$=29.17)은 SS 체장−성숙률 관계식[식 (2)]의 모수 ω_1, ω_2에 입력 값으로 지정했다.

$$Mat_l = \frac{1}{1 + e^{\omega_1(L'_l - \omega_2)}}; \quad (2)$$

Mat_l는 각 체장 계급 구간 l의 중앙값 L'_l에 대응하는 성숙률을 나타낸다. 모수 ω_1은 성숙률 로지스틱 함수의 경사도를 의미하며, 모수 ω_2는 성숙률이 50%일 때, 이에 대응되는 체장(cm)을 의미한다. 각 체장 계급에 해당하는 암컷 고등어가 산란할 수 있는 알의 수는 각 체장 계급에 대응하는 체중에 비례하도록 가정했으며, 다음 식의 모수 ω_3, ω_4를 각각 1과 0으로 가정했다.

$$Eggs_l = \omega_3 + W_l \cdot \omega_4; \quad (3)$$

ω_3는 체장 계급이 0 일 때, 체중에 대응되는 알의 수에 대한 함수의 절편 값을 의미하고, ω_4는 체중(kg) 당 알의 수에 대한 함수의 기울기를 의미한다. 연령별 평균 산란력은 다음 식[식 (4)]과 같이 계산된다.

$$f_a = \sum_{l=0}^{52} \phi_{a,l} \left(Mat_l \cdot Eggs_l \cdot W_l\right); \quad (4)$$

$\phi_{a,l}$는 각 연도 초기의 연령−체장 전이 행렬(age-length transition matrix)을 의미하여, 이는 연령별 체장 계급에 대한 확률 질량

함수(probability mass function)를 표로 나타낸 연령−체장 상관표(Age-Length Key; ALK)로부터 구성된다. 연령−체장 상관표를 유도하는 과정은 **부록 1** 에 제시하였다. f_a는 연령 a에 대응하는 평균 산란력(fecundity)을 의미하며, 단위는 (kg/마리)이다. 산란력 f_a는 각 연도의 초기에 연령별로 추정되는 고등어 개체수에 곱해져 산란가능한 암컷의 생물량(Stock Spawning Biomass; SSB)이 유도된다[식 (5)].

$$SSB_y = frac_{fem} \cdot \sum_{a=0}^{A} N_{y,a} \cdot f_a \,; \tag{5}$$

SSB_y는 각 연도의 초기에 추정되는 산란가능한 암컷의 생물량을 나타내고, $N_{y,a}$ 는 각 연도의 초기에 연령별로 추정되는 마릿수를 나타내며, $frac_{fem}$은 암컷의 비율을 나타낸다.

II. 3. 다. (4). 성장 모델

성장 모델은 Schnute & Fournier (1980)의 본 버틀란피 성장 모델(von Bertalanffy growth model)을 따른다고 가정했다. 기존의 본 버틀란피 성장 모델과 다른 점은, 시간에 대한 모수 t_0 대신, 체장에 대한 모수 L_0를

사용한다는 점이다. 각 연도 초기의 연령별 평균 체장은 다음 식[식 (6)]에 의하여 계산된다.

$$L_a = L_{inf} + \left(L_0 - L_{inf}\right) e^{-K \cdot a}; \qquad for\ 0 \leq a < A \tag{6}$$

L_a는 각 연도 초기의 연령 a에 대응하는 평균 체장을 의미하고, L_0는 0 세에 대응하는 평균 체장, L_{inf}는 이론적 최대 평균 체장을 의미하며, K는 성장 계수(growth coefficient), A는 마지막 연령을 나타낸다. 식 (6)에서 추정되는 모수들(L_0, L_{inf}, K)에 의하여 마지막 연령을 제외한 연령별 평균 체장을 유도할 수 있다. C1 의 경우 성장 모델의 모수를 전부 추정한 반면, C2 의 경우 수치 최적화가 이루어지지 않아 L_{inf}값을 선행 연구(Shiraishi et al., 2008)를 참고하여 입력 값(40.6 cm)으로 지정했다. +그룹으로 고려한 마지막 연령에 대응하는 평균 체장은 다음 식[식 (7)]에 의해 유도된다.

$$L_A = \frac{\sum_{a=A}^{2A}\left(e^{-0.2(a-A)}\right)\left(L_A'' + \left(\frac{a-A}{A}\right) \cdot (\mathrm{L_{inf}} - L_A'')\right)}{\sum_{a=A}^{2A} e^{-0.2(a-A)}}; \tag{7}$$

L_A는 각 연도 초기의 마지막 연령 A에 대응하는 평균 체장을 나타내며, L_A''는 마지막 연령 A에 대응하는 성장 모델의 평균 체장값 즉, 식 (6)에서 A를 대입했을 때 얻을 수 있는 평균 체장값을 의미한다. +그룹의 평균 체장은 최대 연령의 두배 수준($2A$)까지의 성장을 고려하여 평균 체장을 구했으며, 0.2 는 연령이 증가함에 따라 적용되는 자연 사망률을 감소 인자로서 적용했다.

II. 3. 다. (5). 평형 상태 연령별 개체수

SS 는 자료가 시작되는 연도의 이전 연도에서, 평형 상태(equilibrium) 및 어획사망이 일어나지 않는 상황을 가정하여 연령별 개체수를 고려한다. 평형 상태를 가정한 연도 초기의 각 연령별 개체수는 다음 식 (8)에 의해서 계산된다.

$$N_{equ,\,a} = R_0\, e^{-aM}; \qquad for\ 1 \le a < 3A \tag{8}$$

$N_{equ,a}$는 어획이 일어나지 않는 평형상태를 가정한 연도 초기의 연령별 개체수를 나타내며, R_0는 자유 모수로서 추정되며, 가입 마릿수(0 세의 마릿수)를 뜻한다. 평형 상태를 가정한 연도 초기의 마지막 연령(A)에서의

개체수 계산을 위해, 연령을 최대 연령인 A의 3 배 수준($3A$)까지를 고려했으며, +그룹 개체수의 계산은 다음 식에 의해 계산된다[식 (9)].

$$N_{equ,A} = \sum_{a=A}^{3A-1} (N_{equ,a}) + \frac{N_{equ,3A-1} \cdot e^{-M}}{1 - e^{-M}}; \quad (9)$$

$N_{equ,A}$는 평형 상태 및 어획이 일어나지 않는 상태를 가정한 연도 초기의 마지막 연령(A)에서의 개체수를 나타낸다.

II. 3. 다. (6). 가입 모델

연도별 초기의 고등어의 가입 마릿수(R_y)는 R_0와 연도별로 추정된 가입 마릿수의 편차($Rdev_y$)에 의하여 계산된다[식 (10)].

$$R_y = R_0 \cdot e^{-0.5\sigma^2_{\log(Rdev)} + Rdev_y}; \quad (10)$$

R_y는 연도별 초기의 가입 마릿수를 나타내며, 로그 단위의 $Rdev_y$는 평균이 0, 표준 편차가 $\sigma_{\log(Rdev)}$인 정규분포를 따르며[식 (11)],

$\sigma_{\log(Rdev)}$ 는 $\log(Rdev_y)$ 의 표준 편차를 나타낸다. $\sigma_{\log(Rdev)}$ 값은 민감도 분석을 실시하여 최적의 값을 입력 값으로 적용했다.

$$\log(Rdev_y) \sim N(0, \sigma^2_{\log(R_{dev})}); \quad (11)$$

II. 3. 다. (7). 어구 선택성

어구 선택성은 체장에 대한 함수로써 로지스틱 모델을 가정했으며, 각 체장 계급에 대응하는 어구 선택성은 다음 식을 통해서 계산된다[식 (12)].

$$S_l = \frac{1}{1 + e^{-\log(19)\cdot(L'_l - L_{50})/L_{95-50}}}; \quad (12)$$

S_l 는 각 체장 계급에 대응하는 어구 선택성을 나타내고, L_{50} 는 어구 선택성이 50%일 때 이에 대응되는 체장을 나타내며, L_{95-50} 는 어구 선택성이 각각 95%, 50%일 때, 이에 대응되는 각 체장들의 차이를 나타낸다. SS 는 체장 및 연령에 대한 어구 선택성을 종합적으로 고려하여, 최종적으로는 연령에 대한 어구 선택성을 고려한다는 특징이 있다. 본 논문에서는 연령에 대한 어구 선택성을 고려하지 않기 때문에, 모든

가상의 연령에 적용되는 어구 선택성을 100%로 동일하게 가정하였다. 최종적으로 한국 고등어에 고려되는 연령에 대한 어구 선택성은 다음 식을 통해 유도된다[식 (13)].

$$S_a = \sum_{l=0}^{52} (\phi_{a,l} \cdot S_l); \quad (13)$$

S_a는 각 연령 계급에 대응하는 어구 선택성을 나타내며, $\phi_{a,l}$는 각 연도 초기에 대응하는 연령−체장 전이 행렬을 나타낸다.

II. 4. 모수 추정

II. 4. 가. 최대 가능도 방법

최대 가능도 방법(maximum likelihood method)은 가능도 함수의 가능도를 최대로 하는 모수를 선택하는 방법으로써, 본 논문에서 추정한 자유 모수는 $\boldsymbol{\theta}$로 표기했으며, 이를 Table 4 에 제시했다. 또한, 본 논문에서 구성한 가능도 함수의 구성 요소는 다음 같다.

$$CPUE_y \sim \log N\left(\log(\widehat{\mathrm{CPUE}_y}), \sigma^2_{\log(CPUE)}\right); \quad (14)$$

$$Length_{y,l} \sim Multinomial(N_{eff}, \hat{p}_{y,l}); \quad (15)$$

$$\log\left(Rdev_y\right) \sim N(0, \sigma^2_{\log(Rdev)}); \quad (16)$$

$CPUE_y$는 연도별 CPUE 자료를 나타내며 로그 정규 분포를 가정하였고, $Length_{y,l}$는 연도별, 각 체장 계급에 대응하는 체장 비율 자료이며 다항 분포를 가정했다. $\sigma_{\log(CPUE)}$는 $\log(CPUE_y)$의 표준 편차를 의미하며, N_{eff}는 다항분포에 고려되는 연도별 체장 자료의 표본 크기를 뜻한다. 식 (14), (15)의 $\widehat{CPUE_y}$와 $\hat{P}_{y,l}$은 다음과 같이 계산된다.

$$\widehat{CPUE_y} = \hat{q} \cdot B_y; \quad (17)$$

$$C_{y,a,l} = S_l \cdot \phi_{a,l} \cdot N_{y,a}; \quad (18)$$

$$B_y = \sum_{l=1}^{52} W_l \cdot \sum_{a=0}^{A} C_{y,a,l}; \quad (19)$$

$\hat{q}$은 자유 모수로 추정된 어획 능률 추정치를 의미하며, $C_{y,a,l}$은 연도별 연령별 체장 계급별 어획 마릿수를 나타내고, B_y는 연도별로 유도되는 어획 가능한 생물량을 뜻한다.

$$\hat{p}_{y,l} = \frac{\sum_{a=0}^{A} C_{y,a,l} + 0.1^{-9}}{\sum_{l=10}^{52}(\sum_{a=0}^{A} C_{y,a,l} + 0.1^{-9})}; \quad (20)$$

$\hat{p}_{y,l}$ 은 연도별 체장 계급에 대응하는 고등어 어획 수의 예측 비율을 의미하는데, 이는 연도별 체장 계급별 어획 마릿수를 연도별 체장 계급별 어획 마릿수를 다 더한 값으로 나누어 줌으로써 유도할 수 있다. 연도별 가입 마릿수의 편차를 나타내는 $Rdev_y$ 는 자료가 없는 상황에서 가능도 함수를 구성하여 로그 정규 분포를 가정하였다. $Rdev_y$ 를 추정할 때는 실질적인 자료가 적용되지 않기 때문에, 목적 함수 값에는 페널티로서 적용이 되었다. 목적 함수는 음의 로그 가능도 함수의 합으로 구성되고, 이는 식 (21)로 구성된다.

$$Obj = -logL(\boldsymbol{\theta}|\boldsymbol{CPUE}) - logL(\boldsymbol{\theta}\,|\boldsymbol{Length}) - logL(\boldsymbol{\theta}|\boldsymbol{Rdev}); \quad (21)$$

Obj 는 목적함수(objective function value)를 나타내며, 이를 최소화하는 최적의 모수들의 조합을 찾았다. 추정할 자유 모수($\boldsymbol{\theta}$)는 Table 2(free

paramters)에 정리해 놓았으며, 식 (21)의 각 가능도 함수는 **부록 2** 에 제시했다.

II. 4. 나. 불확실한 모수에 대한 민감도 분석

민감도 분석은 모델에서 어떤 모수를 추정할 수 없거나, 정보가 불확실할 때 적용할 수 있는 방법으로, 모수가 취할 수 있는 값들을 일일이 입력 값으로 대입했을 때, 가장 좋은 결과를 내는 최적의 입력 값을 찾는 방법을 말한다. 가능도 함수를 적용한 각 분포들[식 (14), (15), (16)]의 분산 즉, (1). $\log(CPUE_y)$자료의 표준 편차 $\sigma_{\log(CPUE)}$, (2). 체장 조성 자료의 표본 크기를 나타내는 N_{eff}, 그리고 (3) $\log(Rdev_y)$의 표준 편차 $\sigma_{\log(Rdev)}$ 의 최적 조합을 찾기 위해 민감도 분석 방법을 적용했고, 수렴이 발생한 결과의 한해서 비교 분석을 수행했다. 수렴의 기준으로는 (1) Hessian matrix 의 생성, (2) 모수 추정치가 기준 범위 내에서 추정이 잘 되었는지, (3) 최대 기울기가 0.001 보다 작은지를 기준으로 세웠다. 수많은 수렴 결과들 중, 입력 값의 최적 조합을 찾기 위해서 추가적인 결과 비교 기준을 세웠다. 모델의 적합도를 비교하는 AIC (Akaike's Information Criterion)값, 자료와 모델 값의 차이를 수치적으로 보는 RMSE (Root Mean Square Error), 그리고 모수 추정치들($\widehat{\boldsymbol{\theta}}$)의 표준 오차와 추정치들 간의

상관 계수 수준을 기준으로 세웠다. 최종적으로, 경우별로 자료를 가장 잘 설명하는 입력 값의 최적 조합을 선정했다.

II. 5. 모델 검증

모델 검증이란, 모델의 성능 및 안정성을 평가하는 하나의 수단이다. 현재 SS는 약 40년의 긴 역사를 가지고 있기 때문에 모델 검증은 불필요할 수 있다. 하지만, SS를 구성하는 모델(성장 모델, 가입 모델, 자연사망률, 어구선택성 등)별로 적게는 3개, 많게는 10개 이상의 사용 가능한 모델 옵션들이 있다. 따라서, 우리 연구팀이 선택한 모델 옵션들로 구성된 SS의 체장-기반 모델을 검증할 필요가 있다고 판단했다. 본 연구에서 모델 검증은 시뮬레이션 연구 및 레트로스펙티브 오차 분석 방법으로 이루어졌다. 이미 SS의 시뮬레이션 연구를 위한 수단으로 R 소프트웨어 시뮬레이션 패키지인 'ss3sim'이 존재하지만, 적용 대상이 연령 구조 모델에 국한되어 있어, 우리 연구팀이 직접 R 소프트웨어를 활용하여 체장-기반 모델의 시뮬레이션 연구를 수행했다. 시뮬레이션 연구에 사용한 R code는 **부록. 4**에 첨부하였다.

II. 5. 가. 시뮬레이션 연구

시뮬레이션 연구는 실제 자료가 아닌 가짜 자료를 생성한 후, 이를 활용하여 모델을 평가하는 방법이며, 가짜 자료를 생성하기 위해 기준이 되는 참값을 가정하고, 참값으로부터 가짜 자료를 생성할 변동성의 수준을 가정한다. 이후, 생성된 가짜 자료를 모델에 적용하고 추정한 뒤, 그 결과를 분석하는 과정을 거친다. 가짜 자료를 생성한 자료는 최대 가능도법을 적용한 연도별 CPUE 자료($CPUE_y$), 연도별 체장 조성 자료($Length_{y,l}$)이며, 실제 자료를 이용하여 추정한 결과들($\widehat{CPUE}_y, \hat{p}_{y,l}$)을 가짜 자료를 생성하기 위한 참값으로 가정했다. 그리고, 가짜 CPUE 자료와 체장 조성 자료는 각각 로그 정규분포, 다항 분포로부터 생성했다. 시뮬레이션 시나리오는 총 세 가지로, CPUE 자료의 변동 계수(coefficient of variation; CV)를 (1) 10%, (2) 30%, (3) 50%로 적용했으며, 이때 생성할 가짜 체장 조성 자료의 표본 크기는 연도별 90 개로 동일한 수준을 적용했다(Table 1).

Table 1. 시뮬레이션 시나리오. 각 시뮬레이션 시나리오(SimS1, SimS2, SimS3)는 CPUE자료에 가정한 로그 정규분포, 체장 조성 자료에 가정한 다항분포로부터 생성할 가짜 자료의 변동성 수준을 세 가지로 가정했다. $\boldsymbol{CV_{CPUE}}$는 CPUE 자료의 변동계수를 나타내며, $\boldsymbol{N_{eff}}$는 다항 분포로부터 생성할 체장 자료의 표본 크기를 뜻한다.

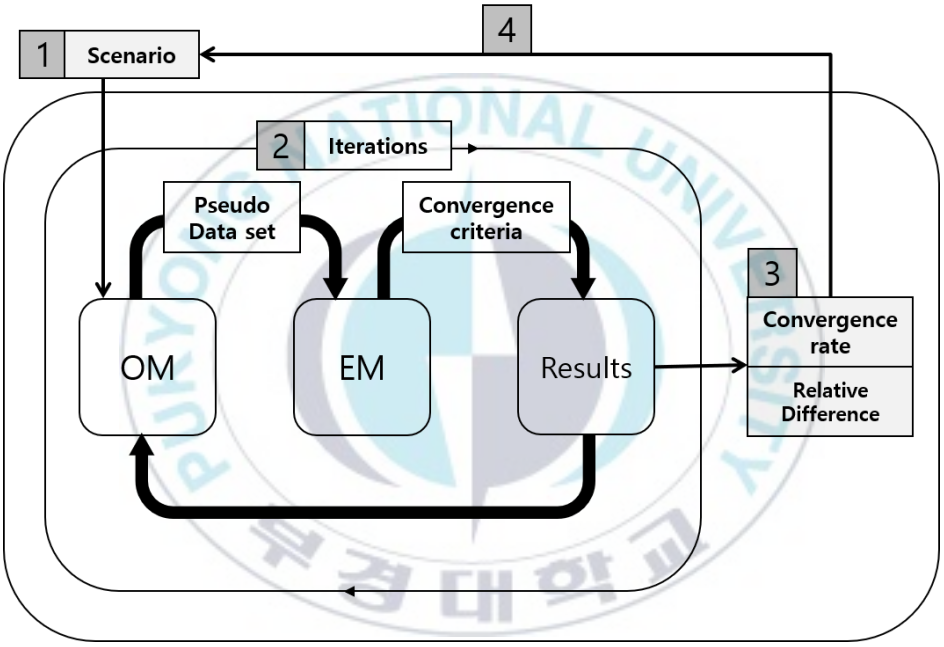
	SimS1	SimS2	SimS3
CV_{CPUE}	10%	30%	50%
N_{eff}	90	90	90

시뮬레이션 연구의 전체적인 과정은 Figure 5 에 제시했다. 먼저, 주어진 시뮬레이션 시나리오에 따라 작동 모델(Operating model; OM)에서 가짜 자료를 생성했고, 생성된 가짜 자료를 추정 모델(Estimation Model; EM)에 입력하고 추정했으며, 이러한 과정을 시뮬레이션 시나리오별로 1,000 번씩 반복했다. 1,000 번의 반복 중에서 수렴 기준을 통과한 결과만을 취합하여 수렴률(convergence rate)과 상대적 오차(Relative Difference; RD)를 제시했으며, 수렴 기준은 민감도 분석의 수렴 기준과 동일하게 적용하였다. 수렴률은 각 시뮬레이션 시나리오의 반복 횟수 중에서 수렴에 성공한 횟수를 백분율로 나타냈고, 수렴에 성공한 결과에 한해서 참값($\boldsymbol{\theta}$)과 추정치($\boldsymbol{\hat{\theta}}$)간의 상대적 오차를 계산했다. 식 (22):

$$\boldsymbol{RDs_{i,j}} = \frac{\widehat{\boldsymbol{\theta}}_{\boldsymbol{i,j}} - \boldsymbol{\theta}}{\boldsymbol{\theta}}; \qquad for\ 1 \le i \le CS_j, j = 1,2,3 \tag{22}$$

상대적 오차(RDs)는 각 시뮬레이션 시나리오별(j) 수렴한 횟수(CS_j)만큼 계산되며, 모수 추정치($\widehat{\boldsymbol{\theta}}_{\boldsymbol{i,j}}$)와 참값($\boldsymbol{\theta}$)의 차를 참값으로 나눈 값으로 정의된다. 시뮬레이션 연구에서 추정할 모수는 Table 2에 제시한 자유 모수 중 $Rdev_y$를 제외한 나머지 모수에 적용했고, 각 시나리오별로 추정된 모수 추정치(Table 3)를 참값으로 가정했다.

Figure 5. 시뮬레이션 모식도. 주어진 시뮬레이션 시나리오(SimS1, SimS2, SimS3)에 따라 작동 모델(Operating model; OM)에서 가짜 자료를 생성한다. 생성된 가짜 자료의 조합은 추정 모델(Estimation model; EM)에 입력되어 추정되고, 추정된 결과는 수렴 기준을 거쳐 저장되거나 누락된다. 이러한 과정을 반복 횟수(1,000)만큼 반복한 후, 수렴 기준에 의해 선별된 결과들로 수렴률 및 상대 오차를 계산한다. 이를 시뮬레이션 시나리오의 경우의 수만큼 실행한다.



II. 5. 나. 레트로스펙티브 오차 분석

레트로스펙티브 오차 분석 방법은 시계열 자료의 마지막 년도에 해당하는 전체 자료를 한 연도씩 제거해가면서 추정된 점추정치 결과를 비교 및 분석하는 방법이다. 특정 연도에서 점추정치간 값의 차이가 크거나, 추세가 체계적으로 편향하는 패턴이 발생했을 경우 모델이 레트로스펙티브 오차를 겪는다고 판단한다(Mohn, 1999). 본 연구에서는 자료의 기간이 동일한 경우(C2)에 한해서 레트로스펙티브 오차 분석을 적용했고, 분석 대상으로는 산란가능한 생물량(SSB_y)과 순간 어획사망률(F_y)을 고려했다. 특정 연도의 추정치 값을 비교하기 위해 자료를 하나씩 제거할 때마다 연도별로 추정된 산란가능한 생물량과 어획사망률의 상대적 차이(RDr)를 식 (23)으로 계산했고, Mohn's (1999) ρ 를 식 (24)으로부터 계산했다.

$$RDr_{y,t}(\theta) = \frac{\hat{\theta}_{y,T-t}}{\hat{\theta}_{y,T}} - 1; \qquad for\ 1 \le t \le 5 \tag{23}$$

$RDr_{y,t}(\theta)$는 사용할 자료의 기간이 각각 $T, T-t$일 때, 특정 연도(y)에서 추정치 간의 상대적 차이를 의미하며, T는 모델에서 고려된 총 연도의 수,

t는 제거한 연도의 수를 나타내며, $\widehat{\boldsymbol{\theta}}_{\boldsymbol{y}}$는 연도별 SSB_y와 F_y의 추정치를 나타낸다.

$$\rho(\theta) = \frac{1}{m}\sum_{t=1}^{m}\frac{\hat{\theta}_{T-t,T-t}}{\hat{\theta}_{T-t,T}} - 1 \qquad for\ 1 \le t \le 5 \tag{24}$$

$\rho(\theta)$은 mohn's ρ값을 나타내며, m은 제거한 총 연도의 수를 나타낸다.

Table 2. 용어의 정의(data, free parameter ($\boldsymbol{\theta}$), input value, and derived quantity)

Symbol	Description	Unit
Data (자료)		
$CPUE_y$	연도별 CPUE 자료	MT/haul
$Length_{y,l}$	연도별, 체장 계급에 대응되는 체장 비율 자료	
Free parameter (자유 모수, $\boldsymbol{\theta}$)		
L_0	0 세의 평균 체장	cm
K	성장 계수(growth coefficient)	
L_{inf}	이론적 최대 평균 체장	cm
R_0	어획이 일어나지 않은 평형상태의 가입 마릿수	number
q	어획 능률(catchability)	
L_{50}	어구 선택성 50%에 대응되는 체장	cm

L_{95-50}	어구 선택성 95%, 50%에 대응되는 체장의 차	cm
$Rdev_y$	연도별 가입 마릿수의 편차	

Input value(입력 값)

A	마지막 연령 즉, 6+	age
l	각 체장 계급 구간의 최젓값(0,1, ···, 52 cm)	cm
M	순간 자연 사망률($year^{-1}$)	year^{-1}
$frac_{fem}$	암컷의 비율	
ω_1	성숙률 로지스틱 함수의 경사도	
ω_2	성숙률 50%일 때, 이에 대응되는 체장(cm)	cm
ω_3	체장 계급이 0 일 때, 체중에 대응되는 알의 수에 대한 함수의 절편값	
ω_4	체중(kg) 당 알의 수에 대한 함수의 기울기	
α	중량 계수(weight coefficient)	

β	중량 지수(weight exponent)	
$\sigma_{\log(R_{dev})}$	$\log(Rdev_y)$의 표준 편차	
$\sigma_{\log(CPUE)}$	$\log(CPUE_y)$의 표준 편차	
N_{eff}	다항분포에 고려되는 연도별 체장 자료의 표본 크기	number
t	Mohn's ρ 계산을 위해 제거한 연도의 수	number
T	모델에서 고려된 총 연도의 수	number
m	제거한 총 연도의 수	number

Derived quantity (유도 값)

a	연령	age
f_a	연령별 평균 산란력(kg /마리)	kg/individual
W_l	L_l'에 대응되는 무게(kg)	kg
L_l'	각 체장 계급 구간의 중앙값	cm

Mat_l	L'_l에 대응되는 성숙률	
$Eggs_l$	L'_l에 대응되는 알의 수	
$\phi_{a,l}$	연도별 초기의 연령-체장 전이 행렬	
SSB_y	연도별 초기의 산란가능한 생물량(MT)	kg
$N_{y,a}$	연도별 초기의 연령별로 추정되는 고등어 개체수	number
$C_{y,a,l}$	연도별 초기의 연령별, 체장 계급별 어획 마릿수	number
B_y	연도별 초기의 어획 가능한 생물량	kg
L_a	연도별 초기의 연령별로 대응되는 평균 체장(cm)	cm
L_A	연도별 초기의 마지막 연령 A에 대응되는 평균 체장	cm
L''_A	연도별 초기의 마지막 연령 A에 대응하는 성장 모델의 평균 체장	cm
$N_{equ,a}$	평형 상태 및 어획이 일어나지 않은 상황을 가정한 연도 초기의 연령별 대응되는 개체수	number
$N_{equ,A}$	평형 상태 및 어획이 일어나지 않은 상황을 가정한 연도 초기의 마지막 연령에 대응되는 개체수	number

R_y	연도별 초기의 가입 마릿수	number
F_y	연도별 순간 어획사망률	$year^{-1}$
S_l	연도별 초기의 L_l'에 대응되는 어구 선택성	
$RDs_{i,j}$	시뮬레이션 시나리오별(j), 해당 반복 횟수(i)에서 추정치와 참값 간의 상대적 차이	
$RDr_{y,t}$	사용할 자료의 기간이 각각 T, $T-t$일 때, 특정 연도(y)에서 추정치 간의 상대적 차이	
$\rho(\theta)$	$Mohn's\ \rho$ 값	

III. 결과

Ⅲ. 1. 한국 고등어 자원 평가

두 경우(C1, C2)에서 입력 값으로 지정한 모수와 그 입력 값을 Table 3에 제시했고, 추정된 모수의 수는 각각 52, 26개로 각 모수의 추정치와 추정치의 표준 오차를 Table 4에 제시했다. 두 경우 간의 모수 $\log(R_0)$, $\log(q)$의 추정치는 유사했지만, 성장 모델의 모수 추정치 간에는 다소 차이를 보였다[C1: L_0= 9.37 cm, K= 0.63, L_{inf}= 33.2 cm, C2: L_0= 15.36 cm, K= 0.17, L_{inf}(입력 값)= 40. 6 cm]. 성장 모델의 모수 추정치 및 입력 값의 차이는 두 경우의 연령-체장 상관표의 차이를 나타냈고, 결과적으로 각 연도 초기의 연령별 체장 계급별 확률 분포의 차이를 나타냈다(Figure 13). 한편, 두 경우의 체장에 대한 어구 선택성 모수 추정치는 C1의 경우 L_{50}= 26.63 cm, L_{95-50}= 6.98 cm, C2의 경우, L_{50}= 27.55 cm, L_{95-50}= 5.46 cm으로 추정되었다. 체장 계급별 어구 선택성을 Figure 11의 상부 패널에 제시했으며, 이후 체장 계급별 어구 선택성과 연령-체장 상관표의 곱으로부터 유도되는 연령별 어구 선택성을 Figure 12에 제시했다. 각 연도별 가입 마릿수는 평형상태 및 어획이 일어나지 않는 상태를 가정한 연도의 가입 마릿수 $\log(R_0)$ 추정치와 연도별 가입 마릿수의 편차($Rdev_y$) 추정치, 그리고

입력 값인 $\sigma_{\log(R_{dev})}$값으로부터 유도되는데, C1에 비해 C2에서 그 변동성이 훨씬 크게 나타났으며(Figure 9), 각 연도별 가입 마릿수의 평균은 C1의 경우 1,874 x 10^6 (마릿수), C2의 경우 3,265 x 10^6 (마릿수)으로 C1에 비해 C2의 평균 가입 마릿수가 약 1.7배 높게 나타났다. 연도별 순간 어획사망률(F_y)은 C1의 경우 1975년도에 0.08 $year^{-1}$로 가장 낮았고, 1996년도에 0.66 $year^{-1}$으로 가장 높았으며, 1996년도부터는 순간 어획사망률의 추세가 우하향하여 2019년도에는 0.14 $year^{-1}$의 수준까지 감소했고, 두 경우가 공통적으로 고려되는 2000년도부터 2019년도까지의 두 경우의 순간 어획사망률의 크기와 추세는 유사하게 나타났다(Figure 10). 두 경우에서 추정된 총 생물량(total biomass) 및 산란가능한 생물량(stock spawning biomass)은 Figure 8에 제시했으며, C1의 경우 1975년도에서 한국 고등어의 총 생물량이 1.2 x 10^6 (MT)으로 가장 많았으며, 1985년도까지 0.2 x 10^6 (MT)의 수준으로 급격히 감소한 이후, 2019년도까지 총 생물량의 추세가 우상향하여 1 x 10^6 (MT)으로 나타났다. 반면, C2의 경우, 2000년도부터 2019년도까지 C1에서 추정된 총생물량의 크기보다 적게는 0.07 x 10^6 (2019년도), 많게는 0.7 x 10^6 (2000년도)의 차이를 보였으며, 평균적으로 C2가 C1에 비해 약 1.5배 높은 총생물량 수준을 나타냈다. 하지만, 2000년도부터 2019년도까지의 두 경우에서 추정된 산란가능한 생물량의 수준은 유사한 추세 및 크기를 나타냈고(Figure 8), 2019년도에는 두 경우에서 각각 0.3 x 10^6 (MT), 0.3 x 10^6 (MT) 수준으로 유사하게 나타났다.

Ⅲ. 2. 모델 검증

가능도 함수를 구성한 연도별 CPUE자료와 체장 조성 자료에 대해 모델의 적합도를 보기위해 각 자료와 모델의 적합선을 그래프로 나타냈으며, 연도별 CPUE 자료와 두 경우의 적합선을 Figure 6에 제시했다. 두 경우 모두에서 모델의 예측값이 CPUE 자료를 잘 관통하는 것으로 나타났지만, C2의 경우 자료가 시작하는 연도인 2000년도와 자료가 끝나는 연도인 2019년도에서 C1에 비해 자료를 잘 설명하지 못했으며, RMSE를 비교했을 때 C1은 3.16, C2는 4.39로 C1이 C2에 비해 자료를 더 잘 설명하는 것으로 나타났다. 한편, 연도별 체장 조성 자료와 모델의 적합선을 Figure 7에 제시했으며, 두 경우 모두에서 모델의 예측값이 히스토그램으로 나타낸 자료를 잘 설명하는 것으로 나타났지만, C2의 경우 특히, 2000, 2001, 2016, 2018년도에서 C1에 비해 자료를 잘 설명하지 못했고, RMSE를 비교했을 때 C1은 0.062, C2는 0.088로 C1이 C2에 비해 자료를 더 잘 설명하는 것으로 나타났다.

모델 검증의 하나의 수단으로 실시한 시뮬레이션 연구의 결과는 다음과 같다. 각 시뮬레이션 시나리오(SimS1, SimS2, SimS3)별로 1,000번의 반복 횟수 중에서 수렴 기준을 통과한 횟수를 백분율로 제시했으며(Table 5), 그 결과 CPUE 자료의 변동성이 10%, 30%, 50%로 커짐에 따라 C1의 경우 수렴률은 91.8%, 66.2%, 34.9%로 현저하게 감소한 반면, C2의 경우 CPUE

자료의 변동성에 상관없이 99.9%, 99.4%, 96.8%의 안정적인 수렴률을 보였다. 또한, 시뮬레이션 연구에서 참값으로 가정한 Table 4의 1-7번 모수 추정치 값과 각 시뮬레이션 시나리오에서 추정된 모수 추정치들 간의 상대적 차이(RDs)를 상자 그림(Figure 14, Figure 15)으로 제시했다. C1의 경우 CPUE 자료의 변동 계수가 10% 수준일 때, 모수 L_0, K를 제외한 나머지 모수에서 상대적 오차(RDs)값들의 중앙값(가로선)이 0에 근사하게 나타났다. 하지만, 모수 L_0, K의 RDs값들의 중앙값은 각각 0으로부터 음의 편향(negative bias), 양의 편향(positive bias)이 나타났다. 한편, C1에서 CPUE 자료의 변동성이 커짐에 따라 특히 성장 모델의 모수 L_0, K와 $\log(q)$의 RDs 값들의 편향이 상대적으로 크게 나타났으며, 모수 $\log(R_0)$, L_{50}의 RDs값들의 편향이 상대적으로 근소하게 발생했다. 반면, 흥미롭게도 성장 모델의 또다른 모수 L_{inf}의 경우에는 RDs값들이 0을 중심으로 분포가 매우 좁게 나타났고, 이는 각 시뮬레이션 시나리오에서 L_{inf} 추정치들이 참값과 근사하게 추정되었다는 것을 나타냈다. 반면, C2의 경우 모든 시뮬레이션 시나리오에서 모수들의 RDs값 분포의 너비 정도는 달랐지만, 중앙값은 0에 근사하게 나타나 참값을 기준으로 분포하는 것으로 나타났다. 결과적으로, C2의 경우 CPUE 자료의 변동성이 커짐에도 불구하고 모수 추정치의 편향이 발생하지 않았지만, C1의 경우 CPUE 자료의 변동성이 커짐에 따라 특히, 모수 L_0, K, $\log(q)$에서 추정치의 편향이 발생했다. 이외에도, RDs값들의 분포가 넓게 나타난 모수 L_0, K, L_{50}이 CPUE 자료의 변동성

에 가장 영향을 많이 미치는 것으로 나타났다.

모델 검증을 위해 실시한 또다른 방법인 레트로스펙티브 오차 분석의 결과는 다음과 같다. C2의 경우, 모델은 레트로스펙티브 오차를 크게 겪지 않는 것으로 나타났다(Figure 16). C2의 마지막 연도부터 자료를 제거하면서 추정된 산란가능한 생물량(SSB_y) 및 순간 어획사망률(F_y)을 상부 패널에 제시했고, 이를 바탕으로 상대적 차이(RDr)를 하부 패널에 제시했으며, $mohn's\ \rho$ 값을 빨간 가로선으로 나타냈다. 한 연도씩 자료를 제거하면서 추정되는 SSB_y, F_y 값과 이들의 RDr 값의 추세는 어떤 체계적인 패턴(systematic pattern)을 나타내지 않았고, 하부 패널에서 점으로 나타낸 값들의 평균값인 $mohn's\ \rho$값 또한 각각 -0.0063, 0.2392으로 레트로스펙티브 오차의 영향을 무시할 수 있는 정도로 나타났다. 하지만, 5개 연도를 제거했을 때의 SSB_y 추정치의 경우에는 자료를 다 사용했을 때의 SSB_y 추정치와 모든 연도에서 차이가 상대적으로 크게 나타났는데, 이에 관한 내용은 고찰에서 다루었다.

Table 3. 각 경우에서 입력 값으로 지정한 모수와 그 입력 값. $\boldsymbol{M}$은 단위가 $\boldsymbol{year^{-1}}$인 순간 자연사망률을 의미하고, $\boldsymbol{L_{inf}}$ 의 경우 C1에서는 추정이 되었으며, C2에서는 추정되지 않아 입력 값으로 지정하였다. $\boldsymbol{CV_0}$, $\boldsymbol{CV_A}$는 0, 6+세에서의 체장 분포의 변동 계수, $\boldsymbol{CV_{Rdev}}$, $\boldsymbol{CV_{CPUE}}$는 가입 마릿수의 편차 및 CPUE 자료에 고려되는 변동계수를 나타낸다. $\boldsymbol{CV_{Rdev}}$, $\boldsymbol{CV_{CPUE}}$ 값은 $\boldsymbol{\sigma_{\log(CPUE)}}$, $\boldsymbol{\sigma_{\log(Rdev)}}$값을 유도하는 데 사용했다.

		C1	C2
No.	Parameters	Input values	
1	M	0.38	
2	L_{inf}	-	40.6
3	CV_0	0.1	
4	CV_A	0.1	
5	α	2.8×10^{-06}	
6	β	3.4428	
7	ω_1	-0.82	
8	ω_2	29.17	
9	ω_3	1	
10	ω_4	0	
11	$frac_{fem}$	0.6	
12	CV_{Rdev}	0.8	
13	CV_{CPUE}	0.1	
14	N_{eff}	90	

Table 4. 자유 모수(free parameters)의 추정치 및 표준 오차(standard error).

		C1		C2	
No.	Parameters	Estimates	S.E.	Estimates	S.E.
1	L_0	9.37	0.88	15.36	0.69
2	L_{inf}	33.2	0.31	-	-
3	K	0.63	0.06	0.17	0.01
4	$log(R_0)$	14.52	0.12	14.61	0.26
5	$log(q)$	-10.24	0.17	-10.33	0.29
6	L_{50}	26.63	0.59	27.55	0.36
7	L_{95-50}	6.98	0.41	5.46	0.23
8	$Rdev_{1975}$	-0.82	0.51	-	-
9	$Rdev_{1976}$	-0.51	0.52	-	-
10	$Rdev_{1977}$	-0.06	0.44	-	-
11	$Rdev_{1978}$	-0.68	0.53	-	-
12	$Rdev_{1979}$	-0.41	0.51	-	-
13	$Rdev_{1980}$	-0.04	0.45	-	-
14	$Rdev_{1981}$	-0.27	0.44	-	-
15	$Rdev_{1982}$	-1.09	0.47	-	-
16	$Rdev_{1983}$	-1.54	0.43	-	-
17	$Rdev_{1984}$	-0.96	0.52	-	-
18	$Rdev_{1985}$	0.6	0.2	-	-
19	$Rdev_{1986}$	-0.24	0.58	-	-
20	$Rdev_{1987}$	0.65	0.28	-	-
21	$Rdev_{1988}$	-0.68	0.54	-	-
22	$Rdev_{1989}$	-0.43	0.53	-	-
23	$Rdev_{1990}$	0.14	0.52	-	-
24	$Rdev_{1991}$	0.36	0.61	-	-

25	$Rdev_{1992}$	0.52	0.46	-	-
26	$Rdev_{1993}$	-0.6	0.6	-	-
27	$Rdev_{1994}$	1.47	0.27	-	-
28	$Rdev_{1995}$	0.3	0.92	-	-
29	$Rdev_{1996}$	-0.15	0.68	-	-
30	$Rdev_{1997}$	0.05	0.44	-	-
31	$Rdev_{1998}$	-0.98	0.49	-	-
32	$Rdev_{1999}$	0.61	0.18	-	-
33	$Rdev_{2000}$	0.7	0.16	1.26	0.22
34	$Rdev_{2001}$	0.19	0.21	0.33	0.49
35	$Rdev_{2002}$	0.1	0.26	0.97	0.26
36	$Rdev_{2003}$	0.44	0.2	-1.25	0.77
37	$Rdev_{2004}$	-0.52	0.35	1.34	0.16
38	$Rdev_{2005}$	0.55	0.2	-1.02	0.89
39	$Rdev_{2006}$	0.4	0.25	1.47	0.15
40	$Rdev_{2007}$	0.65	0.2	-1.5	0.68
41	$Rdev_{2008}$	0.1	0.26	1.45	0.14
42	$Rdev_{2009}$	0.55	0.2	-0.91	0.94
43	$Rdev_{2010}$	0.43	0.24	1.17	0.21
44	$Rdev_{2011}$	0.21	0.27	-0.7	0.88
45	$Rdev_{2012}$	0.29	0.28	1.45	0.16
46	$Rdev_{2013}$	0.38	0.28	-1.14	0.8
47	$Rdev_{2014}$	-0.14	0.38	-1.27	0.75
48	$Rdev_{2015}$	-0.31	0.38	1.76	0.13
49	$Rdev_{2016}$	0.88	0.21	-1.03	0.9
50	$Rdev_{2017}$	0.69	0.27	0.33	0.6
51	$Rdev_{2018}$	-0.4	0.43	-1.24	0.76
52	$Rdev_{2019}$	-0.44	0.63	-1.47	0.68

Figure 6. 연도별 단위노력당 어획량 자료($\boldsymbol{CPUE_y}$)와 두 경우(C1, C2)의 적합선(fitted value). 자료는 점으로 표시했으며, C1과 C2의 적합선은 각각 실선과 파선으로 나타냈다. (C1_RMSE= 3.16, C2_RMSE= 4.39)

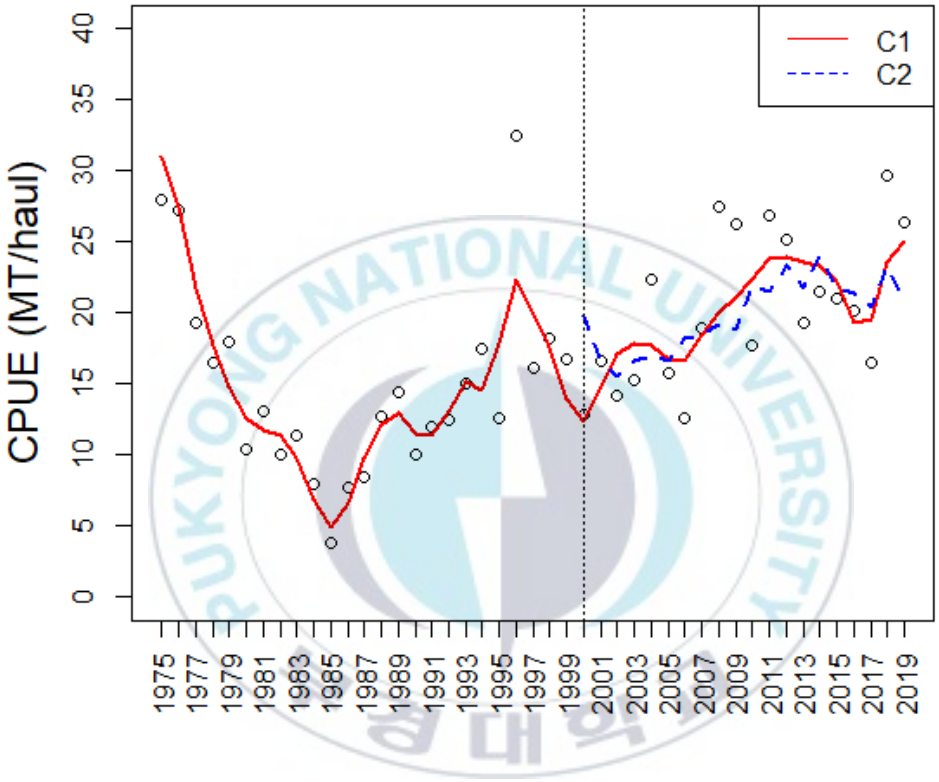


Figure 7. 연도별 체장 계급별 관측 비율 자료와 두 경우(C1, C2)의 적합선. 자료는 히스토그램으로 나타냈으며, C1과 C2의 적합선은 각각 실선과 파선으로 나타냈다. (C1_RMSE= 0.062, C2_RMSE= 0.088)

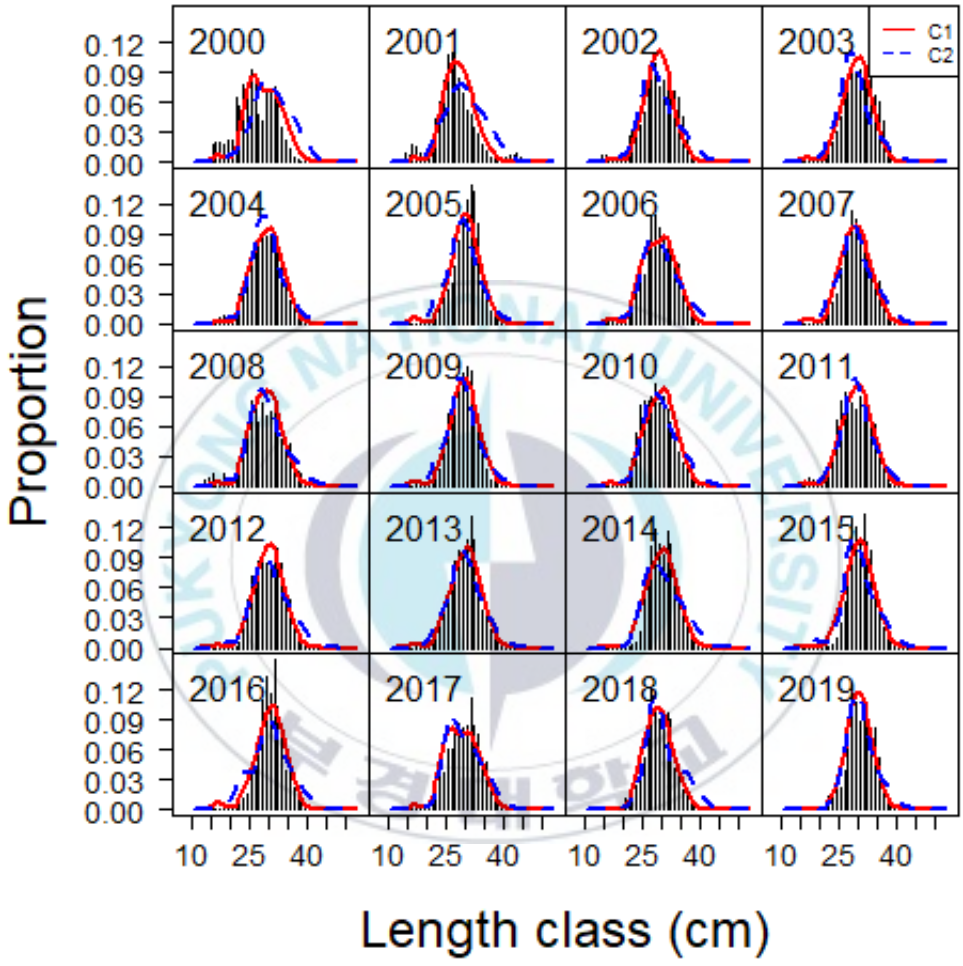


Figure 8. 두 경우(C1, C2)의 각 연도 초기에 추정된 총 생물량(total biomass)과 산란가능한 생물량(stock spawning biomass). C1의 연도별로 추정된 총 생물량을 실선(C1_totB)으로 나타냈고, 산란가능한 생물량(C1_SSB)을 점선으로 나타냈다. C2의 연도별로 추정된 총 생물량을 파선(C2_totB)으로 제시했고, 산란가능한 생물량을 1점 쇄선(C2_SSB)으로 나타냈다.

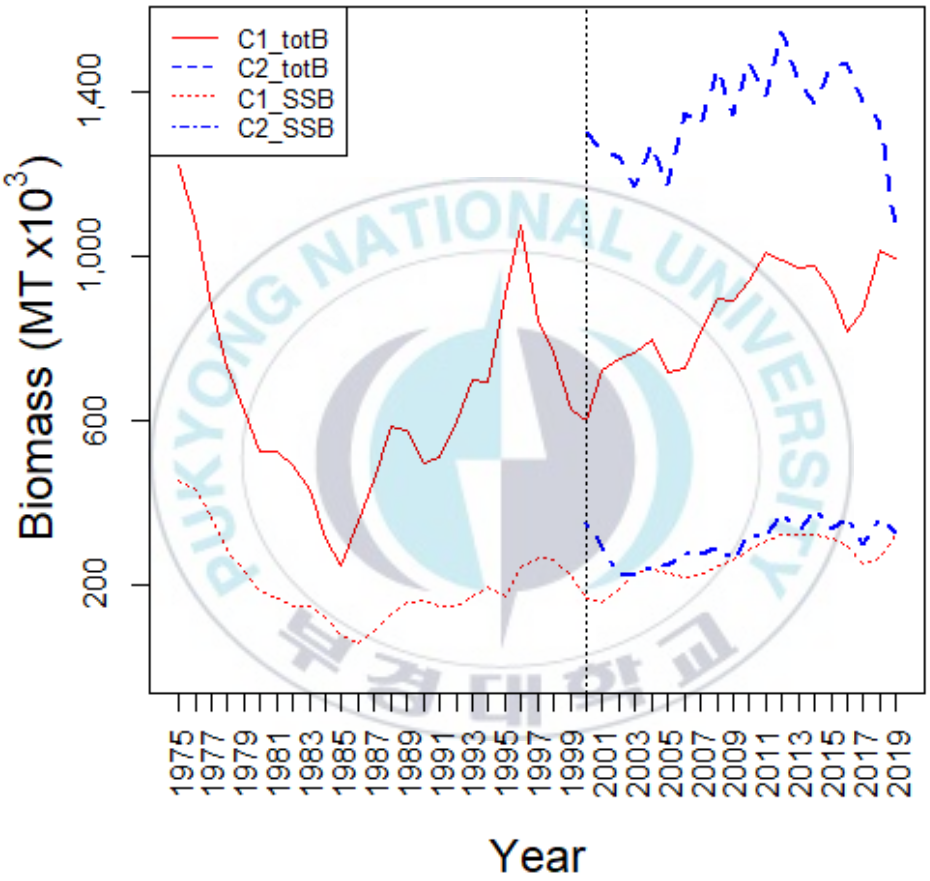


Figure 9. 각 연도 초기에 추정된 가입 마릿수. C1과 C2의 연도별로 추정된 가입 마릿수를 각각 실선과 파선으로 나타냈다. C1의 평균 가입 마릿수를 가로 실선으로 나타냈고, C2의 평균 가입 마릿수를 가로 파선으로 나타냈다.

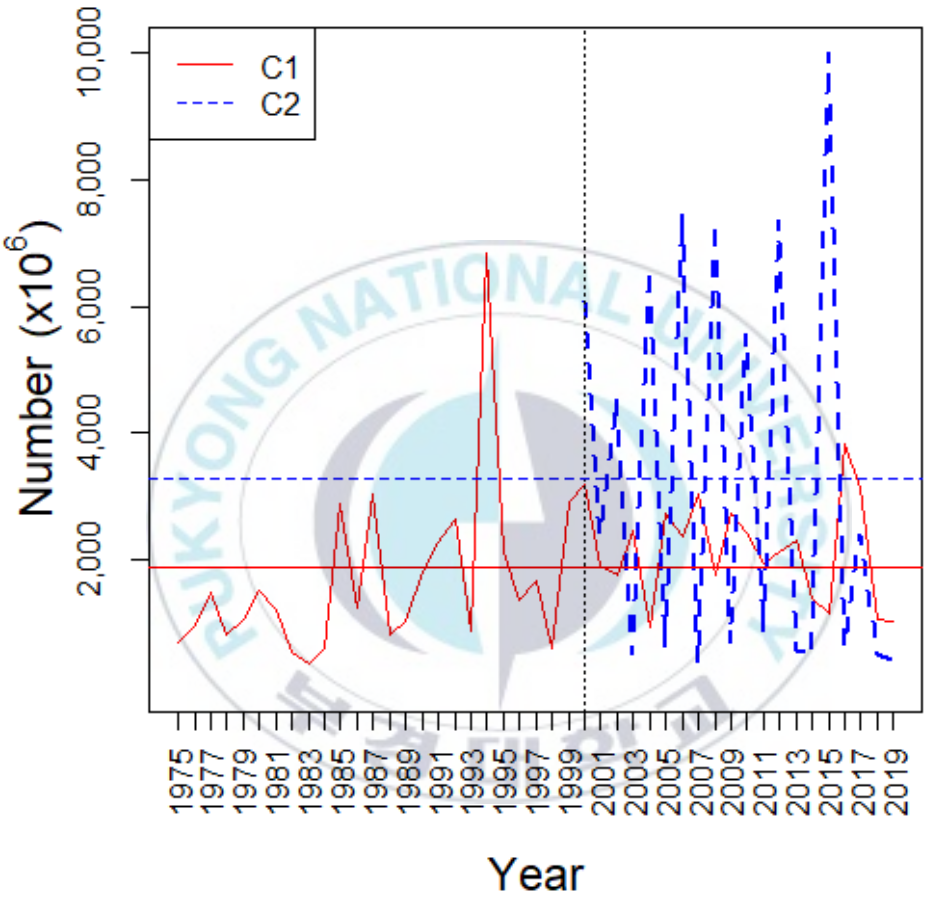


Figure 10. 연도별 순간 어획사망률(F_y). C1과 C2의 연도별 순간 어획사망률을 각각 실선과 파선으로 제시했다.

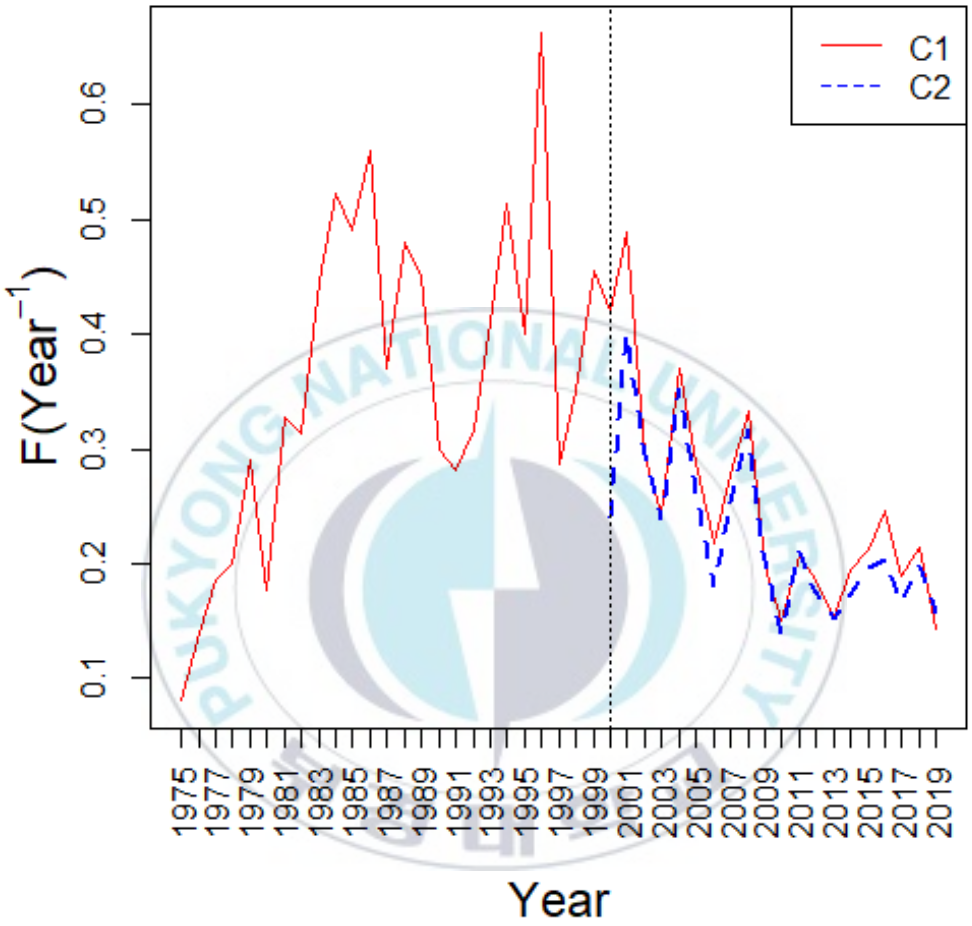


Figure 11. 체장 계급 및 연령에 따른 어구 선택성. C1과 C2의 체장 계급(중앙값)에 대응하는 어구 선택성을 각각 실선과 파선으로 나타냈고(상부 패널), 두 경우에 동일하게 적용되는 연령별 어구 선택성을 점과 실선으로 나타냈다(하부 패널). 상부 패널의 가로 점선은 어구 선택성이 0.5인 지점을 나타냈다.

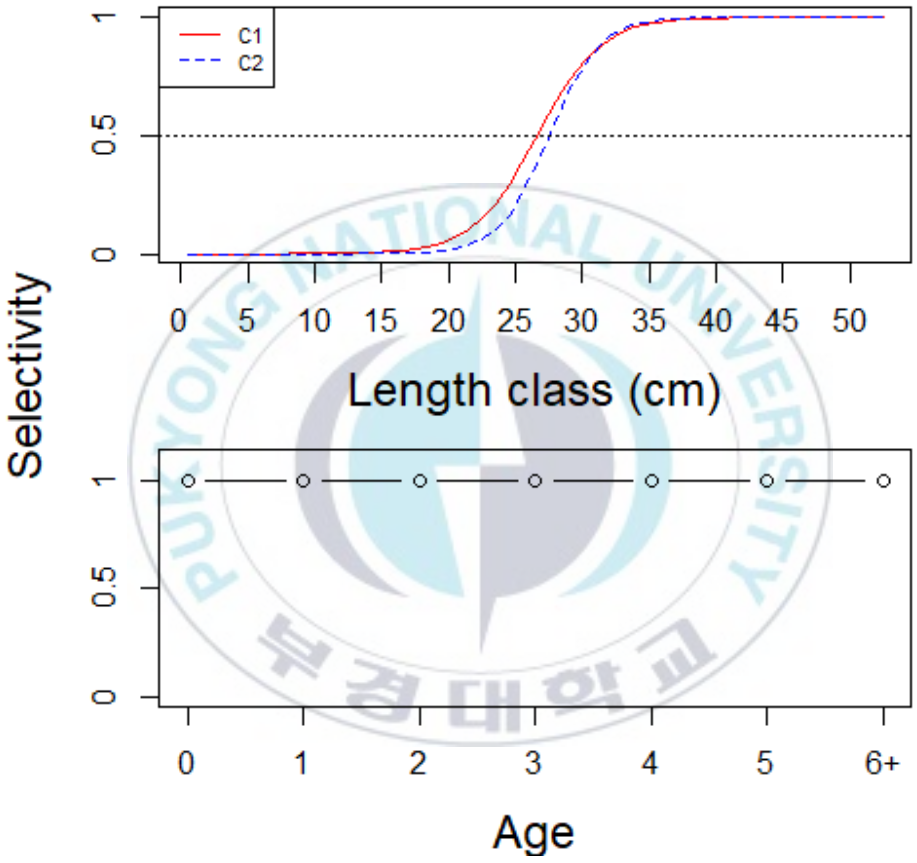


Figure 12. 체장-기반 어구 선택성과 각 연도 초기의 연령-체장 전이행렬 ($\phi_{a,l}$)의 곱으로 유도된 연령에 대한 어구 선택성. C1과 C2를 각각 점과 실선, 점과 파선으로 나타냈다.

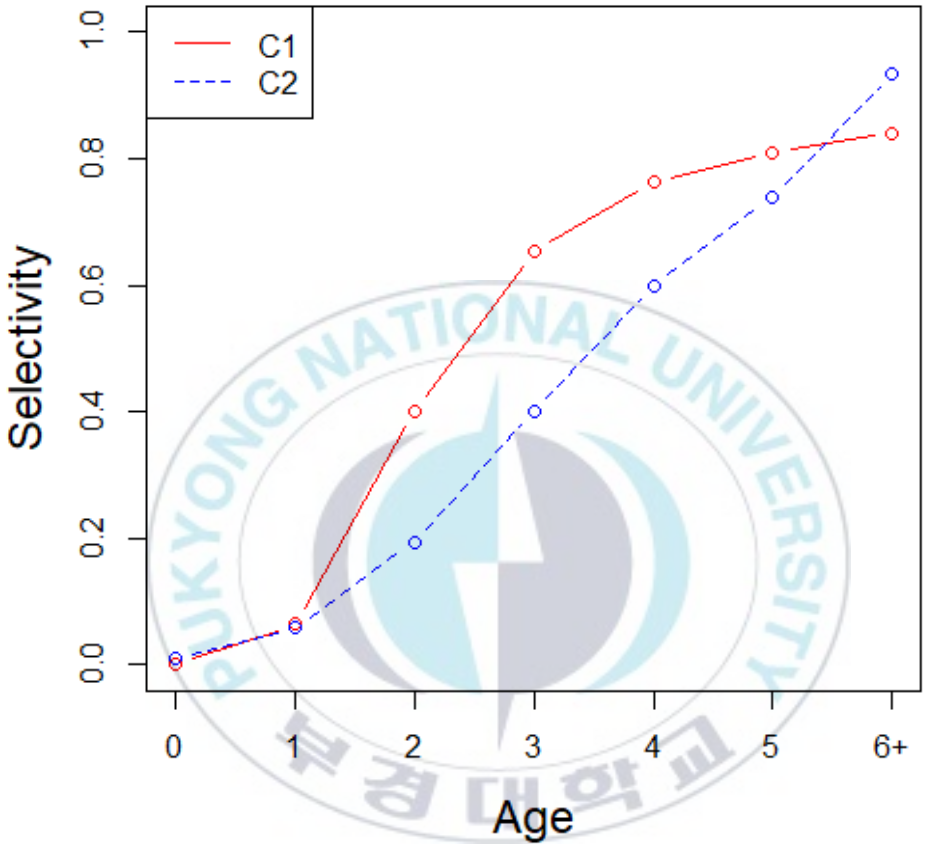


Figure 13. 두 경우(C1, C2)에서 유도된 각 연도 초기의 연령별 체장 계급별 확률 분포. 좌측부터 0세, 1세, ····, 6+세의 체장 분포를 나타낸다.

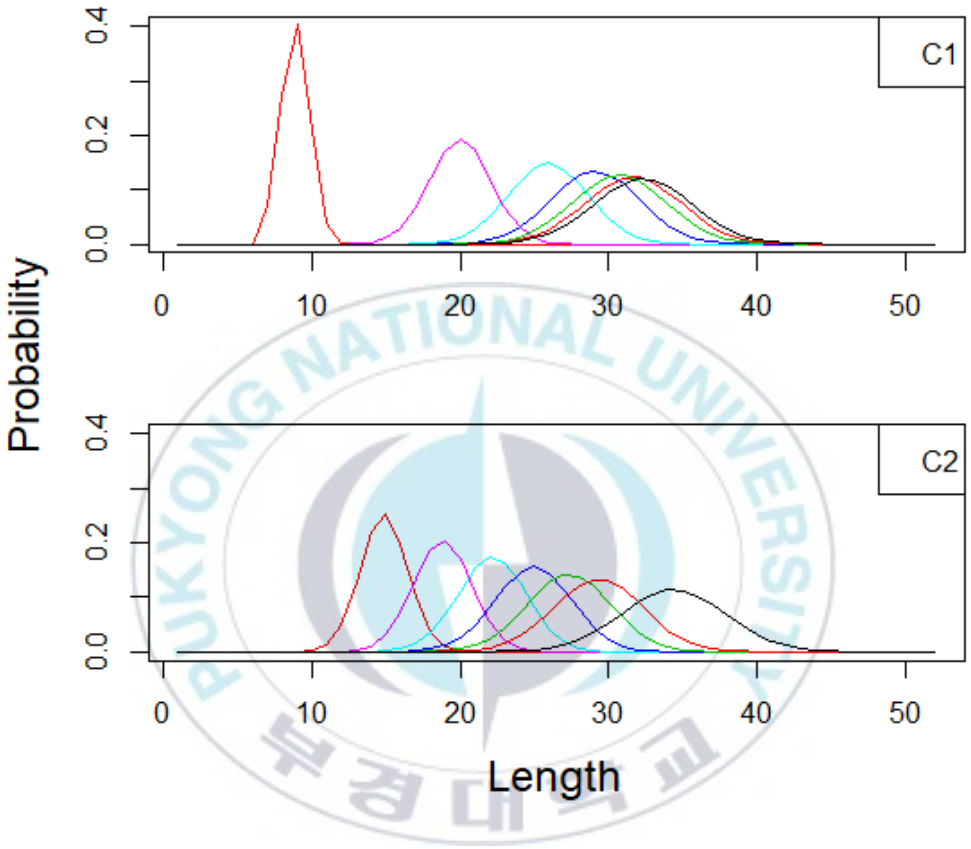


Table 5. 각 시뮬레이션 시나리오에서 두 경우(C1, C2)의 수렴률(convergence rate). 시뮬레이션 시나리오별 1,000번의 반복 횟수 중에서 수렴 기준을 통과한 횟수를 백분율로 제시했다.

	SimS1	SimS2	SimS3
C1	91.8%	66.2%	34.9%
C2	99.9%	99.4%	96.8%

Figure 14. 각 시뮬레이션 시나리오에서 생성된 가짜 자료로부터 추정된 주요 모수 추정치들[$L_0, L_{inf}, K, \log(R_0)$]의 상대적 차이(Relative difference) 상자그림. 시뮬레이션 시나리오(CV10%, CV30%, CV50%)는 CPUE 자료의 변동계수(10%, 30%, 50%)에 대응된다. 박스 내부의 가로선은 중앙값을 나타내고, 박스는 제3사분위수와 제1사분위수의 차이인 사분범위(interquartile range)를 뜻하며, 아랫수염은 중앙값 ±1.5·(사분범위)로 구해진다. Figure 15에 계속.

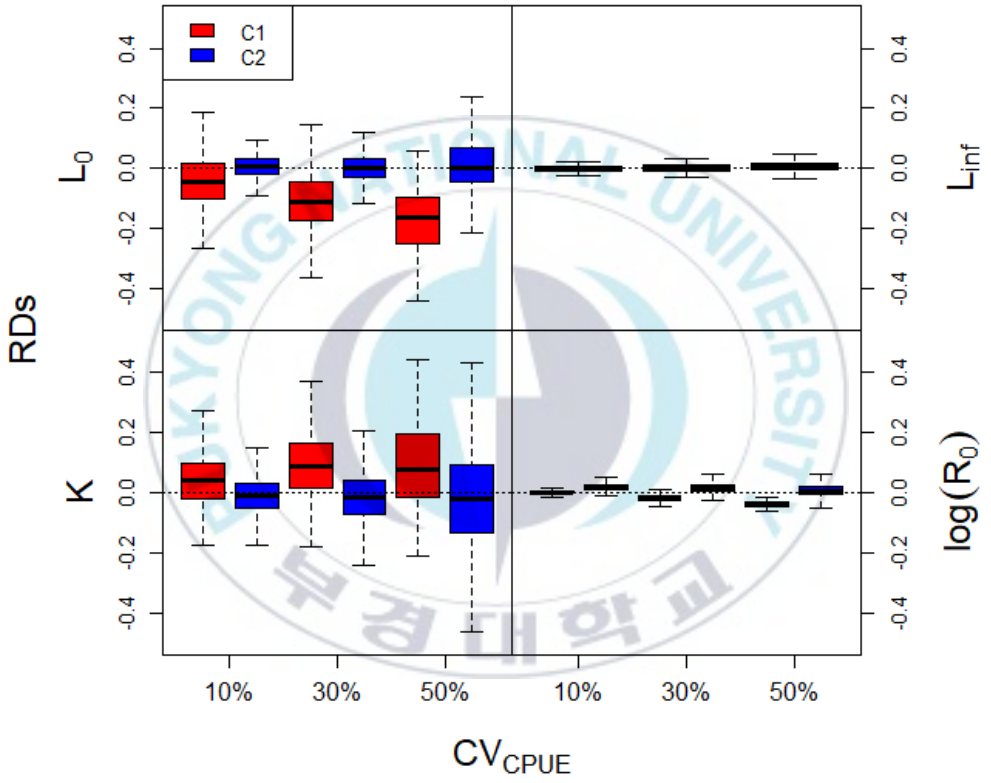


Figure 15. 각 시뮬레이션 시나리오에서 생성된 가짜 자료로부터 추정된 모수 추정치들[$L_{50}, L_{95-50}, \log(q)$]의 상대적 차이 상자 그림.

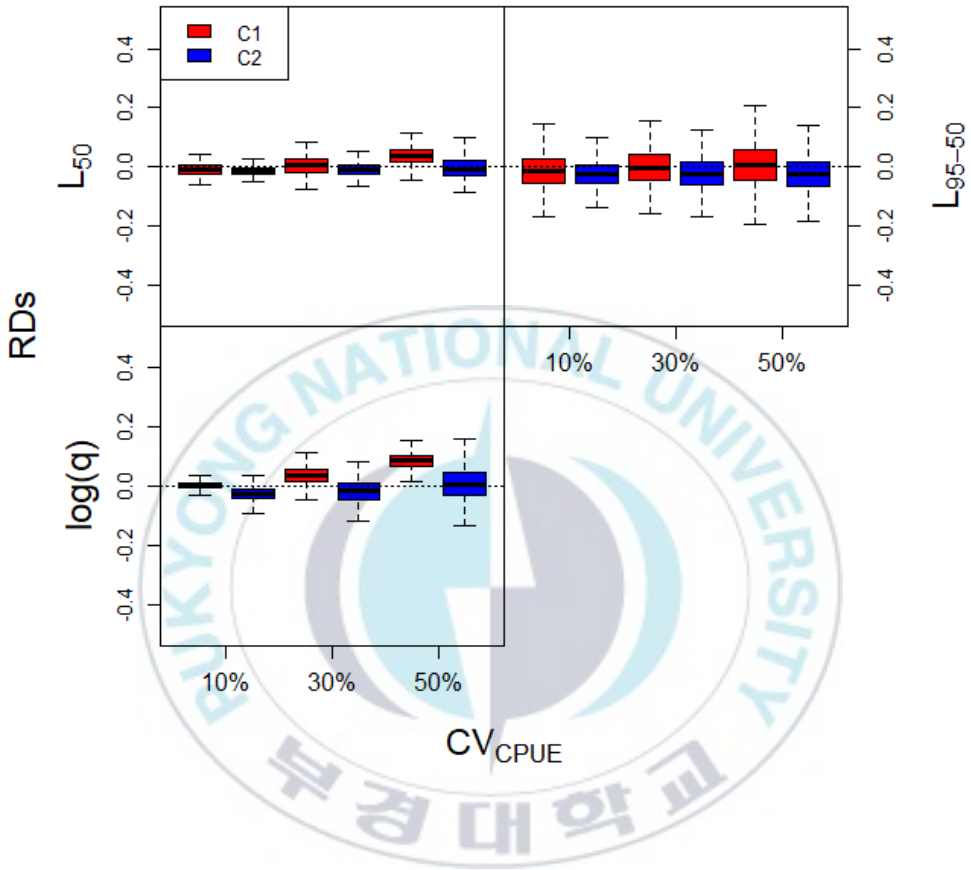
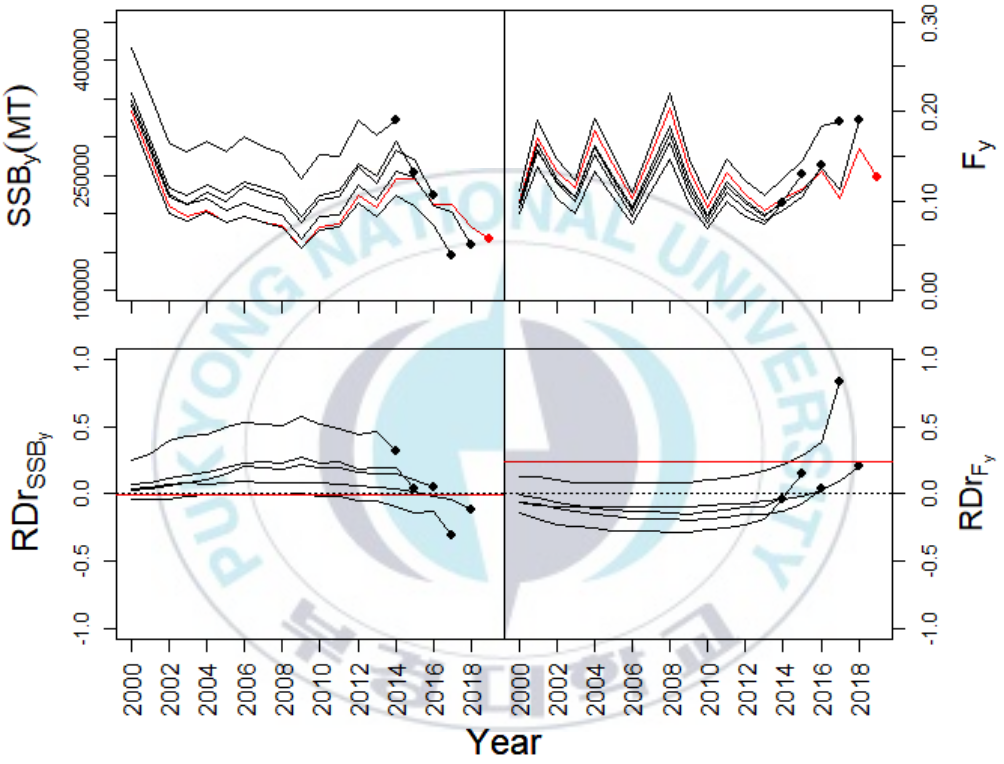


Figure 16. C2의 마지막 년도부터 자료를 제거하면서 추정된 산란가능한 생물량($\boldsymbol{SSB_y}$) 및 순간 어획사망률($\boldsymbol{F_y}$)과 이들의 상대적 차이($\boldsymbol{RDr}$). 빨간 점과 선은 자료를 모두 사용하여 추정한 결과를 나타내며(상부 패널), 빨간 가로선은 $\boldsymbol{Mohn's\ \rho}$ 값을 나타낸다(하부 패널). ($\boldsymbol{Mohn's\ \rho}$ 값: $\boldsymbol{\rho_{SSB}}$ = −0.0063, $\boldsymbol{\rho_F}$= 0.2392).



IV. 고찰

IV. 1. 가용한 자료에 따른 결과 고찰

가용한 자료를 모두 고려한 C1의 경우 C2에 비해 적합도 면에서 자료를 더 잘 설명하는 것으로 나타났으며(Figure 6, Figure 7), C2에서 추정되지 않는 모수 L_{inf}를 추정했다(Table 4). 또한, 실질적인 자료가 고려되지 않는 연도별 가입 마릿수의 편차 추정치는 C2에 비해 변동성이 크지 않은 것으로 나타났다(Figure 9). 이는 동일한 기간의 체장 조성 자료를 사용하더라도 총 어획량 자료 및 CPUE 자료를 추가적으로 고려하는 경우 모델이 자료를 더 잘 설명할 수 있고, 더 많은 모수를 추정할 수 있으며, 연도별 가입 마릿수의 편차를 안정적으로 추정하는 데 영향을 미치는 것으로 해석할 수 있다. 한편, 시뮬레이션 연구에 의하면 C1의 경우는 CPUE 자료의 변동성이 커짐에 따라 수렴률이 현저히 낮아졌고(Table 5), 일부 모수[L_0, K, $\log(q)$]에서 편향이 발생했지만, C2의 경우는 각 시뮬레이션 시나리오에서 높은 수렴률을 나타냈고, 모든 모수의 추정치가 참값을 기준으로 분포하는 것으로 나타났다(Figure 14, Figure 15). 따라서, 모델의 안정성 및 모수 추정치의 정확성 면에서는 C2가 C1에 비해 좋은 것으로 나타났다. C2에서 높은 수렴률을 나타내는 원인으로는 세 가지 정도를 생각해볼 수 있는데, (1).

L_{inf}를 입력 값으로 고려한 점, (2). 추정할 모수의 수가 C1에 비해 상대적으로 적다는 점(연도별 가입 마릿수의 편차), (3). 생성할 가짜 CPUE 자료의 기간이 C1에 비해 상대적으로 짧다는 점 등의 영향으로 생각된다. 결과적으로, 본 논문에서는 가용한 자료에 따라서 적합도 및 추정할 수 있는 모수의 수와 모델의 안정성 및 모수 추정치의 정확성 간에는 트레이드오프(trade-off) 관계를 나타낸다고 볼 수 있다.

IV. 2. 성장 모델의 모수(L_{inf})

본 논문에서는 C1의 경우 가용한 자료를 모두 사용함으로써 성장 모델의 모수(L_{inf})가 33.2 cm로 추정되었지만, C2의 경우에는 추정이 되지 않아 선행 연구를 참고하여 40.6 cm (Shiraishi et al., 2008)값을 입력 값으로 고려했다. 한국 고등어의 이론적 최대 평균 체장(L_{inf})에 대한 선행 연구에 따르면 각각 38.05 cm (Ouchi, 1978), 51.67 cm (Choi et al., 2000), 40.4 cm (Li et al., 2008), 39.2 cm (Crone et al., 2015) 등으로 나타났으며, 이는 C1에서 추정된 L_{inf}값인 33.2 cm와 다소 큰 차이를 보였다. 이러한 차이를 보이는 원인으로는 두 가지 정도를 생각해볼 수 있다. 첫번째로는 선행 연구에서 제시한 L_{inf}값들은 체장 조성 자료 및 연령 자료를 모두 고려하여 본 버틀란피 성장 모델에 적합하여 얻어낸 추정치인 반면, 본 연구에서는 연령 자료를 사용하지 않았다는 점, 그리고 본 버틀란피 성장 모델을 사용했지만 그 외

의 다양한 모델을 복합적으로 고려하여 추정한 결과라는 점을 들 수 있다. 또한, 연구를 진행하는 과정에서 경험적으로 얻은 결론으로는 주어진 자료(총 어획량 자료, CPUE 자료, 체장 조성 자료)에 한해서 SS의 체장-기반 모델을 적용하는 경우 L_{inf} 추정치가 40 cm 이상을 넘지 않는다는 것이다. 이러한 것들을 종합적으로 고려했을 때, L_{inf}의 추정치 33.2 cm값은 주어진 자료 및 적용한 모델에 대해서 수치 최적화를 수행했을 때 가장 적합한 값이라는 판단을 할 수 있다. 만약, 한국 고등어의 연령자료가 추가적으로 고려된다면 한국 고등어 개체군의 L_{inf} 추정치 값은 선행 연구와 비슷한 수준으로 추정될 것으로 예상된다. 반면, C2의 경우 L_{inf}가 추정되지 않아 선행 연구를 참고하여 세 가지의 입력 값을 가정했다. 세 가지 고려한 입력 값은 (1). 51.67 cm (Choi et al., 2000), (2). 40. 6 cm (Shiraishi et al., 2008), (3). 33.2 cm (본 논문)으로 고려했으며, 그 결과 (1), (3)의 경우 추정이 되지 않았으며, (2)의 경우만 성공적으로 추정이 되었다. 따라서, C2의 L_{inf} 값으로 40.6 cm를 고려했다.

IV. 3. 순간 자연사망률(M)

본 연구에서는 순간 자연사망률을 추정하거나 민감도 분석을 통해 적절한 값을 찾으려는 노력에도 불구하고 번번이 수용할 수 없는 결과를 나타냈다. 이미 순간 자연사망률의 추정이 어렵다는 것은 선행 연구(Vetter,

1988)를 통해 잘 알려져 있었기 때문에, 한국 고등어의 순간 자연사망률에 관한 선행 연구(Gim et al., 2020)를 바탕으로 0.38 (yr^{-1})을 입력 값으로 고려했다. 하지만, 자칫 적절치 못한 값을 입력 값으로 적용하는 경우, 추정하고자 하는 생물량 추정치의 편향이 발생할 수 있기 때문에, 이 값에 대한 고찰이 필요하다(Szuwalski, 2016). 따라서, 고등어(Chub mackerel)를 대상으로 순간 자연사망률에 대한 선행 연구 사례를 검토했다. 북태평양 수산 위원회(The North Pacific Fisheries Commission; NPFC)에서 북태평양 고등어(Chub mackerel)의 순간 자연사망률에 대한 연구에서는 적정한 순간 자연사망률 값의 범위를 0.3-0.5 (yr^{-1})로 제시하고, 0.41 (yr^{-1}) 값을 기준 값으로 제시하고 있으며(Takashashi et al., 2019), 또 다른 연구에서는 0.3-0.5 (yr^{-1})의 범위를 제시하고, 0.5 (yr^{-1}) 값을 제시하고 있고(Parrish et al., 1978), 이외에도 순간 자연사망률 값의 범위를 0.4-0.6 (yr^{-1})로 제시하고 있다(Beverton et al., 1959). 이러한 선행 연구들을 바탕으로 0.38 (yr^{-1}) 값은 적절한 수준이라고 판단했다. 반면, 연도 및 연령별로 동일한 순간 자연 사망률 값을 가정하는 경우, 추정된 자원량 크기에 편향(bias)이 나타날 수 있는 문제점이 있다. 일반적으로, 알부터 가입까지의 생활사 기간 동안에는 높은 순간 자연 사망률이 적용된다. 이러한 실정을 무시한 채 0.38 $year^{-1}$ 값이 동일하게 적용된다면, 모든 연령에서 개체 수가 높게 추정되고, 결론적으로 자원량 크기가 실제 자원량 크기보다 높게 추정되어 편향이 발생할 수 있다.

IV. 4. 각 가능도 함수에 고려되는 가중치

부록 2에 제시한 각 가능도 함수에 고려되는 $\log(\sigma_{CPUE})$, N_{eff}, $\log(\sigma_{Rdev})$ 값은 각 가능도 함수의 가중치로서 고려된다.

Sharma et al., 2014의 연구에 의하면, 지수(index) 자료와 체장 조성 자료 간의 가중치는 상대적으로 작용하며, 이는 자원량 추정치에 상당히 영향을 미치는 것으로 나타났다. 또한, Francis, 2011의 연구에 의하면 수산 자원 평가 모델에서 도출되는 결론은 서로 다른 자료에 할당된 상대적 가중치에 의해 크게 좌우될 수 있다는 연구 결과가 있으며, 체장 조성 자료보다는 지수(index) 자료에 더 가중치를 두는 것을 제안했다. 이러한 연구들을 바탕으로 $\log(\sigma_{CPUE})$, N_{eff}, $\log(\sigma_{Rdev})$값을 구하기 위해 민감도 분석의 기준을 다음과 같이 세웠다. 자료가 고려되지 않는 $\log(\sigma_{Rdev})$에 대한 가중치를 상대적으로 가장 낮게 주었고, $\log(\sigma_{CPUE})$, N_{eff}값에 대한 가중치 중에서 지수 자료인 CPUE 자료에 가중치를 상대적으로 높게 주었다. 또한, 앞서 언급한 연구 결과에 따라 CPUE 자료, 체장 조성 자료 간의 가중치를 상대적으로 조절하면서 각 경우의 전반적인 결과를 비교 및 분석하여 최적의 값을 구했다. 결과적으로, 경우별로 AIC값, RMSE값, 추정치의 표준 오차 수준, 모수 추정치 간의 상관 계수 등을 비교한 후, 최적의 값으로 나타난 0.1(CV_{CPUE}), 90(N_{eff}), 80(CV_{Rdev})을 입력 값으로 적용했다.

Ⅳ. 5. 레트로스펙티브 오차 분석

레트로스펙티브 오차 분석은 자료가 추가 혹은 제거될 때, 시계열 추정치인 자원량 크기 또는 어획 사망률의 연속적인 추정치들에 대한 분석을 의미한다(Mohn, 1999). 이 분석을 통해 모델이 레트로스펙티브 오차를 겪는 것을 판단하며, 이는 특정 연도에서 추정치 간의 차이가 크거나, 체계적 패턴의 발생을 기준으로 한다. 체계적 패턴이란 자원 평가 모델에서 시계열 자료를 추가하거나 제거함에 있어 추정되는 자원량의 크기 및 모델 유도값들의 체계적인 패턴을 말하며, 이러한 패턴이 발생하는 경우 추정된 자원량 크기의 심각한 오류를 야기할 수 있다(Hurtado-Ferro et al., 2015). 이 분석을 하는 목적으로는 순간 자연사망률, 어구 선택성 등에 대해 시간 변이성을 고려하지 않고 연도별로 동일하게 가정하는 것에 의해서 발생할 수 있는 자원량 크기 추정치의 편향 문제를 파악하기 위함이다. 본 논문에서 실시한 레트로스펙티브 오차 분석에 의하면, 모델이 체계적 패턴을 나타내지 않았으며, $Mohn's\ \rho$값 또한 레트로스펙티브 오차를 무시할 수 있는 정도로 나타났다(Figure 16). 반면, Figure 16의 상부패널 좌측에서 5개의 연도를 제거했을 때 추정되는 연도별 산란 가능한 자원량(SSB_y) 추정치가 자료를 모두 사용했을 때 추정되는 연도별 SSB_y 추정치(상부패널 좌측 빨간 점과 선)와 다소 차이를 보였다. 이 원인으로 생각해볼 수 있는 점으로는 총 가용한 자료의 기간이 20개의 연도밖에 되지 않기 때문에(C2의 경우

추정할 수 있는 모수의 수가 적고 적합도 면에서 C1에 비해 좋지 않은 점) 5개의 연도를 제거하는 경우 가용한 자료가 15개의 연도밖에 되지 않기 때문에 추정 자체가 쉽지 않고, 추정이 된다 하더라도 20개의 연도를 사용했을 때의 추정치와 상당한 차이를 보일 수 있다고 판단했다.

IV. 6. 한국 고등어 자원량 추정치에 대한 고찰

두 경우의 결과에 따르면 연도별 산란가능한 자원량(SSB_y)은 점차적으로 증가하는 것으로 나타났으며(Figure 8), 연도별 어획사망률은 두 경우에서 비슷한 수준으로 꾸준히 감소하는 추세를 나타냈다(Figure 10). 한편, 자료에 의하면 어획량은 1996년도에 가장 높은 수준으로 나타났지만, 이후 점차 감소한 후 일정한 추세를 나타냈으며(Figure 2), CPUE 자료의 경우 1997년부터 점차 증가하는 추세를 나타냈다(Figure 3). 이는 어획량은 점점 감소하지만 단위노력당 어획량은 증가하고 있는 것으로 해석할 수 있다. 이러한 원인으로는 1999년도부터 고등어를 주요 어획 대상으로 하는 대형 선망 어업에 적용되는 TAC 제도를 포함하여 정부 차원에서 수산 자원 고갈의 우려로 실시하고 있는 다양한 어업 규제 제도인 금어기, 휴어기, 금지체장 지정, 감척 사업 등의 영향으로 볼 수 있다. 따라서, 어획량 및 단위노력당 어획량 자료의 추세 및 본 논문의 자원량 추정치를 종합적으로 고려했을 때, 한국 고등어 개체군의 자원 관리는 효과적으로 이루어지고

있으며, 점차 고등어 자원량이 증가하는 추세를 보이고 있다고 결론 지을 수 있다. 하지만, 이 연구에서 고려하지 못한 점이 있다. 고등어는 회유성 어종이기 때문에 비단 한국 연근해에서만 머무는 것이 아니고 동중국해, 태평양, 동해, 황해를 회유하는 특징을 가진다는 점이다. 본 연구에서는 고등어 자원량이 점차 증가하는 추세를 나타내지만, 한국 연안 외에서 일본·중국에 의해 어획되는 고등어 어획량 수준을 고려하지 못한 한계가 있다. 최근 현대 해양(現代 海洋) 뉴스 기사에 의하면 한·일 어업협정이 수년째 타결되지 않고 있으며, 중국이 한·일 중간 수역에서 불법 호망 어업으로 회유성 어종을 어획하여 일반 선망 어업의 5배 정도를 어획하고 있다는 점을 언급하고 있다(정상원, 2021). 또한, 고등어를 포함한 회유성 어종에 대해 중국과 일본이 함께 참여하는 국제공조를 해야 한다고 주장하고 있다(정석근, 2020). 결론적으로, 본 논문의 결과에 의하면 현재 국내에서 고등어를 대상으로 자원 관리는 잘 이루어지고 있지만, 더 나아가 중국 및 일본과의 국제 공조를 통해 회유성 어종에 대한 효과적인 자원 관리가 수행되어야 한다고 생각한다.

Ⅴ. 참고문헌

Beverton, R., & Holt, S. J. (1959). A review of the lifespans and mortality rates of fish in nature, and their relation to growth and other physiological characteristics. Paper presented at the *CIBA Foundation Colloquia on Ageing, , 5* 142-180.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., . . . Sibert, J. (2012). AD model builder: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Null, 27*(2), 233-249. doi:10.1080/10556788.2011.597854

Fournier, D., & Archibald, C. P. (1982). A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences, 39*(8), 1195-1207. doi:10.1139/f82-157

Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences, 68*(6), 1124-1138. doi:10.1139/f2011-025

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., . . . Punt, A. E. (2015). Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock

assessment models. *ICES Journal of Marine Science, 72*(1), 99-110. doi:10.1093/icesjms/fsu198

Jinwoo Gim. (2019). *A length-based model for korean chub mackerel (scomber japonicus) stock*

Jinwoo Gim, Sang-Yoon Hyun, & Jae Bong Lee. (2020). Management reference points for korea chub mackerel *scomber japonicus* stock. *Korean J Fish Aquat Sci, 53*(6), 942-953. Retrieved from https://doi.org/10.5657/KFAS.2020.0942

Li, G., Chen, X., & Feng, B. (2008). Age and growth of chub mackerel (xcomber japonicus) in the east china and yellow seas using sectioned otolith samples. *Journal of Ocean University of China, 7*, 439-446. doi:10.1007/C11802-008-0439-9

Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research, 142*, 61-74. doi:https://doi.org/10.1016/j.fishres.2012.07.025

Methot, R. D. (1986). *Synthetic estimates of historical abundance and mortality for*

northern anchovy, engraulis mordax. ().

Methot, R. D. (1989). Synthetic estimates of historical and current biomass of northern anchovy, *engraulis mordax*. *Am. Fish. Soc. Sympos.*, , 6,66-82.

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research, 142*, 86-99. doi:https://doi.org/10.1016/j.fishres.2012.10.012

Min Gyou Park. (2021). *A length-based assessment model for the common squid (todarodes pacificus) population*

Mohn, R. (1999). The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES Journal of Marine Science, 56*(4), 473-488. doi:10.1006/jmsc.1999.0481

NFRDI (National Fisheries Research and Development Institute). (2010). *Scomber japonicus. in: Ecology and fishing ground of fisheries resources in korean waters.*

Norio Takahashi, Yasuhiro Kamimura, Momoko Ichinokawa, Shota Nishijima, Ryuji Yukami, Chikako Watanabe & Hiroshi Okamura. (2019). A range of estimates of natural mortality rate for chub mackerel in the north pacific ocean. Retrieved from https://www.npfc.int/range-estimates-natural-mortality-rate-chub-mackerel-north-pacific-ocean

Ouchi, A. (1978). Studies on the age and growth of common mackerel, scomber japonicus, in the waters west of kyushu and east of tsushima islands. *Bull.Seikai Reg.Fish.Res.Lab., 51*, 97-110.

P. R. Crone, & K. T. Hill. (2015). *PACIFIC MACKEREL (scomber japonicus) STOCK ASSESSMENT FOR USA MANAGEMENT IN THE 2015-16 FISHING YEAR.* ().

Parrish, R. H., & MacCall, A. D. (1978). *Climatic variation and exploitation in the pacific mackerel fishery* State of California, Resources Agency, Department of Fish and Game.

Schnute, J., & Fournier, D. (1980). A new approach to Length–Frequency analysis: Growth structure. *Canadian Journal of Fisheries and Aquatic Sciences, 37*(9), 1337-1351. doi:10.1139/f80-172

Sharma, R., Langley, A., Herrera, M., Geehan, J., & Hyun, S. (2014). Investigating the influence of length–frequency data on the stock assessment of indian ocean bigeye tuna. *Fisheries Research; SI: Selectivity, 158*, 50-62. doi:https://doi.org/10.1016/j.fishres.2014.01.012

Shiraishi, T., Okamoto, K., Yoneda, M., Sakai, T., Ohshimo, S., Onoe, S., . . . Matsuyama, M. (2008). Age validation, growth and annual reproductive cycle of chub mackerel scomber japonicus off the waters of northern

kyushu and in the east china sea. *Fisheries Science, 74*(5), 947-954. doi:10.1111/j.1444-2906.2008.01612.x

So Ra Kim, Jung Jin Kim, Hee Won Park, Su Kyung Kang, Hyung Kee Cha, & Hea-Ja Baek. (2020). Maturity and spawning of the chub mackerel *scomber japonicus* in the korean waters. *Korean J Fish Aquat Sci, 53*(1), 9-18.

Szuwalski, C. S. (2016). Biases in biomass estimates: The effect of bin width in size-structured stock assessment methods. *Fisheries Research; Growth: Theory, Estimation, and Application in Fishery Stock Assessment Models, 180*, 169-176. doi:https://doi.org/10.1016/j.fishres.2015.06.023

Vetter, E. F. (1988). Estimation of natural mortality in fish stocks: A review. *Collected Reprints,*

Young Min Choi, Jong Hwa Park, Hyung kee Cha, & Kang Seok Hwang. (2000). Age and growth of common mackerel, *scomber japonicus* houttuyn, in korean waters. *The Korean Society of Fisheries Resource,* , 1-8.

정상원. (2021,). 대형선망수산업협동조합 – 우리나라 연근해어업의 대표주자. *현대 해양*Retrieved from http://www.hdhy.co.kr/news/articleView.html?idxno=13695

정석근. (2020,). ⑪ 선진국 흉내 내는 tac. *현대 해양*Retrieved from http://www.hdhy.co.kr/news/articleView.html?idxno=13511

Ⅵ. 부록

부록 1. 연령−체장 상관표의 유도 과정.

연령-체장 상관표(Age-Length Key; ALK)는 해양과학용어사전에 의하면, “소수의 표본에서 각 개체의 체장과 연령의 관계를 통해 각 체장 계급에 대한 연령분포를 나타낸 표”를 의미한다. 하지만, SS의 체장-기반 모델에서 연령-체장 상관표는 그 의미가 다르다. SS의 체장-기반 모델에서의 연령-체장 상관표는 ‘각 가상의 연령 계급에 대한 체장의 분포를 나타낸 표’라는 의미로 사용된다. 이를 통해 각 가상의 연령 계급에 해당하는 체장 분포를 알 수 있고, 체장-기반 모델을 연령 구조 모델화할 수 있다. 연령-체장 상관표를 유도하기 위해서 연령별 체장 분포에 대해 두 가지 가정이 적용되었다. 첫 번째로는 연령별 체장은 정규 분포를 따른다는 가정이다[식 (1.1)]. 두 번째로, 연령별 체장의 변동 계수(Coefficient of Variation; CV)가 선형적으로 변할 때, 연령별 체장 분포의 표준편차는 식 (1.2)를 통해 계산된다는 것이다. SS에서는 연령별 평균 체장과 변동 계수의 정보를 이용하여 연령별 체장 분포의 표준 편차를 유도한다.

$$X_a \sim N(L_a, \sigma_a^2); \quad for\ 0 \le a \le A \tag{1.1}$$

체장에 대한 확률변수 X_a는 성장 모델[식 (6)]로부터 계산된 연령별 평균 체장(L_a)을 평균으로 하고 식 (1.2)로부터 계산된 연령별 표준 편차 값(σ_a)을 표준 편차로 하는 정규 분포를 따른다.

$$\begin{aligned} \sigma_0 &= CV_0 \cdot L_0; && for\ a = 0 \\ \sigma_a &= L_a \cdot \left(CV_0 + \frac{L'_a - L_0}{L_A - L_0} \cdot (CV_A - CV_0) \right); \\ & && for\ 0 < a < A \\ \sigma_A &= CV_A \cdot L_A; && for\ a = A \end{aligned} \tag{1.2}$$

CV_0는 0세 체장의 변동 계수를 의미하며, CV_A는 6+세 체장의 변동 계수를 의미한다. 이 값들은 추정하거나 입력 값으로 지정함으로써 나머지 연령별 체장 분포의 표준 편차를 유도할 수 있지만, 본 논문에서는 두 값에 대해 민감도 분석을 실시하여 두 값의 최적의 값(CV_0= 0.1, CV_A= 0.1)을 입력 값으로 지정했다. 결과적으로, 연령별 체장 분포의 표준 편차는 연령별 평균 체장에 연령별로 동일한 CV값 0.1이 곱해진 값으로 유도해낼 수 있다. 연령별 체장 분포의 평균과 표준 편차가 구해지면, 연령별 체장에 대한 확률

밀도함수(probability density function)를 얻을 수 있다. 이후, 연령-체장 상관표를 구하기 위해서는 앞서 식 (1.1), (1.2)로부터 구한 연령별 체장에 대한 확률밀도함수를 확률질량함수(probability mass function)로 변환하는 과정이 필요하다. 이 과정은 우선, 연령별 체장에 대한 연속형 확률변수 X_a를 이산형 확률 변수로 취급하고, 연령별 체장 분포를 정규화(normalization)하는 것으로부터 시작된다. 연령별 체장 분포의 정규화는 다음 식 (1.3)을 통해서 계산된다.

$$Z_a = \frac{X - L_a}{\sigma_a} \sim N(0,1^2); \quad for \;\; X = 0, 1, \cdots, \infty \tag{1.3}$$

확률 변수 Z_a는 체장에 대한 이산형 확률변수 X에 대한 함수를 뜻한다. 이렇게 정규화 된 식 (1.3)에서 확률변수 X가 $l(0, 1, \cdots, 52)$일 때, 각 체장 계급에 대응하는 누적 확률 값은 다음 주어진 누적 분포 함수로부터 유도할 수 있다. 식 (1.4):

$$\Pr\left(Z_a \le Z_a(l)\right) = \int_{-\infty}^{Z_a(l)} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy; \quad for\ l = 0, 1, \cdots, 52 \tag{1.4}$$

$Z_a(l)$은 식 (1.3)에서 확률변수 $X = l(0, 1, \cdots, 52)$ 일 때, 유도되는 값이며, $F_Z(Z_a(l))$ 또는 $\Pr(Z_a \le Z_a(l))$는 식 (1.3)에서 계산된 $Z_a(l)$을 식 (1.4)에 대입했을 때 연령별 체장 계급별로 계산되는 누적 확률 값을 뜻한다. 이후, 연령별 각 체장 계급에 대응하는 확률 값은 다음 두 누적 확률 값의 차이로 구할 수 있다. 식 (1.5):

$$P_a(l) = Pr(Z_a \le Z_a(l+1)) - Pr(Z_a \le Z_a(l)); \quad for\ l = 0,1,\cdots,52 \tag{1.5}$$

$P_a(l)$은 연령별 각 체장 계급에 대응하는 확률값을 나타내며, 마지막 체장 계급에서의 누적 분포 확률값의 차를 구하기 위해 $P_a(53)$은 1로 가정했다 (누적 분포 함수의 특징). 마지막으로, 계산된 각 확률값에 0.0001이라는 기준을 세워 기준 이상의 값 만을 수용했다. 0.0001이라는 값은 SS에서 제시하는 기준 값으로, 본 논문에서도 이를 적용하였다. 두 경우(C1, C2)에서 추정된 성장 모델의 모수 추정치를 통해 유도되는 각 연도 초기의 연령-체장 상관표를 그래프로 제시했다(Figure 13).

부록 2. 목적 함수를 구성하는 각 가능도 함수.

목적 함수(식 21)를 구성하는 세 가지 음의 로그 가능도 함수는 다음과 같이 구성된다. 먼저, 로그 정규 분포를 가정한 $CPUE$자료의 음의 로그 가능도 함수는 다음 식 (2.1)으로 구성된다.

$$-\log L(\boldsymbol{\theta}\,|\boldsymbol{CPUE}) =$$

$$\frac{1}{2}\log(\sigma_{CPUE}) + \sum_{y=styr}^{endyr} \frac{\left(\log(CPUE_y) - \log\left(\widehat{CPUE_y}\right)\right)^2}{2\sigma_{CPUE}^2}; \qquad (2.1)$$

$CPUE_y$는 연도별 CPUE 자료이며, CPUE 자료의 표준편차(σ_{CPUE})는 목적 함수를 구성하는 각 가능도 함수의 가중치 와도 관련되어 있기 때문에, 민감도 분석을 통해 선정된 최적의 값을 입력 값으로 고려했다. $styr$은 시작 년도를, $endyr$은 마지막 년도를 뜻한다. 다항분포를 고려한 체장 조성 자료의 음의 로그 가능도 함수는 다음 식 (2.2)으로 구성된다.

$$-\log L(\boldsymbol{\theta}\,|\boldsymbol{Length}) =$$

$$\sum_{y=styr}^{endyr} \sum_{l=10}^{52} N_{eff} \cdot p_{y,l} \cdot \log\left(p_{y,l}/\hat{p}_{y,l}\right); \qquad (2.2)$$

$P_{y,l}$은 연도별 체장 계급별 비율 자료를 나타내며, 다항분포에 고려되는 연도별 체장 자료의 표본 크기(N_{eff})는 연도별로 동일하게 적용하였다. N_{eff}는 목적 함수를 구성하는 각 가능도 함수의 가중치와도 관련되어 있기 때문에, 민감도 분석을 통해 최적의 값을 선정하여 입력 값으로 고려했다. 로그 정규 분포를 가정한 연도별 가입 마릿수의 편차($Rdev_y$)의 음의 로그 가능도 함수는 다음 식 (2.3)으로 구성된다.

$$-\log L(\boldsymbol{\theta}\,|\boldsymbol{Rdev}) = \frac{1}{2}\log(\sigma_{Rdev}) + \sum_{y=styr}^{endyr}\frac{Rdev_y^2}{2\sigma_{Rdev}^2}; \quad (2.3)$$

연도별 가입 마릿수의 편차($Rdev_y$)는 자유 모수로 추정되며, σ_{Rdev}는 목적 함수를 구성하는 각 가능도 함수의 가중치 와도 관련되어 있기 때문에, 민감도 분석을 통해 최적의 값을 선정하여 입력 값으로 고려했다. 식 (2.1), (2.2), (2.3)의 합으로 구성된 목적 함수(Obj)값이 최소가 되게 하는 모수들($\boldsymbol{\theta}$)을 추정했다.

부록 3. 한국 고등어 자원 평가를 위해 실제로 사용된 SS의 네 가지 구성 파일(starter.ss, data.ss, control.ss, forecast.ss).

국립수산과학원(NIFS)에서 제공받은 CPUE 자료 및 체장 조성 자료는 사정상 비공개로 했고, 국가 통계 포털(KOSIS)에서 제공하는 총 어획량(total catch in MT) 자료만을 data.ss 파일에 제시했다. starter.ss 파일과 forecast.ss 파일의 경우, 두 경우(C1, C2)에 동일하게 적용되었으며, data.ss 파일과 control.ss 파일의 경우, 사용한 자료의 기간만 다르게 적용된 점, C2에서 L_{inf}를 추정하지 않고 40.6을 입력 값으로 지정한 점 외에는 다른 것이 없기 때문에, 실제로 사용된 SS의 네 가지 구성 파일은 C1에 해당하는 것만을 제시했다.

starter.ss

#V3.30.16.00;_2020_09_03;_safe;_Stock_Synthesis_by_Richard_Methot_(NOAA)_using_ADMB_12.2
#Stock Synthesis (SS) is a work of the U.S. Government and is not subject to copyright protection in the United States.
#Foreign copyrights may apply. See copyright.txt for more information.
#_user_support_available_at:NMFS.Stock.Synthesis@noaa.gov
#_user_info_available_at:https://vlab.ncep.noaa.gov/group/stock-synthesis
#C starter comment here
data.ss
control.ss
0 # 0=use init values in control file; 1=use ss.par
0 # run display detail (0,1,2)
1 # detailed output (0=minimal for data-limited, 1=high (w/ wtatage.ss_new), 2=brief, 3=custom)
custom report options: -100 to start with minimal; -101 to start with all; -number to remove, +number to add, -999 to end
0 # write 1st iteration details to echoinput.sso file (0,1)
0 # write parm values to ParmTrace.sso (0=no,1=good,active; 2=good,all; 3=every_iter,all_parms; 4=every,active)
1 # write to cumreport.sso (0=no,1=like&timeseries; 2=add survey fits)
0 # Include prior_like for non-estimated parameters (0,1)
1 # Use Soft Boundaries to aid convergence (0,1) (recommended)
#
2 # Number of datafiles to produce: 1st is input, 2nd is estimates, 3rd and higher are bootstrap, 0 turns off all *.ss_new output
6 # Turn off estimation for parameters entering after this phase
#
0 # MCeval burn interval
1 # MCeval thin interval
0 # jitter initial parm value by this fraction
-1 # min yr for sdreport outputs (-1 for styr); #_1995
-1 # max yr for sdreport outputs (-1 for endyr+1; -2 for endyr+Nforecastyrs); #_2019
0 # N individual STD years
#vector of year values

0.0001 # final convergence criteria (e.g. 1.0e-04)
0 # retrospective year relative to end year (e.g. -4)
1 # min age for calc of summary biomass
1 # Depletion basis: denom is: 0=skip; 1=rel X*SPB0; 2=rel SPBmsy; 3=rel X*SPB_styr; 4=rel X*SPB_endyr
1 # Fraction (X) for Depletion denominator (e.g. 0.4)
4 # SPR_report_basis: 0=skisp; 1=(1-SPR)/(1-SPR_tgt); 2=(1-SPR)/(1-SPR_MSY); 3=(1-SPR)/(1-SPR_Btarget); 4=rawSPR

1 # Annual_F_units: 0=skip; 1=exploitation(Bio); 2=exploitation(Num); 3=sum(Apical_F's); 4=true F for range of ages; 5=unweighted avg. F for range of ages #COND 0 6 #_min and max age over which average F will be calculated with F_reporting=4 or 5
0 # F_std_basis: 0=raw_annual_F; 1=F/Fspr; 2=F/Fmsy; 3=F/Fbtgt; where F means annual_F; values >=11 invoke multiyr with 10's digit
0 # MCMC output detail: integer part (0=default; 1=adds obj func components); and decimal part (added to SR_LN(R0) on first call to mcmc)
0.0001 # ALK tolerance (example 0.0001)
-1 # random number seed for bootstrap data (-1 to use long(time) as seed): # 1599073571
3.30 # check value for end of file and for version control

data.ss

#V3.30.15.00-
safe;_2020_03_26;_Stock_Synthesis_by_Richard_Methot_(NOAA)_using_ADMB_
12.0
#Stock Synthesis (SS) is a work of the U.S. Government and is not subject to copyright protection in the United States.
#Foreign copyrights may apply. See copyright.txt for more information.
#_user_support_available_at:NMFS.Stock.Synthesis@noaa.gov
#_user_info_available_at:https://vlab.ncep.noaa.gov/group/stock-synthesis
#_Start_time: Fri Mar 05 16:36:25 2021
#_Number_of_datafiles: 2

#_observed data:
#V3.30.15.00-
safe;_2020_03_26;_Stock_Synthesis_by_Richard_Methot_(NOAA)_using_ADMB_
12.0
#Stock Synthesis (SS) is a work of the U.S. Government and is not subject to copyright protection in the United States.
#Foreign copyrights may apply. See copyright.txt for more information.
1975 #_StartYr
2019 #_EndYr
1 #_Nseas
 12 #_months/season
2 #_Nsubseasons (even number, minimum is 2)
1 #_spawn_month
-1 #_Ngenders: 1, 2, -1 (use -1 for 1 sex setup with SSB multiplied by female_frac parameter)
6 #_Nages=accumulator age, first age is always age 0
1 #_Nareas
1 #_Nfleets (including surveys)
#_fleet_type: 1=catch fleet; 2=bycatch only fleet; 3=survey; 4=ignore
#_sample_timing: -1 for fishing fleet to use season-long catch-at-age for observations, or 1 to use observation month; (always 1 for surveys)
#_fleet_area: area the fleet/survey operates in
#_units of catch: 1=bio; 2=num (ignored for surveys; their units read later)
#_catch_mult: 0=no; 1=yes
#_rows are fleets
#_fleet_type fishery_timing area catch_units need_catch_mult fleetname
 1 -1 1 1 0 0 # 1
#Bycatch_fleet_input_goes_next
#a: fleet index
#b: 1=include dead bycatch in total dead catch for F0.1 and MSY optimizations and forecast ABC; 2=omit from total catch for these purposes (but still include the mortality)
#c: 1=Fmult scales with other fleets; 2=bycatch F constant at input value; 3=bycatch

F from range of years
#d: F or first year of range
#e: last year of range
#f: not used
a b c d e f
#_Catch data: yr, seas, fleet, catch, catch_se
#_catch_se: standard error of log(catch)
#_NOTE: catch data is ignored for survey fleets
-999 1 1 0 0.05
1975 1 1 70123 0.001
1976 1 1 107382 0.001
1977 1 1 113051 0.001
1978 1 1 99519 0.001
1979 1 1 120283 0.001
1980 1 1 62690 0.001
1981 1 1 108082 0.001
1982 1 1 99447 0.001
1983 1 1 122883 0.001
1984 1 1 101714 0.001
1985 1 1 68479 0.001
1986 1 1 103511 0.001
1987 1 1 101337 0.001
1988 1 1 162828 0.001
1989 1 1 163617 0.001
1990 1 1 96297 0.001
1991 1 1 89738 0.001
1992 1 1 115619 0.001
1993 1 1 174684 0.001
1994 1 1 210442 0.001
1995 1 1 200481 0.001
1996 1 1 415003 0.001
1997 1 1 160448 0.001
1998 1 1 172925 0.001
1999 1 1 177540 0.001
2000 1 1 145908 0.001
2001 1 1 203717 0.001
2002 1 1 141751 0.001
2003 1 1 122044 0.001
2004 1 1 184274 0.001
2005 1 1 135596 0.001
2006 1 1 101427 0.001
2007 1 1 143776 0.001
2008 1 1 187240 0.001
2009 1 1 117960 0.001
2010 1 1 94331 0.001

2011 1 1 138729 0.001
2012 1 1 125143 0.001
2013 1 1 102114 0.001
2014 1 1 127452 0.001
2015 1 1 131735 0.001
2016 1 1 133200 0.001
2017 1 1 103870 0.001
2018 1 1 141530 0.001
2019 1 1 101121 0.001
-9999 0 0 0 0
#
 #_CPUE_and_surveyabundance_observations
#_Units: 0=numbers; 1=biomass; 2=F; 30=spawnbio; 31=recdev; 32=spawnbio*recdev; 33=recruitment; 34=depletion(&see Qsetup); 35=parm_dev(&see Qsetup)
#_Errtype: -1=normal; 0=lognormal; >0=T
#_SD_Report: 0=no sdreport; 1=enable sdreport
#_Fleet Units Errtype SD_Report
1 1 0 0 # FISHERY
#_yr month fleet obs stderr
1975 1 1 CPUE 0.0997513 #_ FISHERY
.
.
.
2019 1 1 CPUE 0.0997513 #_ FISHERY
-9999 1 1 1 1 # terminator for survey observations
#
0 #_N_fleets_with_discard
#_discard_units (1=same_as_catchunits(bio/num); 2=fraction; 3=numbers)
#_discard_errtype: >0 for DF of T-dist(read CV below); 0 for normal with CV; -1 for normal with se; -2 for lognormal; -3 for trunc normal with CV
note, only have units and errtype for fleets with discard
#_Fleet units errtype
-9999 0 0 0.0 0.0 # terminator for discard data
#
0 #_use meanbodysize_data (0/1)
#_COND_0 #_DF_for_meanbodysize_T-distribution_like
note: type=1 for mean length; type=2 for mean body weight
#_yr month fleet part type obs stderr
-9999 0 0 0 0 0 0 # terminator for mean body size data
set up population length bin structure (note - irrelevant if not using size data and using empirical wtatage
2 # length bin method: 1=use databins; 2=generate from binwidth,min,max below; 3=read vector
1 # binwidth for population size comp

1 # minimum size in the population (lower edge of first bin and size at age 0.00)
52 # maximum size in the population (lower edge of last bin)
1 # use length composition data (0/1)
#_mintailcomp: upper and lower distribution for females and males separately are accumulated until exceeding this level.
#_addtocomp: after accumulation of tails; this value added to all bins
#_males and females treated as combined gender below this bin number
#_compressbins: accumulate upper tail by this number of bins; acts simultaneous with mintailcomp; set=0 for no forced accumulation
#_Comp_Error: 0=multinomial, 1=dirichlet
#_Comp_Error2: parm number for dirichlet
#_minsamplesize: minimum sample size; set to 1 to match 3.24, minimum value is 0.001
#_mintailcomp addtocomp combM+F CompressBins CompError ParmSelect minsamplesize
-999 1e-10 0 0 0 0 0 1 #_fleet:1_0
sex codes: 0=combined; 1=use female only; 2=use male only; 3=use both as joint sexxlength distribution
partition codes: (0=combined; 1=discard; 2=retained
43 #_N_LengthBins; then enter lower edge of each length bin
10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
#_yr month fleet sex part Nsamp datavector(female-male)
// p: proportion
2000 1 1 3 0 90 p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12
p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25 p26 p27 p28 p29 p30
p31 p32 p33 p34 p35 p36 p37 p38 p39 p40 p41 p42 p43
.
.
.
2019 1 1 3 0 90 p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12
p13 p14 p15 p16 p17 p18 p19 p20 p21 p22 p23 p24 p25 p26 p27 p28 p29 p30
p31 p32 p33 p34 p35 p36 p37 p38 p39 p40 p41 p42 p43
-9999 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
#
0 #_N_age_bins
0 #_N_ageerror_definitions
#_mintailcomp: upper and lower distribution for females and males separately are accumulated until exceeding this level.
#_addtocomp: after accumulation of tails; this value added to all bins
#_males and females treated as combined gender below this bin number
#_compressbins: accumulate upper tail by this number of bins; acts simultaneous with mintailcomp; set=0 for no forced accumulation
#_Comp_Error: 0=multinomial, 1=dirichlet

#_Comp_Error2: parm number for dirichlet
#_minsamplesize: minimum sample size; set to 1 to match 3.24, minimum value is 0.001
#_mintailcomp addtocomp combM+F CompressBins CompError ParmSelect minsamplesize
#_fleet:1_0
#_Lbin_method_for_Age_Data: 1=poplenbins; 2=datalenbins; 3=lengths
sex codes: 0=combined; 1=use female only; 2=use male only; 3=use both as joint sexxlength distribution
partition codes: (0=combined; 1=discard; 2=retained
#_yr month fleet sex part ageerr Lbin_lo Lbin_hi Nsamp datavector(female-male)
#
0 #_Use_MeanSize-at-Age_obs (0/1)
#
0 #_N_environ_variables
#Yr Variable Value
#
0 # N sizefreq methods to read
#
0 # do tags (0/1)
#
0 # morphcomp data(0/1)
Nobs, Nmorphs, mincomp
yr, seas, type, partition, Nsamp, datavector_by_Nmorphs
#
0 # Do dataread for selectivity priors(0/1)
Yr, Seas, Fleet, Age/Size, Bin, selex_prior, prior_sd
feature not yet implemented
#
999

control.ss

#Stock Synthesis (SS) is a work of the U.S. Government and is not subject to copyright protection in the United States.
#Foreign copyrights may apply. See copyright.txt for more information.
#_user_support_available_at:NMFS.Stock.Synthesis@noaa.gov
#_user_info_available_at:https://vlab.ncep.noaa.gov/group/stock-synthesis
#_data_and_control_files: data.ss // control.ss
0 # 0 means do not read wtatage.ss; 1 means read and use wtatage.ss and also read and use growth parameters
1 #_N_Growth_Patterns (Growth Patterns, Morphs, Bio Patterns, GP are terms used interchangeably in SS)
1 #_N_platoons_Within_GrowthPattern
#_Cond 1 #_Platoon_within/between_stdev_ratio (no read if N_platoons=1)
#_Cond 1 #vector_platoon_dist_(-1_in_first_val_gives_normal_approx)
#
4 # recr_dist_method for parameters: 2=main effects for GP, Area, Settle timing; 3=each Settle entity; 4=none (only when N_GP*Nsettle*pop==1)
1 # not yet implemented; Future usage: Spawner-Recruitment: 1=global; 2=by area
1 # number of recruitment settlement assignments
0 # unused option
#GPattern month area age (for each settlement assignment)
 1 1 1 0
#
#_Cond 0 # N_movement_definitions goes here if Nareas > 1
#_Cond 1.0 # first age that moves (real age at begin of season, not integer) also cond on do_migration>0
#_Cond 1 1 1 2 4 10 # example move definition for seas=1, morph=1, source=1 dest=2, age1=4, age2=10
#
0 #_Nblock_Patterns
#_Cond 0 #_blocks_per_pattern
begin and end years of blocks
#
controls for all timevary parameters
1 #_env/block/dev_adjust_method for all time-vary parms (1=warn relative to base parm bounds; 3=no bound check)
#
AUTOGEN
 1 1 1 1 1 # autogen: 1st element for biology, 2nd for SR, 3rd for Q, 4th reserved, 5th for selex
where: 0 = autogen time-varying parms of this category; 1 = read each time-varying parm line; 2 = read then autogen if parm min==-12345
#
#_Available timevary codes

#_Block types: 0: P_block=P_base*exp(TVP); 1: P_block=P_base+TVP; 2: P_block=TVP; 3: P_block=P_block(-1) + TVP
#_Block_trends: -1: trend bounded by base parm min-max and parms in transformed units (beware); -2: endtrend and infl_year direct values; -3: end and infl as fraction of base range
#_EnvLinks: 1: P(y)=P_base*exp(TVP*env(y)); 2: P(y)=P_base+TVP*env(y); 3: null; 4: P(y)=2.0/(1.0+exp(-TVP1*env(y) - TVP2))
#_DevLinks: 1: P(y)*=exp(dev(y)*dev_se; 2: P(y)+=dev(y)*dev_se; 3: random walk; 4: zero-reverting random walk with rho; 21-24 keep last dev for rest of years
#
#_Prior_codes: 0=none; 6=normal; 1=symmetric beta; 2=CASAL's beta; 3=lognormal; 4=lognormal with biascorr; 5=gamma
#
setup for M, growth, maturity, fecundity, recruitment distibution, movement
#
0#_natM_type:_0=1Parm;1=N_breakpoints;_2=Lorenzen;_3=agespecific;_4=agespec_withseasinterpolate
  #_no additional input for selected M option; read 1P per morph
#
1 # GrowthModel: 1=vonBert with L1&L2; 2=Richards with L1&L2; 3=age_specific_K_incr; 4=age_specific_K_decr; 5=age_specific_K_each; 6=NA; 7=NA; 8=growth cessation
0 #_Age(post-settlement)_for_L1;linear growth below this
999 #_Growth_Age_for_L2 (999 to use as Linf)
-999 #_exponential decay for growth above maxage (value should approx initial Z; -999 replicates 3.24; -998 to not allow growth above maxage)
0 #_placeholder for future growth feature
#
0 #_SD_add_to_LAA (set to 0.1 for SC2 V1.x compatibility)
0 #_CV_Growth_Pattern: 0 CV=f(LAA); 1 CV=F(A); 2 SD=F(LAA); 3 SD=F(A); 4 logSD=F(A)
#
1 #_maturity_option: 1=length logistic; 2=age logistic; 3=read age-maturity matrix by growth_pattern; 4=read age-fecundity; 5=disabled; 6=read length-maturity
2 #_First_Mature_Age
1 #_fecundity option:(1)eggs=Wt*(a+b*Wt);(2)eggs=a*L^b;(3)eggs=a*Wt^b; (4)eggs=a+b*L; (5)eggs=a+b*W
0 #_hermaphroditism option: 0=none; 1=female-to-male age-specific fxn; -1=male-to-female age-specific fxn
1 #_parameter_offset_approach (1=none, 2= M, G, CV_G as offset from female-GP1, 3=like SC2 V1.x)
#
#_growth_parms
#_ LO HI INIT PRIOR PR_SD PR_type PHASE env_var&link dev_link dev_minyr dev_maxyr dev_PH Block Block_Fxn

```
# Sex: 1  BioPattern: 1  NatMort
 0.1 1.5 0.38 0 0 0 -3 0 0 0 0 0 0 0 # NatM_p_1_Fem_GP_1
# Sex: 1  BioPattern: 1  Growth
 5 18 9.37417 0 0 0 2 0 0 0 0 0 0 0 # L_at_Amin_Fem_GP_1
 25 50 33.1963 0 0 0 4 0 0 0 0 0 0 0 # L_at_Amax_Fem_GP_1
 0.01 2 0.629513 0 0 0 2 0 0 0 0 0 0 0 # VonBert_K_Fem_GP_1
 0.05 0.25 0.1 0 0 0 -4 0 0 0 0 0 0 0 # CV_young_Fem_GP_1
 0.05 0.25 0.1 0 0 0 -4 0 0 0 0 0 0 0 # CV_old_Fem_GP_1
# Sex: 1  BioPattern: 1  WtLen
 -3 3 2.8278e-06 0.003 0.8 0 -3 0 0 0 0 0 0 0 # Wtlen_1_Fem_GP_1
 -3 4 3.4428 3.34694 0.8 0 -3 0 0 0 0 0 0 0 # Wtlen_2_Fem_GP_1
# Sex: 1  BioPattern: 1  Maturity&Fecundity
 25 35 29.1704 30 0.8 0 -3 0 0 0 0 0 0 0 # Mat50%_Fem_GP_1
 -3 3 -0.81689 -0.25 0.8 0 -3 0 0 0 0 0 0 0 # Mat_slope_Fem_GP_1
 -5 5 1 0 0 0 -3 0 0 0 0 0 0 0 # Eggs/kg_inter_Fem_GP_1
 -5 5 0 0 0 0 -3 0 0 0 0 0 0 0 # Eggs/kg_slope_wt_Fem_GP_1
# Hermaphroditism
#  Recruitment Distribution
#  Cohort growth dev base
 0.1 10 1 1 1 0 -1 0 0 0 0 0 0 0 # CohortGrowDev
#  Movement
#  Age Error from parameters
#  catch multiplier
#  fraction female, by GP
 1e-06 0.999999 0.6 0.5 0.5 0 -99 0 0 0 0 0 0 0 # FracFemale_GP_1
#
#_no timevary MG parameters
#
#_seasonal_effects_on_biology_parms
 0  0  0  0  0  0  0  0  0  0
#_femwtlen1,femwtlen2,mat1,mat2,fec1,fec2,Malewtlen1,malewtlen2,L1,K
#_ LO HI INIT PRIOR PR_SD PR_type PHASE
#_Cond -2 2 0 0 -1 99 -2 #_placeholder when no seasonal MG parameters
#
3 #_Spawner-Recruitment; Options: 1=NA; 2=Ricker; 3=std_B-H; 4=SCAA; 5=Hockey; 6=B-H_flattop; 7=survival_3Parm; 8=Shepherd_3Parm; 9=RickerPower_3parm
0  # 0/1 to use steepness in initial equ recruitment calculation
0  #  future feature:  0/1 to make realized sigmaR a function of SR curvature
#_  LO  HI  INIT  PRIOR  PR_SD  PR_type  PHASE  env-var  use_dev  dev_mnyr  dev_mxyr  dev_PH  Block  Blk_Fxn #  parm_name
3  31  14.5189  0  0  0  1  0  0  0  0  0  0  0 # SR_LN(R0)
0.2  1  1  0  0  0  -4  0  0  0  0  0  0  0 # SR_BH_steep
0  2  0.703346 0  0  0  -4  0  0  0  0  0  0  0 # SR_sigmaR
```

-5 5 0 0 1 0 -4 0 0 0 0 0 0 0 # SR_regime
0 0 0 0 0 0 -99 0 0 0 0 0 0 0 # SR_autocorr
#_no timevary SR parameters
1 #do_recdev: 0=none; 1=devvector (R=F(SSB)+dev); 2=deviations (R=F(SSB)+dev); 3=deviations (R=R0*dev; dev2=R-f(SSB)); 4=like 3 with sum(dev2) adding penalty
1975 # first year of main recr_devs; early devs can preceed this era
2019 # last year of main recr_devs; forecast devs start in following year
2 #_recdev phase
0 # (0/1) to read 13 advanced options
#_Cond 0 #_recdev_early_start (0=none; neg value makes relative to recdev_start)
#_Cond -4 #_recdev_early_phase
#_Cond 0 #_forecast_recruitment phase (incl. late recr) (0 value resets to maxphase+1)
#_Cond 1 #_lambda for Fcast_recr_like occurring before endyr+1
#_Cond 975 #_last_yr_nobias_adj_in_MPD; begin of ramp
#_Cond 1969 #_first_yr_fullbias_adj_in_MPD; begin of plateau
#_Cond 2019 #_last_yr_fullbias_adj_in_MPD
#_Cond 2020 #_end_yr_for_ramp_in_MPD (can be in forecast to shape ramp, but SS sets bias_adj to 0.0 for fcast yrs)
#_Cond 1 #_max_bias_adj_in_MPD (typical ~0.8; -3 sets all years to 0.0; -2 sets all non-forecast yrs w/ estimated recdevs to 1.0; -1 sets biasadj=1.0 for all yrs w/ recdevs)
#_Cond 0 #_period of cycles in recruitment (N parms read below)
#_Cond -5 #min rec_dev
#_Cond 5 #max rec_dev
#_Cond 0 #_read_recdevs
#_end of advanced SR options
#
#_placeholder for full parameter lines for recruitment cycles
read specified recr devs
#_Yr Input_value
#
all recruitment deviations
1975R 1976R 1977R 1978R 1979R 1980R 1981R 1982R 1983R 1984R 1985R 1986R 1987R 1988R 1989R 1990R 1991R 1992R 1993R 1994R 1995R 1996R 1997R 1998R 1999R 2000R 2001R 2002R 2003R 2004R 2005R 2006R 2007R 2008R 2009R 2010R 2011R 2012R 2013R 2014R 2015R 2016R 2017R 2018R 2019R
-0.822743 -0.505766 -0.063263 -0.677131 -0.412869 -0.0445883 -0.26676 -1.08792 -1.53822 -0.955114 0.602746 -0.237221 0.648728 -0.682834 -0.42979 0.135067 0.361908 0.518473 -0.599505 1.46583 0.299991 -0.149598 0.0507184 -0.977841 0.608665 0.697295 0.19159 0.10123 0.43923 -0.515205 0.546178 0.397163 0.65315 0.104765 0.545757 0.427426 0.208603 0.289195 0.383026 -0.13515 -0.310462 0.881607 0.686663 -0.397588 -0.435434
#
#Fishing Mortality info
0.3 # F ballpark value in units of annual_F

-2001 # F ballpark year (neg value to disable)
3 # F_Method: 1=Pope; 2=instan. F; 3=hybrid (hybrid is recommended)
3.9 # max F or harvest rate, depends on F_Method
no additional F input needed for Fmethod 1
if Fmethod=2; read overall start F value; overall phase; N detailed inputs to read
if Fmethod=3; read N iterations for tuning for Fmethod 3
4 # N iterations for tuning F in hybrid method (recommend 3 to 7)
#
#_initial_F_parms; count = 0
#_ LO HI INIT PRIOR PR_SD PR_type PHASE
#2019 2070
F rates by fleet
Yr: 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019
seas: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 0.0804866 0.139591 0.184997 0.200339 0.289824 0.177258 0.328208 0.312831 0.447199 0.523336 0.490055 0.560223 0.370047 0.479731 0.450478 0.298991 0.280964 0.316221 0.411119 0.514161 0.399501 0.662053 0.2866 0.351056 0.454881 0.420144 0.488708 0.294858 0.242969 0.370111 0.288948 0.217851 0.277622 0.33383 0.198406 0.150044 0.207267 0.186492 0.154503 0.194452 0.211578 0.245565 0.189739 0.213805 0.144124
#
#_Q_setup for fleets with cpue or survey data
#_1: fleet number
#_2: link type: (1=simple q, 1 parm; 2=mirror simple q, 1 mirrored parm; 3=q and power, 2 parm; 4=mirror with offset, 2 parm)
#_3: extra input for link, i.e. mirror fleet# or dev index number
#_4: 0/1 to select extra sd parameter
#_5: 0/1 for biasadj or not
#_6: 0/1 to float
#_ fleet link link_info extra_se biasadj float # fleetname
1 1 0 0 0 0 # FISHERY
-9999 0 0 0 0 0
#
#_Q_parms(if_any);Qunits_are_ln(q)
#_ LO HI INIT PRIOR PR_SD PR_type PHASE env-var use_dev dev_mnyr dev_mxyr dev_PH Block Blk_Fxn # parm_name
-50 5 -10.2413 0 0 0 4 0 0 0 0 0 0 0 # LnQ_base_0(1)
#_no timevary Q parameters
#
#_size_selex_patterns

#Pattern:_0; parm=0; selex=1.0 for all sizes
#Pattern:_1; parm=2; logistic; with 95% width specification
#Pattern:_5; parm=2; mirror another size selex; PARMS pick the min-max bin to mirror
#Pattern:_15; parm=0; mirror another age or length selex
#Pattern:_6; parm=2+special; non-parm len selex
#Pattern:_43; parm=2+special+2; like 6, with 2 additional param for scaling (average over bin range)
#Pattern:_8; parm=8; New doublelogistic with smooth transitions and constant above Linf option
#Pattern:_9; parm=6; simple 4-parm double logistic with starting length; parm 5 is first length; parm 6=1 does desc as offset
#Pattern:_21; parm=2+special; non-parm len selex, read as pairs of size, then selex
#Pattern:_22; parm=4; double_normal as in CASAL
#Pattern:_23; parm=6; double_normal where final value is directly equal to sp(6) so can be >1.0
#Pattern:_24; parm=6; double_normal with sel(minL) and sel(maxL), using joiners
#Pattern:_25; parm=3; exponential-logistic in size
#Pattern:_27; parm=3+special; cubic spline
#Pattern:_42; parm=2+special+3; // like 27, with 2 additional param for scaling (average over bin range)
#_discard_options:_0=none;_1=define_retention;_2=retention&mortality;_3=all_discarded_dead;_4=define_dome-shaped_retention
#_Pattern Discard Male Special
 1 0 0 0 # 1 0
#
#_age_selex_patterns
#Pattern:_0; parm=0; selex=1.0 for ages 0 to maxage
#Pattern:_10; parm=0; selex=1.0 for ages 1 to maxage
#Pattern:_11; parm=2; selex=1.0 for specified min-max age
#Pattern:_12; parm=2; age logistic
#Pattern:_13; parm=8; age double logistic
#Pattern:_14; parm=nages+1; age empirical
#Pattern:_15; parm=0; mirror another age or length selex
#Pattern:_16; parm=2; Coleraine - Gaussian
#Pattern:_17; parm=nages+1; empirical as random walk N parameters to read can be overridden by setting special to non-zero
#Pattern:_41; parm=2+nages+1; // like 17, with 2 additional param for scaling (average over bin range)
#Pattern:_18; parm=8; double logistic - smooth transition
#Pattern:_19; parm=6; simple 4-parm double logistic with starting age
#Pattern:_20; parm=6; double_normal,using joiners
#Pattern:_26; parm=3; exponential-logistic in age
#Pattern:_27; parm=3+special; cubic spline in age
#Pattern:_42; parm=2+special+3; // cubic spline; with 2 additional param for scaling (average over bin range)

#Age patterns entered with value >100 create Min_selage from first digit and pattern from remainder
#_Pattern Discard Male Special
 11 0 0 0 # 1 0
#
#_ LO HI INIT PRIOR PR_SD PR_type PHASE env-var use_dev dev_mnyr dev_mxyr dev_PH Block Blk_Fxn # parm_name
1 0 LenSelex
20 35 26.628 0 0 0 2 0 0 0 0 0 0 0 0 # Size_inflection_0(1)
0.01 10 6.9816 0 0 0 2 0 0 0 0 0 0 0 0 # Size_95%width_0(1)
1 0 AgeSelex
 0 40 0 5 99 0 -99 0 0 0 0 0 0 0 # minage@sel=1_0(1)
 0 40 6 6 99 0 -99 0 0 0 0 0 0 0 # maxage@sel=1_0(1)
#_no timevary selex parameters
#
0 # use 2D_AR1 selectivity(0/1)
#_no 2D_AR1 selex offset used
#
Tag loss and Tag reporting parameters go next
0 # TG_custom: 0=no read and autogen if tag data exist; 1=read
#_Cond -6 6 1 1 2 0.01 -4 0 0 0 0 0 0 0 #_placeholder if no parameters
#
no timevary parameters
#
#
Input variance adjustments factors:
 #_1=add_to_survey_CV
 #_2=add_to_discard_stddev
 #_3=add_to_bodywt_CV
 #_4=mult_by_lencomp_N
 #_5=mult_by_agecomp_N
 #_6=mult_by_size-at-age_N
 #_7=mult_by_generalized_sizecomp
#_Factor Fleet Value
 -9999 1 0 # terminator
#
4 #_maxlambdaphase
1 #_sd_offset; must be 1 if any growthCV, sigmaR, or survey extraSD is an estimated parameter
read 1 changes to default Lambdas (default value is 1.0)
Like_comp codes: 1=surv; 2=disc; 3=mnwt; 4=length; 5=age; 6=SizeFreq; 7=sizeage; 8=catch; 9=init_equ_catch;
10=recrdev; 11=parm_prior; 12=parm_dev; 13=CrashPen; 14=Morphcomp; 15=Tag-comp; 16=Tag-negbin; 17=F_ballpark; 18=initEQregime

```
#like_comp fleet  phase  value  sizefreq_method
#
# lambdas (for info only; columns are phases)
#  1 1 1 1 #_CPUE/survey:_1
#  1 1 1 1 #_lencomp:_1
#  1 1 1 1 #_init_equ_catch
#  1 1 1 1 #_recruitments
#  1 1 1 1 #_parameter-priors
#  1 1 1 1 #_parameter-dev-vectors
#  1 1 1 1 #_crashPenLambda
#  0 0 0 0 # F_ballpark_lambda
0 # (0/1/2) read specs for more stddev reporting: 0 = skip, 1 = read specs for reporting stdev for selectivity, size, and numbers, 2 = mortality in addition to values in option 1
 # 0 2 0 0 # Selectivity: (1) fleet, (2) 1=len/2=age/3=both, (3) year, (4) N selex bins
 # 0 0 # Growth: (1) growth pattern, (2) growth ages
 # 0 0 0 # Numbers-at-age: (1) area(-1 for all), (2) year, (3) N ages
 # -1 # list of bin #'s for selex std (-1 in first bin to self-generate)
 # -1 # list of ages for growth std (-1 in first bin to self-generate)
 # -1 # list of ages for NatAge std (-1 in first bin to self-generate)
999
```

forecast.ss

#V3.30.16.00;_2020_09_03;_safe;_Stock_Synthesis_by_Richard_Methot_(NOAA)_using_ADMB_12.2
#Stock Synthesis (SS) is a work of the U.S. Government and is not subject to copyright protection in the United States.
#Foreign copyrights may apply. See copyright.txt for more information.
#C generic forecast file
for all year entries except rebuilder; enter either: actual year, -999 for styr, 0 for endyr, neg number for rel. endyr
1 # Benchmarks: 0=skip; 1=calc F_spr,F_btgt,F_msy; 2=calc F_spr,F0.1,F_msy
2 # MSY: 1= set to F(SPR); 2=calc F(MSY); 3=set to F(Btgt) or F0.1; 4=set to F(endyr)
0.4 # SPR target (e.g. 0.40)
0.342 # Biomass target (e.g. 0.40)
#_Bmark_years: beg_bio, end_bio, beg_selex, end_selex, beg_relF, end_relF, beg_recr_dist, end_recr_dist, beg_SRparm, end_SRparm (enter actual year, or values of 0 or -integer to be rel. endyr)
 1995 2019 1995 2019 1995 2019 1995 2019 1995 2019
2001 2001 2001 2001 2001 2001 1971 2001 1971 2001
1 #Bmark_relF_Basis: 1 = use year range; 2 = set relF same as forecast below
#
-1 # Forecast: -1=none; 0=simple_1yr; 1=F(SPR); 2=F(MSY) 3=F(Btgt) or F0.1; 4=Ave F (uses first-last relF yrs); 5=input annual F scalar
where none and simple require no input after this line; simple sets forecast F same as end year F
10 # N forecast years
0.2 # Fmult (only used for Do_Forecast==5) such that apical_F(f)=Fmult*relF(f)
#_Fcast_years: beg_selex, end_selex, beg_relF, end_relF, beg_mean recruits, end_recruits (enter actual year, or values of 0 or -integer to be rel. endyr)
 0 0 -10 0 -999 0
2001 2001 1991 2001 1971 2001
0 # Forecast selectivity (0=fcast selex is mean from year range; 1=fcast selectivity from annual time-vary parms)
1 # Control rule method (1: ramp does catch=f(SSB), buffer on F; 2: ramp does F=f(SSB), buffer on F; 3: ramp does catch=f(SSB), buffer on catch; 4: ramp does F=f(SSB), buffer on catch)
0.4 # Control rule Biomass level for constant F (as frac of Bzero, e.g. 0.40); (Must be > the no F level below)
0.1 # Control rule Biomass level for no F (as frac of Bzero, e.g. 0.10)
0.75 # Buffer: enter Control rule target as fraction of Flimit (e.g. 0.75), negative value invokes list of [year, scalar] with filling from year to YrMax
3 #_N forecast loops (1=OFL only; 2=ABC; 3=get F from forecast ABC catch with allocations applied)
3 #_First forecast loop with stochastic recruitment
1 #_Forecast recruitment: 0= spawn_recr; 1=value*spawn_recr_fxn;

2=value*VirginRecr; 3=recent mean from yr range above (need to set phase to -1 in control to get constant recruitment in MCMC)
1 # value is multiplier of SRR
0 #_Forecast loop control #5 (reserved for future bells&whistles)
2010 #FirstYear for caps and allocations (should be after years with fixed inputs)
0 # stddev of log(realized catch/target catch) in forecast (set value>0.0 to cause active impl_error)
0 # Do West Coast gfish rebuilder output (0/1)
1999 # Rebuilder: first year catch could have been set to zero (Ydecl)(-1 to set to 1999)
2002 # Rebuilder: year for current age structure (Yinit) (-1 to set to endyear+1)
1 # fleet relative F: 1=use first-last alloc year; 2=read seas, fleet, alloc list below
Note that fleet allocation is used directly as average F if Do_Forecast=4
2 # basis for fcast catch tuning and for fcast catch caps and allocation (2=deadbio; 3=retainbio; 5=deadnum; 6=retainnum); NOTE: same units for all fleets
Conditional input if relative F choice = 2
enter list of: season, fleet, relF; if used, terminate with season=-9999
1 1 1
-9999 0 0 # terminator for list of relF
enter list of: fleet number, max annual catch for fleets with a max; terminate with fleet=-9999
-9999 -1
enter list of area ID and max annual catch; terminate with area=-9999
-9999 -1
enter list of fleet number and allocation group assignment, if any; terminate with fleet=-9999
-9999 -1
#_if N allocation groups >0, list year, allocation fraction for each group
list sequentially because read values fill to end of N forecast
terminate with -9999 in year field
no allocation groups
2 # basis for input Fcast catch: -1=read basis with each obs; 2=dead catch; 3=retained catch; 99=input apical_F; NOTE: bio vs num based on fleet's catchunits
#enter list of Fcast catches or Fa; terminate with line having year=-9999
#_Yr Seas Fleet Catch(or_F)
-9999 1 1 0
#
999 # verify end of input

부록 4. 시뮬레이션 연구에 사용된 R code.

제시한 R code는 효율적인 시뮬레이션 연구를 목적으로 작성했다. 먼저, R 파일을 저장한 작업 공간내에 작동모델(OM)과 추정모델(EM), 결과(sim_results) 폴더를 생성했고, 생성된 OM폴더 안에 SS의 구현에 필요한 네 가지 파일과 실행 파일을 저장했다. OM 폴더에 저장한 data.ss 파일을 'read.csv()' 함수를 이용해 R 소프트웨어 환경으로 가져오고, 참값으로부터 가짜 자료를 생성하기 위해 R 내장 함수인 'rlnorm()' 함수와 'rmultinorm()'을 이용했다. 생성된 가짜 자료(연도별 CPUE, 체장 조성 자료)를 data.ss 파일에 덮어쓰고 난 후, 작업공간을 EM 폴더로 변경하고 덮어쓴 data.ss 파일을 'write.csv()' 함수로 저장했다. EM 폴더안에는 data.ss 파일을 제외한 나머지 파일을 갖춰야 하며, 생성된 가짜 자료로 구성된 data.ss 파일을 기반으로 수치 최적화를 수행했다. 추정된 결과를 저장 또는 누락시키기 위해 수렴 기준(convergence criteria)을 설정했고, 이 과정을 시뮬레이션 시나리오마다 1,000번 반복했다. 시뮬레이션 과정이 끝난 후, 작업 공간을 sim_results 폴더로 변경한 후 결과들을 저장했다.

R file

```
# R code for simulation study of length-based SS model.
#
# 2021 01 15
# edited by Lee SeungJoon
#
install.packages("r4ss")
#
library(r4ss) #
#
simulation_analysis<-1 # # (0,1,2,3) Input 1 when using M and CV as fixed values, 2 when estimating CV with fixed M, and 3 when estimating both M and CV. Input 0 when not performing simulation analysis.
save<-1 # (0,1) If you want to save the results, enter 1 here, otherwise enter 0.
#
### set working directory
setwd((dirname(rstudioapi::getActiveDocumentContext()$path)))
getwd()
#
for(Create_directory in 1:1){
  if(dir.exists("OM")==FALSE){
    dir.create("OM") # Operating model
  }
  if(dir.exists("EM")==FALSE){
    dir.create("EM") # Estimation model
  }
  if(dir.exists("sim_results")==FALSE){
    dir.create("sim_results") # results of simulation study
  }
}# OM, EM, results

##################
# simulation study #
##################
setwd(paste0(dirname(rstudioapi::getActiveDocumentContext()$path),"/OM"))
getwd()
#
for(read_OM_file in 1:1){
  ### read data file
  OM_dat_file<-read.table("data.ss", sep="")
  if(length(which(is.na(OM_dat_file)))!=0){
    OM_dat_file<-OM_dat_file[-(which(is.na(OM_dat_file)))]
  }
  OM_dat_file<-unlist(OM_dat_file)
```

```
   OM_dat_file<-as.numeric(OM_dat_file)
   ### read control file
   OM_ctl_file<-read.table("control.ss", sep="")
   OM_ctl_file<-unlist(OM_ctl_file)
   OM_ctl_file<-as.numeric(paste(OM_ctl_file))

   ### read expected value of CPUE, Len. comp.
   OM_new_dat_file<-read.table("data.ss_new", sep="")
   OM_report<-read.table("Report.sso",header= FALSE, sep="",fill=TRUE) # read
report file
   ### read par file
   OM_par<-read.table("ss.par",sep="")
   OM_par<-unlist(OM_par)
   OM_par<-as.numeric(paste(OM_par))

   year<-
(as.numeric(paste(OM_new_dat_file[1,])):as.numeric(paste(OM_new_dat_file[2,])))
   N_years<-length(year)
   len_years<-2000:2019 # length of year (len. comp.)
   N_len_years<-length(len_years)
   Ndat_bins<-as.numeric(paste(OM_new_dat_file[500,]))
}

## expected value of CPUE data
for(read_true_value in 1:1){
OM_exp_cpue_1975<-numeric(N_years)
for(i in 1:N_years){
   OM_exp_cpue_1975[i]<-as.numeric(paste(OM_new_dat_file[i*5+1834,]))
}
## expected value of Length composition data
OM_exp_Len_1975<-matrix(data=NA,nrow=N_len_years,ncol=Ndat_bins)
for(i in 1:N_len_years){
   OM_exp_Len_1975[i,]<-
as.numeric(paste(OM_new_dat_file[(i*49+2081):(i*49+2123),]))
}
N_eff<-as.numeric(paste(OM_new_dat_file[2129,]))
OM_exp_Len_1975<-OM_exp_Len_1975/N_eff
#
if(simulation_analysis<2){
   true_para_1975<-OM_par[c(3,4,5,16,66,67,68)];
   # L0,L_inf,K,SR_LN(R0),LN_Q,sel(1),sel(2)
}else if(simulation_analysis==2){
   true_para_1975<-OM_par[c(3,4,5,6,7,16,66,67,68)];
   # L0,L_inf,K,CV0,CV6,SR_LN(R0),LN_Q,sel(1),sel(2)
} else{
```

```
   true_para_1975<-OM_par[c(2,3,4,5,6,7,16,66,67,68)];
   # M,L0,L_inf,K,CV0,CV6,SR_LN(R0),LN_Q,sel(1),sel(2)
} # get true parameters
}

true_para_1975 # check parameters

###
cpue_CV<-c(0.1,0.3,0.5) # CPUE CV scenarios.
eff_samsize<-as.numeric(paste(OM_new_dat_file[2129,])) # effective sample size of length composition likelihood function.
iterations<-1000 # iterations of each scenario.
#
N_case<-length(cpue_CV) # total number of cases.
logN_shape<-sqrt(log(1+(cpue_CV)^2)) # approximation of standard deviation.
convergence_rate<-numeric(N_case) # results of model validation (%).
#
if(simulation_analysis>0){
   #
   start_time<-Sys.time()
   #
   for(s in 1:N_case){
   setwd(paste0(dirname(rstudioapi::getActiveDocumentContext()$path),"/EM"))
      # generating result_matrix
      length_par<-length(OM_par)
      #
      EM_check_grad<-numeric(iterations)
      EM_N_bound_hit<-numeric(iterations)
      EM_Hessian_results<-numeric(iterations)
      para_results<-matrix(data=NA,nrow=length_par,ncol=iterations,byrow=FALSE
                           ,dimnames =
list(c("dummy_parm","M","L0","L_inf","K","CV0","CV6",

"Wtlen_1","Wtlen_2","Mat50%","Mat_slope",

"eggs/kg_inter","eggs/kg_slope","CohortGrowDev",

"FracFemale","SR_LN(R0)","steepness","SR_sigmaR",

"SR_regime","SR_autocorr",paste0("RecrDev_",year),

"LnQ","Size_inflection","Size_95%width","minage_sel",

"maxage_sel","checksum999")))
```

```
for(iter in 1:iterations){

  if(file.exists("ss.cor")==TRUE){
    file.remove("ss.cor")
  }
  ### generating pseudo data/ CPUE data
  sam_cpue<-numeric(N_years)
  sam_p_hat<-matrix(data=NA, nrow= N_len_years, ncol= Ndat_bins)

  #
  for(i in 1:N_years){
    sam_cpue[i]<-rlnorm(n=1, meanlog = log(OM_exp_cpue_1975[i]), sdlog =
logN_shape[s])
  }
  #
  for(i in 1:N_len_years){
    sam_p_hat[i,]<-
rmultinom(1,size=eff_samsize,prob=OM_exp_Len_1975[i,])/eff_samsize
  }

  # modifying the data file/ CPUE data
  for(i in 1:N_years){
    OM_dat_file[5*i+254]=sam_cpue[i]
  }
  # modifying the data file/ proportion of the length data
  for(i in 1:N_len_years){
    OM_dat_file[(49*i+501):(49*i+543)]=sam_p_hat[i,1:43]
  }
  # save the modified data file
   write.table(OM_dat_file,file="data.ss",row.names = FALSE,col.names =
FALSE)

   write.table(OM_ctl_file,file="control.ss",row.names = FALSE,col.names =
FALSE)
   #
  #
  #
  system(get_bin) # run
  #
  #
  #
  # free_estimates of parameters
  EM_par<-read.table("ss.par",sep="")
  EM_par<-unlist(EM_par)
```

```
        EM_par<-as.numeric(paste(EM_par))

        # checking the maximum gradient
        EM_max_grad<-read.table("gradient.dat",sep="")
        EM_max_grad<-EM_max_grad[-c(1),-c(1,2)]
        EM_max_grad<-unlist(EM_max_grad)
        EM_max_grad<-as.numeric(paste(EM_max_grad))
        EM_max_grad<-max(abs(EM_max_grad))

        # checking the bound hit
        EM_report<-read.table("Report.sso",header= FALSE, sep="",fill=TRUE)

        #write.table(EM_report ,file="Report.sso",row.names = FALSE,col.names =
 FALSE)
        EM_N_bound<-unlist(EM_report[290,2])
        EM_N_bound<-as.numeric(paste(EM_N_bound))

        #
        EM_Hessian<-file.exists("ss.cor")

        #check points for model validation
        para_results[,iter]<-EM_par
        EM_check_grad[iter]<-EM_max_grad
        EM_N_bound_hit[iter]<-EM_N_bound
        EM_Hessian_results[iter]<-EM_Hessian
        #
        EM_N_run<-length(cpue_CV)*length(eff_samsize)*iterations
        for(i in 1:20){
           print(paste0("========================= "
                        ,round((((s-1)*iterations)+iter)/EM_N_run*100,digits = 2)
                        ,"% ========================="))
        }
        #
        list<-list(para_results,EM_check_grad,EM_N_bound_hit,EM_Hessian_results)
        names(list)<-
c("para_results","EM_check_grad","EM_N_bound_hit","EM_Hessian_results")
      }
      ### getting convergence iterations
      #
      ref_1<-sort(which(EM_N_bound_hit > 0),decreasing = TRUE)
      ref_2<-sort(which(EM_check_grad > 0.001),decreasing = TRUE)
      ref_3<-sort(which(EM_Hessian_results == FALSE),decreasing = TRUE)
      #
      ref_final<-unique(c(ref_1,ref_2,ref_3))
      length(ref_final)
```

```
      #
      if (length(ref_final)==0){
        Filtered_para<-para_results
        convergence_rate[s]<-paste0(100," %")
      }else{
        Filtered_para<-para_results[,-c(ref_final)]
        convergence_rate[s]<-paste0(dim(Filtered_para)[2]/iterations*100," %")
      }
      #

setwd(paste0(dirname(rstudioapi::getActiveDocumentContext()$path),"/sim_results")
);
      save.image(file=paste0(s,"_",Sys.Date(),"_simulation.RData"));
      write.csv(Filtered_para, file=paste0(s,"_para.csv"), row.names = TRUE)
      # save results as csv file
    }
    end_time<-Sys.time()
    print(c(paste0("start time: ",start_time," ----> end time: ",end_time,"    * TOTAL
RUN NUMBER: ",EM_N_run)))
    print(convergence_rate)
}
```

VII. 사사

많이 부족한 저에게 기회를 주시고 수산 자원 관리학 분야에 대해 심도 있는 공부를 할 수 있게 지도해주시고 격려해주신 현상윤 교수님께 감사의 말씀을 드립니다.

2 년이라는 짧지 않은 시간 동안 힘든 시간과 즐거운 시간을 함께 보냈던 수산자원관리학 실험실 동료 연구원들께 감사의 말씀을 드립니다. 특히, 석사 과정 내내 많은 도움을 준 김진우 선배, 김도율 형에게 감사의 말씀을 전합니다. 덕분에 쉽지 않았던 연구를 성공적으로 마무리할 수 있었습니다.

연구를 진행함에 있어 학문적으로 많은 도움을 주신 국립수산과학원 서영일 박사님, 강희중 박사님께 감사의 말씀을 드립니다. 또한, 재정적으로 지원해준 국립수산과학원 및 연구재단 관계자분들께 감사드립니다.

국립수산과학원(연구과제번호: R2020022)

연구재단(연구과제번호: NRF-2019R1I1A2A01052106)

누구보다도 사랑하는 가족 분들께 감사의 말씀을 전합니다. 앞으로 제가 가는 모든 길에 부모님과 동생의 행복이 있었으면 좋겠습니다. 먼 곳에서 언제나 저를 지켜 봐주시고 도움을 주시는 조부모님께 감사의 말씀을 드립니다. 그 외 모든 친척분들께 감사의 말씀을 전합니다.

석사과정을 진행함에 있어 개인적인 고민들을 들어주고 항상 격려해주고 응원해줬던 고향 친구들인 손영덕, 안수길, 정지원, 한상혁 친구들에게 감사의 말씀을 전합니다.