



Ollscoil
Teicneolaíochta
an Atlantaigh

Atlantic
Technological
University

**Addressing cohort uncertainty through
advanced length frequency and
stage-based assessment models with
application to anglerfish**

Luke Batts

Submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

Supervisors: Dr. Cólín Minto, Dr. Hans Gerritsen and Dr.
Deirdre Brophy

Submitted to
Atlantic Technological University (formerly Galway–Mayo Institute of Technology)
February 2022

This PhD (Cullen Fellowship: CF/16/03) was carried out with the support of the Marine Institute and is funded under the Marine Research Sub-Programme by the Irish Government.





Ollscoil
Teicneolaíochta
an Atlantaigh

Atlantic
Technological
University

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____  _____ Candidate

ID No. _____ G00352861 _____

Date: _____ 4 February 2022 _____

Contents

List of Figures	ix
List of Tables	xii
Abstract	xiii
Acknowledgements	xv
1 General Introduction	1
1.1 Stock assessment and age-estimation	2
1.1.1 Stock assessment background	2
1.1.2 Importance of growth estimates and ageing data in stock assessment	3
1.1.3 Challenges to age-estimation	4
1.1.4 Alternative methods for age assignment	4
1.1.5 Implications for assessing fish stocks	5
1.2 Anglerfish fisheries and assessment	5
1.2.1 Anglerfish fisheries around the world	5
1.2.2 Historical anglerfish fisheries in the North Atlantic . . .	6
1.2.3 Anglerfish stocks in the North-East Atlantic	6
1.2.4 Current knowledge and state of anglerfish assessment and management	7
1.2.5 Ageing difficulties for anglerfish	9
1.3 Stock assessment in the presence of ageing uncertainty	10
1.3.1 Choosing a stock assessment model	10
1.3.2 Aggregated biomass stock assessment models	10
1.3.3 Integrated stock assessment models	11
1.3.4 Stage-based stock assessments as an alternative	11
1.4 Testing the effect of ageing uncertainty and bias on stock as- sessment and management	12
1.4.1 Origins of management strategy evaluation	12
1.4.2 Components and attributes of management strategy evaluations	13
1.5 Gaps in our understanding	15
1.6 Thesis aims and objectives	16
2 Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models	34
2.1 Introduction	36

2.2	Methods	39
2.2.1	Growth function	39
2.2.2	LFEM model and algorithm	40
2.2.3	Hierarchical extensions	41
2.2.4	Model parameters, inputs and outputs	42
2.2.5	Model selection criteria and standard errors	43
2.2.6	Validation for haddock	44
2.2.7	Software	44
2.2.8	Data	45
2.3	Results	46
2.3.1	Haddock	46
2.3.1.1	Haddock basic LFEM model	46
2.3.1.2	Haddock hierarchical LFEM model with cohort specific l and k	49
2.3.2	White-bellied anglerfish	52
2.3.2.1	White-bellied anglerfish basic LFEM model	52
2.3.2.2	White-bellied anglerfish hierarchical LFEM model with cohort specific l and L	56
2.3.2.3	Other White-bellied anglerfish hierarchical models	58
2.4	Discussion	58
2.4.1	Overview of LFEM	58
2.4.2	Haddock	58
2.4.3	White-bellied anglerfish	59
2.4.4	Caveats	61
2.4.5	Further work	62
2.5	Conclusions	63
2.6	Supplementary material	64
2.7	Acknowledgements	64
3	Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real- life stocks	71
3.1	Introduction	73
3.2	Methods	76
3.2.1	Schnute model	76
3.2.1.1	Population dynamics and notation	76
3.2.1.2	Deterministic equations	78

3.2.1.3	Model variants	79
3.2.1.4	Model summaries	80
3.2.1.5	Likelihood	81
3.2.1.6	Estimation in this study	82
3.2.2	CSA	83
3.2.2.1	Population dynamics and notation	83
3.2.2.2	Deterministic equations and model summary	85
3.2.2.3	Likelihood	85
3.2.2.4	Estimation for this study	86
3.2.3	Data	86
3.2.3.1	Simulation framework	86
3.2.3.2	Real anglerfish stock	88
3.2.4	Model performance	89
3.2.4.1	Self-tests	90
3.2.5	Software	90
3.3	Results	91
3.3.1	Self-tests	91
3.3.2	Simulation framework	92
3.3.3	Real white-bellied anglerfish stock	98
3.4	Discussion	101
3.4.1	Model performance	101
3.4.2	Model developments	103
3.4.3	Growth parameter estimation and stage mean fish weights configuration	104
3.4.4	Convergence and identifiability	105
3.4.5	Limitations and further work	105
3.5	Conclusions	106
3.6	Acknowledgements	107
4	sbar vignette: An R package for implementing stage-based assessment models	114
4.1	Introduction	116
4.1.1	General details	117
4.2	Installation, data and other useful packages	117
4.2.1	Installation	117
4.2.2	Data	118
4.2.3	Setup	118
4.3	Introduction to CSA	119

4.3.1	CSA - quick start	119
4.3.2	Simple plot of CSA results	123
4.4	Introduction to Schnute models	124
4.4.1	Schnute models - quick start	125
4.4.1.1	<i>Original Schnute Process Error model</i>	128
4.4.1.2	<i>Adapted Schnute Observation Error mode</i>	131
4.5	Schnute models in more detail	134
4.5.1	Growth and estimating growth parameters	135
4.5.2	Version and the relative importance of recruit or previously-exploited biomass	139
4.5.3	Comparison of Schnute model types	140
5	Comparative impacts of ageing bias on stage and age-based fisheries stock assessment models	146
5.1	Introduction	148
5.2	Methods	152
5.2.1	Overview	152
5.2.2	General MSE settings	152
5.2.3	Operating model	152
5.2.4	Observation Error model	153
5.2.4.1	Bias growth	154
5.2.5	Stock assessment methods	158
5.2.6	a4a biological reference points	160
5.2.7	CSA biological reference points	162
5.2.7.1	F_{msy} and N_{msy}	162
5.2.8	Harvest Control Rules and implementation	164
5.2.9	Performance statistics	165
5.3	Results	166
5.3.1	Initial assessments	166
5.3.2	Effect of age-slicing bias in MSE simulations	168
5.3.3	Performance statistics comparison	172
5.4	Discussion	176
5.4.1	Effect of age-slicing bias	176
5.4.2	General effectiveness of CSA	178
5.4.3	Conclusions	179
6	General Discussion	189
6.1	Summary of main thesis results	190

6.2	Applications to fisheries stock assessment	192
6.2.1	Acquiring reliable growth information from length frequency data	192
6.2.2	Stage-based assessment models: a diverse and useful middle ground	195
6.2.3	Applications to anglerfish stock assessment	198
6.3	Applications to fisheries stock assessment and management over time	199
6.3.1	Implications of biased age-composition data for management advice and performance	200
6.3.2	Performance of stage-based management procedures	200
6.3.3	Performance of age-based management procedures	201
6.4	Future directions	202
Appendix A Published manuscript I		213
Appendix B Supplementary material for chapter 2		228
Appendix C Working document to ICES Benchmark Workshop on Anglerfish Stocks in the ICES Area (WKANGLER), 2018244		
Appendix D Supplementary material for chapter 3		253
Appendix E Working document to ICES working group for the Bay of Biscay and The Iberian Waters Ecoregion (WGBIE), 2020		275
Appendix F sbar R package manual		293
Appendix G Supplementary material for chapter 5		307

List of Figures

Figure 1.1	ICES subareas and survey coverage of the anglerfish stocks	8
Figure 1.2	Schematic of a typical MSE simulation framework	14
Figure 2.1	Survey coverage for haddock and white-bellied anglerfish	45
Figure 2.2	Length-at-age data and estimated von Bertalanffy growth curves for haddock	48
Figure 2.3	Length-at-age data and estimated von Bertalanffy growth curves for cohorts of haddock	49
Figure 2.4	Length frequency distributions of haddock with fitted distributions from models	51
Figure 2.5	von Bertalanffy growth curves estimated for white-bellied anglerfish	53
Figure 2.6	Length frequency distributions of white-bellied anglerfish with fitted distributions	55
Figure 2.7	Estimated random effects on first (l) and last (L) component means of cohorts for white-bellied anglerfish	57
Figure 3.1	Self-tests for a simulated stock	91
Figure 3.2	Time series of median estimated stock numbers and real values	94
Figure 3.3	Time series of median estimated fishing mortality and real values	95
Figure 3.4	Relative error and relative standard error of estimated stock numbers	96
Figure 3.5	Relative error and relative standard error of estimated fishing mortality at the start and end of time series . . .	97
Figure 3.6	Model fits for the white-bellied anglerfish stock	99
Figure 3.7	Estimated stock numbers and fishing mortality of model fits for the white-bellied anglerfish stock	100
Figure 4.1	CSA model fit for black-bellied anglerfish	123
Figure 4.2	Estimated stock numbers and fishing mortality from the CSA model fit for black-bellied anglerfish	124
Figure 4.3	Mean fish weights of stages and entire assessed population of black-bellied anglerfish	126
Figure 4.4	Mean fish weights of stages and entire assessed population of black-bellied anglerfish with missing values approximated	127
Figure 4.5	<i>Original Schnute Process Error model</i> fit for black-bellied anglerfish	130

Figure 4.6	Estimated stock numbers and fishing mortality from the Original Schnute Process Error model fit for black- bellied anglerfish	131
Figure 4.7	<i>Adapted Schnute Observation Error model</i> fit for black- bellied anglerfish	133
Figure 4.8	Estimated stock numbers and fishing mortality from the <i>Adapted Schnute Observation Error model</i> fit for black- bellied anglerfish	134
Figure 4.9	Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the survey and the catch	135
Figure 4.10	Linear model for growth	137
Figure 4.11	Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the combined IBTS survey and approximate linear growth	137
Figure 5.1	Effect of bias on growth curves, numbers-at-length and numbers-at-age	157
Figure 5.2	Effect of bias on maturity ogives and catch weight-at-age	158
Figure 5.3	a4a reference points estimation	161
Figure 5.4	CSA reference points estimation	163
Figure 5.5	a4a stock assessment estimates across three OEMs for initial OM period	167
Figure 5.6	CSA stock assessment estimates across three OEMs for the initial OM period	168
Figure 5.7	MSE simulations for a4a assessment MPs across three OEMs	169
Figure 5.8	a4a assessment bias of spawning stock biomass (SSB) and fishing mortality across three OEMs	170
Figure 5.9	MSE simulations for CSA assessment MPs across three OEMs	171
Figure 5.10	CSA assessment bias of total stock numbers and fishing mortality across three OEMs	172
Figure 5.11	CSA MP outputs of fishing mortality from the HCR and implemented TAC across three OEMs	172
Figure 5.12	Distribution of stock status of MSE simulations in the short term and mid-term for both MPs and across three OEMs	173

Figure 5.13 Kobe plot for final year of MSE simulations for both MPs
and across three OEMS 174

Figure 5.14 Yearly trade-offs between risk or sustainability and catch 175

List of Tables

Table 2.1	Summary of the starting values tested for the model parameters of the basic model	42
Table 2.2	Summary of model parameters, and other relevant inputs/outputs of models. N.B. Not all hierarchical model parameters are used in each model	43
Table 2.3	von Bertalanffy growth parameters model runs on haddock	47
Table 2.4	Component/cohort means for haddock	48
Table 2.5	von Bertalanffy growth parameters from runs on white-bellied anglerfish	52
Table 2.6	Component/cohort means for white-bellied anglerfish . .	54
Table 3.1	Summary of the Schnute model notation	77
Table 3.2	Summary of the steps within Schnute models	81
Table 3.3	Summary of Schnute model and growth/mean weight configurations	82
Table 3.4	Summary of the CSA model notation	84
Table 3.5	Summary of the steps within the CSA model	85
Table 3.6	Simulation experimental design factor levels	87
Table 3.7	Convergence of models	92
Table 3.8	Assessment model performance	93
Table 5.1	Summary of steps to bias age-slicing with age-transition matrices	156
Table 5.2	a4a reference points	162
Table 5.3	CSA reference points	164

Addressing cohort uncertainty through advanced length frequency and stage-based assessment models with application to anglerfish

Luke Batts

Abstract

Fisheries stock assessments are important tools for successful management of fisheries. Reliability of a stock assessment model is often determined by the data available and accounting for key uncertainties in the data is an important aspect of stock assessments and management. Central to many assessment models is the tracking of cohorts through the population, but uncertainty in the identification of cohorts, in the form of uncertainty in age-estimation and fish growth, can impact model performance. Overall, the aim of this thesis was to develop methods to address this cohort uncertainty, as well as assess the impact of biased age-composition data on the fisheries advice process. I focus on anglerfish as the main case study species, as their assessment and management is affected by the challenges touched upon earlier.

To provide context, Chapter 1 gives an overview of stock assessment and management, focussing on the challenges of uncertainty in age-estimation and growth, as well as the approaches used to account for them. Anglerfish fisheries are also summarised and issues affecting their assessment and management discussed.

Chapter 2 addresses the topic of cohort uncertainty by developing a new method of length frequency analysis. The model is a maximum likelihood-based procedure that uses Gaussian mixture models and the Expectation-Maximisation algorithm to estimate von Bertalanffy growth parameters from length frequency data from fisheries surveys. The model was applied to length data from the white-bellied anglerfish stock in the Celtic Sea and Bay of Biscay. The basic model estimated a single set of growth parameters, whereas the hierarchical extension to the model was able to model some of the natural variability in fish growth between cohorts or years with bivariate random effects on key parameters.

Chapter 3 approaches the issue of cohort uncertainty from a different perspective, implementing and testing the performance of stage-based stock assessment models. Stage-based assessment models have less data requirements and simpler population dynamics than more complex assessment models, so are likely to be more robust to cohort uncertainty in the data. The stage-based assessment models implemented were: Catch-Survey Analysis (CSA), and a model first described in a theoretical paper by Schnute (1987). The perfor-

mance of these two theoretically different stage-based assessment models was assessed with a simulation-testing framework and on a real anglerfish stock. The findings showed both models are useful stock assessment models, with CSA more robust but less precise than the Schnute model. The Schnute model was more precise than CSA but required growth and mean fish weight data unaffected by selectivity.

As part of the work conducted for Chapter 3, the R package ‘sbar’ was developed. This is a fully documented R package that contains the functions to run the stage-based assessment models. Chapter 4 outlines the key assumptions and data requirements of the models, as well as demonstrating use with data from a real anglerfish stock. Versions of the Schnute model not described or tested in Chapter 3 were also detailed here. A goal of this chapter was to allow new users to begin running stage-based assessment models with relevant background information provided.

Chapter 5 addresses the impact of using age-composition data generated with a biased growth function in stock assessment and management over time. A management strategy evaluation framework was used, with both stage-based and age-based management procedures tested. A method for estimating CSA reference points was also developed. Both management procedures were affected by the biased sampling data, but in different ways. Performance statistics indicated that it is important to consider the uncertainty and potential bias in growth estimates when generating age-composition data.

Overall, the research presented in this thesis has developed and implemented techniques that aim to further advance the field of fisheries stock assessment and management when cohorts are uncertain. The thesis focussed on anglerfish for the majority of case studies due to the fisheries’ reported issues, however the methods implemented here are useful in a wider context and there are many species to which the techniques could be applied.

Acknowledgements

This thesis is dedicated to the memory of Nick Batts, whose enthusiasm and support are greatly missed

Firstly, I would like to thank the Marine Institute Cullen Fellowship Programme, as well as the HEA support for Covid-19 related Research Costed Extensions for providing the funding for this project.

A huge thanks goes to my supervisors, who have provided guidance and support throughout. I would like to thank my main PhD supervisor, Dr C oil n Minto, for his seemingly infinite patience every time I popped my head into his office with a ‘quick’ question in the early stages of the PhD and his well-timed encouragement when I needed it in the latter stages. I am so grateful to C oil n for the committent and warmth he has shown me over the last five years. His wealth of knowledge, expertise and passion for science is an inspiration. Go raibh m ile maith agat. Moreover, I would like to thank Dr Hans Gerritsen for co-supervising me during the PhD. His insights into the the world of fisheries stock assessment and management have been invaluable. The constant good advice and openness to involving me his work at the Marine Institute have been very much appreciated. I would also like to thank Dr Deirdre Brophy for co-supervising me during the PhD. Deirdre’s perspective and calm approach has been particularly helpful as the PhD has progressed.

There have been many people that I have had the good fortune to meet, chat to, and learn from over the time of the PhD. To be involved with the Marine Institute and the Cullen Fellowship programme has been a great aspect of the project. I am grateful for the opportunities I have had to meet many of the scientists that work or were visiting the Marine Institute, whether that was at meetings, Cullen events, or at sea. Special thanks goes to the Irish fishing industry representatives who engaged with some of my research. Thanks for the welcome extended by the ICES community on the trips I made over to Copenhagen. Thanks also to Iago, Ernesto and the team at the JRC for an enjoyable and productive trip to Ispra.

I’d also like to thank the fishing community and my former colleagues from my time in Shetland, and Dr Paul Macdonald in particular. The experience in practical and applied fisheries research that I gained whilst working up in the Shetland Isles was an unforgettable one, and gave me the solid grounding in fisheries that a formerly landlocked Yorkshire lad needed. Thanks also goes to

to Dr Paul Fernandes for first introducing me to the murky world of monkfish during my Masters degree.

To my friends and colleagues at the MFRC, thank you for your companionship, support, extended tea breaks, stimulating conversations, nonsense chats, and everything in between. It's been tough not being in the office with you all over the last two years but I'm grateful for the times we enjoyed! Special thanks to Dr Ghassen Halouani and Paula Silvar-Viladomiu, it has been a pleasure to work on other projects with you as friends as well as colleagues. Thanks also to Dr Paul Dolder, for his help with formatting the thesis. Particular thanks to Dr Andrew Power, Dr Elena Pagter, Dr James O'Connor, María Perez Tadeo, Christie O'Donnell, Ryan McKenna, Dr Adon McFarlane, Dr Joao Frias and Yvonne Boyd. I'd also like to say thanks to all my friends over the years that have had to listen to me talk about fish too much, but more importantly thanks for the conversations about anything but fish!

Thanks to my family, for their reassurance and belief in me throughout. Special thanks to my mum for her constant encouragement, love and support. Thanks to my sister Alice for her positivity and faith in me. Thanks to David for his encouragement into the field of fisheries, and a massive thanks to Kate, Tansy, Ronan, Laura, Vaila, and Steve and the family.

Finally, I would like to thank my partner Maya. Throughout this PhD, her ability to distract my thoughts away from work and cheer me up have been unparalleled, as well as much welcomed. Thank you for your unwavering support and love. I look forward to having more free time and can't wait to embark on some more adventures together.

Chapter 1

General Introduction

1.1 Stock assessment and age-estimation

1.1.1 Stock assessment background

Fish are an important source of food globally and reliant fisheries often have a complex dynamic of social, economic and ecological aspects (Hilborn and Walters, 1992; FAO, 2020). Since the 1970's there has been a growing realisation and concern about the overexploitation of global fish stocks, leading to the concept of sustainable fisheries management growing in prominence (Hilborn et al., 2020). In 2015, The United Nations Sustainable Development Goals were published, with conservation and sustainable use of the marine environment being one of the key goals (SDG 14) (United Nations, 2015). A primary target of SDG 14 is to regulate fishing practices and implement management that restores fish stocks to 'levels that can produce maximum sustainable yield' (United Nations, 2015).

Successful and sustainable fisheries management (i.e. stock abundance within biologically sustainable levels) typically requires a reliable stock assessment (Worm et al., 2009; Costello et al., 2012; Bianchi et al., 2014; Hilborn et al., 2020). Fisheries stock assessments utilise a range of different data to estimate relative size of a fish stock and the rate at which the stock has been harvested. Results of a stock assessment often determine the management advice for a fishery (Hilborn and Walters, 1992).

Uncertainty is an integral part of stock assessment and management, and can generally be split into five components: process uncertainty, observation uncertainty, model uncertainty, estimation uncertainty and implementation uncertainty (Rosenberg and Restrepo, 1994; Privitera-Johnson and Punt, 2020). The first four listed concern the stock assessment, whilst implementation is the uncertainty associated with the implementing of management procedures (Privitera-Johnson and Punt, 2020).

A need to reduce and account for different types of uncertainty has often driven research and development of methods in fisheries science. Reliability and choice of a particular stock assessment is often determined by the available data (Hilborn and Walters, 1992; ICES, 2012; Maunder and Punt, 2013). In many cases sparseness or uncertainty in data are key issues that hinder development of assessments for fish stocks (Chen, 2003; Maunder and Piner, 2015; Chrysafi et al., 2019).

Stock assessment models range from data-poor to data-rich approaches. At the data-poor end of the spectrum there are approaches such as CMSY (Catch-Maximum Sustainable Yield) (Martell and Froese, 2013; Froese et al., 2017) and length-based methods such as LBSPR (Length-based Spawning Potential Ratio) (Hordyk et al., 2015) and LIME (Length-based Integrated Mixed Effects) (Rudd and Thorson, 2018). Surplus production models such as SPiCT (Surplus Production in Continuous Time) (Pedersen and Berg, 2017) or ASPIC (A Surplus-Production model Incorporating Covariates) (Prager, 1992) are an example of a data-moderate approach, generally requiring a time-series of catch and an index of biomass to run. Stage-based assessment models such as biomass-based delay-difference models (Deriso, 1980; Schnute, 1987; Meyer and Millar, 1999) or CSA (Catch-Survey Analysis) (Collie and Sissenwine, 1983; Mesnil, 2003) can also be considered data-moderate approaches. As stock assessment models increase in complexity to age-structured models (Shepherd, 1999; Jardim et al., 2014; Nielsen and Berg, 2014) and further still to integrated models (Bull et al., 2005; Methot and Wetzel, 2013), spatial models (Cadigan et al., 2017) or ecosystem models (Begley and Howell, 2004), they are considered data-rich approaches to fisheries stock assessment. Generally, the majority of contemporary stock assessments are based on age-based models, when sufficient data are available (Punt et al., 2013).

1.1.2 Importance of growth estimates and ageing data in stock assessment

Generating reliable age-composition data and/or estimates of growth are often fundamental aspects of developing reliable stock assessments, as they facilitate tracking of the progress cohorts of fish within a fish stock (Hilborn and Walters, 1992; Quinn and Deriso, 1999). More broadly, fish growth is an important branch of fish biology and is well studied due to its importance for many aspects of fisheries science, including life-history theory (Denney et al., 2002), ecosystem modelling (Smith et al., 2015; Halouani et al., 2019) and investigations of climate change impacts (Blanchard et al., 2012; Denechaud et al., 2020). Fish growth in the form of a length-at-age relationship is most commonly modelled with the well-known von Bertalanffy growth function (von Bertalanffy, 1938).

1.1.3 Challenges to age-estimation

A common method for obtaining growth information for teleost fish is to utilise otoliths (or some other hard part that lays down incremental growth rings e.g. fin rays, scales, vertebrae) to estimate the age of individual fish (Campana, 2001). An age-length key (ALK) can then be used to convert catch-at-length data to catch-at-age estimates. However, this method can often be challenging in terms of time and resources, or if age-estimates are considered unreliable (Hilborn and Walters, 1992). Unreliable age estimates are a common issue in tropical and deep sea marine fish where seasonal changes in growth are not thought to be substantial enough to leave clear growth increments in the calcified structures of fish (Maunder and Piner, 2015), although there is evidence that this may not be the case for some species (Morales-Nin and Panfili, 2005; McMillan et al., 2021). Even with species for which the ageing method is well established, there can be issues due to the environmental complexities affecting incremental growth rings (Folkvord et al., 2000; De Pontual et al., 2006; Hüsey et al., 2016).

1.1.4 Alternative methods for age assignment

Alternative methodologies are needed to generate age-composition data for age-based stock assessments in situations where age-estimates are unreliable, require validation or are not feasible (Sparre and Venema, 1998; Chang et al., 2012; Maunder and Piner, 2015). Pre-assessment processing of length frequency data is commonly conducted either through age-slicing (cohort-slicing) (Hilborn and Walters, 1992; Ailloud et al., 2015) or length frequency analysis (Hasselblad, 1966; Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017) where an ALK is not available. Age-slicing requires at least basic information on growth, to separate numbers-at-length data into length intervals according to an assumed distribution of lengths-at-age (Hilborn and Walters, 1992; Ailloud et al., 2015). Length frequency analysis is another typical approach to estimating growth parameters and assigning fish age prior to an assessment. This approach makes use of the distinct modes found in length frequency data and their progression through periodic data (Hasselblad, 1966; Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017; Batts et al., 2019).

A range of different methodologies have been developed for length frequency analysis, from relatively simple approaches such as the Bhattacharya method (Bhattacharya, 1967), which is a graphical method for separating length

frequency distributions into Gaussian components, to more complex model-based procedures such as ELEFAN (Electronic Length Frequency Analysis) and MULTIFAN (Pauly, 1987; Fournier et al., 1990; Taylor and Mildenerger, 2017). MULTIFAN is a widely used maximum likelihood-based procedure, whereas ELEFAN is a non-parametric procedure. Both models require fish lengths to be grouped into length bins prior to running the procedure and are typically constrained to estimating a single growth curve (although this is addressed to an extent with the development of the integrated stock assessment MULTIFAN-CL) (Pauly, 1987; Fournier et al., 1990, 1998). Recent developments of ELEFAN have improved performance with modern optimisation algorithms and have also provided methods for estimating uncertainty in parameter estimates via bootstrapping (Mildenerger et al., 2017; Taylor and Mildenerger, 2017; Schwamborn et al., 2019).

1.1.5 Implications for assessing fish stocks

Ageing difficulties are common to many fish stocks around the world and anglerfish fisheries also fall into this category. Anglerfish stocks face a number of challenges in their assessment and management, many of which arise from key uncertainties in data observation and processing, often specifically related to growth and age estimation (Perez et al., 2005a; Farina et al., 2008; Maguire et al., 2008; Richards, 2016; ICES, 2018). Other key uncertainties include aspects of anglerfish life-history that are not well-known (Farina et al., 2008).

1.2 Anglerfish fisheries and assessment

1.2.1 Anglerfish fisheries around the world

Commercially caught anglerfish, often referred to as monkfish, are a family of bottom-dwelling teleost fish (*Lophidae*) that are found all over the world and sustain fisheries in many of these regions (Wilson, 1937; Farina et al., 2008). These fisheries range from the historical and profitable fisheries found in the north-west (Haring and Maguire, 2008; Richards et al., 2008; Richards, 2016; Charbonneau et al., 2020) and north-east Atlantic (Laurenson, 2003; Farina et al., 2008; ICES, 2021a,e); to more recently exploited stocks in the south-eastern (Maartens and Booth, 2001; Walmsley et al., 2005; Glazer and Butterworth, 2013) and south-western Atlantic (Perez et al., 2005a). There is also a fishery in the Pacific (Yoneda et al., 1997; Sun et al., 2021). The majority

of these fisheries are conducted with bottom-trawling or gillnets and anglerfish are usually a by-catch or a component of a mixed fishery, although for some fleets and certain times of year they are the primary target (Perez et al., 2005a; Farina et al., 2008; Haring and Maguire, 2008; Glazer and Butterworth, 2013; ICES, 2018).

1.2.2 Historical anglerfish fisheries in the North Atlantic

In the earlier part of the 20th century anglerfish were mostly considered a by-catch species for many of the historical groundfish fisheries and their potential as a lucrative product had not been fully explored (Farina et al., 2008; Maguire et al., 2008; Richards et al., 2008). In the north-east and north-west Atlantic there were reports of widespread discarding of the species due to the lack of commercial value (Dobby et al., 2008; Haring and Maguire, 2008). However, from the 1980's onwards anglerfish fisheries in the North Atlantic grew rapidly (Farina et al., 2008).

In the US, this expansion of the fishery saw goosefish (*Lophius americanus*) become the highest value groundfish in the north-eastern United States (Richards et al., 2008). By the late 1990's, this fishery began to show signs of over-exploitation and a fisheries management plan was developed (Haring and Maguire, 2008). Some of the European anglerfish fisheries saw similar patterns in exploitation, with landings in the southern Bay of Biscay, as well as the North Sea, Rockall, West of Scotland, Skagerrak and Kattegat, reaching their peak in the 1980's or early 1990's respectively and decreasing since (ICES, 2021a,b,d). However, the anglerfish fishery in the Celtic Seas ecoregion has shown relatively steady catches since the mid-eighties (ICES, 2021c,e).

1.2.3 Anglerfish stocks in the North-East Atlantic

In the north-eastern Atlantic there are two species of exploited anglerfish; white-bellied anglerfish *Lophius piscatorius*; and black-bellied anglerfish *Lophius budegassa* (Farina et al., 2008; ICES, 2018; Vasilakopoulos and Maravelias, 2016). Distribution of each species differs somewhat across the region, with white-bellied anglerfish relatively more abundant in more northern waters and black-bellied more abundant in the south (Farina et al., 2008). For these two species, there are six stocks that the International Council for the Exploration of the Seas (ICES) provides advice for (ICES, 2018). Assessment of stocks is species-specific in both the Celtic Seas and Northern Bay of Biscay region

(ICES subarea 7 and divisions 8.a-b and 8.d (Figure 1.1))(ICES, 2021c,e), as well as the Cantabrian Sea and Atlantic Iberian waters (ICES divisions 8.c and 9.a) (ICES, 2021b,d). However, as anglerfish are not separated by species in the landings data, each of these regions are managed by a combined total allowable catch (TAC) for the two species in their respective management areas (ICES, 2018). Northern shelf stocks (anglerfish in the North Sea, Rockall, West of Scotland, Skagerrak and Kattegat - ICES subareas 4, 6 and 3.a; and anglerfish in the North-East Arctic - ICES subareas 1-2) jointly assess the population of both species and manage under a joint TAC (ICES, 2018, 2021a).

1.2.4 Current knowledge and state of anglerfish assessment and management

Collectively, anglerfish species have been considered particularly susceptible to over-exploitation, due to a combination of rapidly developing fisheries, lack of data and uncertainty in the life-history characteristics (Maguire et al., 2008). Nevertheless, there has been a considerable amount of work focussing on *Lophidae* species that has greatly improved our understanding of their biology and distribution (Farina et al., 2008). Focussing on northern Atlantic stocks, assessment and management of anglerfish is still somewhat problematic due to a multitude of data issues (e.g. catch separation, survey coverage, etc.) and a considerable degree of uncertainty in key aspects of the biology (e.g. age estimates, growth, maturity, stock structure etc.) (Farina et al., 2008; Maguire et al., 2008; ICES, 2018; Richards, 2016).

For example, the assessment of anglerfish in the US had adopted a SCALE model (a statistical catch-at-length model) in 2007 to assess the status of the population (NESFC, 2007). Whilst the assessment was based on catch-at-length data, it also relied upon a growth curve estimated from the vertebral ageing estimates of sampled anglerfish (NESFC, 2007; Richards, 2016). In 2016, it was found that the vertebral ageing method was not valid, leading to the abandonment of the SCALE model and an approach using the rate/direction of change in survey indices to assess the anglerfish stock in the previous two assessments (NESFC, 2007; Richards, 2016; Bank et al., 2020). Ability to incorporate accurate catch data is central to stock assessments as they represent observed removals.

On the other side of the Atlantic there are similar issues with ageing and growth, as well other uncertainties in the data (such as no separation of

species in the landings data) which make assessment problematic (Laurenson et al., 2005; ICES, 2011; Landa et al., 2013; ICES, 2018). Currently, of the six stocks that ICES provides advice for, only two (white-bellied anglerfish in: subarea 7, 8.a-b and 8.d (Figure 1.1); and in 8.c and 9.a) are defined as category 1 assessments (i.e., stocks with quantitative assessments, including those with full analytical assessments and forecasts as well as stocks with quantitative assessments based on production models (ICES, 2019)) (ICES, 2021d,e). Black-bellied anglerfish in 8.c and 9.a is a category 2 stock (analytical assessment treated qualitatively) and other stocks are classified as category 3 stocks (use relative abundance trends from assessments or survey trends to assess the population) or category 4 (only reliable catch available) in the case of Arctic anglerfish (ICES, 2018; ICES, 2019; ICES, 2021a,b,c).

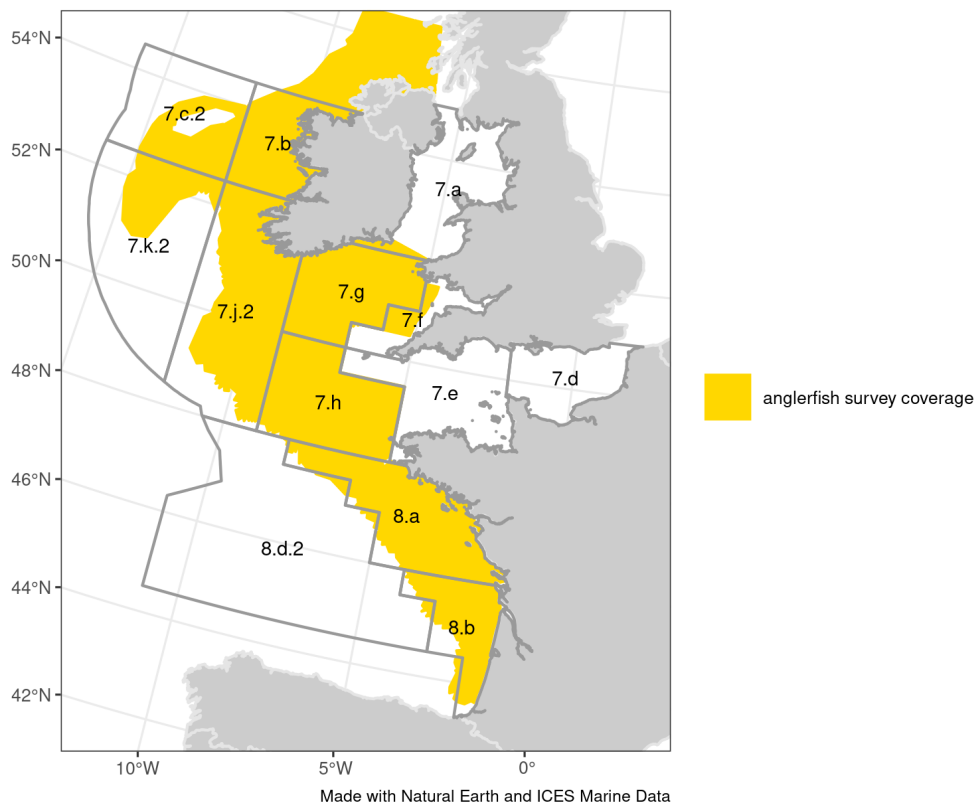


Figure 1.1: ICES subareas of the anglerfish stocks in the Celtic Sea and Northern Bay of Biscay and survey coverage (French EVHOE survey, Irish groundfish survey IE-IGFS and Spanish Porcupine Bank survey SP-PORC).

In an Irish context, anglerfish are an important component of the catch for the Irish demersal fishing industry (Marine Institute, 2020). The fleet mainly fish the Celtic Seas and Northern Bay of Biscay stocks and despite their relatively low quota share, anglerfish catch is one of the most valuable (Marine Institute, 2020). At the time this project began catch advice for both the white-bellied and black-bellied stocks were based on survey trends, with no estimation of stock levels, fishing mortality or MSY reference points (ICES, 2012). Many of the key uncertainties mentioned above affect these stocks but the main issue hampering assessments was inconsistency and uncertainty in age estimates and growth (ICES, 2018). An interesting development that occurred during the period of this study was the acceptance of a statistical catch-at-age assessment for the white-bellied anglerfish stock at a benchmark meeting, where age-composition data was constructed through model-assisted cohort-slicing partly informed by work done for this thesis Batts et al. (2019) (ICES, 2018).

This thesis focusses mainly on the white-bellied anglerfish stock in the Celtic Sea and Northern Bay of Biscay as a case study for the methods developed throughout (Figure 1.1).

1.2.5 Ageing difficulties for anglerfish

Structures used for ageing *Lophius* species differ between countries and there have also been numerous comparisons between vertebrae, otoliths and illicia (dorsal fin spine) over the years in regions where anglerfish are fished (Yoneda et al., 1997; Maartens et al., 1999; Laurenson et al., 2005; ICES, 2011; Landa et al., 2013; Ofstad et al., 2013). Anglerfish have historically been difficult to age with accuracy and precision (Farina et al., 2008; Bank et al., 2020). For anglerfish in the north Atlantic, the validity of age estimates is a key issue for the reliability of the stock assessments (Farina et al., 2008; ICES, 2018; Richards, 2016). In the US, anglerfish have traditionally been aged using vertebrae of the fish (Armstrong et al., 1992; Richards, 2016). Using chemical marking a recent study showed that growth increments in cross-sectioned vertebrae were not annual increments and therefore vertebrae did not give valid age estimates (Bank et al., 2020). In Europe, an ICES workshop compared age estimates from otoliths and illicia, finding the latter to be a more reliable structure in terms of precision and reliability, particularly when the reader is less experienced (ICES, 2011). However, there were still considerable

disparities in the age estimates between structures and between readers (ICES, 2011). Following this workshop there has been further validation of age estimation methods (Landa et al., 2013; Brophy, 2019; Brophy et al., 2021), however there is still concern about their robustness for use in assessments (ICES, 2018).

1.3 Stock assessment in the presence of ageing uncertainty

1.3.1 Choosing a stock assessment model

In many cases, it is the availability and reliability of data that drives the use of a particular stock assessment model (Hilborn and Walters, 1992). One approach to dealing with uncertainty in age estimates and individual growth of fish is to use alternative assessments to the typical statistical catch-at-age models (Maunder and Punt, 2013). Stock assessment models that are aggregated, stage-based, compositional size-based or integrated are often preferred in these situations (Smith and Addison, 2003; Punt et al., 2013; Kathena et al., 2018; Wang et al., 2020).

1.3.2 Aggregated biomass stock assessment models

Aggregated surplus-production stock assessment models such as ASPIC (A Surplus-Production model Incorporating Covariates) (Prager, 1992) and SPiCT (Surplus-Production in Continuous Time) (Pedersen and Berg, 2017) are one option that has been explored for anglerfish stocks. These models typically have less data requirements than catch-at-age models and often less assumptions. However, their performance can warrant caution for management use when there is a lack of contrast in available data (Hilborn and Walters, 1992; Bouch et al., 2021) and they are generally limited by biomass indices that should be from the exploited biomass thus do not contain information on incoming recruits. ASPIC is currently used for the black-bellied anglerfish in ICES areas 8.c and 9.a (ICES, 2021b) and SPiCT was investigated as a “data moderate” comparison for the age-based stock assessment of *Lophius vomerinus* in Namibia (Kathena et al., 2018). Aggregated biomass estimates have also been used to assess the anglerfish stock in Brazilian waters (Perez et al., 2005b). Elsewhere, a length-based estimator (LBB) was used to assess

Lophius litulon (Wang et al., 2020) and as mentioned previously, the US assessment for *Lophius americanus* was based on a statistical catch-at-length model (SCALE) until issues with age estimates were highlighted (NESFC, 2007; Richards, 2016). In compositional size-based models such as LBB and SCALE, uncertainty in age estimates can often be circumvented somewhat by the use of raw length frequency data. However, sometimes these models still require information on growth (i.e. mean length-at-age) (NESFC, 2007).

1.3.3 Integrated stock assessment models

In recent years there has been an increase in the use of integrated stock assessments such as Stock Synthesis (Methot and Wetzel, 2013), CASAL (C++ Algorithmic Stock Assessment Laboratory) (Bull et al., 2005) or MULTIFAN-CL (Fournier et al., 1998). These models are relatively complex and can often incorporate different forms of data. Integrated stock assessments also often have the added advantage of incorporating uncertainty in growth or age-composition data directly into the estimation of the model (Maunder and Piner, 2015; Punt et al., 2021), however there is a trade-off as they require more data and more expert experience than simpler models (Ailloud et al., 2015). In the context of anglerfish stocks, a Stock Synthesis assessment is currently implemented for the white-bellied anglerfish stock in ICES areas 8.c and 9.a (ICES, 2021d). At the same stock assessment benchmark where the assessment for white-bellied anglerfish in ICES areas 8.c and 9.a was developed Stock Synthesis was also applied to other stocks but these were not considered suitable for candidate assessments due to the lack of time to fully develop the models at the meeting (ICES, 2018).

1.3.4 Stage-based stock assessments as an alternative

Stage-based models are a somewhat under-utilised form of assessment that can be considered an alternative middle ground between aggregate and complex compositional models. Generally, these models require considerably less data than compositional stock assessments along with relatively simple population dynamics and assumptions, whilst still incorporating information on recruitment (Hilborn and Walters, 1992). Two theoretically different approaches to stage-based assessment models are: biomass-based delay-difference models (Deriso, 1980; Schnute, 1985; Fournier and Doonan, 1987; Schnute, 1987; Meyer and Millar, 1999; Kienzle et al., 2014); and numbers-based depletion

models, the most prevalent of which is Catch-Survey Analysis (CSA) (Collie and Sissenwine, 1983; Cadrin, 2000; Smith and Addison, 2003; Cook and Heath, 2018; Li et al., 2019). Stage-based models have generally been used for species that are difficult to age (Cadrin, 2000; Smith and Addison, 2003; Zhou et al., 2011; Kienzle et al., 2014), are short-lived or lacking in data (e.g. by-catch species) (Cook and Heath, 2018). One of the recommendations from the ICES stock assessment benchmark for anglerfish was to develop stage-based assessment models for the species (ICES, 2018).

1.4 Testing the effect of ageing uncertainty and bias on stock assessment and management

It is important to understand how uncertainty and bias in data may impact stock assessment and fisheries management over time (Polacheck et al., 1999; Punt et al., 2014; Privitera-Johnson and Punt, 2020). The traditional approach of testing certain sensitivities of the “best assessment” does not always give a complete picture, as a full range of uncertainties are rarely taken into account, there is a lack of feedback control and risk is not properly evaluated (Butterworth, 2007). Management Strategy Evaluation (MSE) offers a comprehensive simulation framework where the effect of key uncertainties such as age-composition and growth uncertainty can be investigated in detail (Smith et al., 1999; Butterworth, 2007; Kell et al., 2007; Punt et al., 2014).

1.4.1 Origins of management strategy evaluation

Management strategy evaluation originated within the Scientific Committee of the International Whaling Commission (IWC). Issues were encountered with the traditional approach of developing a “best assessment” for a resource and so a method was sought to investigate the appropriateness of values used within harvest control rules and to account for their uncertainty (de la Mare, 1986; Kirkwood, 1992; Butterworth, 2007; Punt and Donovan, 2007). Around this time the Precautionary Approach (PA) was beginning to be adopted and an FAO Technical Consultation on the Precautionary Approach to Capture Fisheries in 1995 explicitly recommended the approach developed by the IWC (Butterworth, 2007). MSEs have been used and developed extensively since the 1990’s, as applied in regions such as South Africa and Australia, who

were early adopters of the MSE approach to manage many of their fisheries (Smith, 1994; Smith et al., 1999; Plagányi et al., 2007; Punt et al., 2014). Generally, the focus has been on single species MSE's but in more recent years there has been extensive development of multispecies MSEs based on various multispecies models such as Atlantis and Gadget (e.g. Howell and Bogstad, 2010; Fulton et al., 2014).

1.4.2 Components and attributes of management strategy evaluations

Management strategies or management procedures (MPs) are a key component of an MSE and are comparable to the “traditional approach” to fisheries management where, an estimator (e.g. abundance or harvest state) is applied to a stock, usually some form of reference points are estimated and a harvest control rule is applied to give future management controls such as a given level of effort or total allowable catch (TAC) (Butterworth, 2007; Holland, 2010) (Figure 1.2). The estimator does not always have to be model-based (i.e. stock assessment) and empirical estimators (e.g., survey trends-based) can be used to good effect in certain circumstances such as when there are absolute abundance estimates or as an interim MP if there is not sufficient data available for a model-based assessment immediately (Rademeyer et al., 2007).

The other main component of an MSE is the operating model (OM), which represents the “true” dynamics of the stock and the fishery (in contrast to the MP which can be thought of as how the stock is perceived)(Holland, 2010). An OM is a simulated population and fishery with associated stochasticity (e.g., recruitment variability). Linking the OM to the MP is the observation error model (OEM), where data that would be available to an MP is generated with associated uncertainty and stochasticity (Punt et al., 2014). The OEM is where uncertainty and/or bias in age estimates can be incorporated. Linking the MP back to the OM is the implementation error model. Performance statistics are often used to evaluate the performance of MPs with respect to management objectives with quantitative results (e.g., probability of stock dropping below a given level or reference point) (Kell et al., 2005; Holland, 2010; Punt et al., 2014) (Figure 1.2).

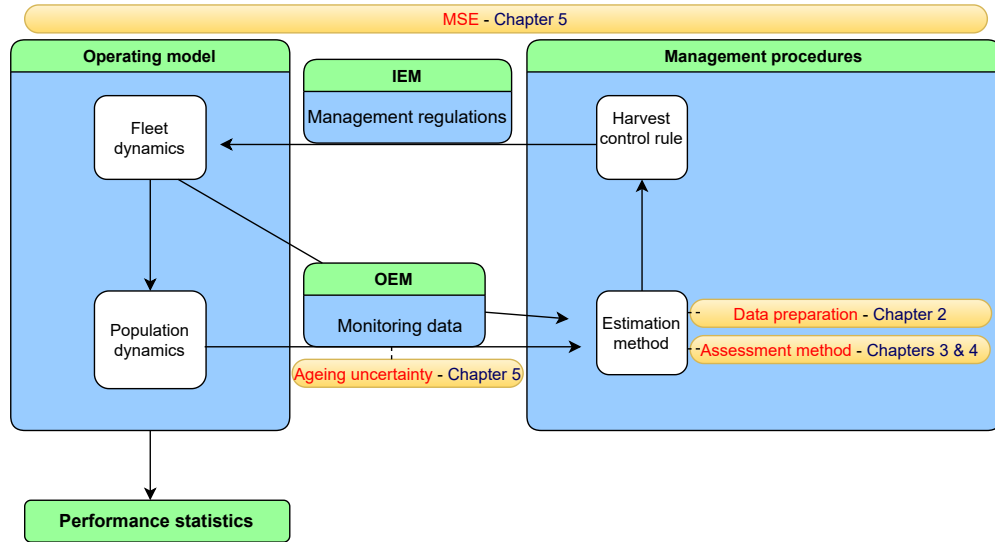


Figure 1.2: Schematic of a typical management strategy evaluation (MSE) simulation framework. Components include: operating model (OM), observation error model (OEM), implementation error model (IEM), management procedure (MP), and performance statistics. Also shown are the chapters of this thesis in relation to the relevant aspects of an MSE framework. Adapted from Punt et al. (2014).

An MSE is a simulation framework that allows users to assess the performance and trade-offs of MPs, given the properties and uncertainty inherent in the system. Users can assess the ability of an MP to achieve management goals (e.g., preventing stock collapse) set for the stock (Holland, 2010; Punt et al., 2014). Where MSEs differ from sensitivity testing within the traditional approach is that they can incorporate a wider range of uncertainties into simulations, including in the OM itself. Furthermore, a key aspect of MSEs is the feedback loop from the simulated management decisions (MP) to the simulated population (OM), which allows for a better understanding for how an MP will perform over a number of years. (Butterworth, 2007; Punt et al., 2014) (Figure 1.2). MSEs have not only been used to assess the potential risk (and reward) of MPs in a range of different contexts (Dichmont et al., 2006; Butterworth, 2007; Szuwalski and Punt, 2013; Kell et al., 2014), the framework can also be utilised to identify future research areas to focus on and allocate resources efficiently to reduce scientific uncertainty, as well as understand how potential biases and uncertainty within the system can affect the status of stock over time (Marasco et al., 2007; Kell et al., 2012; Tyszko and Pritt, 2017; Carruthers et al., 2015).

1.5 Gaps in our understanding

Current approaches to tracking cohorts from length frequency analysis leave scope for further development of new methods. Both MULTIFAN and ELEFAN require pre-defined length bins, potentially introducing subjectivity. Both models are also constrained to a single growth curve and do not incorporate inter-annual or inter-cohort variability in growth (although this is addressed to some extent in the development of MULTIFAN-CL) (Fournier et al., 1998). In this thesis I address these limitations by developing an approach that removes the need to pre-define length classes and offers the flexibility to include variability in growth.

As mentioned previously, stage-based stock assessment models are a set of useful data-moderate stock assessment models that are somewhat underutilised. In many instances these models have been tested with simulations and on real stocks (Collie and Sissenwine, 1983; Meyer and Millar, 1999; Cadrin, 2000; Mesnil, 2003; Kienzle et al., 2014; Zhou et al., 2011), however there has been a lack of a comprehensive simulation-testing framework where numbers-based and biomass-based stage-based assessment models were compared. I address this shortcoming in this thesis and add further evidence to the usefulness of stage-based stock assessment in the wider context of fisheries stock assessment and also specifically for anglerfish stocks (which was recommended at the stock assessment benchmark (ICES, 2018)).

Stage-based assessment models may limit the need to disentangle all cohorts but how do these methods perform in an applied management context? CSA has not been tested in a management strategy evaluation framework, nor has the current assessment and management approach (an age-based assessment and the ICES advice rule) been tested for an OM based on the life-history of anglerfish. Engagement with stakeholders is a key aspect of a true MSE (Butterworth, 2007; Punt et al., 2014) and an early consultation with Irish fishing industry stakeholders was part of the MSE process in this thesis. These stakeholders expressed a desire to test the effectiveness of the ICES advice rule as a starting point. There is also still much to be learned about the effect of ageing uncertainty and bias on stock assessment and fisheries management over time (Lai and Gunderson, 1987; Tyler et al., 1989; Bertignac and De Pontual, 2007). The full-feedback system of a management strategy evaluation framework is an ideal setting to address these knowledge gaps in this thesis.

1.6 Thesis aims and objectives

Key uncertainties prevail in stock assessment and fisheries management. The main objective of this thesis was to improve our knowledge of how these key uncertainties (specifically focussing on age estimates and growth) potentially impact the fish populations and fisheries, as well as investigating methods to account for them. The methods developed in this thesis focus on anglerfish as a case study.

The overall aim of this thesis was to explore and develop methods used to address cohort uncertainty in fisheries assessment and management (Chapters 2-4), and to developing a framework for testing how uncertainty in the form of biased age-composition data may influence stock status and management goals over time (Chapter 5). The white-bellied anglerfish *Lophius piscatorius* stock in the Celtic Seas and Bay of Biscay region (ICES subareas 7, 8.a-b and 8.d) was used throughout as the main case study stock but these methods are broadly applicable to many other fish stocks. Firstly, a new method for length frequency analysis was developed (Chapter 2). Next, stage-based stock assessments were implemented and tested in a factorial simulation experiment covering a range of population and fishery features/factors. (Chapter 3), in addition to an associated R package with vignettes (Chapter 4). Finally, a management strategy evaluation framework was constructed to test the impact of uncertainty in growth estimates of white-bellied anglerfish (Chapter 5).

Specific objectives for this thesis are presented by chapter below. These are linked via the background of the General Introduction and placed in context in the General Discussion:

Chapter 1: General Introduction

Provides an overview of: stock assessment and the challenges of age-estimation; anglerfish fisheries; estimating fish growth when uncertainty in age-estimates is present; stock assessment when uncertainty in ageing estimates is present and testing the effect of ageing uncertainty on stock assessment and fisheries management over time.

Chapter 2: Estimating growth parameters and growth variability

from length frequency data using hierarchical mixture models.

A new length frequency analysis procedure (LFEM) that uses mixture models and an Expectation-Maximisation algorithm was developed to estimate growth parameters by maximum likelihood. The basic model estimated a single set of von Bertalanffy growth parameters that determined component means and standard deviations that best fit length frequency distributions from multiple surveys and years. The LFEM model was then extended into a hierarchical framework, which allows for bivariate random effects to be estimated and variability in fish growth between cohorts and/or years to be modelled. Testing on haddock and white-bellied anglerfish showed that the basic model provides a useful alternative to available methods, whilst the hierarchical extension enables modelling of some of the natural variation in growth of fish.

This Chapter has been published in the *ICES Journal of Marine Science* (Manuscript I):

Batts, L., Minto, C., Gerritsen, H., & Brophy, D. (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science*, 76(7), 2150-2163.

Chapter 3: Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real life stocks.

Stage-based stock assessment models are an alternative middle ground between aggregate and compositional stock assessment models. Two contrasting stage-based stock assessment models, the numbers-based Catch-Survey Analysis (CSA); and a biomass-based delay-difference model (Schnute, 1987), were implemented and extended. Model performance was compared in a comprehensive simulation testing framework and in comparison to the accepted age-based assessment on the real white-bellied anglerfish stock in ICES subareas 7, 8.a-b and 8.d. In the simulation-testing

framework, estimates from the biomass-based two-stage models were close to the true values in certain scenarios, but were sensitive to selectivity assumptions and configuration of growth within the model. CSA was more robust to selectivity assumptions, performing well in all simulated stock scenarios. CSA estimates of stock numbers were remarkably similar to the values estimated in the accepted age-based assessment for the real anglerfish stock.

This Chapter has been published in the *Canadian Journal of Fisheries and Aquatic Sciences* (Manuscript II):

Batts, L., Minto, C., Gerritsen, H., & Brophy, D. (2022). Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real life stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, Just-In

Chapter 4: sbar vignette: An R package for implementing stage-based assessment models

An R package was developed for the stage-based assessment models implemented in Chapter 3. Key functions were documented and were demonstrated in the vignette. The stage-based assessment models: Catch-Survey Analysis (CSA), and a delay-difference model (Schnute, 1987), were introduced and their key assumptions outlined. Simple and more complex examples of the models in use were given, as well as detailing some of the models (versions not implemented in Chapter 3) in more detail. Black-bellied anglerfish in the Celtic Seas and Bay of Biscay, which is not currently analytically assessed with an assessment model, was used as an example stock.

The package can be found at: <https://github.com/lbatts/sbar>.

Chapter 5: Comparative impacts of ageing bias on stage and age-based fisheries stock assessment models.

Accounting for uncertainty in age-composition data can be difficult in

age-based stock assessment models. A management strategy evaluation framework was used to investigate the impact of bias in age-composition data on the assessment and management of white-bellied anglerfish. Both stage-based and age-based management procedures were investigated. Bias in age-composition data was found to impact the stock in opposite ways for each of the MPs. The perception of the stock was directly affected in the age-based MP, whereas for the stage-based MP biased age-composition data affected the management indirectly through implementation. The nature of bias (high or low growth) on age-composition data had contrasting effects on the stock status and stock trajectory depending on the MP. Non-proxy reference points for CSA were also developed as part of this research and in combination with the CSA assessment model results showed that a CSA MP could be used to successfully assess and manage a stock over time.

This chapter is in preparation for submission to the *ICES Journal of Marine Science*:

Comparative impacts of ageing bias on stage and age-based fisheries stock assessment models. Batts, L., Minto, C., Gerritsen, H., & Brophy, D.

Chapter 6: General Discussion

Here, I provide a summary of the findings within this thesis and place them in the context of existing knowledge. Application of the methods developed are discussed broadly, as well as specifically in the context of anglerfish stocks. Implications of cohort uncertainty on stock assessment and fisheries management over time are addressed. I also explore future extensions of the work.

Bibliography

- Ailloud, L. E., M. W. Smith, A. Y. Then, K. L. Omori, G. M. Ralph, and J. M. Hoenig (2015). Properties of age compositions and mortality estimates derived from cohort slicing of length data. *ICES Journal of Marine Science* 72(1), 44–53.
- Armstrong, M. P., J. Musick, and J. A. Colvocoresses (1992). Age, growth, and reproduction of the goosefish *Lophius-americanus* (*Pisces, Lophiiformes*). *Fishery Bulletin* 90(2), 217.
- Bank, C. M., K. Oliveira, S. J. Sutherland, M. P. Armstrong, J. Landa, and S. X. Cadrin (2020). Age validation for goosefish (*Lophius americanus*) in the northeastern United States. *Fishery Bulletin* 118(1).
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science* 76(7), 2150–2163.
- Begley, J. and D. Howell (2004). An overview of GADGET, the globally applicable area-disaggregated general ecosystem toolbox. ICES CM 2004/FF:13.
- Bertignac, M. and H. De Pontual (2007). Consequences of bias in age estimation on assessment of the northern stock of European hake (*Merluccius merluccius*) and on management advice. *ICES Journal of Marine Science* 64(5), 981–988.
- Bhattacharya, C. (1967). A simple method of resolution of a distribution into Gaussian components. *Biometrics*, 115–135.
- Bianchi, M., F. Chopin, T. Farme, N. Franz, C. Fuentevilla, L. Garibaldi, and A. Laurenti (2014). FAO: The State of World Fisheries and Aquaculture. *Food and Agriculture Organization of the United Nations: Rome, Italy*.

- Blanchard, J. L., S. Jennings, R. Holmes, J. Harle, G. Merino, J. I. Allen, J. Holt, N. K. Dulvy, and M. Barange (2012). Potential consequences of climate change for primary production and fish production in large marine ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1605), 2979–2989.
- Bouch, P., C. Minto, and D. G. Reid (2021). Comparative performance of data-poor CMSY and data-moderate SPiCT stock assessment methods when applied to data-rich, real-world stocks. *ICES Journal of Marine Science* 78(1), 264–276.
- Brophy, D. (2019). Validating age-determination of anglerfish and hake. Technical report, EASME/EMFF/2016/1.3.2.7/SI2.762036. Final Report. <https://doi.org/10.2826/748632>.
- Brophy, D., S. Pérez-Mayol, R. Duncan, K. Hüsey, A. J. Geffen, H. D. Gerritsen, M. C. Villanueva, and B. Morales-Nin (2021). Elemental composition of illicia and otoliths and their potential application to age validation in white anglerfish (*Lophius piscatorius* linnaeus, 1758). *Estuarine, Coastal and Shelf Science* 261, 107557.
- Bull, B., R. Francis, A. Dunn, A. McKenzie, D. Gilbert, M. Smith, R. Bian, and D. Fu (2005). CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.
- Butterworth, D. S. (2007). Why a management procedure approach? some positives and negatives. *ICES Journal of Marine Science* 64(4), 613–617.
- Cadigan, N. G., E. Wade, and A. Nielsen (2017). A spatiotemporal model for snow crab (*Chionoecetes opilio*) stock size in the southern Gulf of St. Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* 74(11), 1808–1820.
- Cadrin, S. X. (2000). Evaluating two assessment methods for Gulf of Maine northern shrimp based on simulations. *Journal of Northwest Atlantic Fishery Science* 27, 119–132.
- Campana, S. (2001). Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of fish biology* 59(2), 197–242.

- Carruthers, T. R., J. F. Walter, M. K. McAllister, and M. D. Bryan (2015). Modelling age-dependent movement: an application to red and gag groupers in the Gulf of Mexico. *Canadian Journal of Fisheries and Aquatic Sciences* 72(8), 1159–1176.
- Chang, Y.-J., C.-L. Sun, Y. Chen, and S.-Z. Yeh (2012). Modelling the growth of crustacean species. *Reviews in Fish Biology and Fisheries* 22(1), 157–187.
- Charbonneau, J. A., D. M. Keith, M. A. MacNeil, J. A. Sameoto, and J. A. Hutchings (2020). Pervasive declines in monkfish (*Lophius americanus*) size structure throughout the northwest Atlantic. *Fisheries Research* 230, 105633.
- Chen, Y. (2003). Quality of fisheries data and uncertainty in stock assessment. *Scientia Marina* 67(S1), 75–87.
- Chrysafi, A., J. M. Cope, and A. Kuparinen (2019). Eliciting expert knowledge to inform stock status for data-limited stock assessments. *Marine Policy* 101, 167–176.
- Collie, J. S. and M. P. Sissenwine (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* 40(11), 1871–1879.
- Cook, R. M. and M. R. Heath (2018). Population trends of bycatch species reflect improving status of target species. *Fish and Fisheries* 19(3), 455–470.
- Costello, C., D. Ovando, R. Hilborn, S. D. Gaines, O. Deschenes, and S. E. Lester (2012). Status and solutions for the world’s unassessed fisheries. *Science* 338(6106), 517–520.
- de la Mare, W. K. (1986). Simulation studies on management procedures. Technical report, Reports of the International Whaling Commission, Volume 36, pg. 429–450.
- De Pontual, H., A. L. Groison, C. Piñeiro, and M. Bertignac (2006). Evidence of underestimation of European hake growth in the Bay of Biscay, and its relationship with bias in the agreed method of age estimation. *ICES Journal of Marine Science* 63(9), 1674–1681.
- Denechaud, C., S. Smoliński, A. J. Geffen, J. A. Godiksen, and S. E. Campana (2020). A century of fish growth in relation to climate change, population dynamics and exploitation. *Global Change Biology* 26(10), 5661–5678.

- Denney, N. H., S. Jennings, and J. D. Reynolds (2002). Life–history correlates of maximum population growth rates in marine fishes. *Proceedings of the Royal Society of London B: Biological Sciences* 269(1506), 2229–2237.
- Deriso, R. B. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* 37(2), 268–282.
- Dichmont, C. M., A. R. Deng, A. E. Punt, W. Venables, and M. Haddon (2006). Management strategies for short lived species: the case of Australia’s Northern Prawn Fishery: 2. Choosing appropriate management strategies using input controls. *Fisheries Research* 82(1-3), 221–234.
- Dobby, H., L. Allan, M. Harding, C. Laurenson, and H. McLay (2008). Improving the quality of information on Scottish anglerfish fisheries: making use of fishers’ data. *ICES Journal of Marine Science* 65(7), 1334–1345.
- FAO (2020). The state of world fisheries and aquaculture. Technical report, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.
- Farina, A., M. Azevedo, J. Landa, R. Duarte, P. Sampedro, G. Costas, M. Torres, and L. Cañás (2008). *Lophius* in the world: a synthesis on the common features and life strategies. *ICES Journal of Marine Science: Journal du Conseil* 65(7), 1272–1280.
- Folkvord, A., G. Blom, A. Johannessen, and E. Moksness (2000). Growth-dependent age estimation in herring (*Clupea harengus L.*) larvae. *Fisheries Research* 46(1-3), 91–103.
- Fournier, D. A. and I. J. Doonan (1987). A length-based stock assessment method utilizing a generalized delay-difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 44(2), 422–437.
- Fournier, D. A., J. Hampton, and J. R. Sibert (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences* 55(9), 2105–2116.
- Fournier, D. A., J. R. Sibert, J. Majkowski, and J. Hampton (1990). MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using

- data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian journal of fisheries and aquatic sciences* 47(2), 301–317.
- Froese, R., N. Demirel, G. Coro, K. M. Kleisner, and H. Winker (2017). Estimating fisheries reference points from catch and resilience. *Fish and Fisheries* 18(3), 506–526.
- Fulton, E. A., A. D. Smith, D. C. Smith, and P. Johnson (2014). An integrated approach is needed for ecosystem based fisheries management: insights from ecosystem-level management strategy evaluation. *PloS one* 9(1), e84242.
- Glazer, J. P. and D. S. Butterworth (2013). An updated assessment of the South African Monkfish resource, *Lophius vomerinus*. Technical report, Report No. FISHERIES/2013/SEPT/SWG-DEM/42 (Capetown, South Africa).
- Halouani, G., F. Le Loc'h, Y.-J. Shin, L. Velez, T. Hattab, M. S. Romdhane, and F. B. R. Lasram (2019). An end-to-end model to evaluate the sensitivity of ecosystem indicators to track fishing impacts. *Ecological Indicators* 98, 121–130.
- Haring, P. and J. Maguire (2008). The monkfish fishery and its management in the northeastern USA. *ICES Journal of Marine Science* 65(7), 1370–1379.
- Hasselblad, V. (1966). Estimation of parameters for a mixture of normal distributions. *Technometrics* 8(3), 431–444.
- Hilborn, R., R. O. Amoroso, C. M. Anderson, J. K. Baum, T. A. Branch, C. Costello, C. L. de Moor, A. Faraj, D. Hively, O. P. Jensen, H. Kurota, L. R. Little, P. Mace, T. McClanahan, M. C. Melnychuk, C. Minto, G. C. Osio, A. M. Parma, M. Pons, S. Segurado, C. S. Szuwalski, J. R. Wilson, and Y. Ye (2020). Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences* 117(4), 2218–2224.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.
- Holland, D. S. (2010). Management strategy evaluation and management procedures: tools for rebuilding and sustaining fisheries.

- Hordyk, A., K. Ono, S. Valencia, N. Loneragan, and J. Prince (2015). A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science* 72(1), 217–231.
- Howell, D. and B. Bogstad (2010). A combined Gadget/FLR model for management strategy evaluations of the Barents Sea fisheries. *ICES Journal of Marine Science* 67(9), 1998–2004.
- Hüssy, K., K. Radtke, M. Plikshs, R. Oeberst, T. Baranova, U. Krumme, R. Sjöberg, Y. Walther, and H. Mosegaard (2016). Challenging ICES age estimation protocols: lessons learned from the eastern Baltic cod stock. *ICES Journal of Marine Science* 73(9), 2138–2149.
- ICES (2011). Report of the Anglerfish (*Lophius piscatorius*) illicia and otoliths exchange 2011, ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS).
- ICES (2012). Report of the Benchmark Workshop on Flatfish Species and Anglerfish (WKFLAT), 1–8 March 2012, Bilbao, Spain. ICES CM 2012/A-COM:46. 283 pp.
- ICES (2018). Report of the Benchmark Workshop on Anglerfish Stocks in the ICES Area (WKANGLER), 12–16 February 2018, Copenhagen, Denmark. ICES CM 2018.
- ICES (2019). Advice basis. In Report of the ICES Advisory Committee, 2019. ICES Advice 2019, section 1.2.
- ICES (2021a). Anglerfish (*Lophius budegassa*, *Lophius piscatorius*) in subareas 4 and 6, and division 3.a (North sea, Rockall and West of Scotland, Sagerrak and Kattegat). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, anf.27.3a46.
- ICES (2021b). Black-bellied anglerfish (*Lophius budegassa*) in divisions 8.c and 9.a (Cantabrian sea, Atlantic Iberian waters). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, ank.27.8c9a.
- ICES (2021c). Black-bellied anglerfish (*Lophius budegassa*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, ank.27.78abd.

- ICES (2021d). White anglerfish (*Lophius piscatorius*) in divisions 8.c and 9.a (Cantabrian sea, Atlantic Iberian waters). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, mon.27.8c9a.
- ICES (2021e). White anglerfish (*Lophius piscatorius*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, mon.27.78abd.
- Jardim, E., C. P. Millar, I. Mosqueira, F. Scott, G. C. Osio, M. Ferretti, N. Alzorriz, and A. Orió (2014). What if stock assessment is as simple as a linear model? the a4a initiative. *ICES Journal of Marine Science* 72(1), 232–236.
- Kathena, J., A. Kokkalis, M. Pedersen, J. Beyer, and U. Thygesen (2018). Data-moderate assessments of Cape monkfish *Lophius vomerinus* and west coast sole *Austroglossus microlepis* in Namibian waters. *African Journal of Marine Science* 40(3), 293–302.
- Kell, L., G. Pilling, G. Kirkwood, M. Pastoors, B. Mesnil, K. Korsbrekke, P. Abaunza, R. Aps, A. Biseau, P. Kunzlik, and C. Needle (2005). An evaluation of the implicit management procedure used for some ICES roundfish stocks. *ICES Journal of Marine Science* 62(4), 750–759.
- Kell, L. T., J.-M. Fromentin, and S. Bonhommeau (2012). An evaluation of the implications of population structure on the current bluefin tuna advice framework. *Collective Volume of Scientific Papers* 68(1), 370–378.
- Kell, L. T., G. Merino, P. De Bruyn, J. O. de Urbina, H. Arrizabalaga, and J. S. H. Murua (2014). An example management strategy evaluation of a harvest control rule. *Collective Volume of Scientific Papers. ICCAT* 70(5), 2096–2110.
- Kell, L. T., I. Mosqueira, P. Grosjean, J.-M. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. Pastoors, J. Poos, and F. Scott (2007). FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science* 64(4), 640–646.
- Kienzle, M., A. J. Courtney, and M. F. O’Neill (2014). Environmental and fishing effects on the dynamics of brown tiger prawn (*Penaeus esculentus*) in Moreton Bay (Australia). *Fisheries research* 155, 138–148.

- Kirkwood, G. (1992). Background to the development of revised management procedures. *Report of the International Whaling Commission 42*, 236–243.
- Lai, H. L. and D. R. Gunderson (1987). Effects of ageing errors on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). *Fisheries Research 5*(2-3), 287–302.
- Landa, J., J. Barrado, and F. Velasco (2013). Age and growth of anglerfish (*Lophius piscatorius*) on the Porcupine Bank (west of Ireland) based on illicia age estimation. *Fisheries research 137*, 30–40.
- Laurenson, C., A. Johnson, and I. Priede (2005). Movements and growth of monkfish *lophius piscatorius* tagged at the Shetland Islands, northeastern Atlantic. *Fisheries Research 71*(2), 185–195.
- Laurenson, C. H. (2003). *Aspects of the Biology and Fishery for Monkfish Lophius Piscatorius Linnaeus 1758 in Waters Around the Shetland Isles, Northeastern Atlantic*. Ph. D. thesis, University of Aberdeen.
- Li, Y., L. M. Lee, and J. Rock (2019). Modeling population dynamics and nonstationary processes of difficult-to-age fishery species with a hierarchical Bayesian two-stage model. *Canadian Journal of Fisheries and Aquatic Sciences 76*(12), 2199–2214.
- Maartens, L. and A. J. Booth (2001). Assessment of the monkfish *Lophius vomerinus* resource off Namibia. *African Journal of Marine Science 23*, 275–290.
- Maartens, L., A. J. Booth, and T. Hecht (1999). The growth of monkfish *Lophius vomerinus* in Namibian waters, with a comparison of otolith and illicia methods of ageing. *Fisheries Research 44*(2), 139–148.
- Maguire, J.-J., P. Pereda, R. Duarte, H. Dobby, and M. Azevedo (2008). Monkfish/anglerfish across the world; common problems and common solutions: an introduction to papers presented at the ices theme session in september 2007. *ICES Journal of Marine Science 65*(7), 1270–1271.
- Marasco, R. J., D. Goodman, C. B. Grimes, P. W. Lawson, A. E. Punt, and T. J. Quinn II (2007). Ecosystem-based fisheries management: some practical suggestions. *Canadian Journal of Fisheries and Aquatic Sciences 64*(6), 928–939.

- Marine Institute (2020). The Stock Book 2020: Annual Review of Fish Stocks in 2020 with Management Advice for 2021. Marine Institute, Galway, Ireland. Technical report.
- Martell, S. and R. Froese (2013). A simple method for estimating MSY from catch and resilience. *Fish and Fisheries* 14(4), 504–514.
- Maunder, M. N. and K. R. Piner (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science* 72(1), 7–18.
- Maunder, M. N. and A. E. Punt (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research* 142, 61–74.
- McMillan, P. J., A. C. Hart, P. L. Horn, D. M. Tracey, C. Ó Maolagáin, B. Finucci, and M. R. Dunn (2021). Review of life history parameters and preliminary age estimates of some New Zealand deep-sea fishes. *New Zealand Journal of Marine and Freshwater Research* 55(4), 565–591.
- Mesnil, B. (2003). The Catch-Survey Analysis (CSA) method of fish stock assessment: an evaluation using simulated data. *Fisheries Research* 63(2), 193–212.
- Methot, R. D. and C. R. Wetzel (2013). Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142, 86–99.
- Meyer, R. and R. B. Millar (1999). Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 56(1), 37–52.
- Mildenberger, T. K., M. H. Taylor, and M. Wolff (2017). Tropfishr: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution* 8(11), 1520–1527.
- Morales-Nin, B. and J. Panfili (2005). Seasonality in the deep sea and tropics revisited: what can otoliths tell us? *Marine and Freshwater Research* 56(5), 585–598.
- NESFC (2007). Northeast Data Poor Stocks Working Group. Monkfish assessment report for 2007. US Department of Commerce NEFSC (Northeast Fisheries Science Center). *Northeast Fisheries Science Center Ref Doc. 07-21; 232 p.*

- Nielsen, A. and C. W. Berg (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158, 96–101.
- Ofstad, L. H., C. Angus, T. Pedersen, and P. Steingrund (2013). Age and growth of anglerfish (*Lophius piscatorius*) in Faroese waters. *Fisheries Research* 139, 51–60.
- Pauly, D. (1987). A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. In *ICLARM conf. Proc.*, Volume 13, pp. 7–34.
- Pedersen, M. W. and C. W. Berg (2017). A stochastic surplus production model in continuous time. *Fish and Fisheries* 18(2), 226–243.
- Perez, J. A. A., P. R. Pezzuto, and H. A. de Andrade (2005a). Biomass assessment of the monkfish *Lophius gastrophysus* stock exploited by a new deep-water fishery in southern Brazil. *Fisheries research* 72(2-3), 149–162.
- Perez, J. A. A., P. R. Pezzuto, and H. A. de Andrade (2005b). Biomass assessment of the monkfish *Lophius gastrophysus* stock exploited by a new deep-water fishery in southern Brazil. *Fisheries research* 72(2), 149–162.
- Plagányi, É. E., R. A. Rademeyer, D. S. Butterworth, C. L. Cunningham, and S. J. Johnston (2007). Making management procedures operational—innovations implemented in South Africa. *ICES Journal of Marine Science* 64(4), 626–632.
- Polacheck, T., N. Klaer, C. Millar, and A. Preece (1999). An initial evaluation of management strategies for the southern bluefin tuna fishery. *ICES Journal of Marine Science* 56(6), 811–826.
- Prager, M. (1992). Aspic-a surplus-production model incorporating covariates. *Collective Volume of Scientific Papers, ICCAT* 28, 218–229.
- Privitera-Johnson, K. M. and A. E. Punt (2020). A review of approaches to quantifying uncertainty in fisheries stock assessments. *Fisheries Research* 226, 105503.
- Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. De Oliveira, and M. Haddon (2014). Management strategy evaluation: best practices. *Fish and Fisheries* 17(2), 303–334.

- Punt, A. E. and G. P. Donovan (2007). Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. *ICES Journal of Marine Science* 64(4), 603–612.
- Punt, A. E., T. Huang, and M. N. Maunder (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70(1), 16–33.
- Punt, A. E., F. Trinnie, T. I. Walker, R. McGarvey, J. Feenstra, A. Linnane, and K. Hartmann (2013). The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face of non-stationary dynamics. *Fisheries Research* 137, 116–128.
- Punt, A. E., G. N. Tuck, J. Day, P. Burch, R. B. Thomson, and P. Bessell-Browne (2021). The impact of alternative age-length sampling schemes on the performance of stock assessment methods. *Fisheries Research* 238, 105904.
- Quinn, T. J. and R. B. Deriso (1999). *Quantitative Fish Dynamics*. Oxford University Press.
- Rademeyer, R. A., É. E. Plagányi, and D. S. Butterworth (2007). Tips and tricks in designing management procedures. *ICES Journal of Marine Science* 64(4), 618–625.
- Richards, R. (2016). Monkfish Operational Assessment. NEFSC (Northeast Fisheries Science Center). *Northeast Fisheries Science Center Ref Doc Ref Doc. 16-09; 109 p.*
- Richards, R. A., P. C. Nitschke, and K. A. Sosebee (2008). Population biology of monkfish *Lophius americanus*. *ICES Journal of Marine Science* 65(7), 1291–1305.
- Rosenberg, A. A. and V. R. Restrepo (1994). Uncertainty and risk evaluation in stock assessment advice for US marine fisheries. *Canadian Journal of Fisheries and Aquatic Sciences* 51(12), 2715–2720.
- Rudd, M. B. and J. T. Thorson (2018). Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. *Canadian Journal of Fisheries and Aquatic Sciences* 75(7), 1019–1035.
- Schnute, J. (1985). A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* 42(3), 414–429.

- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences* 44(5), 924–940.
- Schwamborn, R., T. Mildenerger, and M. Taylor (2019). Assessing sources of uncertainty in length-based estimates of body growth in populations of fishes and macroinvertebrates with bootstrapped ELEFAN. *Ecological Modelling* 393, 37–51.
- Shepherd, J. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science* 56(5), 584–591.
- Smith, A. (1994). Management strategy evaluation: the light on the hill. *Population dynamics for fisheries management*, 249–253.
- Smith, A., K. Sainsbury, and R. Stevens (1999). Implementing effective fisheries-management systems—management strategy evaluation and the Australian partnership approach. *ICES Journal of Marine Science* 56(6), 967–979.
- Smith, M., E. Fulton, R. Day, L. Shannon, and Y.-J. Shin (2015). Ecosystem modelling in the southern Benguela: comparisons of Atlantis, Ecopath with Ecosim, and OSMOSE under fishing scenarios. *African Journal of Marine Science* 37(1), 65–78.
- Smith, M. T. and J. T. Addison (2003). Methods for stock assessment of crustacean fisheries. *Fisheries Research* 65(1-3), 231–256.
- Sparre, P. and S. C. Venema (1998). *Introduction to Tropical Fish Stock Assessment-Part 1: Manual*. FAO.
- Sun, Y., C. Zhang, Y. Tian, and Y. Watanabe (2021). Age, growth, and mortality rate of the yellow goosfish *Lophius litulon* (Jordan, 1902) in the Yellow Sea. *Journal of Oceanology and Limnology* 39(2), 732–740.
- Szuwalski, C. S. and A. E. Punt (2013). Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. *ICES Journal of Marine Science* 70(5), 955–967.
- Taylor, M. and T. K. Mildenerger (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology* 24(4), 330–338.

- Tyler, A., R. Beamish, and G. McFarlane (1989). Implications of age determination errors to yield estimates. *Canadian Special Publication of Fisheries and Aquatic Sciences* 108, 27–35.
- Tyszko, S. M. and J. J. Pritt (2017). Comparing otoliths and scales as structures used to estimate ages of Largemouth Bass: consequences of biased age estimates. *North American Journal of Fisheries Management* 37(5), 1075–1082.
- United Nations (2015). Transforming our world: The 2030 agenda for sustainable development. Technical report, Adopted by the General Assembly on 25 September 2015. Seventieth Session, Agenda Items 15 and 116. A/RES/70/1. UN, New York, New York, USA.
- Vasilakopoulos, P. and C. Maravelias (2016). Application of data-limited assessment methods on black anglerfish (*Lophius budegassa*; Spinola, 1807) stocks in the Mediterranean Sea. *Journal of Applied Ichthyology* 32(2), 277–287.
- von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii). *Human biology* 10(2), 181–213.
- Walmsley, S., R. Leslie, and W. Sauer (2005). The biology and distribution of the monkfish *Lophius vomerinus* off South Africa. *African Journal of Marine Science* 27(1), 157–168.
- Wang, Y., Y. Wang, S. Liu, C. Liang, H. Zhang, and W. Xian (2020). Stock assessment using LBB method for eight fish species from the Bohai and Yellow Seas. *Frontiers in Marine Science* 7, 164.
- Wilson, D. P. (1937). The habits of the angler-fish, *Lophius piscatorius* L., in the plymouth aquarium. *Journal of the Marine Biological Association of the United Kingdom* 21(2), 477–496.
- Worm, B., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings, and O. Jensen (2009). Rebuilding global fisheries. *science* 325(5940), 578–585.
- Yoneda, M., M. Tokimura, H. Fujita, N. Takeshita, K. Takeshita, M. Matsuyama, and S. Matsuura (1997). Age and growth of anglerfish *Lophius litulon* in the East China Sea and the Yellow Sea. *Fisheries science* 63(6), 887–892.

- Zhou, S., A. E. Punt, R. Deng, and J. Bishop (2011). Estimating multifleet catchability coefficients and natural mortality from fishery catch and effort data: comparison of Bayesian state–space and observation error models. *Canadian Journal of Fisheries and Aquatic Sciences* 68(7), 1171–1181.

Chapter 2

Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models

This chapter is a verbatim reproduction from the following published paper. The published version is found in Appendix A, Supplementary Material in Appendix B.

Batts, L., Minto, C., Gerritsen, H., & Brophy, D. Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science*, 76(7), 2150-2163.

Abstract

Analysis of length frequency distributions from surveys is one well known method for obtaining growth parameter estimates where direct age estimates are not available. We present a likelihood-based procedure that uses mixture models and the Expectation- Maximisation (EM) algorithm to estimate growth parameters from length frequency data. A basic LFEM model estimates a single set of growth parameters that produce one set of component means and standard deviations that best fits length frequency distributions over all years and surveys. The hierarchical extension incorporates bivariate random effects into the model. A hierarchical framework enables inter-annual or inter-cohort variation in some of the growth parameters to be modelled, thereby accommodating some of the natural variation that occurs in growth of fish. Testing on two fish species, haddock (*Melanogrammus aeglefinus*) and white-bellied anglerfish (*Lophius piscatorius*), we were able to obtain reasonable estimates of growth parameters, as well as successfully model growth variability. Estimated growth parameters showed some sensitivity to the starting values and occasionally failed to converge on biologically realistic values. This was dealt with through model selection and was partly addressed by the addition of the hierarchical extension.

Keywords: EM algorithm, von Bertalanffy growth, LFEM, bivariate random effects, anglerfish *Lophius piscatorius*, haddock *Melanogrammus aeglefinus*

2.1 Introduction

Fish growth is a widely studied aspect of fish biology (Pardo et al., 2013), reflecting its importance for understanding life histories (Denney et al., 2002), the effects of changing environmental conditions (Jobling, 2002; Baudron et al., 2014) as well as ecosystem and population dynamics (Quinn and Deriso, 1999; Travers et al., 2007; Smith et al., 2015). Growth parameters are a fundamental component of fisheries stock assessments (Quinn and Deriso, 1999) and their accurate estimation can be key to ensuring that assessments are reliable (Hilborn and Walters, 1992). Estimation of other important population dynamics parameters such as fishing mortality and selectivity depend on the accurate estimation of growth (Hoggarth, 2006; Maunder et al., 2016).

For teleost fish, growth information is usually obtained by estimating age from incremental growth layers laid down within the otoliths (fish ear bones), scales or fin rays. When age estimation using calcified structures is too costly to implement, not practical or is suspected to give unreliable estimates then other methods such as tagging or length frequency analysis provide an alternative means of obtaining growth estimates (Hilborn and Walters, 1992). Clear, unambiguous growth increments are often lacking in the calcified structures of tropical or deep sea species that are not exposed to marked seasonal temperature changes, while direct ageing of crustaceans using hard parts still requires further validation (Kilada and Driscoll, 2017). As a consequence, assessments of fisheries for these species often rely on alternative methods for estimating age (Sparre and Venema, 1998; Chang et al., 2012).

The most widely used fish growth model is the von Bertalanffy growth function (VBGF), which relates length to age using three parameters; L_∞ (asymptotic mean length), K (rate of growth) and t_0 (the theoretical time at which fish length equals zero) (von Bertalanffy, 1938). Various modifications of the traditional VBGF appear in the literature. For example, Schnute and Fournier (1980) re-parameterised the VBGF to include more biologically intuitive parameters (i.e. the mean lengths of the first and last ages of the sampled population and a parameter that controls growth between ages). Another example is the incorporation of seasonal oscillation into the VBGF (Pitcher and MacDonald, 1973; Lloyd-Jones et al., 2016). Other growth functions that have been used to model fish growth include the Gompertz (Winsor, 1932), allometric (Quinn and Deriso, 1999), logistic (Cormon et al.,

2016) and bi-phasic models (Minte-Vera et al., 2016); the most appropriate choice of model depends on the species (Katsanevakis and Maravelias, 2008). Here, the re-parameterised VBGF was chosen as the underlying model because of the widespread use of the VBGF and its appropriateness for length frequency analysis (Schnute and Fournier, 1980).

Growth can often vary spatially, temporally, between cohorts, between individuals and within individuals (Quinn and Deriso, 1999; Wang, 1999; Morrongiello and Thresher, 2015). Cohort effects can vary to a lesser or greater extent depending on the species and have been linked to environmental changes (Baudron et al., 2014; Morrongiello and Thresher, 2015). Accounting for cohort growth variability can also improve stock assessment models (Whitten et al., 2013). Methods for accounting for individual growth variation in length-at-age data have also been explored, for example Pilling et al. (2002) used a non linear random effects model to model growth parameter variability between individuals length-at-age data. Cadigan et al. (2016) developed a hierarchical mixed-effects model to account for between stock and between individual variability in length-at-age data. Growth models incorporating random effects have also been applied to analysis of tagging data, where between-individual growth variability is often considerable (Eveson et al., 2015).

In situations where fish cannot be aged directly, length frequency analysis offers an alternative means of estimating fish growth parameters, providing there is a distinct spawning season (Hasselblad, 1966; Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017). Perhaps the most prominent length-frequency method in the literature is the robust maximum likelihood-based estimation procedure MULTIFAN, which is now incorporated into a length-based statistical catch at age stock assessment model (Fournier et al., 1990, 1998). The non-parametric ELEFAN procedure (Pauly, 1987) is also widely used and has recently been extended into the R environment where modern optimisation algorithms are used to improve its performance (Mildenerger et al., 2017; Taylor and Mildenerger, 2017).

Currently available approaches to length frequency analysis present some limitations. Both MULTIFAN and ELEFAN require that length bins are defined prior to the analysis, potentially introducing subjectivity to the identification of cohorts. For ELEFAN it is recommended that L_∞ be fixed in the initial analysis and is then re-calculated post optimisation. Both models are also largely constrained to a single growth curve and do not incorporate

inter-annual or inter-cohort variability in growth (although this is addressed to some extent in the development of MULTIFAN-CL) (Fournier et al., 1998). Current limitations could be addressed by developing a model and using an expectation-maximisation (EM) algorithm to estimate growth parameters (Dempster et al., 1977). This approach would remove the need to pre-define length classes and offers the flexibility to include variability in growth.

Whilst the EM algorithm is widely used to model length frequency distributions, it has rarely been used to estimate mixture model parameters with an incorporated von Bertalanffy growth structure. However, the methodology of MULTIFAN is similar to the EM algorithm in its application of maximum likelihood theory (Fournier et al., 1990). The “mixdist” package in the R environment enables the user to fit a mixture model with component means constrained to a growth curve and this utilises an EM algorithm (Macdonald and Du, 2011). In addition, a recent paper by Lloyd-Jones et al. (2016) used a minorisation-maximisation (MM) algorithm (the EM algorithm is a form of MM algorithm) to estimate seasonally oscillating von Bertalanffy growth over a year on monthly length frequency data of a crab species (Lloyd-Jones et al., 2016), where L_∞ is considered a random variable, although it is not modelled explicitly. These examples leave much scope for expanding this area of research further.

This study aims to develop a new method, based on the EM algorithm, to model fish growth from length frequency data whilst incorporating von Bertalanffy growth structure and accounting for variability in growth between cohorts or years. Gaussian mixture models are developed and fit to length frequency distributions, where means of components within a cohort (represented by normal distributions in the mixture models) are restricted to von Bertalanffy growth. Firstly, a basic version (analogous to MULTIFAN) of the methodology is introduced and tested using length frequency data for two fish species; haddock *Melanogrammus aeglefinnus* and white-bellied anglerfish *Lophius piscatorius*. Secondly, hierarchical models that introduce various bivariate random effects, are introduced and tested with the intention of modelling some of the variation in the recruitment and growth of the fish species. The overall aim of the study is to introduce an alternative approach to obtaining a range of credible estimates of growth parameters from length frequency data and account for variability in growth by cohort or year.

2.2 Methods

A description of the growth function used within the model is first introduced, followed by a description of the model itself. A summary of model parameters, inputs and outputs is then given, as well as a description of the model selection criteria and software used for development. Finally the two data sources that the LFEM model is applied to are described.

2.2.1 Growth function

Length frequency distributions were modelled using mixture models, where the means of the cohorts (normal distributions within the mixture models) follow a von Bertalanffy growth function. Schnute and Fournier (1980) re-parameterised the classical von Bertalanffy (CVBG) equation into one more appropriate to length frequency analysis, represented by

$$\mu_i = l + (L - l) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M \quad (2.1)$$

where μ_i is the mean length of fish at age a_i ; l and L are the first and final component or cohort mean lengths, i.e. μ_1 and μ_M where M is the maximum age class; k is equivalent to $\exp(-K)$ and controls the distance between two successive component means (Equation 2.1). The CVBG parameters can then be calculated from these parameters using

$$L_\infty = \frac{L - lk^{M-1}}{1 - k^{M-1}}, \quad (2.2)$$

$$K = -\ln(k), \quad (2.3)$$

$$t_0 = a_1 - \frac{1}{\ln(k)} \ln \left\{ \frac{L - l}{L - lk^{M-1}} \right\} \quad (2.4)$$

where L_∞ is the asymptotic length of the fish, K controls the rate at which the asymptote is approached and t_0 is the theoretical time when an individual would have length zero. a_1 is the assumed age of the l component. Standard deviation (SD) of the normal distribution of the component (σ_i) is either estimated as constant (CSD) or a linear (LSD) function of component means (Schnute and Fournier, 1980). The equation for LSD is

$$\sigma_i = s + (S - s) \frac{\mu_i - l}{L - l}; \quad i = 1, \dots, M \quad (2.5)$$

where s is the SD of the first component and S is the SD of the final component.

2.2.2 LFEM model and algorithm

This methodology summarises the basic model developed in this study where a single set of growth parameters that produce one set of component means and standard deviations are estimated (i.e. if l , L and k are identical for every cohort then component means in each annual mixture for a particular survey model are identical).

Based on the observed data log-likelihood for a single mixture model (McLachlan and Peel, 2004), the observed data log-likelihood for a set of finite mixture models on length frequency distributions of fish where cohort component means are restricted to the re-parameterised von Bertalanffy growth function (RVBG) is

$$\ln L(\Psi) = \sum_{v=1}^V \sum_{r=1}^R \sum_{j=1}^{n_{vr}} \ln \left\{ \sum_{i=1}^M \lambda_{vri} f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) \right\} \quad (2.6)$$

$$f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) = N(y_{jvr}; \mu_{vri}, \sigma_{vri}, a_{1v}) \quad (2.7)$$

$$\mu_{vri} = l_v + (L_v - l_v) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V \quad (2.8)$$

where the components of the mixture models are normally distributed and their means, μ_{vri} , are given by equation 2.8. Component SDs are either assumed constant or constrained by equation 2.5. Indices denote survey (v), year (r), component (i) and observation (j). Ψ is a vector that contains all unknown parameters in the model, V is the number of surveys, R is the number of years, M is the number of components and n_{vr} is the total number of fish in any given survey and year. λ_{vri} is the mixing proportion for the i th component in the v th survey in the r th year, y_{jvr} is the length of fish j in the v th survey in the r th year and θ_{vri} are the parameters of the i th component in the v th survey in the r th year. a_{1v} determines the assumed age of first component of the v th survey and is used to incorporate surveys of differing timings by back

or forward projecting component means using equations 2.2, 2.3 and 2.4.

The complete data log-likelihood can be given as

$$\log L_c(\Psi) = \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \sum_{j=1}^{n_{vr}} z_{jvri} \{ \log \lambda_{vri} + \log f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) \} \quad (2.9)$$

where equations 2.7 and 2.8 apply. z_{jvr} is a data label vector distributed according to a multinomial distribution of length of M with $z_{jvri} = (z_{jvr})_i = 1$ or 0, according to whether that particular fish or y_j belongs to the i th component in the v th survey in the r th year (McLachlan and Peel, 2004). However, as we do not know what component each fish belongs to the expectation-maximisation (EM) algorithm was used to treat z_{jvri} as missing data and estimate Ψ iteratively. Direct estimation of the observed log-likelihood is possible but not practical in this case and the EM algorithm offered a simple framework to implement the model in (Minto et al., 2017). For details on the EM algorithm see Appendix B.

2.2.3 Hierarchical extensions

Hierarchical models were developed, incorporating bivariate random effects into the model formulation and allowing inter-annual or inter-cohort variation to be modelled. In contrast to the basic model that applies a single set of component means to the data, these models allow cohorts to be modelled through years and surveys. For the three hierarchical models both CSD and LSD models were implemented. To avoid over-parameterisation these hierarchical models were not extended beyond bivariate random effects and in each model either the key parameter for the asymptotic length (L) or the growth parameter (k) is random, not both. The logic behind this is also discussed by Eveson et al. (2007) and Lloyd-Jones et al. (2016), suggesting that random effect on either K or L_∞ incorporates sufficient variability but also has less bias than if both parameters were random.

Hierarchical models were fitted using the same number of components as the best fitting basic model (based on the model selection criteria described below). Three different hierarchical model structures were tested: 1) bivariate random effects on cohort specific l and L ; 2) bivariate random effects on cohort specific l and k ; 3) bivariate random effects on cohort specific l and

yearly k . Correlation between the bivariate random effects in each scenario is also modelled. A random effect on l was specifically chosen to account for variability in the mean length of the first component in each cohort. Other random effects were chosen to investigate and incorporate growth variability. For details on the model formulation see Appendix B.

2.2.4 Model parameters, inputs and outputs

A wide range of starting values for parameters of the basic model were tested (Table 2.1), amounting to over 1000 model runs for each species and each SD type. In order to test the suitability of the model selection criteria for identifying the number of identifiable age classes in the length frequency data, the number of components tested ranged from 3-14 for haddock and 6-14 for white-bellied anglerfish. Standard deviations for the random effects of l and k for the haddock hierarchical model were fixed at $exp(-5)$ in order to give the model stability but still retain the random effect. More information on model stability in the haddock hierarchical models and general model inputs can found in Appendix B.

Table 2.1: Summary of the starting values tested for the model parameters of the basic model

Species	parameter	Starting values tested		
		From	To	By
haddock	l	10	10	NA
	L	40	70	10
	k	0.6	0.9	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	3	14	1
	λ_i	1/No. comp	NA	NA
white-bellied anglerfish	l	16	16	NA
	L	100	130	10
	k	0.7	0.99	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	6	14	1
	λ_i	1/No. comp	NA	NA

Table 2.2: Summary of model parameters, and other relevant inputs/outputs of models. N.B. Not all hierarchical model parameters are used in each model

Notation	Type	Description
Basic model		
L	parameter	mean of final component
l	parameter	mean of first component
k	parameter	controls rate of growth
s	parameter	standard deviation of first component
S	parameter	standard deviation of final component
j	index	individual fish length
i	index	component number
v	index	survey number
r	index	number of year
λ_{vri}	parameter(s)	mixing proportion of i th component in survey(v) and year(r)
μ_i	output	mean of i th component
σ_i	output	standard deviation i th component
L_∞	output	CVBG parameter
K	output	CVBG parameter
t_0	output	CVBG parameter
a_{1v}	input	used to calculate t_0
M	input	number of components/ages assumed present in length frequency data
n_{vr}	input	number of observations(fish) in a given survey(v) and year(r)
Hierarchical models		
L_c	parameter	mean of final component of a cohort
l_c	parameter	mean of first component of a cohort
k_c	parameter	mean of random effect that controls rate of growth for a cohort
k_t	parameter	mean of random effect that controls rate of growth for a particular year
σ_L	parameter	standard deviation of random effect on L
σ_l	parameter	standard deviation of random effect on l
σ_k	parameter	standard deviation of random effect on k
ρ_{lk}	parameter	correlation of bivariate random effects on l, k for example
μ_{vri}	output	mean of i th component in survey(v) and year(r)

2.2.5 Model selection criteria and standard errors

Through exploratory simulation analysis it was found that AIC performed best from a range of model selection criteria. Model selection for this study was performed by selecting the model with highest observed log-likelihood value for each “number of components” tested. A variation on AIC we call “sub-AIC” was used to select the best fitting model. “sub-AIC” is similar to AIC, however the observed log-likelihood used is only contributed to by a subset of the length frequency data where reasonable numbers of fish were observed (i.e. *haddock* $\leq 40cm$ & *angler fish* $\leq 100cm$). This method ensures model selection is based on the majority of the data and not confounded by low fish numbers at high lengths. A similar method where model deviance is calculated from a subset of length classes when numbers in some length classes are low was used in model selection for fish selectivity curves (Millar and Fryer, 1999).

The EM algorithm does not directly provide estimates of uncertainty of

parameters (McLachlan and Peel, 2004). Standard errors were approximated by first fitting the model to convergence with the EM algorithm, then using the maximum likelihood parameter estimates to optimise the observed data log-likelihood function (Equation 2.6) (Minto et al., 2017). Standard errors were approximated via the delta method with the R package “TMB” (Kristensen et al., 2016).

2.2.6 Validation for haddock

Age data on the North Sea International Bottom Trawl Survey (NS-IBTS) are collected on a length-stratified basis. In order to obtain unbiased length-at-age data for haddock the length-stratified age data was used to construct annual age length keys (ALKs). Annual length frequency distributions (i.e. the same used in the length frequency analysis) were applied to their respective ALKs using the method described by Isermann and Knight (2005), resulting in an assigned age for each individual fish of the length frequency distribution. To enable comparison with the growth parameters estimated from the length frequency data alone, a VBGF was fit to this length-at-age data using maximum likelihood and mean length-at-age was also calculated.

2.2.7 Software

We develop a hybrid EM algorithm that uses the R package “TMB” within the “maximisation” step (Kristensen et al., 2016). Specifically, complete data log-likelihood functions were written in C++ using the TMB library class, compiled and automatically differentiated (Kristensen et al., 2016). These objective functions, including gradients, were then optimised at each iteration with the statistical software R using the `nloptr` function within base R (R Core Team, 2018). The remaining computations of the algorithm and other analysis were conducted with R (R Core Team, 2018). The “FSA” R package (Ogle, 2018) was used to apply ALKs to the un-aged length frequency data, using the semi-random method from Isermann and Knight (2005).

2.2.8 Data

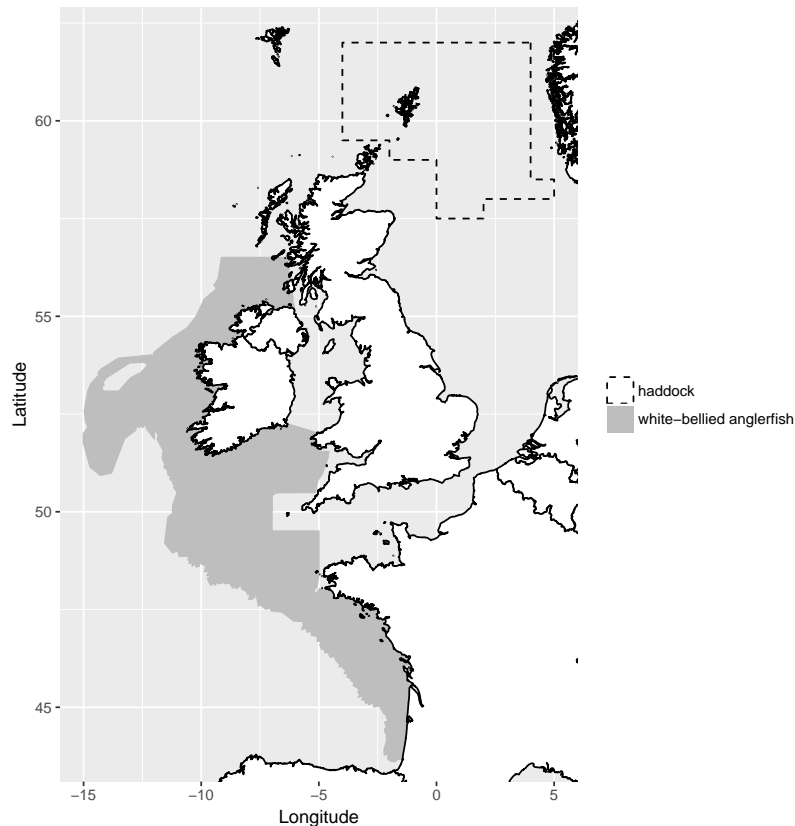


Figure 2.1: Areas of survey coverage from where length frequency data was obtained for haddock (standard roundfish area one of the NS-IBTS) and white-bellied anglerfish (EVHOE, IE-IGFS and SP-PORC)

Haddock was chosen to validate the method as otolith age readings and growth parameters derived from lengths at age are reliable. White-bellied anglerfish was chosen because age readings are considered unreliable (ICES, 2011) and there is need for a cohort analysis to build an assessment model on. Both species are characterised by fast growth and a discrete recruitment season, leading to distinct length cohorts that can be tracked over time. Combined sex data was used for both species as is convention for haddock and white-bellied anglerfish in the most recent ICES assessment of the stock (ICES, 2018). Survey catch data were obtained from the ICES DATRAS database (ICES, 2018b). Haddock length frequency data were from standard roundfish area one (Figure 2.1) of the North Sea International Bottom Trawl Survey (NS-IBTS). Only data collected in the third quarter of the year were used. The corresponding age-length key for this area and time of year was also extracted from DATRAS. Length frequency data for white-bellied anglerfish

were obtained from three surveys that cover the anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d; the French EVHOE groundfish survey, the Irish groundfish survey (IE-IGFS) and the Spanish Porcupine groundfish survey (SP-PORC) (Figure 2.1). The French and Irish surveys were conducted mainly in the fourth quarter of the year and the Spanish survey in September. A Grande Ouverture Verticale (GOV) fishing gear was used on all surveys with the exception of the Spanish Porcupine bank survey which used "porcupine baca" fishing gear (ICES, 2018a).

2.3 Results

2.3.1 Haddock

2.3.1.1 Haddock basic LFEM model

The basic model (single set of growth parameters estimated) was sensitive to starting parameter values. Some of the model runs tended to converge on k parameter values very close to 1. Whilst these gave good or even the best model fits according to the sub-AIC model selection criteria, the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). A number of model runs were discarded (406 of 1488 CSD models and 42 of 1488 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for "biologically unreasonable" L_∞ estimates. To further filter out any spurious models the results were further refined by identifying and removing any models that had a second component mean ≤ 15 cm. An example of why refining is important when a wide range of starting parameters are tested can be found in Appendix B.

A nine component model was the best fitting (lowest sub-AIC) for CSD and no models above nine components had reasonable fits for this SD type. LSD models gave generally lower sub-AIC values than CSD models and the best fitting model overall had 12 components and LSD (Table 2.3). Within both SD types, but particularly in LSD models, parameter estimates and component means were consistent across the range of number of components tested (Table 2.3 and 2.4). When compared to the VBGF estimated from length-at-age data ($\widehat{K} = 0.499$, $\widehat{L}_\infty = 39.5$ and $\widehat{t}_0 \approx 0$), the VBGF from the best fitting basic LSD model is similar, although it does slightly underestimate early age mean lengths (Figure 2.2). The best fitting basic CSD model follows a similar growth pattern until the fifth component, the following component means

Table 2.3: von Bertalanffy growth parameters from (lowest sub-AIC) model runs on haddock length frequency data after refining valid model runs. Number of components is the number of ages assumed to be present in the data. Rows highlighted in grey are the models with the lowest sub-AIC model selection criteria values. Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameter estimates are given to 3 s.f. and CVBG to 2 s.f.

SD type	No. components	RVBG parameters			CVBG parameters			sub AIC	
		\hat{l} (cm)	\hat{L} (cm)	\hat{k}	\hat{L}_∞ (cm)	\hat{K}	\hat{t}_0		
CSD	3	10.5 (0.008)	32.4 (0.007)	0.816 (0.0012)	76	0.20	-0.10	6167506	
	4	10.3 (0.006)	37.1 (0.011)	0.780 (0.0007)	61	0.25	-0.12	6031048	
	5	10.3 (0.005)	39.7 (0.014)	0.754 (0.0006)	53	0.28	-0.13	5989126	
	6	10.3 (0.005)	41.8 (0.018)	0.750 (0.0005)	51	0.29	-0.15	5974158	
	7	10.3 (0.005)	43.0 (0.021)	0.752 (0.0005)	50	0.29	-0.18	5957923	
	8	10.4 (0.004)	44.0 (0.022)	0.761 (0.0004)	49	0.27	-0.23	5946292	
	9	10.4 (0.004)	45.0 (0.026)	0.773 (0.0004)	50	0.26	-0.27	5941101	
	LSD	3	10.1 (0.003)	31.3 (0.008)	0.854 (0.0013)	88	0.16	-0.14	5982480
		4	10.1 (0.003)	34.9 (0.012)	0.720 (0.0007)	49	0.33	-0.07	5918972
5		10.1 (0.003)	37.0 (0.015)	0.689 (0.0006)	44	0.37	-0.06	5901698	
6		10.1 (0.003)	37.5 (0.016)	0.665 (0.0005)	41	0.41	-0.06	5889986	
7		10.1 (0.003)	38.4 (0.018)	0.661 (0.0005)	40	0.41	-0.06	5886624	
8		10.1 (0.003)	39.0 (0.020)	0.658 (0.0004)	40	0.42	-0.06	5885749	
9		10.1 (0.003)	39.5 (0.021)	0.658 (0.0004)	40	0.42	-0.06	5885240	
10		10.1 (0.003)	40.0 (0.023)	0.661 (0.0004)	40	0.41	-0.06	5885145	
11		10.1 (0.003)	40.4 (0.024)	0.663 (0.0004)	40	0.41	-0.07	5885031	
12		10.1 (0.003)	40.6 (0.025)	0.665 (0.0004)	40	0.41	-0.07	5885004	
13		10.1 (0.003)	40.8 (0.026)	0.666 (0.0004)	41	0.41	-0.07	5885005	
14		10.1 (0.003)	40.9 (0.027)	0.667 (0.0004)	41	0.41	-0.07	5885023	

then appeared to over estimate the mean length of fish in comparison to the LSD model and the length-at-age estimated model fit (Figure 2.2). Estimated standard deviation was very similar within SD types (Table 2.4).

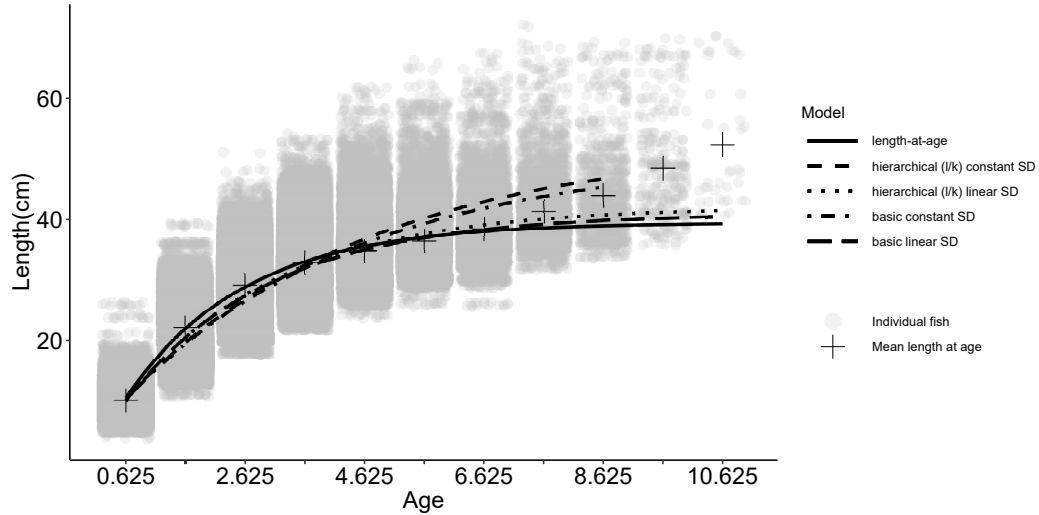


Figure 2.2: Length-at-age data and estimated von Bertalanffy growth curves for haddock. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data, growth curves from best fitting basic length frequency models and overall growth curves (i.e. means of RE parameters) from hierarchical l/k models. Ages are jittered for visualisation

Table 2.4: Component/cohort means and their standard deviation parameters (SD) for haddock, corresponding to the basic models presented in Table 2.3. In the case of constant SD models $S = s$ but values are only given in the final (S) column.

SD type	No. components	Component										\hat{s}	\hat{S}				
		1	2	3	4	5	6	7	8	9	10			11			
CSD	3	10.5	22.6	32.4												3.66	
	4	10.3	21.5	30.2	37.1											2.89	
	5	10.3	21.0	29.0	35.1	39.7										2.65	
	6	10.3	20.6	28.4	34.2	38.5	41.8									2.50	
	7	10.3	20.2	27.7	33.3	37.5	40.7	43.0								2.39	
	8	10.3	19.8	27.0	32.4	36.6	39.8	42.2	44.0							2.29	
	9	10.4	19.4	26.4	31.9	36.1	39.3	41.8	43.8	45.3						2.21	
	3	10.1	21.5	31.3												1.63	4.93
	4	10.1	21.1	29.1	34.9											1.65	4.59
LSD	5	10.1	20.9	28.3	33.5	37.0									1.66	4.42	
	6	10.1	20.7	27.7	32.4	35.5	37.5								1.67	4.28	
	7	10.1	20.6	27.5	32.1	35.1	37.1	38.4							1.67	4.23	
	8	10.1	20.6	27.4	32.0	34.9	36.9	38.2	39.0						1.68	4.22	
	9	10.1	20.5	27.4	31.9	34.9	36.8	38.1	39.0	39.5					1.68	4.21	
	10	10.1	20.5	27.4	31.9	34.9	36.9	38.2	39.1	39.6	40.0				1.68	4.20	
	11	10.1	20.5	27.3	31.9	34.9	36.9	38.3	39.1	39.7	40.1	40.4			1.68	4.19	
	12	10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.4	40.6		1.68	4.18	
	13	10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.5	40.7	40.8	1.68	4.18	
	14	10.1	20.4	27.3	31.9	35.0	37.0	38.3	39.3	39.9	40.3	40.5	40.7	40.8	40.9	1.69	4.18

2.3.1.2 Haddock hierarchical LFEM model with cohort specific l and k

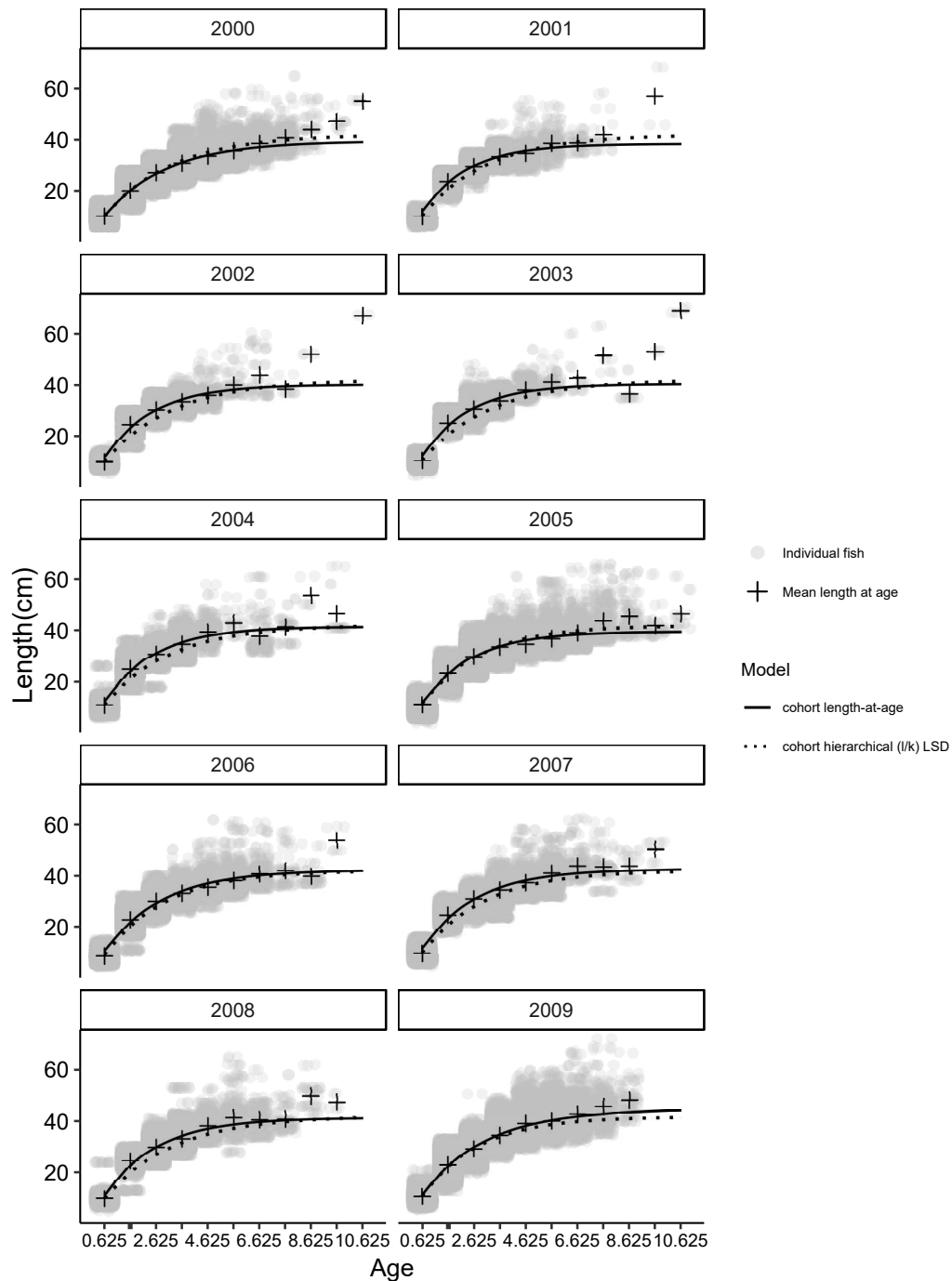
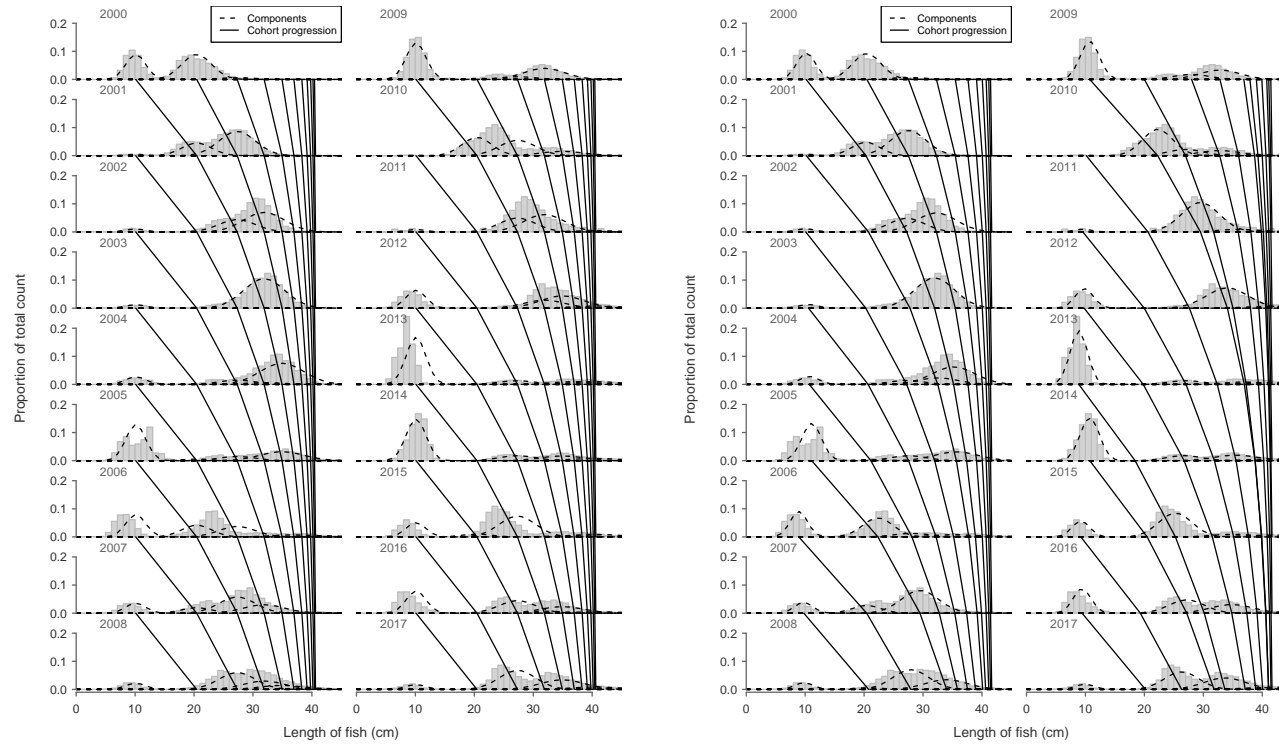


Figure 2.3: Length-at-age data and estimated von Bertalanffy growth curves for cohorts of haddock observed up to at least their 9th component. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data of the cohort and cohort specific growth curves estimated within the linear SD hierarchical model that varies l and k . Ages are jittered for visualisation

Overall CVBG parameters for the hierarchical CSD model ($\widehat{K} = 0.241$, $\widehat{L}_\infty = 52.9$ cm and $\widehat{t}_0 = -0.31$) were similar to those given by the basic CSD model and followed a similar trajectory. Hierarchical LSD model parameters ($\widehat{K} = 0.394$, $\widehat{L}_\infty = 42.1$ cm and $\widehat{t}_0 = -0.06$) were also similar to estimates from its corresponding basic model (Table 2.3 and Figure 2.2).

The performance of the hierarchical l and k LSD model (corresponding to the best fitting basic model) was examined in more detail. Ranges of the cohort specific L_∞ estimates were 41.8 - 42.5 cm for \widehat{L}_∞ and 0.334 - 0.463 for \widehat{K} . Correlation parameter $\widehat{\rho}$ for the random effects l and k was estimated as -0.23. Cohort specific random effects gave the model flexibility in its estimated growth trajectories for each cohort. When examining the cohorts that were observed at least till the ninth component over the period analysed (2000 - 2009) the estimated cohort growth curves from this study's hierarchical model appear to fit reasonably well in some cohorts (e.g cohorts 2002, 2003 and 2004) and very well in others (e.g. cohorts 2000, 2005 and 2009) (Figure 2.3). When cross-referenced with Figure 2.4 the hierarchical model cohort curves that are very similar to length-at-age estimated cohort curves are those cohorts where a strong signal can be seen through a number of years (i.e. cohort 2000) (Figures 2.3 and 2.4). Deviance (difference between LFEM model length-at-age and ML estimated length-at-age) of these cohorts showed that in the early ages of some cohorts the hierarchical LFEM model underestimated the mean length-at-age in comparison to the ML fit on the length-at-age data Appendix B. In the later years of the length frequency distributions the hierarchical model begins to severely underestimate growth (Figure 2.4). Also of note is the flexibility the random effect on the mean of the first component (l) gives when modelling the shifting position of the first component over the years, such as in 2006 or 2013 (Figure 2.4).



(a) lowest sub-AIC basic LFEM model fit

(b) corresponding hierarchical LFEM model to (a) with bivariate random effects on l and k

Figure 2.4: Length frequency distributions of haddock with fitted distributions from: (a) the lowest sub-AIC basic model (12 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and k . Also shown is the progression of the cohort mean values. Note that the maximum length of fish was 72 cm but only fish up to 45 cm were plotted here as there was little data above this length, all estimated cohorts are displayed.

2.3.2 White-bellied anglerfish

2.3.2.1 White-bellied anglerfish basic LFEM model

As with haddock, the basic LFEM model was sensitive to starting parameters, particularly k and some of the model runs tended to converge on k parameter values very close to 1. These models often gave good or even the best model fits according to the sub-AIC model selection criteria, however the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). A large number of model runs were discarded (720 of 1080 CSD models and 740 of 1080 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for “biologically unreasonable” L_∞ estimates.

Overall, the nine component model for LSD was best fitting (lowest sub-AIC) and LSD models consistently had lower sub-AIC values than their counterpart CSD models. For LSD models as the number of components assumed was increased the final component means (L) increased, as did the CVBG parameter K , while L_∞ decreased. CSD models showed no such pattern (Table 2.5 and 2.6).

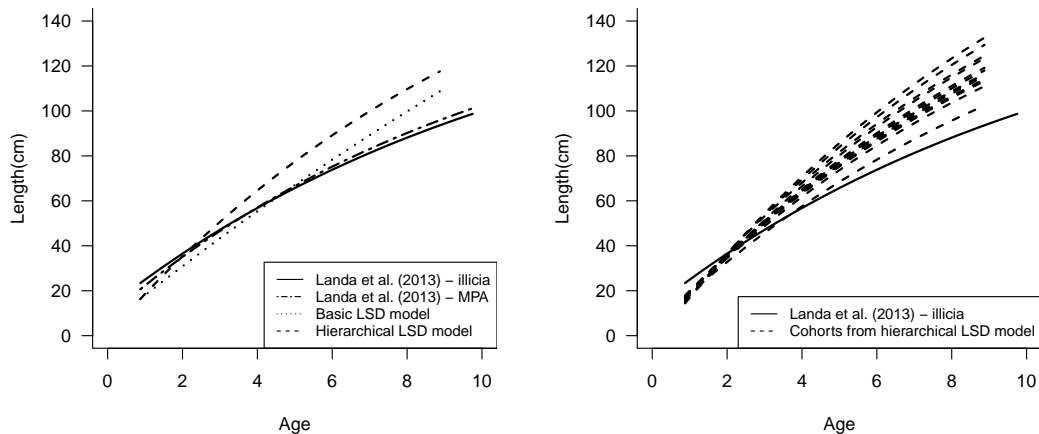
Table 2.5: von Bertalanffy growth parameters from LFEM model runs that gave the lowest sub-AIC value on white-bellied anglerfish length frequency data. Number of components is the number of ages assumed to be present in the data. Parameters l and L are those estimated for the surveys that occurred in the fourth quarter of the year (EVHOE and IE-IGFS). Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameters estimates are given to 3 s.f. and CVBG to 2 s.f. (with the exception of L_∞). The Spanish survey estimates of can be found in Appendix B.

SD type	No. components	RVBG parameters			CVBG parameters			sub AIC
		\hat{l} (cm)	\hat{L} (cm)	\hat{k}	\hat{L}_∞ (cm)	\hat{K}	\hat{t}_0	
CSD	8	17.1 (0.090)	114.4 (0.572)	0.966 (0.003)	469	0.035	-0.20	99430
	9	17.0 (0.087)	118.6 (0.621)	0.945 (0.003)	294	0.057	-0.17	99397
	10	17.0 (0.085)	119.9 (0.621)	0.930 (0.003)	232	0.072	-0.17	99419
	11	16.6 (0.082)	125.3 (0.653)	0.974 (0.002)	489	0.026	-0.45	99341
	12	16.6 (0.081)	128.4 (0.699)	0.969 (0.002)	401	0.031	-0.48	99367
	13	16.6 (0.080)	130.1 (0.734)	0.951 (0.002)	267	0.050	-0.40	99419
	14	16.5 (0.078)	126.8 (0.653)	0.940 (0.002)	215	0.062	-0.41	99444
LSD	8	16.5 (0.076)	107.3 (0.919)	0.941 (0.005)	276	0.061	-0.13	99112
	9	16.3 (0.075)	108.7 (0.997)	0.967 (0.005)	412	0.033	-0.34	99102
	10	16.3 (0.076)	113.1 (0.999)	0.974 (0.004)	471	0.027	-0.45	99110
	11	16.4 (0.076)	114.9 (1.017)	0.951 (0.004)	264	0.051	-0.39	99173
	12	16.4 (0.076)	116.3 (0.992)	0.938 (0.003)	213	0.064	-0.37	99237
	13	16.4 (0.076)	120.1 (1.022)	0.930 (0.003)	195	0.072	-0.34	99312
	14	16.4 (0.076)	122.1 (1.077)	0.921 (0.003)	177	0.082	-0.31	99393

Component means were consistent between and within SD types up to the

fourth component, particularly for models with an assumed number of components of 11 or above. LSD model component means did not change with increasing number of components assumed and the best fitting model across all models (linear SD, nine components) had very similar component means to the best fitting constant SD model (Table 2.6). When plotted on the raw length frequency data the best fitting basic model appears to fit the data reasonably well in some years but not in others and cohort progression can be seen up until the seventh component in some cohorts (Figure 2.6a).

The growth curve of the parameters estimated from the best fitting basic LFEM model differs somewhat from the growth curves estimated in a previous study (Landa et al., 2013). Growth curves were reasonably similar up to age six, although the present study's VBGF shows marginally faster growth. After age six the growth curves begin to diverge more, with the present study estimates of length-at-age markedly higher (Figure 2.5a).

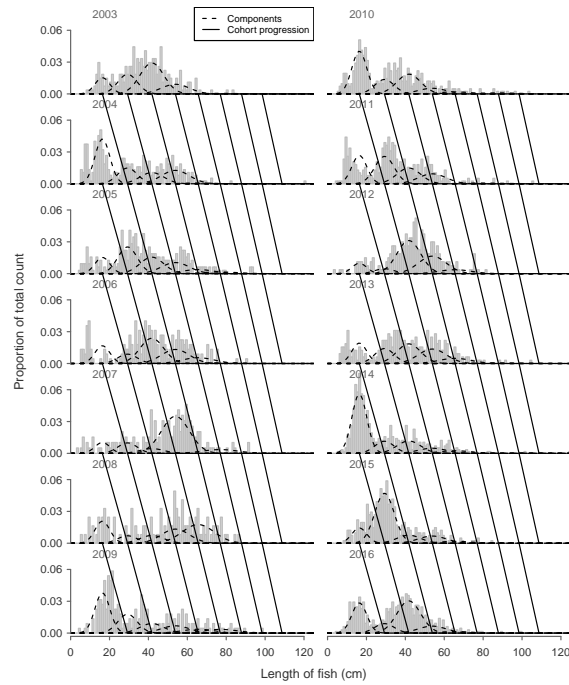


(a) Estimated von Bertalanffy growth curves for the present study and a previous study (b) Estimated cohort growth curves compared to a previous study

Figure 2.5: von Bertalanffy growth curves estimated for white-bellied anglerfish from: (a) the best fitting basic model and overall parameters from the corresponding hierarchical model with a cohort specific random effect on l and L , as well as growth curves (estimated from illicia and model progression analysis (MPA)) from a previous study (Landa et al., 2013), (b) estimated cohort growth curves from the hierarchical LSD model with cohort specific random effects on l and L .

Table 2.6: Component/cohort means and their standard deviation parameters for white-bellied anglerfish, corresponding to the basic models presented in Table 2.5. In the case of constant SD models $S = s$ but values are only given in the final (S) column.

SD type	No. components	Component														\hat{s}	\hat{S}		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14				
CSD	8	17.1	32.5	47.3	61.7	75.5	88.9	101.9	114.4									4.91	
	9	17.0	32.4	46.9	60.6	73.6	85.8	97.4	108.3	118.6								4.75	
	10	16.9	32.0	45.9	58.9	71.0	82.3	92.7	102.5	111.5	119.9							4.62	
	11	16.6	28.8	40.7	52.3	63.5	74.5	85.2	95.6	105.8	115.7	125.3						4.34	
	12	16.6	28.4	39.8	50.9	61.7	72.1	82.3	92.1	101.6	110.8	119.7	128.4					4.23	
	13	16.6	28.9	40.6	51.7	62.2	72.3	81.9	90.9	99.6	107.8	115.6	123.0	130.1				4.20	
	14	16.5	28.5	39.8	50.4	60.3	69.7	78.5	86.7	94.5	101.8	108.6	115.1	121.1	126.8			4.12	
LSD	8	16.5	32.0	46.5	60.2	73.1	85.2	96.6	107.3									3.96	11.15
	9	16.3	29.3	41.8	53.9	65.6	76.9	87.9	98.5	108.7								3.83	12.03
	10	16.3	28.2	39.9	51.2	62.2	72.9	83.4	93.5	103.5	113.1							3.85	9.90
	11	16.3	28.6	40.2	51.3	61.8	71.8	81.4	90.4	99.0	107.2	114.9						3.86	9.28
	12	16.4	28.6	40.1	50.9	61.1	70.5	79.4	87.8	95.6	103.0	109.9	116.3					3.87	8.55
	13	16.4	28.9	40.5	51.3	61.3	70.6	79.3	87.4	94.9	101.9	108.4	114.5	120.1				3.90	7.72
	14	16.4	29.1	40.8	51.6	61.5	70.6	79.1	86.8	94.0	100.5	106.6	112.2	117.3	122.1			3.91	7.60



(a) basic LFEM fit on IE-IGFS

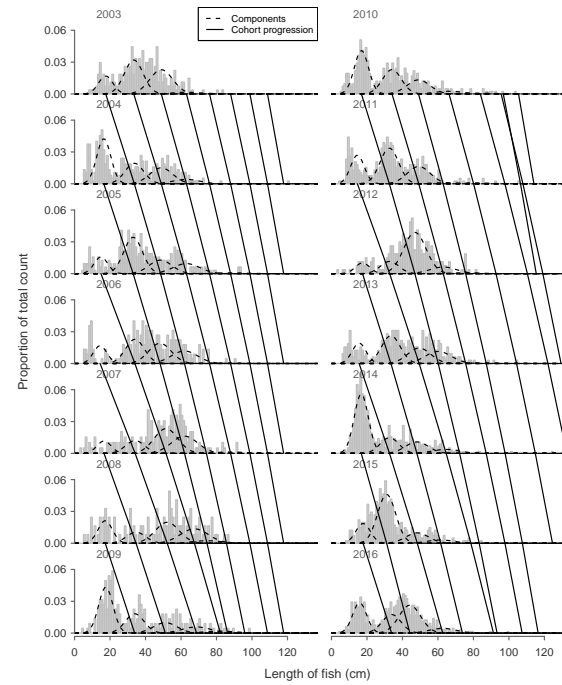
(b) corresponding hierarchical LFEM model fit to (a) with bivariate random effects on l and L on IE-IGFS

Figure 2.6: Length frequency distributions of white-bellied anglerfish from the quarter four Irish groundfish survey (IE-IGFS) with fitted distributions from: (a) the lowest sub-AIC basic LFEM model (9 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and L . Also shown is the progression of the cohort mean values (black solid lines). Similar plots for the Spanish Porcupine Bank survey (SP-PORC) on the French groundfish survey (EVHOE) can be found in Appendix B.

2.3.2.2 White-bellied anglerfish hierarchical LFEM model with cohort specific l and L

Overall CVBG parameters for the hierarchical LSD model with cohort specific l and L corresponding to the best fitting basic model ($\widehat{K} = 0.089$, $\widehat{L}_\infty = 215.2$ cm and $\widehat{t}_0 = -0.14$) differed to those given by the basic model. \widehat{K} was increased and \widehat{L}_∞ reduced dramatically, \widehat{t}_0 was also slightly reduced. The resulting VBGF curve was different from the basic model curve as well as the growth curves from a previous study (Landa et al., 2013). The hierarchical model showed faster growth than the other curves and estimated component means somewhat different to the best fitting basic LFEM model (Figure 2.5a). Cohort specific growth curves from the hierarchical LFEM model showed a range of trajectories and cohort \widehat{L}_∞ estimates ranged from 184.9 - 246.6 cm (Figure 2.5b). The slowest growing cohort with the lowest estimated L_∞ (2014 cohort) followed a similar trajectory to the basic LFEM model (Figure 2.5).

There was a negative correlation estimated ($\widehat{\rho} = -0.69$) between the bivariate random effects l and L of each cohort (Figure 2.7). The bivariate cohort specific random effects show a moderate degree of variation over a few centimetres for \widehat{l} and approximately 30 cm for \widehat{L} . There is a also negative trend in L (Figure 2.7). There is a high degree on uncertainty around estimates of L for the last few cohorts (Figure 2.7b).

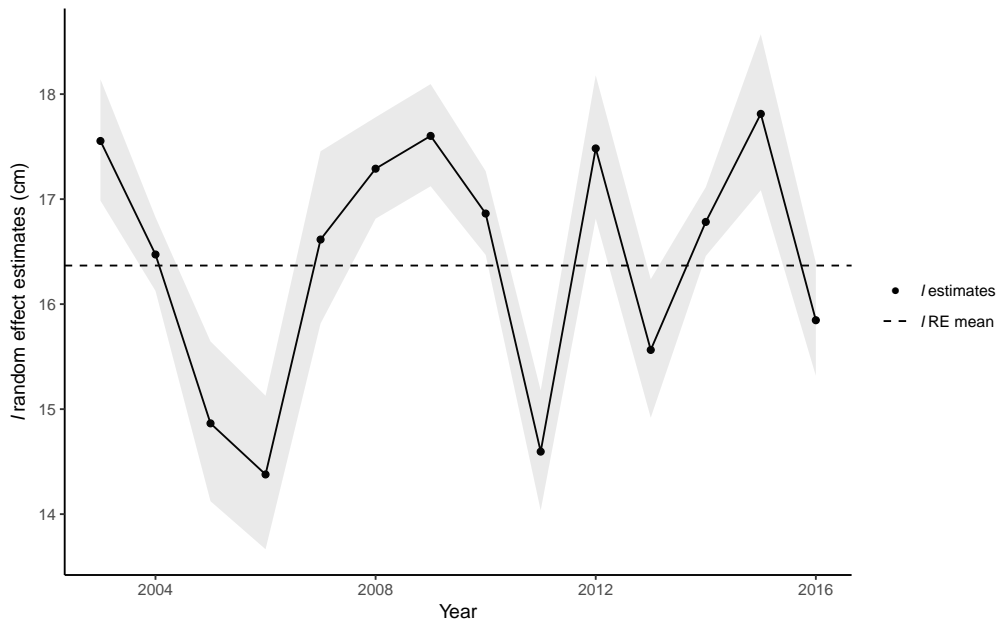
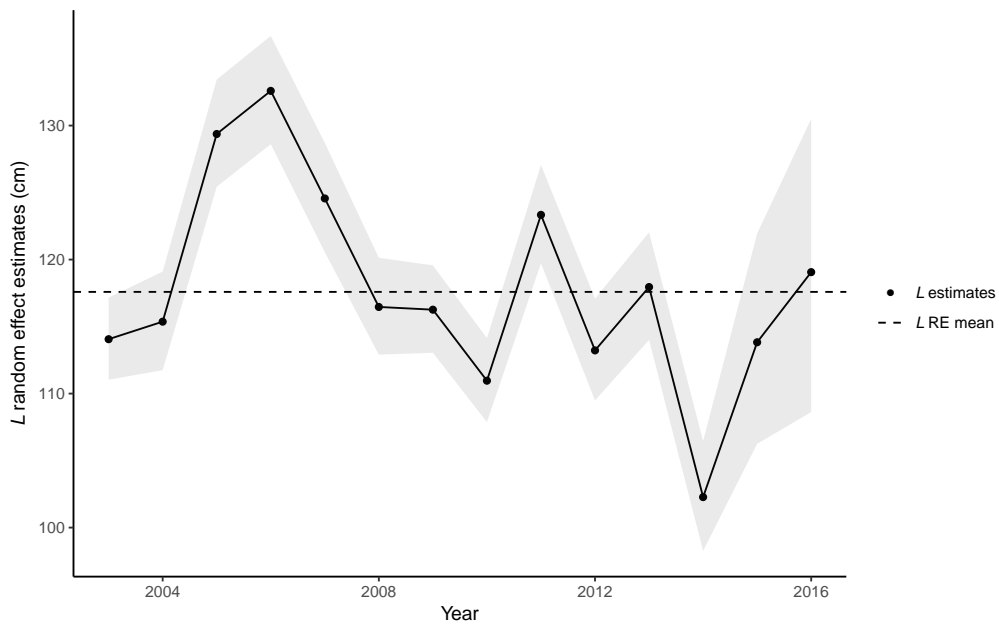
(a) Cohort l random effect(b) Cohort L random effect

Figure 2.7: Estimated random effects on first (l) and last (L) component means of cohorts from a nine component hierarchical LSD model on length frequency distributions of white-bellied anglerfish. Shaded area denotes $2 \times$ standard error of the random effect estimates and dashed lines denote the mean of the random effect. Parameter estimates are for quarter four surveys (EVHOE and IE-IGFS).

2.3.2.3 Other White-bellied anglerfish hierarchical models

Hierarchical models with a cohort specific k or yearly varying k were also fit to the anglerfish data. In both cases the RVBG parameter K varies to a large extent between cohorts, however L_∞ also varies and the resulting cohort curves differ but not dramatically. Parameter estimates can be found in Appendix B.

2.4 Discussion

2.4.1 Overview of LFEM

This study aimed to improve upon currently available approaches for growth estimation of fish through length frequency analysis. A basic method utilising the EM algorithm and analogous to MULTIFAN (Fournier et al., 1990) was produced. The methodology was then extended further to incorporate bivariate random effects in order to capture some of the growth variation in length frequency data.

Few studies have modelled growth variation in length frequency data (Fournier et al., 1998; Roa-Ureta, 2010; Lloyd-Jones et al., 2016). The approach of using a VBGF to constrain mixture model component means through a time series (i.e. along a cohort) whilst explicitly estimating cohort or year specific random effects on two growth parameters is a worthwhile extension to this area of research (see Lloyd-Jones et al. (2016) for an alternative). The aim of this integrated approach was to aid identification of growth parameters in variable length frequency distributions obtained from multiple annual surveys. Furthermore, the associated uncertainty of the parameter estimates is also approximated. Overall both versions of the methodology produced reasonable estimates of von Bertalanffy growth parameters and their variability between cohorts/years. However there were problems with convergence to biologically unreasonable estimates of classical von Bertalanffy parameters particularly in the basic models, as well as issues with model stability in the hierarchical models where information was intermittently sparse (i.e weak or low abundance cohorts).

2.4.2 Haddock

Basic and overall hierarchical parameter estimates for haddock were within or close to previously reported ranges ($K \approx 0.1 - 0.5$ and $L_\infty \approx 45\text{cm} -$

55cm) (Baudron et al., 2011). Parameter estimates from the best fitting basic model (12 components and LSD) and overall parameter estimates from the l/k hierarchical LSD model gave very similar growth curves to that of a maximum likelihood fit on length-at-age data for fish from the same region and time of year (Figure 2.2). CSD models overestimated growth somewhat in older fish indicating the importance of identifying the correct standard deviation. When applying their method to data for pike and abalone Schnute and Fournier (1980) also found that standard deviation treatment was particularly important when deriving growth from observed length frequencies. The number of components identified by the best fitting basic LFEM model (12 components) is close to the number of ages present in the ALK (11 ages). The model selection criterion was also particularly useful for identifying the more suitable SD type. The effect of misidentifying component number is also negated somewhat by the consistency of the component means with differing number of components (Table 2.4).

From the hierarchical model (random effects on l and k) we can see a similar variation in cohort specific growth parameters as Baudron et al. (2011) found in their study. Hierarchical model and length-at-age estimated cohort growth curves were similar in cohorts observed up until the ninth component. However outside of these cohorts growth was often underestimated, indicating the effect of the lack of information in later ages. Cohort growth curves that showed the most similar fits to the data between the two methods also tended to be for cohorts where strong signals were present in sequential years (Figure 2.3 and 2.4).

2.4.3 White-bellied anglerfish

Parameter estimates for white-bellied anglerfish (both sexes combined) were broadly similar to those in the literature, although differences are more pronounced in the later ages and estimates of L_∞ from this study were consistently higher than previously reported (Farina et al., 2008; Landa et al., 2013).

The basic and hierarchical models produced different growth curves, despite having the same starting parameters. The 2014 cohort from the hierarchical model follows a similar trajectory to the best fitting basic LFEM model growth curve, indicating that 2014-2016 (the years 2014 cohort is observed) may be dominating in the estimation of the basic model. Flexibility within the hierarchical model was able estimate a slower growth for the 2014 cohort, whilst

estimating a slightly faster growth for other cohorts. Further investigations showed that when years 2014 to 2016 (these years had a higher number of fish than other years) were removed from the data the basic LFEM model component means were more alike to the hierarchical model estimates for the majority of years. Subjectively, fitted distributions of the hierarchical model seem to fit better than the basic LFEM model (Figure 2.6). Estimated growth curves predicted faster growth, particularly in the older fish, compared to anglerfish age validation studies (Figure 2.5) (Farina et al., 2008; Landa et al., 2013). Faster growth in the later ages was less pronounced in models that assumed a higher number of components, however these were not selected as the best fitting model (Table 2.5).

Model selection based on sub-AIC indicated that a nine component model with LSD provided the best fit to the data. Although *L. piscatorius* is thought to live frequently to over ten years of age (Farina et al., 2008), fish of this age are not sufficiently well represented in the length frequency data to support the selection of a large number of components. The nine component model produced a higher estimate of L_∞ and a lower estimate of K compared to other studies (Farina et al., 2008; Landa et al., 2013), however, the estimates of components means and standard deviations fit the length frequency distributions reasonably well (Figure 2.6).

As with haddock SD treatment was important. Including a linear increasing component standard deviation produced components with very wide length distributions and mixing proportions appeared to be biased towards a single component in some cases. Constant standard deviation allowed for more overlap between components and more balanced mixing estimations but not necessarily a better fitting model (Table 2.5).

Overall, discrepancies between parameter estimates from this study and those from studies that use direct ageing methods may reflect modelling challenges with the data or the difficulties of obtaining reliable age estimates for anglerfish (ICES, 2011). Landa et al. (2013) used length-frequency analysis techniques to validate direct ageing of illicia and found good agreement between the parameter estimates produced by both methods. However that study was based on survey data from the Porcupine Bank while the present study also included data from two other surveys as model inputs. Further difficulty in estimating growth may have been introduced here by combining data for males and females as white-bellied anglerfish show sexual dimorphism in growth

(Farina et al., 2008).

2.4.4 Caveats

The EM algorithm is an accepted method for estimating latent variables but is sensitive to the choice of starting parameters (Biernacki et al., 2003), as was evident from the model outputs presented here. This was accounted for by conducting sensitivity runs on some of the main parameters (k , L and number of components), producing a wide range of values for each parameter estimate. Some other starting parameters were fixed, such as the mean of the first component (l) and the starting mixing proportions, which was arbitrarily set as $1/No. components$. It is likely that the starting mixing proportions in conjunction with a k growth rate parameter drove the variation in parameter estimates at convergence. Lloyd-Jones et al. (2016) modelled length frequency distributions using a methodology similar to the one employed here, but used a minorisation-maximisation algorithm in order to estimate parameters.

The number of sensitivity runs was large for this study (>1000 for each SD type) and a wide range of starting parameter values was used. Many k parameter starting values were very different to known estimates, which may partially account for the number of model runs that were excluded before further analysis. The present study's approach to parameter sensitivity by multiple model runs followed by exclusion of models with "biologically unreasonable" parameter estimates is a robust method for dealing with the sensitivity, although caution should be used when setting a threshold (e.g. this study used a L_∞ of 500 cm, so as not to introduce subjectivity into model selection). Haddock required a further refining of models by identifying and removing any models with a second component mean ≤ 15 cm. Second component means ≤ 15 cm were caused by certain (slow growth) starting parameters and the second component mixing proportion tending to zero.

In this study the sub-AIC model selection criteria appeared to work reasonably well (details of sub-AIC are given in subsection 2.5). sub-AIC was chosen as the model selection criteria in part due to the performance of AIC in exploratory simulation analysis with the LFEM model. AIC was able to identify the correct number of components when the standard deviation of components was small (~ 1), as did BIC. With increasing standard deviation of length frequency data AIC was more robust, still under-fitting the data somewhat but much less severely than other model selection criteria. However, model selection is often

problematic in studies such as this, hence should be conducted cautiously using expert knowledge of the species in question (Schnute and Fournier, 1980; Fournier et al., 1990; Taylor and Mildenerger, 2017).

Correlation between parameters could contribute to the observed variation in parameter estimates. Classical von Bertalanffy parameters (K and L_∞) have been shown to covary strongly (Gallucci and Quinn, 1979; Pilling et al., 2002). Here, the re-parameterised von Bertalanffy growth parameters appeared strongly correlated, with many combinations of k and L parameters resulting in very similar model fits. The same issue arises with other length frequency analysis methods such as MULTIFAN and ELEFAN and is usually dealt with by fixing either K or L_∞ in the initial stages of the analysis (Fournier et al., 1990; Taylor and Mildenerger, 2017). Variation in estimated maximum likelihood parameters could also be driven by the substantial variability that exists in the input data.

Hierarchical models appeared to be less sensitive to the choice of starting parameters, although due to computational demands these models were not subjected to the same level of sensitivity testing. Model testing using the haddock data showed that when length frequency data indicates very variable recruitment (i.e. in some years the first component is almost not present) the hierarchical model initially struggled to estimate these components, shifting the first component to the second component position and thus confounding estimates of cohort progression. It seemed that there was not sufficient information within the estimated distribution of the random effect to prevent this. This issue is a form of the label switching problem that is well-known in mixture models (Yao, 2015). Fixing the standard deviation of the bivariate random effects at a suitably low value (i.e. $\exp(-5)$) as detailed in section 2.5, retained enough flexibility in the model to estimate variable random effects but sufficient constraint to keep l and k random effects estimates within reasonable bounds, thus preventing label switching.

2.4.5 Further work

A comprehensive comparison of the basic and hierarchical models with other methodologies such as MULTIFAN, the updated ELEFAN procedure and the more subjective Bhattacharya method would be a useful avenue for further research (Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017). In MULTIFAN-CL density dependence can be included in the structure

of the model as abundance of a cohort can effect growth rate in some fish species (Fournier et al., 1998). The hierarchical model presented in this study is currently designed to model growth variability between cohorts independently of cohort abundance but could be modified to include cohort abundance as a covariate, increasing complexity and modelling the connection between mixing probabilities of a cohort. Laslett et al. (2004) and Roa-Ureta (2010) devised stage-based approaches to model fitting. A comparison of these stage based approaches with hierarchical model presented here would be useful in order to the assess the effectiveness of integrating constraints and random effects into the mixture model parameter estimation.

Currently the components of the mixture models are assumed to be normally distributed. Further development of the method could enable the user to specify alternative distributions such as log normal, gamma or other, thus allowing non-normality and skewness to be modelled (Macdonald and Pitcher, 1979; Lloyd-Jones et al., 2016).

2.5 Conclusions

Use of length frequency analysis to obtain growth parameters is a well developed area of research (Hasselblad, 1966; Fournier et al., 1990, 1998; Taylor and Mildenerger, 2017). These techniques provide growth information when age of individuals cannot be determined otherwise. Parameter estimates, their associated uncertainty and model fits presented here demonstrate the usefulness of our method for obtaining growth information from length frequency data.

Growth parameters are estimated whilst simultaneously fitting finite mixture models using the EM algorithm. The use of the EM algorithm removes the need to define length bins, allowing a more objective set up of the modelling procedure. Models can take inputs from multiple surveys from different times of year and offer an alternative framework to other well-known approaches for estimating growth parameters from length frequency data. Furthermore, within the hierarchical framework we explicitly model bivariate random effects on growth parameters, successfully modelling growth variability and allowing trends to be identified in the estimates.

Where age data is not readily available modelling cohort growth variability through the present study's methodology can improve information available for stock assessments (i.e. cohort specific growth curves for age slicing. In addition,

estimates from these length frequency models give useful credible estimates of the boundaries of growth parameters that could be tested for stock assessments in a simulation context, such as in a management strategy evaluation (Punt et al., 2013). Finally, length frequency distribution derived cohort growth parameter estimates could also be used to investigate environmental covariates such as sea temperature (Baudron et al., 2014; Barrow et al., 2018).

2.6 Supplementary material

The following supplementary material is available at ICESJMS online. Further details on the model derivation and some additional results can be found here. Code and examples of model runs can be found at <https://github.com/lbatts/LFEM>.

2.7 Acknowledgements

We thank the Cullen fellowship programme of the Marine Institute, Ireland for the funding for this postgraduate research. We'd also like to thank ICES, the reviewers and editor for their comments and those that provided and collated the data that was used in the study.

Bibliography

- Barrow, J., J. Ford, R. Day, and J. Morrongiello (2018). Environmental drivers of growth and predicted effects of climate change on a commercially important fish, *Platycephalus laevigatus*. *Marine Ecology Progress Series* 598, 201–212.
- Baudron, A., C. Needle, and C. Marshall (2011). Implications of a warming North Sea for the growth of haddock *Melanogrammus aeglefinus*. *Journal of Fish Biology* 78(7), 1874–1889.
- Baudron, A. R., C. L. Needle, A. D. Rijnsdorp, and C. Tara Marshall (2014). Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes. *Global change biology* 20(4), 1023–1031.
- Bhattacharya, C. (1967). A simple method of resolution of a distribution into Gaussian components. *Biometrics*, 115–135.
- Biernacki, C., G. Celeux, and G. Govaert (2003). Choosing starting values for the em algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis* 41(3-4), 561–575.
- Cadigan, N. G., S. E. Campana, and H. editor: Jan Jaap Poos (2016). Hierarchical model-based estimation of population growth curves for redfish (*Sebastes mentella* and *Sebastes fasciatus*) off the Eastern coast of Canada. *ICES Journal of Marine Science* 74(3), 687–697.
- Chang, Y.-J., C.-L. Sun, Y. Chen, and S.-Z. Yeh (2012). Modelling the growth of crustacean species. *Reviews in Fish Biology and Fisheries* 22(1), 157–187.
- Cormon, X., B. Ernande, A. Kempf, Y. Vermard, and P. Marchal (2016). North sea saithe *Pollachius virens* growth in relation to food availability, density dependence and temperature. *Marine Ecology Progress Series* 542, 141–151.

- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1–38.
- Denney, N. H., S. Jennings, and J. D. Reynolds (2002). Life–history correlates of maximum population growth rates in marine fishes. *Proceedings of the Royal Society of London B: Biological Sciences* 269(1506), 2229–2237.
- Eveson, J. P., J. Million, F. Sardenne, and G. Le Croizier (2015). Estimating growth of tropical tunas in the Indian Ocean using tag-recapture data and otolith-based age estimates. *Fisheries research* 163, 58–68.
- Eveson, J. P., T. Polacheck, and G. M. Laslett (2007). Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. *Canadian Journal of Fisheries and Aquatic Sciences* 64(4), 602–617.
- Farina, A., M. Azevedo, J. Landa, R. Duarte, P. Sampedro, G. Costas, M. Torres, and L. Cañás (2008). *Lophius* in the world: a synthesis on the common features and life strategies. *ICES Journal of Marine Science: Journal du Conseil* 65(7), 1272–1280.
- Fournier, D. A., J. Hampton, and J. R. Sibert (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences* 55(9), 2105–2116.
- Fournier, D. A., J. R. Sibert, J. Majkowski, and J. Hampton (1990). MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian journal of fisheries and aquatic sciences* 47(2), 301–317.
- Gallucci, V. F. and T. J. Quinn (1979). Reparameterizing, fitting, and testing a simple growth model. *Transactions of the American Fisheries Society* 108(1), 14–25.
- Hasselblad, V. (1966). Estimation of parameters for a mixture of normal distributions. *Technometrics* 8(3), 431–444.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.

- Hoggarth, D. D. (2006). *Stock assessment for fishery management: A framework guide to the stock assessment tools of the fisheries management and science programme*. Number 487. Food & Agriculture Org.
- ICES (2011). Report of the Anglerfish (*Lophius piscatorius*) illicia and otoliths exchange 2011, ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS).
- ICES (2018a). DATRAS online database. https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx. Accessed: 2018-05-09.
- ICES (2018b). DATRAS Survey descriptions. <https://datras.ices.dk/home/descriptions.aspx>. Accessed: 2018-05-09.
- ICES (2018). ICES Advice on fishing opportunities, catch, and effort Bay of Biscay and the Iberian Coast, Celtic Seas, Greater North Sea, and Oceanic Northeast Atlantic ecoregions - mon.27.78abd. Technical report.
- Isermann, D. A. and C. T. Knight (2005). A computer program for age-length keys incorporating age assignment to individual fish. *North American Journal of Fisheries Management* 25(3), 1153–1160.
- Jobling, M. (2002). Environmental factors and rates of development and growth. *Handbook of fish biology and fisheries* 1, 97–122.
- Katsanevakis, S. and C. D. Maravelias (2008). Modelling fish growth: multi-model inference as a better alternative to a priori using von Bertalanffy equation. *Fish and fisheries* 9(2), 178–187.
- Kilada, R. and J. G. Driscoll (2017). Age determination in crustaceans: a review. *Hydrobiologia* 799(1), 21–36.
- Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* 70(5), 1–21.
- Landa, J., J. Barrado, and F. Velasco (2013). Age and growth of anglerfish (*Lophius piscatorius*) on the Porcupine Bank (west of Ireland) based on illicia age estimation. *Fisheries research* 137, 30–40.
- Laslett, G. M., J. P. Eveson, and T. Polacheck (2004). Fitting growth models to length frequency data. *ICES Journal of Marine Science* 61(2), 218–230.

- Lloyd-Jones, L. R., H. D. Nguyen, G. J. McLachlan, W. Sumpton, and Y.-G. Wang (2016). Mixture of time-dependent growth models with an application to blue swimmer crab length-frequency data. *Biometrics* 72(4), 1255–1265.
- Macdonald, P. and J. Du (2011). *mixdist: finite mixture distribution models*. R package version 0.5-4.
- Macdonald, P. and T. Pitcher (1979). Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. *Journal of the Fisheries Board of Canada* 36(8), 987–1001.
- Maunder, M. N., P. R. Crone, A. E. Punt, J. L. Valero, and B. X. Semmens (2016). Growth: Theory, estimation, and application in fishery stock assessment models. *Fisheries Research* (180), 1–3.
- McLachlan, G. and D. Peel (2004). *Finite Mixture Models*. John Wiley & Sons.
- Mildenberger, T. K., M. H. Taylor, and M. Wolff (2017). Tropfishr: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution* 8(11), 1520–1527.
- Millar, R. B. and R. J. Fryer (1999). Estimating the size-selection curves of towed gears, traps, nets and hooks. *Reviews in Fish Biology and Fisheries* 9(1), 89–116.
- Minte-Vera, C. V., M. N. Maunder, J. M. Casselman, and S. E. Campana (2016). Growth functions that incorporate the cost of reproduction. *Fisheries Research* 180, 31–44.
- Minto, C., J. Hinde, and R. Coelho (2017). Including unsexed individuals in sex-specific growth models. *Canadian Journal of Fisheries and Aquatic Sciences* 75(2), 282–292.
- Morrongiello, J. R. and R. E. Thresher (2015). A statistical framework to explore ontogenetic growth variation among individuals and populations: a marine fish example. *Ecological Monographs* 85(1), 93–115.
- Ogle, D. H. (2018). *FSA: Fisheries Stock Analysis*. R package version 0.8.20.
- Pardo, S. A., A. B. Cooper, and N. K. Dulvy (2013). Avoiding fishy growth curves. *Methods in Ecology and Evolution* 4(4), 353–360.

- Pauly, D. (1987). A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. In *ICLARM conf. Proc.*, Volume 13, pp. 7–34.
- Pilling, G. M., G. P. Kirkwood, and S. G. Walker (2002). An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. *Canadian Journal of Fisheries and Aquatic Sciences* 59(3), 424–432.
- Pitcher, T. and P. MacDonald (1973). Two models for seasonal growth in fishes. *Journal of applied ecology*, 599–606.
- Punt, A. E., F. Trinnie, T. I. Walker, R. McGarvey, J. Feenstra, A. Linnane, and K. Hartmann (2013). The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face of non-stationary dynamics. *Fisheries Research* 137, 116–128.
- Quinn, T. J. and R. B. Deriso (1999). *Quantitative Fish Dynamics*. Oxford University Press.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roa-Ureta, R. H. (2010). A likelihood-based model of fish growth with multiple length frequency data. *Journal of agricultural, biological, and environmental statistics* 15(3), 416–429.
- Schnute, J. and D. Fournier (1980). A new approach to length–frequency analysis: growth structure. *Canadian Journal of Fisheries and Aquatic Sciences* 37(9), 1337–1351.
- Smith, M., E. Fulton, R. Day, L. Shannon, and Y.-J. Shin (2015). Ecosystem modelling in the southern Benguela: comparisons of Atlantis, Ecopath with Ecosim, and OSMOSE under fishing scenarios. *African Journal of Marine Science* 37(1), 65–78.
- Sparre, P. and S. C. Venema (1998). *Introduction to Tropical Fish Stock Assessment-Part 1: Manual*. FAO.
- Taylor, M. and T. K. Mildenerger (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology* 24(4), 330–338.

- Travers, M., Y.-J. Shin, S. Jennings, and P. Cury (2007). Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. *Progress in oceanography* 75(4), 751–770.
- von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii). *Human biology* 10(2), 181–213.
- Wang, Y.-G. (1999). Estimating equations for parameters in stochastic growth models from tag–recapture data. *Biometrics* 55(3), 900–903.
- Whitten, A. R., N. L. Klaer, G. N. Tuck, and R. W. Day (2013). Accounting for cohort-specific variable growth in fisheries stock assessments: a case study from south-eastern Australia. *Fisheries Research* 142, 27–36.
- Winsor, C. P. (1932). The gompertz curve as a growth curve. *Proceedings of the national academy of sciences* 18(1), 1–8.
- Yao, W. (2015). Label switching and its solutions for frequentist mixture models. *Journal of Statistical Computation and Simulation* 85(5), 1000–1012.

Chapter 3

Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real- life stocks

This chapter is a verbatim reproduction of a paper accepted and in the *Canadian Journal of Fisheries and Aquatic Sciences*. Supplementary material can be found in Appendix D.

Batts, L., Minto, C., Gerritsen, H., & Brophy, D. Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real- life stocks. *Canadian Journal of Fisheries and Aquatic Sciences*

Abstract

Stage-based assessment models are a type of fisheries stock assessment model that offer an alternative middle ground between aggregate and compositional models. We compare the capabilities of two theoretically different stage-based assessment approaches: an implementation of a biomass-based delay-difference model first described in a theoretical paper by Schnute (1987), and an implementation of the well-known numbers-based two-stage model Catch-Survey Analysis (CSA). Models were tested within a simulation framework as well as on the real stock of white-bellied anglerfish in the Celtic Seas and Northern Bay of Biscay. For the simulated stocks, estimates from the biomass-based two-stage models were close to the true values in certain scenarios, but were sensitive to selectivity assumptions and configuration of growth within the model. CSA was more robust to selectivity assumptions, performing well in all simulated stock scenarios. Key estimated values from CSA were remarkably similar to the values estimated in the accepted age-based assessment for the real anglerfish stock. Overall, results indicated that CSA was a robust stock assessment model but with relatively low precision, whereas the Schnute model was precise but required growth and mean fish weight data unaffected by selectivity.

Keywords: anglerfish *Lophius piscatorius*, data-limited, stage-based, stock assessment, TMB

3.1 Introduction

Fisheries stock assessments are an important tool for fisheries management (Hilborn and Walters, 1992; Dichmont et al., 2016). Using a variety of data sources, stock assessments determine the size of a fish population and the rate at which the stock is being fished. There are many different forms of fisheries stock assessment models, ranging from biomass dynamic models (e.g. Prager, 1992; Pedersen and Berg, 2017), to age-based assessment models (e.g. Shepherd, 1999; Jardim et al., 2014; Nielsen and Berg, 2014), to more complex spatial (e.g. Cadigan et al., 2017), ecosystem (e.g. Begley and Howell, 2004) or integrated assessment models (e.g. Methot and Wetzell, 2013). The choice of a suitable assessment model and its reliability is often determined by the available data (Hilborn and Walters, 1992). For example, in stocks where there is large uncertainty in age and growth estimates, more complex compositional age-based models may not be possible and a different approach using aggregated, stage-based, size-based or integrated models is preferred (Smith and Addison, 2003; Punt et al., 2013).

Two well known but theoretically different approaches to stage-based assessment models are biomass-based delay-difference models, first described by Deriso (1980) and developed by Schnute (1985, 1987) as well as Fournier and Doonan (1987), and numbers-based depletion models, the most well-known of which is the Catch-Survey Analysis (CSA) (Collie and Sissenwine, 1983; Smith and Addison, 2003). Standard implementations of these two approaches consist of two stages; fish recruited to the exploited population in a given year, and fish that recruited in previous years. Delay-difference models can be generally considered as biomass dynamic models but with different parameters and a structure that allows for time lags in growth and recruitment (Hilborn and Walters, 1992). Schnute (1987) described a size-based delay-difference model linking population size structure and mean weights. In its simplest form, this consists of two stages of biomass (recruits and previously exploited biomass) and assumes deterministic growth of all individuals in the exploited stock. The model also allows for flexibility in the assumptions regarding the relative importance of recruitment and previously exploited biomass that underlie estimation of total biomass. CSA is a relatively simple two-stage model (numbers of recruits and post-recruits), which has changed somewhat from the original model first described in Collie and Sissenwine (1983). The most recent version can be found in the NOAA Fish and Fisheries Integrated Toolbox

(NFFT), where population dynamics centre on Baranov's catch equation and estimation is through maximum likelihood (NOAA, 2019).

Unreliable age-composition data or uncertainty in data that is not suitably accounted for either in the assessment model or in the processing of data prior to an assessment model, can lead to bias in numbers-at-age estimates (Kell and Kell, 2011; Maunder and Punt, 2013; Ailloud et al., 2015). Integrated assessment models such as Stock Synthesis (Methot and Wetzel, 2013) or MULTIFAN-CL (Fournier et al., 1998) have been used widely in recent years to address a lack of reliable age-composition data. However, these assessment models are often relatively complex, requiring more data and expertise than other simpler assessments (Maunder and Punt, 2013; Punt et al., 2013; Ailloud et al., 2015). In these situations, stage-based models can be useful as they generally require fewer data and parameters to estimate, and are likely to be more robust to large uncertainties in the data.

Stage-based assessment models such as the Schnute (1987) model or CSA (NOAA, 2019) offer an alternative middle ground between aggregate and compositional models, as they incorporate information on recruitment where aggregate models do not and also have more limited data requirements than compositional models (Hilborn and Walters, 1992; Cadrin, 2000; Ailloud et al., 2015; Cook and Heath, 2018; Li et al., 2019). Generally, stage-based models have been used in situations where species are difficult to age (e.g. crustacean fisheries) (Cadrin, 2000; Smith and Addison, 2003; Zhou et al., 2011; Kienzle et al., 2014) or are short lived (e.g. squid fisheries) (Roel and Butterworth, 2000; Ibaibarriaga et al., 2008).

While the importance of two-stage models has long been known and their theoretical properties well understood, the two approaches (numbers-based and biomass-based) have not been compared in detail. We attribute this primarily to a lack of implementations of the general biomass-based model. Here, we implemented the Schnute (1987) delay-difference and NOAA (2019) CSA assessment models within a common estimation framework (Template Model Builder, TMB) (Kristensen et al., 2016), taking advantage of automatic differentiation of the likelihood during optimisation.

Our aims are to: 1) fully describe the comparative implementation of the models; 2) assess the ability to recover known properties of realistically simulated stocks under a range of data scenarios; and 3) compare against a fully

age-based stock assessment model for the white-bellied anglerfish (*Lophius piscatorius*) in the Celtic Seas and Northern Bay of Biscay. A simulated stock framework was used to determine the reliability of the assessment models, as well as the accuracy of their parameter estimation. The white-bellied anglerfish stock gave a useful test case to evaluate the performance of assessment models on a real stock with more complex population dynamics. Here, a stage-based modelling approach could be of use in the future assessment and management of anglerfish, as well as similar species with prominent uncertainty regarding growth.

3.2 Methods

A description of the theoretical model from Schnute (1987) is first introduced, followed by implementation details. A summary of the CSA model and its implementation is then given. Finally, the simulated and real stocks are described. We use the same notation as Schnute (1987) to describe both models where possible. An asterisk is used to distinguish between biomass and numbers (e.g. N^* and N are population biomass and numbers respectively).

3.2.1 Schnute model

Schnute (1987) described a size-based delay-difference model that links population size structure and biomass with the mean weights of size classes. Many of the following equations are taken directly from Schnute (1987) in order to explain the model and its implementation. Where possible, we have used the same terminology, in addition, the reader should note that the term weight is used to refer to mass throughout. The prime is used to denote a weight after a year of growth (e.g. V' is a function of starting weight V). The reader should refer to Schnute (1987) for a full description of the theoretical underpinnings.

3.2.1.1 Population dynamics and notation

Time is discrete in this model and there is an implicit assumption that growth occurs at the end of the year. Total survival of the population from one year to the next is also assumed to be weight-independent (analogous to age-structured delay-difference models (Schnute, 1985)). These fundamentals lead to the key dynamic equations for a weight-structured population model described in Schnute (1987) and summarised in *Appendix D*.

In its simplest form, the Schnute (1987) model defines two weight intervals that correspond to the two stages in the delay-difference model. V is the recruitment weight at which fish are assumed to be fully selected and the weight range $[V, V')$ encompasses all fish that are large enough to be exploited for the first time in the current year (i.e., recruit stage). Fish larger than V' are assumed to have been large enough to be exploited for at least one year previous to the current year (i.e., previously-exploited biomass stage). The theoretical weight limit, V^∞ , was the upper bound of the interval, i.e., weight interval $[V, V^\infty)$ encompasses the entire range of fish weights in the assessed population. Schnute (1987) elaborates this methodology further and extends

the theory to a full cohort analysis, however, implementation of this was beyond the scope of this study and our implementation follows the two-stage model.

Table 3.1: Summary of the Schnute model notation. Hat operator denotes estimated parameters (for this study) or predicted values from the model. Parameters without hats were treated as known values. ‘S1 specific’ indicates notation specific to this form of the Schnute model.

Notation	Type	Description
General model		
N^*	output	Population biomass
P^*	output	previously exploited biomass
R^*	output	Recruitment biomass
σ	parameter	proportion of population that survives natural mortality
\hat{q}_s	parameter	catchability coefficient of survey indices s
$\hat{\delta}_s$	parameter	standard deviation of log-normal survey indices
λ	parameter	standard deviation of log-normal catch biomass
ρ	parameter	coefficient of linear growth model
W	parameter	intercept of linear growth model
θ_s	parameter	timing of survey
\bar{Z}_t	input	mean weight of fish in previously exploited biomass stage at time t
\bar{X}_t	input	mean weight of fish in entire biomass at time t
\bar{Y}_t	input	mean weight of fish in recruit biomass stage at time t
ω_t	output	fraction of total biomass in year t due to newly recruited fish
C_t^*	input	observed catch biomass at time t
I_{ts}^*	input	observed survey biomass at time t for survey s
μ	input	fraction of the catch removed before natural mortality
\hat{I}_{ts}^*	output	predicted survey biomass at time t for survey s
s	index element	survey number
t	index element	number of year
S1 specific		
\hat{N}_1^*	parameter	initial biomass of population at $t = 1$
\hat{C}_t^*	output	predicted catch biomass at time t
\hat{F}_t	parameter	fishing mortality at time t
$\hat{\mu}_t$	output	fraction of the catch removed before natural mortality

Growth of the surviving population biomass from one year to the next relies on the assumption that fish are fully available to the fishery above the specified recruitment weight V (knife-edged selection). The growth model for a weight structured population is assumed to be linear here, where weight this year (w') is a linear function of the weight in the previous year (w) with intercept and slope a function of the von Bertalanffy K and W_∞ parameters (Schnute, 1987). This growth model is typical of delay-difference models (usually with weights-at-age) (Hilborn and Walters, 1992) and implies that growth does not vary over time.

Average weights are also a key component for the weight-structured model proposed in Schnute (1987). The model implemented in this study requires

the following inputs: mean fish weight of the assessed population \bar{X}_t , recruit mean fish weight \bar{Y}_t and the previously exploited biomass mean fish weight \bar{Z}_t at time t (Table 3.1). Mean weights are assumed to stay constant through the year.

We implement two forms of the proposed two-stage model from Schnute (1987). Models $S0$ and $S1$ are presented (Tables 3.1 and 3.2) and detailed below.

3.2.1.2 Deterministic equations

The basic delay-difference equation can be simply described as predicting biomass at time $t+1$ from three basic components: recruitment biomass at $t+1$, surviving biomass from time t and growth of the surviving fish (Deriso, 1980; Schnute, 1985; Hilborn and Walters, 1992). Details of how this process is specifically presented in Schnute (1987) can be found in *Appendix D*. In summary, biomass in year $t + 1$ minus recruitment, or previously exploited biomass P_{t+1}^* , is derived from the surviving biomass from year t multiplied by a growth factor of $\frac{X'_t}{X_t}$, where:

$$X'_t = W + \rho X_t \quad (3.1)$$

X'_t is the average weight of a fish in the entire surviving population after a year of growth and equivalent to Z_{t+1} , W is the intercept of the linear growth model and ρ is the slope.

An intriguing aspect of the model proposed in Schnute (1987) is that there are three model versions, where a predicted total biomass index can be calculated a number of ways. Two of these model versions use all three time series of mean fish weights (X , Z and Y) to calculate ω_t , where ω_t is defined as the fraction of total biomass in year t due to newly recruited fish (Table 3.1),

$$\omega_t = \frac{R_t^*}{N_t^*} \quad (3.2)$$

where R_t^* is recruitment biomass at time t and N_t^* is population biomass at time t . Schnute (1987) demonstrates that ω_t can be derived from mean weights

alone:

$$\omega_t = \left(\frac{Y_t}{X_t} \right) \left(\frac{Z_t - X_t}{Z_t - Y_t} \right) \quad (3.3)$$

To compute estimates of N_t^* , ω_t values can then be used with either the recruit stage R_t^* or the previously-exploited population stage P_t^* .

Here, we focus on the model version that estimates N_t^* from P_t^* (P_t^* being calculated from N_{t-1}^*) at each time step. The more classical model version where N_t^* is the sum of estimated recruits and previously exploited biomass, as well as the model version that calculates N_t^* from R_t^* , were also implemented and are available within the developed R package (<https://github.com/lbatts/sbar>) but are not investigated in this study. The particular model version used in this study was chosen due to its relative simplicity compared to the other versions. Details for the model version focused on in this study can be found in *Appendix D*.

3.2.1.3 Model variants

Schnute (1987) presented an autoregressive process error model (referred to as Original Schnute Process Error Model (*S0*) subsequently) that takes catch biomass as known and μ (fraction of the catch removed prior to natural mortality) (Table 3.1) is specified by the user. We also develop a model where fishing mortality rate is estimated and catch biomass is predicted using the Baranov equation (*S1*).

When demonstrating how their theory relates to more classical theory, Schnute (1987) shows (but does not include in the composite model) that μ is a function of instantaneous rates of fishing mortality and natural mortality, we index fishing mortality by time to give:

$$\mu_t = \frac{F_t(1 - e^{-M}) - Me^{-M}(1 - e^{-F_t})}{F_t(1 - e^{-M})(1 - e^{-M-F_t})} \quad (3.4)$$

where F_t is the fishing mortality at time t . Survival σ in the Schnute (1987) model is equivalent to e^{-M} where M is instantaneous natural mortality rate

(Table 3.1).

Using the indexed equation we are able to estimate yearly fishing mortality within the model, whilst maintaining the majority of the the original equations of Schnute (1987). μ_t computed from the estimated F_t can be thought of as the fraction of the catch taken before natural mortality occurs that mimics the application of exponential total mortality ($-M - F_t$) in the classical Baranov catch equation:

$$C_t^* = \frac{F_t}{M + F_t} (1 - e^{-M - F_t}) N_t^* \quad (3.5)$$

where C_t^* is catch biomass at time t .

We develop a model that uses the above to predict catch biomass. This new adaptation of the Schnute (1987) model also estimates biomass in the first time step and is an observation-error-only model (referred to subsequently as the Adapted Schnute Observation Error Model (*S1*)) (Tables 3.1 and 3.2) .

3.2.1.4 Model summaries

We implement two forms of the proposed two-stage model from Schnute (1987) (Schnute model version where entire biomass is derived from previously exploited biomass as stated in section 3.2.1.2) that can be distilled into a set of steps representing the processes within the models. Model variants *S0* and *S1* are presented (Table 3.2). *S0* is faithful to Schnute (1987) with an autoregressive process error, *S1* estimates initial biomass (N_1^*) as a stand-alone parameter and with observation-error-only.

Table 3.2: Summary of the steps within the Original Schnute Process Error Model (*S0*) and Adapted Schnute Observation Error Model (*S1*). Hat operator denotes predicted states assuming measurement error only within the model.

Step	Model	
	<i>S0</i>	<i>S1</i>
1	$N_t^* = \frac{I_{t,s}^* + q_s \theta_s [(1-\sigma)C_t^*]}{q_s [1-\theta_s (1-\sigma)]}$	\hat{N}_1^* estimated within the model
2	$P_{t+1}^* = \frac{X_t'}{X_t} [\sigma [N_t^* - \mu C_t^*] - (1-\mu)C_t^*]$	$\hat{P}_{t+1}^* = \frac{X_t'}{X_t} [\sigma [\hat{N}_t^* - \mu_t \hat{C}_t^*] - (1-\mu_t)\hat{C}_t^*]$
3	$R_{t+1}^* = \frac{P_{t+1}^* \omega_{t+1}}{1-\omega_{t+1}}$	$\hat{R}_{t+1}^* = \frac{\hat{P}_{t+1}^* \omega_{t+1}}{1-\omega_{t+1}}$
4	$N_{t+1}^* = \frac{P_{t+1}^*}{1-\omega_{t+1}}$	$\hat{N}_{t+1}^* = \frac{\hat{P}_{t+1}^*}{1-\omega_{t+1}}$
5	$\hat{I}_{t+1,s}^* = q_s ([1-\theta_s (1-\sigma)]N_{t+1}^* - \theta_s [(1-\sigma)C_{t+1}^*])$	$\hat{I}_{ts}^* = q_s ([1-\theta_s (1-\sigma)]\hat{N}_t^* - \theta_s [(1-\sigma)\hat{C}_t^*])$

¹ Where the subscript *s* denotes survey specific values. μ_t and \hat{C}_t^* for *S1* are calculated through Equations 3.4 and 3.5 respectively (conversely μ is a fixed constant scalar and C_t^* is not predicted in *S0*). ω_t is calculated through Equation 3.3.

Table 3.2 contains the same key equations within Schnute (1987) with the exception of the estimation of N_1^* in *S1*, which this study has introduced as a stand-alone parameter in the adapted model. This was done to allow the removal of the autoregressive process error and allow catch and surveys to be predicted for the first time step. In *S0*, when there are multiple survey indices, a geometric mean of biomasses calculated for each q_s is used for N_t^* . *S0* cycles through steps 1-5 at each time step *t* and the autoregressive process comes from the recalculation of biomass (step 1) from the observed survey index.

3.2.1.5 Likelihood

Schnute (1987) does not detail parameter estimation. Observed log-likelihood can be simply split into two key components: log-likelihood for the catch $\ln L_c$ (only relevant for *S1*) and log-likelihood for the survey indices $\ln L_I$. We assume that the observed catches C^* and observed survey indices I^* are log-normally distributed, with measurement error on catches with variance λ^2 and survey-specific error variances for indices δ_s^2 . Likelihood components are given as:

$$\ln L_c(\Psi|\mathbf{C}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \lambda^2 - \frac{1}{2\lambda^2} \sum_{t=T_1}^n (\ln C_t^* - \ln \hat{C}_t^*)^2 \quad (3.6)$$

$$\ln L_I(\Psi|\mathbf{I}) = \sum_{s=1}^G \left[-\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \delta_s^2 - \frac{1}{2\delta_s^2} \sum_{t=T1}^n (\ln I_{ts}^* - \ln \hat{I}_{ts}^*)^2 \right] \quad (3.7)$$

where G is the number of surveys and n is the number of years, and Ψ the full set of parameters. Total log likelihood is the sum of these components for $S1$ (where $T1=1$) and only the indices component (where $T1=2$) for $S0$, assuming independence in the model system.

3.2.1.6 Estimation in this study

Typically, growth parameters W and ρ are taken as inputs to the model (as suggested by Schnute (1987)), with no error associated with them, although in some circumstances they may be estimable. We estimated these parameters prior to running assessment models through estimation of linear models on mean weights in order to simplify model testing. Stage mean fish weights from sampling are also taken into the model without any error (i.e., assuming these are the true mean weights of the population).

We explore two growth/mean weight configurations for the Schnute model: 1) growth parameters were estimated with mean weights-at-age (i.e., akin to a Ford-Walford plot and detailed in *Appendix D*) and stage mean fish weights are taken from the survey (i.e., not skewed by selection) and 2) Stage mean fish weights are taken from the catch (i.e., potentially biased due to selection pattern) and growth parameters are estimated from the catch overall mean fish weights \bar{X}_t and previously exploited stage mean fish weights \bar{Z}_t (i.e., \bar{Z}_{t+1} vs. \bar{X}_t). When fitting these linear models prior to running the assessment models, residuals were assumed to be normally distributed.

This study implements both $S0$ and $S1$ with the two different growth/mean weight configurations (Table 3.3). These configurations aim to cover possible differences in the data available for an assessment.

Table 3.3: Summary of Schnute model and growth/mean weight configurations.

Notation	Model	Stage mean weights sampling	Growth parameter estimation
$S0_c$	Original Schnute process error model	catch	using mean fish weights of stages from catch
$S0_{waa}$	Original Schnute process error model	survey	using weights-at-age
$S1_c$	Adapted Schnute observation error model	catch	using mean fish weights of stages from catch
$S1_{waa}$	Adapted Schnute observation error model	survey	using weights-at-age

Survival σ (i.e., e^{-M}) was also fixed at the correct value for the stock. Pre-

analysis testing indicated models were not sensitive to starting values in estimated parameters. A table of starting values for key parameters in both the simulation framework as well as the real fish stock is given in *Appendix D*.

For simulated stocks *Original Schnute Process Error models* (S0) estimated two parameters (q_s and δ_s), whereas *Adapted Schnute Observation Error models* (S1) estimated either 23 or 43 parameters depending on the length of time series (N_1^*, q_s, δ_s and $20/40 F_t$). For the real stock, Schnute models ($S0_{waa}$ and $S1_{waa}$) were fit to three total biomass indices as well as catch biomass. $S1_{waa}$ estimated 23 parameters ($16 \hat{F}_t, 3 \hat{q}_s, 3 \hat{\delta}_s$ and \hat{N}_1^*). $S0_{waa}$ estimated six parameters ($3 \hat{q}_s$ and $3 \hat{\delta}_s$). All parameters were estimated in log-space, except q_s , which was logit transformed (based on the assumption $0 \leq q_s \leq 1$).

3.2.2 CSA

Our implementation of CSA is similar to that found in the NOAA Fish and Fisheries Toolbox (NFFT), with a few minor alterations (i.e., options for a stand-alone post-recruit index and estimating survey indices standard deviation). Notable differences between the current version of CSA and the original model described by Collie and Sissenwine (1983) are that the latter model does not consider catch error and catch does not occur continuously through the year but is taken instantaneously. The original model also has a process error component to the likelihood.

3.2.2.1 Population dynamics and notation

To project the population, the Baranov equation (Equation 3.5, here catch biomass and population biomass are replaced by catch numbers C_t and population numbers N_t) are used (Table 3.4). Instantaneous fishing mortality F_t and natural mortality rate M are constant across stages. Natural mortality rate does not vary through time, whereas fishing mortality is estimated at each time step. Population numbers N_t are subject to a total mortality ($F_t + M$) at each time step, which exponentially decreases to give post-recruit numbers at the next time step P_{t+1} (Table 3.5). Recruit numbers R_t are estimated at each time step as parameters and post-recruit numbers at the first time step P_1 is also estimated as a parameter to initialise population dynamics (Tables 3.4 and 3.5).

The recruit stage R_t is the group of fish that enter the fishery at or during the

current time step, the post-recruit stage P_t is comprised of fish recruited in previous years. Theoretically these stages are analogous to the recruits stage and previously exploited stage of the proposed model of Schnute (1987). Both stages are assumed to be fully available to the fishery in this implementation. Similarly, survey indices are assumed to have constant catchability between stages (knife-edged selection); however, there is an option (not used in our study) where this assumption can be relaxed and users can enter a proportion of recruits available to the recruit survey index. Stages can be defined by age, length or any other grouping that offers distinct stages that move from one to the other over a given time period.

Table 3.4: Summary of the CSA model notation. Hat operator denotes estimated parameters (for this study) or predicted values from the model. Parameters without hats were treated as known values.

Notation	Type	Description
General model		
N_t	output	Population numbers at time t
P_t	output	Post-recruit numbers at time t
M	parameter	natural mortality rate
\hat{q}_s	parameter	catchability coefficient of survey indices s
$\hat{\tau}_s$	parameter	standard deviation of log-normal survey numbers
ω	parameter	standard deviation of log-normal catch numbers
\hat{P}_1	parameter	initial numbers of post-recruits at $t = 1$
\hat{R}_t	parameter	Recruitment numbers at time t
\hat{F}_t	parameter	fishing mortality at time t
θ_s	parameter	timing of survey
h_{ts}	input	selectivity for recruit survey s
C_t	input	observed catch numbers at time t
I_{ts}^R	input	observed survey recruit numbers at time t for survey s
I_{ts}^P	input	observed survey post-recruit numbers at time t for survey s
I_{ts}	input	observed survey population numbers at time t for survey s
\hat{C}_t	output	predicted catch numbers at time t
\hat{I}_{ts}^R	output	predicted survey recruit numbers at time t for survey s
\hat{I}_{ts}^P	output	predicted survey post-recruit numbers at time t for survey s
\hat{I}_{ts}	output	predicted survey population numbers at time t for survey s
s	index element	survey number
t	index element	number of year

Surveys can be undivided surveys (i.e., catch both stages) or split into a recruit index and a post-recruit index or a stand-alone post-recruit survey index (a new addition to NFFT implementation). Survey catchability coefficients q_s are survey specific (not index specific), thus for split surveys, one coefficient is estimated across the recruit and post-recruit indices. Similarly, standard deviations are survey specific, not index specific. At least one split survey is required for a CSA model run (Table 3.4).

The GUI (Graphical User Interface) provided for CSA does not allow for estimating survey standard deviation; however, the version implemented in this study allows for estimation of this parameter. Another adaption here is that stand-alone post-recruit survey indices can be input allowing some flexibility in indices used (NOAA, 2019).

3.2.2.2 Deterministic equations and model summary

The CSA GUI helpfiles contain the underlying code and additional information that was used to implement CSA (NOAA, 2019). Key dynamic equations and model steps in mathematical notation are summarised in Table 3.5. The model was an observation-error-only model and parameters were estimated by maximum likelihood.

Table 3.5: Summary of the steps within the CSA model. Hat operator denotes predicted values from the model that are found in the likelihood.

	Step
1	\hat{P}_1 is a key parameter estimated within the model
2	$N_t = P_t + \hat{R}_t$
3	$\hat{C}_t = \hat{F}_t \left(\frac{1 - e^{-M - \hat{F}_t}}{M + \hat{F}_t} \right) N_t$
4	$P_{t+1} = N_t e^{-M - \hat{F}_t}$
5	$\hat{I}_{ts}^R = (\hat{q}_s \hat{R}_t h_{ts}) e^{\theta_s(-M - \hat{F}_t)}$
6	$\hat{I}_{ts}^P = (\hat{q}_s P_t) e^{\theta_s(-M - \hat{F}_t)}$
7	$\hat{I}_{ts} = (\hat{q}_s N_t) e^{\theta_s(-M - \hat{F}_t)}$

3.2.2.3 Likelihood

The observed log-likelihood is comprised of two key components: log-likelihood for the catch $\ln L_c$, and log-likelihood for the survey indices $\ln L_I$. We assume that the observed catches C and observed survey indices I, I^R and I^P are log-normally distributed, with measurement error on catches with variance ω^2 and survey-specific measurement error variances for indices τ_s^2 . Likelihood equations for catch numbers and survey indices were the same as those used in the Schnute model (Equations 3.6 and 3.7), where G is the number of surveys and n is the number of years, and Ψ the full set of parameters. The notation

I_{ts} represents any survey index. The total log-likelihood was the sum of these components, assuming independence in the model system.

3.2.2.4 Estimation for this study

Natural mortality rate was a fixed parameter and was at the correct value for the stock. Recruit survey index selectivity h_{ts} is assumed to be 1 (i.e., recruits fully selected in survey) for all CSA models in our study. Starting values for key parameters for both the simulated stocks and the real fish stock are found in Table 2 in *Appendix D*. Pre-analysis testing indicated CSA was not sensitive to starting values in estimated parameters.

For simulated stocks, CSA estimated either 43 or 83 parameters depending on the length of time series ($\hat{q}_s, \hat{\tau}_s, \hat{P}_1, 20/40 \hat{F}_t$ and $20/40 \hat{R}_t$). For the real stock, CSA estimated 39 parameters (of which 16 \hat{F}_t , 16 \hat{R}_t , 3 \hat{q}_s , 3 $\hat{\tau}_s$ and \hat{P}_1) and was fit to one recruit numbers index and three post-recruit numbers indices as well as catch numbers. All parameters were estimated in log-space, except \hat{q}_s , which was logit transformed (based on the assumption $0 \leq q_s \leq 1$).

3.2.3 Data

Self-tests were used initially to check consistency and implementation. To test model capabilities a simulation framework was developed. A real fish stock was then used to test models in a real world scenario.

3.2.3.1 Simulation framework

Simulated stocks were used to test the assessment models in different data scenarios. Stocks were age-based, with explicit modelling of age-structured population dynamics and fisheries selectivity. The population was projected forward using a stock-recruit model for incoming recruitment in the following year and all other age class subject to exponential mortality (Quinn and Deriso, 1999). We developed a simulation framework based on Rosenberg et al. (2014) using the “FLife” package in R (Kell, 2018) to simulate fish stocks and generate data under a variety of conditions (Kell et al., 2017). A quasi-factorial simulation framework included three harvest dynamics scenarios, three selectivity patterns, two life-history strategies, two time-series lengths, two recruitment variability scenarios and two recruitment autocorrelation scenarios (Table 3.6). FLife was used to create 144 unique age-structured stock scenarios with 100 replicates for each of the 144 scenarios. Variability among replicates

came from a combination of recruitment variability, catch error and survey index error. Stocks were created at equilibrium when fishing mortality was zero and there was no initial depletion before applying the harvest dynamics (HD) scenario. The stock-recruit model used was a Beverton-Holt model parameterised with virgin biomass and steepness of the life-history strategy (Equations in *Appendix D*).

Assessment models were tested across two contrasting life-history strategies. These were a) small pelagic fish — similar to Atlantic herring *Clupea harengus harengus* and b) large demersal fish — similar to white-bellied anglerfish *Lophius piscatorius* (Table 3.6). The latter life-history strategy was chosen as this was the species of the case study (ICES, 2019b,a). The small pelagic life-history strategy was chosen for contrast. Details of parameters used for each life-history can be found in *Appendix D*.

The harvest dynamics (HD) scenarios were chosen to test the assessment model’s ability to model stocks with different fisheries contrasts. HD scenarios were: 1) constant (“c”), 2) one-way trip (“ow”) and 3) roller-coaster (“rc”) (Table 3.6). Details of of these scenarios can be found in *Appendix D*.

Table 3.6: Simulation experimental design factor levels.

Factor	Factor label	Level 1	Level 2	Level 3
Harvest dynamics	HD	Constant harvest rate (c)	One way trip (ow)	Roller-coaster (rc)
Time-series length	TS	20	40	-
life-history	LH	small pelagic	large demersal	
Selectivity	SEL	constant/knife-edged	logistic	dome-shaped
Recruitment error	SR	0.1	0.4	
Recruitment autocorrelation	AR	0	0.6	

All assessment models tested assume a knife-edged selection pattern. To test how these models perform under a different selectivity pattern, simulation scenarios were set up with three different selectivity patterns: 1) constant/knife-edged, 2) logistic and 3) dome-shaped (Table 3.6). Knife-edged selection pattern assumed that all ages of the fish stock were fully selected in the fishery. Exact values for the other selectivity patterns differed between life-history strategies. Details of logistic and dome-shaped selectivity patterns used in the simulated stocks can be found in *Appendix D*.

Time-series lengths varied between 20 and 40. Recruitment error was random log-normal noise around the S-R function, set at either relatively low (sd = 0.1) or relatively high (sd = 0.4). Recruitment of the stock was either set up with no autocorrelation or high autocorrelation (rho = 0.6) (Table 3.6).

A fully selected survey index of numbers-at-age was generated for each stock as the product of stock numbers at the beginning of each year, a catchability coefficient of $1E - 6$ and a random log-normal error ($sd = 0.3$) on numbers-at-age. A random log-normal error was also applied to catch numbers-at-age ($sd = 0.1$ for all ages).

For simplicity, stages for the assessment models were defined by age. It was assumed that fish at age zero were in $[V, V')$ (i.e., recruit stage) and fish of ages 1+ were in $[V', V^\infty)$ (i.e., previously exploited/post-recruit stage). This implies that there was no significant overlap in lengths between age zero and age 1.

3.2.3.2 Real anglerfish stock

Data for the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d was collated from the 2019 ICES assessment (ICES, 2019b,a). This consisted of catch data and three survey indices that cover the anglerfish stock: a combined IBTS index of the French EVHOE survey (Q4) and Irish Groundfish survey (IE-IGFS, Q4) spanning 2003-2018, Spanish Porcupine Groundfish survey (SP-PORC, Q3/Q4) spanning 2003-2018, Irish Anglerfish and Megrim Survey (IAMS, Q1) in years 2007-2008 and 2016-2018. The SP-PORC survey has a relatively small spatial coverage compared to the other surveys. The stock is currently assessed with an age-based analytical assessment using the a4a Stock Assessment Modelling framework (Jardim et al., 2014; ICES, 2019b) and was used for comparison with the two-stage models implemented here. The purpose of comparison for the case study was not to investigate the potential bias of the age-based a4a assessment model but to explore whether stage-based assessment models with less data requirements were capable of giving comparable results to the age-based assessment model. Natural mortality rate was fixed at 0.25 for all model runs, as this was the value used for all ages for the official a4a assessment. Similar to the simulated stocks, stages for the assessment models were defined by age, with the same assumptions as above. A key assumption was that the procedure used to compile age-composition data for the age-based assessment identifies and defines recruits without bias (age 0 fish in catch and surveys).

In the official assessment model that is used for comparison, the number of ages to which each survey index is tuned to is specific, i.e., ages 0-2 for IBTS combined survey, 1-5 for IAMS and 2-6 for SP-PORC. Schnute and CSA

models did not allow for this subtlety. Four survey indices were used for the CSA model; both recruit and post-recruit indices from the IBTS combined survey as well as post-recruit indices of IAMS and SP-PORC. Schnute models used the three surveys (combined IBTS, IAMS and SP-PORC) as biomass survey indices.

3.2.4 Model performance

To compare between the Schnute and CSA approaches, stock numbers were used to compare across the different models. As Schnute is a biomass-based model with a time-series of stage mean fish weights, stock numbers were calculated by dividing the estimated biomass by the corresponding mean weight at a given time step. For comparison (for the simulated stocks and age-based assessment of anglerfish) overall fishing mortality F was calculated from fishing mortality-at-age as an average F weighted by population numbers-at-age. Convergence of assessment models was determined in two steps. After convergence was indicated by the optimiser, parameters were checked for identifiability with the second-derivatives of the marginal likelihood (Thorson, 2020). Only model runs where all parameters were identifiable were used in the results.

For the simulation framework, performance metrics were applied to key outputs from the assessment models (i.e., fishing mortality at the start and end and stock numbers). Performance of assessment models was evaluated using Relative Error ($RE = (\hat{\theta} - \theta)/\theta$) to assess bias and Relative Standard Error ($RSE = \sigma_{\hat{\theta}}/\hat{\theta}$) to assess precision ($\hat{\theta}$ is the estimated value from an assessment model, θ is the true value and $\sigma_{\hat{\theta}}$ is the standard error of $\hat{\theta}$). Standard error was the approximated asymptotic SE estimated within TMB via the delta-method. Medians were calculated for relative errors giving performance metrics MRE_N for stock numbers in all years and MRE_{F1} or MRE_{Fend} for fishing mortality at the start and end of a time-series respectively. Median Absolute Relative error (MARE) was also calculated for estimated stock numbers in the final year $MARE_{Nend}$ and estimated fishing mortality in the final year $MARE_{Fend}$. These values at the end of the time series were chosen as they would likely have the most uncertainty as well as arguably being the most important predictions from an assessment model. MARE accounts for both bias and variability and can be used to identify the best overall model by finding the model configuration that minimised the maximum MARE (min-max solution)

across scenarios (Johnson et al., 2015).

3.2.4.1 Self-tests

Self-tests were used to check the consistency of assessment models on a simulated stock scenario and white-bellied anglerfish under the structural assumptions and conditions of the models (Deroba et al., 2015). This process involved fitting the models and computing or simulating model predictions (100 replicates) of survey indices and catch (if relevant) from the models directly, with the same settings and estimated parameters from the original fit. Models were then fit to the simulated data (Nielsen and Berg, 2014; Cadigan, 2015; Deroba et al., 2015).

The simulated stock used was a replicate from the simulation framework (life-history = large demersal, harvest dynamics = rollercoaster, time series length = 40, selection pattern = knife-edged, recruitment error = 0.1 and autocorrelation = 0). In this case, self-tests were conducted with (same error on survey and catch as simulation framework) and without error on observations.

3.2.5 Software

Both models in this study were implemented using the R package "TMB" (Kristensen et al., 2016). Specifically, complete data log-likelihood functions were written in C++ using the TMB library class, compiled and automatically differentiated (Kristensen et al., 2016). All other analysis as well as optimisation (using `nlminb`) was conducted within the statistical software R (R Core Team, 2019). R package "TMBhelper" was used to check identifiability of parameters of models after optimisation for the real fish stock (Thorson, 2020). As mentioned previously, the simulated stock was created with the "FLife" R package (Kell, 2018) and code for the simulations that is loosely based on Mosqueira (2016) can be found at : https://github.com/lbatts/sim_sbar. An R package was also developed for these assessment models (as well as other forms of the Schnute model) and can be found at: <https://github.com/lbatts/sbar>.

3.3 Results

Results are first presented for self-tests on a simulated stock. Followed by the five assessment models (the two forms of Schnute (1987) with two growth/mean weight configurations and CSA) on the simulated stocks. Stock assessment models on the white-bellied anglerfish stock are then presented.

3.3.1 Self-tests

Self-tests demonstrated that CSA and *Adapted Schnute Observation Error model* $S1_c$ were self-consistent and able to estimate parameters without bias from data simulated under the same conditional assumptions as the models with or without observation error. In contrast, the *Original Schnute Process Error model* $S0_c$ showed a small degree of bias (mainly at the latter end of the time series) when there was no observation error but showed considerable bias when observation error was present (Figure 3.1).

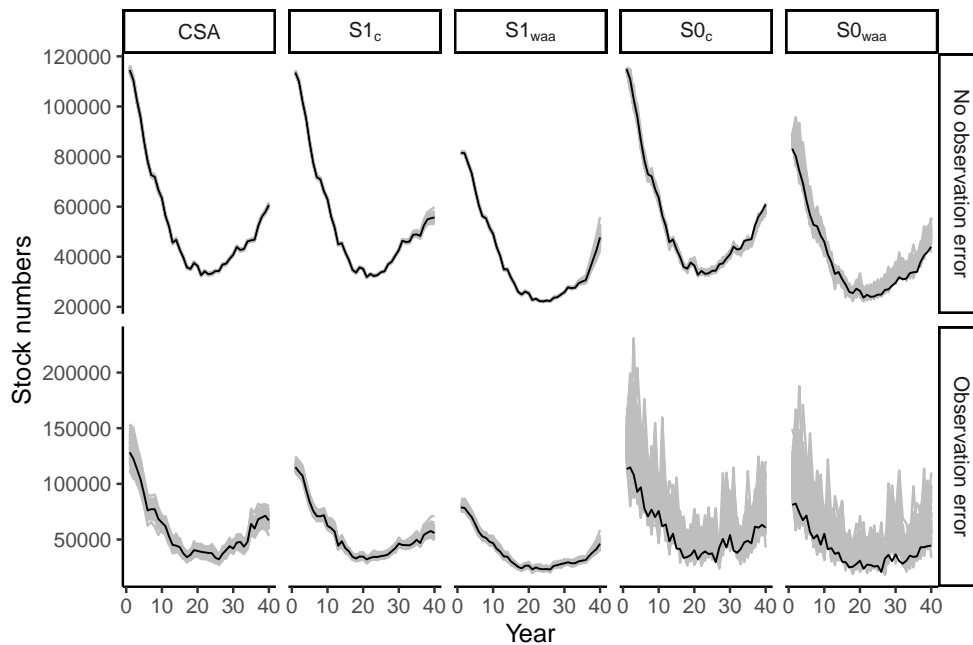


Figure 3.1: Estimated population numbers (\hat{N}) and “self-test” replicates for three assessment models; CSA, $S1_c$ and $S0_c$ on a simulated stock replicate (LH = large demersal, HD = rc, TS = 40, SEL = knife-edged, SR = 0.1 and AR = 0) with or without error on observations. Estimate from original data (bold black line) and re-estimated stock numbers from conditional simulations (thin grey lines).

3.3.2 Simulation framework

Across all assessment models, 73% (52961) of runs converged and had identifiable estimated parameters. 15% (10713 runs) converged but at least one parameter was not identifiable and 12% (8326 runs) did not converge. The most important factor causing unidentifiable parameters or non-convergence was selectivity pattern (Table 3.7). Across all other factors of the simulation framework, the percentages of model runs in each category were roughly equal (Table 1 in *Appendix D*). All assessment models (with the exception of a small number of runs for $S1_{waa}$) converged with identifiable parameters on stocks created with knife-edged selection (Table 3.7). In addition, CSA, $S0_{waa}$, and $S1_{waa}$ assessment models also converged reliably across the other selectivity patterns (Table 3.7). Models $S0_c$ and $S1_c$ consistently did not converge with identifiable parameters on stocks with logistic or dome-shaped selectivity patterns.

Table 3.7: Percentage (to two decimal points) of model runs across the five assessment models and three selectivity patterns that: converged, did not converge, or converged but with unidentifiable parameters.

Assessment model	Classification	Selectivity		
		knife-edged	logistic	dome
CSA	All parameters are identifiable	100.00	99.90	99.96
	Converged but not all parameters identifiable		0.04	0.02
	Did not converge		0.06	0.02
$S1_c$	All parameters are identifiable	100.00	1.42	2.17
	Converged but not all parameters identifiable		11.38	12.98
	Did not converge		87.21	84.85
$S1_{waa}$	All parameters are identifiable	99.85	99.83	99.92
	Converged but not all parameters identifiable			
	Did not converge	0.15	0.17	0.08
$S0_c$	All parameters are identifiable	100.00	0.10	0.21
	Converged but not all parameters identifiable		99.38	99.40
	Did not converge		0.52	0.40
$S0_{waa}$	All parameters are identifiable	100.00	100.00	100.00
	Converged but not all parameters identifiable			
	Did not converge			

Performance of assessment models was variable across factors and between models. Generally, assessment models performed similarly across factor levels for time series length, recruitment error and recruitment autocorrelation. Selection pattern, life-history strategy and harvest dynamics, as well as the assessment model and its configuration were the primary drivers of differences

in the results on simulated stocks (Figures 3.2, 3.3, 3.4, 3.5 and Table 3.8, *Appendix D*). For this reason, the following figures focus on one recruitment error, recruitment autocorrelation and times series length combination, figures showing the full set of scenarios are in *Appendix D*.

Table 3.8: Assessment model performance (model runs where all parameters were identifiable) in terms of the metric Median Absolute Relative Error (MARE) of estimated fishing mortality in the final year $MARE_{F_{end}}$ (F) and estimated stock numbers in the final year $MARE_{N_{end}}$ (N). Highlighted cells are the min-max solution (i.e., the minimum maximum MARE across assessment models indicating the model which would be best placed to reduce bias and variability in different situations) for each life-history scenario by selectivity scenario block. Min-max solutions for $MARE_{N_{end}}$ across all selectivity patterns for each life-history strategy are models highlighted in bold (only comparing models with values across all scenarios).

LH	Assessment model	Knife-edged						Logistic						Dome						
		Constant		One-way		Roller		Constant		One-way		Roller		Constant		One-way		Roller		
		F	N	F	N	F	N	F	N	F	N	F	N	F	N	F	N	F	N	
large demersal	CSA	0.22	0.22	0.17	0.14	0.12	0.11	0.27	0.23	0.36	0.14	0.29	0.12	0.22	0.22	0.24	0.16	0.16	0.12	
	S1 _c	0.10	0.09	0.08	0.08	0.08	0.08					0.97	17.37					0.95	16.55	
	S1 _{waa}	0.98	0.48	0.70	0.34	0.31	0.22	0.88	0.43	0.63	0.29	0.27	0.17	1.14	0.46	0.99	0.33	0.55	0.20	
	S0 _c		0.15		0.11		0.10												27.31	44.24
	S0 _{waa}		0.43		0.23		0.22		0.37		0.17		0.15		0.41		0.20			0.18
small pelagic	CSA	0.25	0.23	0.23	0.19	0.20	0.17	0.61	0.33	0.64	0.22	0.62	0.18	0.50	0.30	0.53	0.21	0.51	0.18	
	S1 _c	0.11	0.09	0.13	0.11	0.11	0.10													
	S1_{waa}	0.15	0.12	0.17	0.12	0.11	0.10	0.28	0.11	0.45	0.13	0.43	0.14	0.14	0.10	0.30	0.12	0.30	0.15	
	S0 _c		0.30		0.25		0.27													
	S0 _{waa}		0.16		0.19		0.22		0.42		0.44		0.45		0.38		0.41		0.45	

Harvest dynamics had a noticeable effect on each assessment model’s performance (Figures 3.2, 3.3, 3.4, 3.5 and Table 3.8). MARE, MRE, and MRSE were generally lower (i.e., less bias and more precision) in one-way trip (HD: ow) and roller-coaster (HD: rc) scenarios than for constant harvest dynamics (HD: c) (Figures 3.4, 3.5 and Table 3.8). Trends in median stock numbers and fishing mortality from assessment models also moved closer to median real values for stocks with harvest dynamic scenarios with some fisheries contrast (HD: ow and HD: rc) (Figures 3.2, 3.3, 3.4 and 3.5).

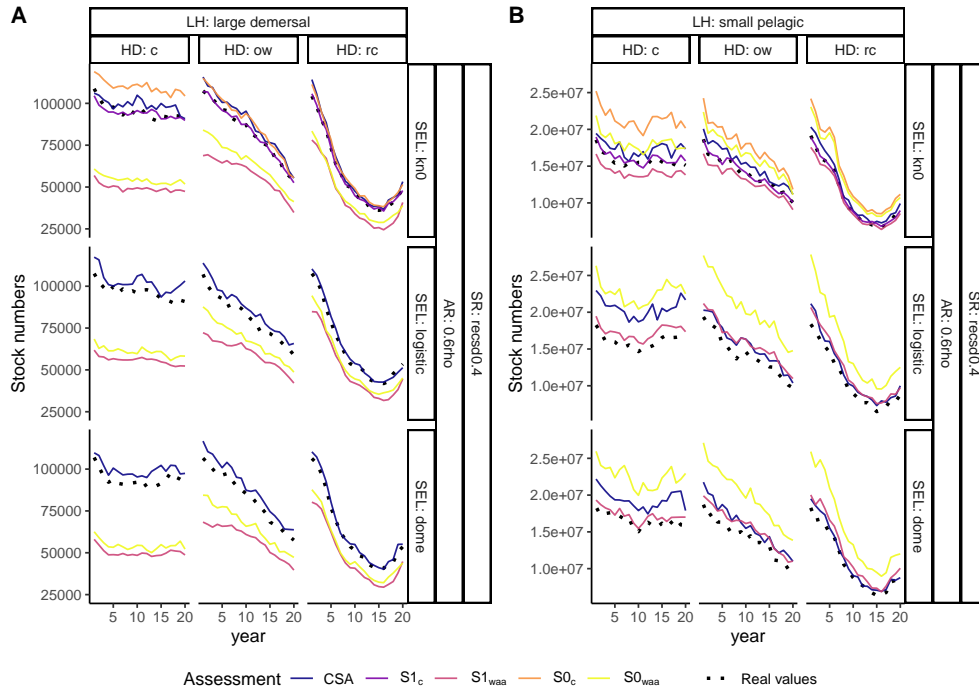


Figure 3.2: Median estimated stock numbers, as well the corresponding median real values for large demersal (A) and small pelagic(B) life-history strategies. Simulated stock scenarios with high recruitment error (SR:recsd0.4) and autocorrelation (AR:0.6rho) across all selectivity patterns and harvest dynamics are shown. Estimates from five different assessment models (CSA, $S1_c$, $S1_{waa}$, $S0_c$ and $S0_{waa}$) are given. As the majority of $S1_c$ and $S0_c$ runs did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios, these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

For knife-edged selection scenarios, $S1_c$ was the min-max solution for both life-histories. $S1_c$ consistently had minimal bias, as well as high precision in stock numbers and fishing mortality estimates relative to other models across the other factor levels (Figures 3.2, 3.3, 3.4, 3.5 and Table 3.8, *Appendix D*). CSA was similarly accurate with MREs close to zero but precision was lower (higher MRSE) and MARE values were higher.

For the small pelagic life-history, CSA and Schnute assessment models set up the with ‘waa’ growth/mean weight configuration ($S0_{waa}$ and $S1_{waa}$) performed largely the same across selectivity scenarios, giving similar MREs across the factor levels and estimating similar median stock numbers (Figures 3.2, 3.3, 3.4, 3.5 and Table 3.8, *Appendix D*). These models were the min-max solutions for $MARE_{Nend}$ and $MARE_{Fend}$ in logistic and dome-shaped selectivity

scenarios with small pelagic life-history (Table 3.8).

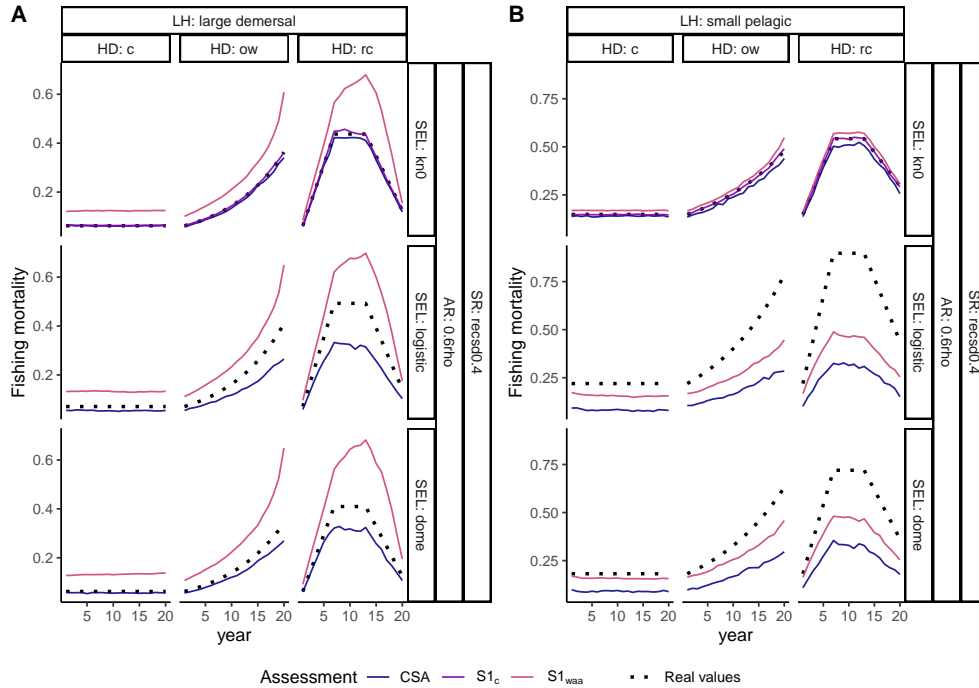


Figure 3.3: Median estimated fishing mortality, as well the corresponding median real values for large demersal (A) and small pelagic(B) life-history strategies. Simulated stock scenarios with high recruitment error (SR:recsd0.4) and autocorrelation (AR:0.6rho) across all selectivity patterns and harvest dynamics are shown. Estimates from three different assessment models (CSA, $S1_c$ and $S1_{waa}$) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

Schnute models with ‘waa’ growth/mean weight configuration ($S0_{waa}$ and $S1_{waa}$) did not perform well (relatively high MARE and MRE) in the large-demersal scenarios, but performed much better in the small pelagic scenarios, particularly $S1_{waa}$ (Figures 3.4, 3.5 and Table 3.8). CSA was the only assessment model to have relatively low MARE and MRE, as well as median estimated values close to the median real values across both life-history strategies and all selectivity patterns, as well as other factors. CSA was only the min-max solution in large-demersal logistic or dome-shaped selection scenarios (Figures 3.2, 3.3, 3.4, 3.5 and Table 3.8, *Appendix D*).

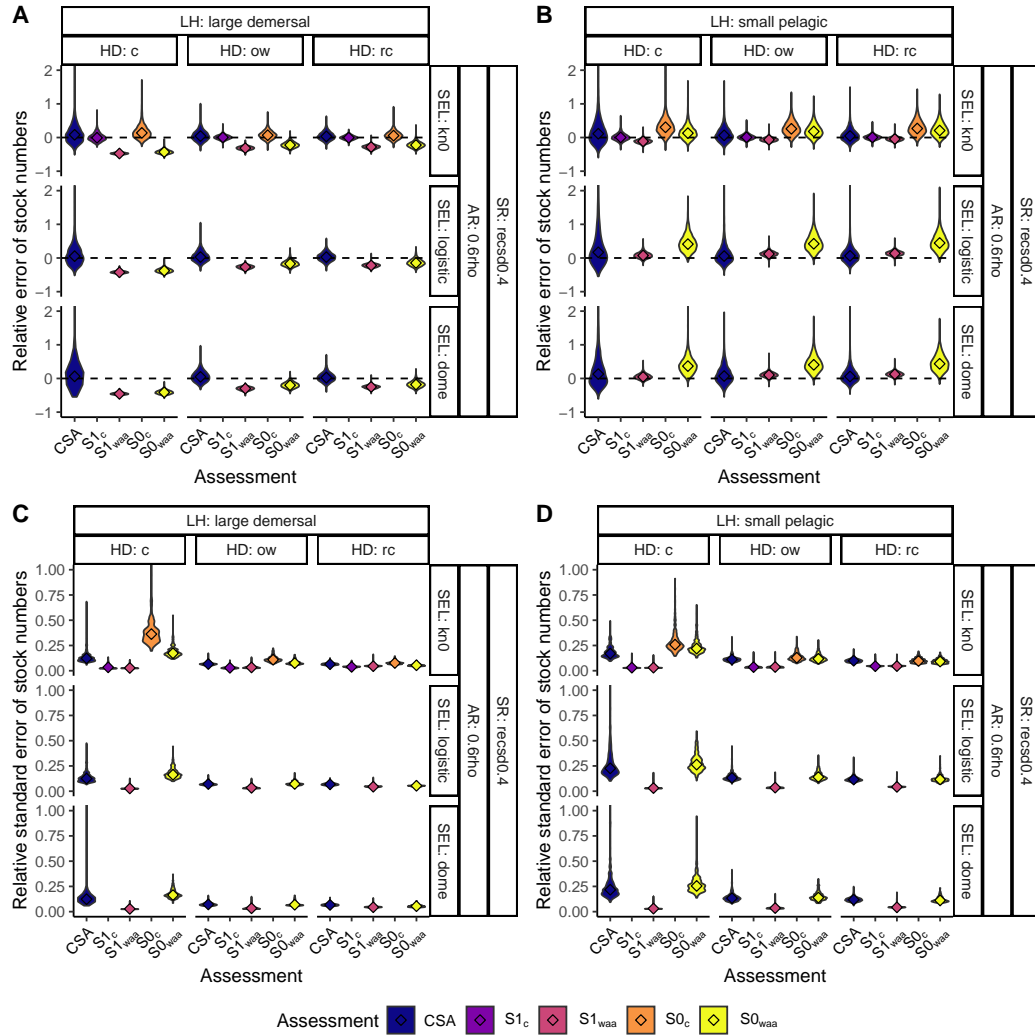


Figure 3.4: Distribution of relative error (A, B) and relative standard error (C, D) of the stock numbers in the current year for large demersal (A, C) and small pelagic (B, D) life-history strategies. Simulated stock scenarios with high recruitment error (SR:recsd0.4) and autocorrelation (AR:0.6rho) across all selectivity patterns and harvest dynamics are shown. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from five different assessment models (CSA, $S1_c$, $S1_{waa}$, $S0_c$ and $S0_{waa}$) are given. As the majority of $S1_c$ and $S0_c$ runs did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios, these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

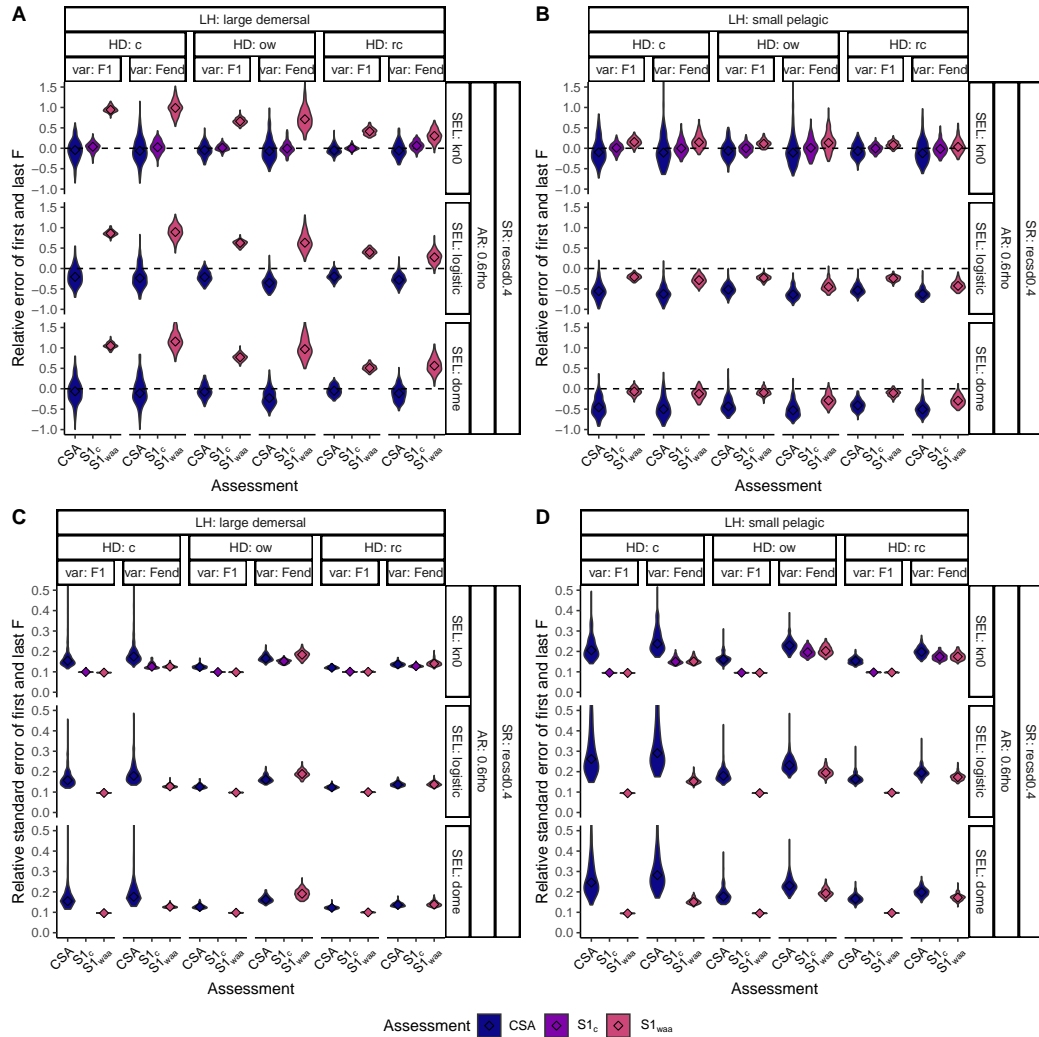


Figure 3.5: Distribution of relative error (A, B) and relative standard error (C, D) of fishing mortality at the start (F1) and end (Fend) of a time series for large demersal (A, C) and small pelagic (B, D) life-history strategies. Simulated stock scenarios with high recruitment error (SR:recsd0.4) and autocorrelation (AR:0.6rho) across all selectivity patterns and harvest dynamics are shown. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from three different assessment models (CSA, $S1_c$ and $S1_{waa}$) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

3.3.3 Real white-bellied anglerfish stock

Model runs for assessment models using the stage mean fish weights and growth parameters derived from the catch ($S0_c$ and $S1_c$) did not have a stable convergence and were not included in the following results. Self-tests for the three assessment models that did converge demonstrated that CSA and *Adapted Schnute Observation Error model* $S1_{waa}$ were self-consistent and able to estimate parameters without bias from data simulated under the same conditional assumptions as the models. In contrast, the *Original Schnute Process Error model* $S0_{waa}$ showed considerable bias in these simulations (*Appendix D*).

CSA fit the survey data relatively well (Figure 3.6B). Overall, the recruit index and post-recruit index from the combined IBTS survey had the closest fit between observations and the predicted values. The model fit for the Spanish Porcupine Bank survey (SP-PORC) index for post-recruits was reasonably close to observations at the start of the time series and then departed markedly at a couple of points near the end of the time series (Figure 3.6). Schnute model fits to survey data differed somewhat between the models. $S1_{waa}$ fit closely to the combined IBTS survey index for the whole time-series, whereas $S0_{waa}$ had a closer fit to the Spanish Porcupine Bank survey (SP-PORC) index for the entire time series (Figure 3.6A).

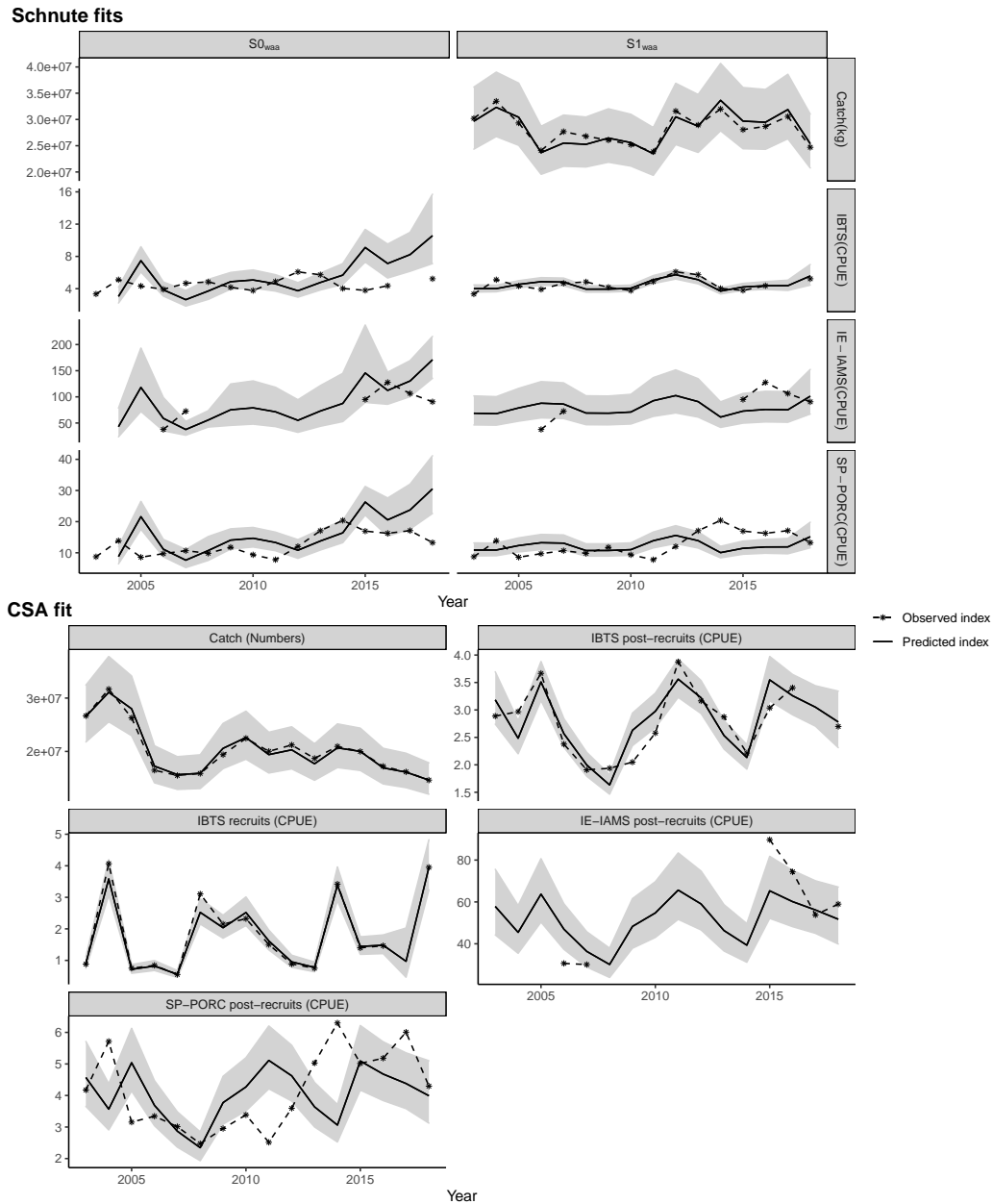


Figure 3.6: Observations and predicted values for survey indices ($S0_{waa}$ and $S1_{waa}$ survey index [kg. unit effort], CSA survey index [numbers of fish unit effort]) and entire catch (biomass (kg) or numbers respectively) for the three assessment models (('Schnute fits' panel) $S0_{waa}$, $S1_{waa}$ and ('CSA fit' panel) CSA) on the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d. Surveys for CSA are both a recruit and post-recruit indices for the combined IBTS survey index (comprising of the Irish groundfish and French EVHOE surveys), a post-recruit index for the Irish monkfish and megrim survey (IE-IAMS) and a post-recruit index for the Spanish Porcupine bank survey (SP-PORC). Surveys for $S0_{waa}$ and $S1_{waa}$ are entire biomass surveys. Shaded area denotes ± 2 asymptotic SE of the predicted values.

Estimates for the entire population, post-recruit and recruit numbers from the CSA assessment model were similar to estimates from the age-based assessment model (Figure 3.7A). Fishing mortality levels and overall trend were similar between the two assessment models, but CSA estimated overall fishing mortality somewhat lower than the age-based model (Figure 3.7B).

Estimates for entire population and previously exploited numbers for $S0_{waa}$ and $S1_{waa}$ were reasonably close to the population estimates from the age-based assessment model and trends were very similar. Recruit numbers were notably closer to the age-based assessment model estimates. Uncertainty around those estimates from $S1_{waa}$ was relatively low compared to CSA, but $S0_{waa}$ estimates were less precise (Figure 3.7A). Fishing mortality levels estimated by model $S1_{waa}$ were not close to those estimated by the age-based assessment model and did not appear to follow the same trend (Figure 3.7B).

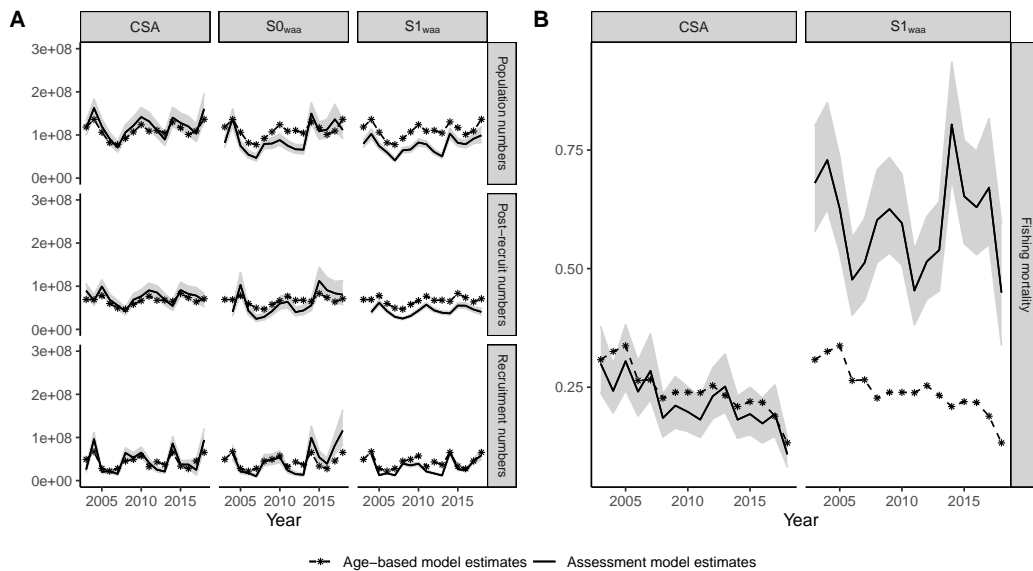


Figure 3.7: Estimated time series for population numbers (\hat{N}), post-recruit numbers (\hat{P}), recruitment numbers (\hat{R}) (Figure A) and fishing mortality (\hat{F}) (Figure B) for three assessment models ($S0_{waa}$, $S1_{waa}$ and CSA) on the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d. Note $S0_{waa}$ does not estimate fishing mortality so is not included in Figure B. Shaded area denotes ± 2 asymptotic SE of the predicted values. Full range of estimates and associated error are not given for $S0_{waa}$ to allow comparison.

3.4 Discussion

We have implemented two different approaches to stage-based stock assessment modelling in a common estimation framework, applied them within a simulation-testing framework and on a real fish stock, and then compared their performance. The first approach was a biomass-based size structured delay-difference model based on the theoretical paper by Schnute (1987), with some alternative adaptations (e.g. observation error instead of process error) and configurations ($S1_c$, $S1_{waa}$, $S0_c$ and $S0_{waa}$). The second was a version of the current CSA model (NOAA Fish and Fisheries Toolbox) with some adjustments (i.e., ability to have stand-alone post-recruit indices) (NOAA, 2019).

Differences in the performance of models on simulated stocks were largely dependent on harvest dynamics, life-history strategy and model assumption violations in selectivity patterns. CSA performed consistently (i.e., converged with relatively low bias on estimates) across all scenarios, although uncertainty was inherently higher than in Schnute models. On the other hand, CSA was the min-max solution for a minority of the compared scenarios and *Adapted Schnute Observation Error models* ($S1$) performed similarly or outperformed CSA in many scenarios. Results were also mixed for the real white-bellied anglerfish stock, reflecting the complexity of a real stock. However, despite their relative simplicity, the *Adapted Schnute Observation Error model* ($S1_{waa}$) estimated population numbers were reasonably close to a4a assessment model estimates and CSA estimated population numbers (and to a lesser extent fishing mortality) were remarkably close to a4a assessment model estimates.

3.4.1 Model performance

Self-tests, followed by simulation-testing showed that both CSA and *Adapted Schnute Observation Error model* $S1$ were robust assessment models that had the ability to fairly accurately estimate population numbers and to a lesser extent fishing mortality under a variety of different data scenarios, including violations of model assumptions in selectivity pattern of the fishery. Schnute model convergence and accuracy was dependent on the growth/mean weight configuration, and how this matched the selectivity and/or life-history of a stock. In contrast, CSA consistently converged and was consistently accurate across all scenarios but lacked precision, which may have prevented the model being the min-max solution more frequently. This was more apparent in

the small-pelagic scenarios, where CSA may have been more sensitive to observation errors in the recruit survey index.

CSA is a well-known stage-based stock assessment model that has proven to be particularly effective in data-limited situations such as where species are difficult to age and our results support these previous findings (Conser, 1995; Collie and Kruse, 1998; Cadrin, 2000; Cook and Heath, 2018; Li et al., 2019). Mesnil (2003) also found in their sensitivity analysis of a prior version of CSA that the model performed remarkably well, given the limited data. Their study also highlighted issues with correctly identifying the recruit stage and variability in survey catchability of discrete groups of fish (i.e., different catchabilities at each stage) (Mesnil, 2003), which could also be an issue for this study's implementation as well.

Given CSA's accurate estimation and $S1_{waa}$'s underestimation of stock numbers in large demersal life-history scenarios with logistic or dome-shaped selectivity patterns, it was somewhat expected that the assessment models would also perform as they did on the real fish stock. CSA estimated population numbers close to estimates from the official assessment model and $S1_{waa}$ estimates were reasonably close. Fishing mortality was less accurate for both models but this was likely due to the selectivity of the fishery, which has been estimated as dome-shaped within the official age-based assessment model (ICES, 2018). Closeness of CSA and $S1_{waa}$ population numbers estimates to the official assessment model estimates is notable given the reduced data, conflicting signals from the survey indices and that at least some surveys indices are unlikely to be fully selected.

Self-tests for both a simple simulated stock scenario and for white-bellied anglerfish indicated that our implementation of the *Original Schnute Process Error model S0* was inherently biased and not a reliable or consistent assessment model, despite the model sometimes estimating close to expected values. Divergence in self-tests is not uncommon in case studies (Deroba et al., 2015); however, the extent of inconsistency under low levels of observation error and for the anglerfish case study is cause for concern. Furthermore, when no observation error was present, there was still a small but obvious divergence at the latter end of the times series. A comprehensive check of code was the first step in these circumstances but no issues could be found, indicating both structural uncertainty, as well as observation error may be causing the inconsistency (Deroba et al., 2015). Further checks with data generated from

a separately coded Schnute population dynamics model are also needed for further clarity.

3.4.2 Model developments

An interesting aspect of the approach described in Schnute (1987) is that the user may adjust the model in accordance with the relative importance of recruitment (not tested here) or previously exploited biomass (Schnute models in our study) or both (also not tested) in the population dynamics of the stock. Our study focussed on the method where future biomass is calculated from previously exploited biomass because it reduced estimated parameters (no recruitment parameters) and for practical reasons (pre-testing indicated recruitment parameters were difficult to estimate without optimisation with a genetic algorithm). The benefit of this flexibility was demonstrated in the results, as Schnute models (in certain scenarios) were able to model the stocks accurately without estimating recruitment parameters or assumptions on the spawning proportion of the stock. Of course, the implication is that population estimates are largely dependent on the relationship between stage mean fish weights, which are currently assumed without error within the model.

Recruitment is still calculated within the model, so it can be extracted post-assessment and fit with a suitable S-R model, rather than attempting to fitting it internally.

Since individual recruitment parameters (\widehat{R}_t) are not estimated in the Schnute models, this greatly reduces the number of parameters of the model compared to CSA or other more complex assessment models. Furthermore, given only total biomass from an annual survey is needed for the model, it lends itself to assessment situations that are more data-limited. The theory behind the Schnute model could be extended further still by introducing recruitment or previously exploited biomass survey indices into the likelihood and the possibility of using only stage-specific survey indices to estimate entire population biomass.

An initial biomass parameter and a switch to an observation error model with additional fishing mortality estimation were the adaptations implemented in this study to the original “composite model” found in Schnute (1987). These additions appear to have worked well for the *Adapted Schnute Observation Error models* (S1) enabling accurate estimation of the simulated stocks with

relatively low uncertainty. Estimating fishing mortality was also an important addition, as it is a useful output from stock assessment models that aids fisheries management (Hilborn and Walters, 1992).

CSA is highly parameterised in comparison to Schnute models (due to CSA estimating \hat{R}_t) and requires more information/data to set up the stock assessment model (i.e., at least one survey needs to be split into numbers of recruits and post-recruits). One minor adaptation of this study's implementation of CSA was that post-recruit surveys can stand alone and do not need a recruit survey as well. This meant that the CSA set up was reasonably similar to the configuration of the age-based assessment model with its age-range-tailored surveys. The nuance of CSA that allows differing trends in recruits and post recruits gives more flexibility within the model, hence the similar fit to the age-based assessment model on the white-bellied anglerfish stock.

3.4.3 Growth parameter estimation and stage mean fish weights configuration

Growth parameters (ρ and W) and stage mean fish weights had a profound effect on the accuracy of Schnute models. Knife-edged selection gave a convenient comparison between the estimation methods for growth parameters, as stage mean fish weights were the same between “c” and “waa” growth/mean weight configurations (i.e., survey mean fish weights = catch mean fish weights). This meant that the differences between $S1_c$ and $S1_{waa}$ were due solely to the growth parameter estimation method. $S1_c$ gave more consistently accurate population estimates than $S1_{waa}$. These results suggest that the ideal Schnute model growth/mean weight configuration should be stage mean fish weights from sampling not skewed by selection (e.g. survey) and growth parameters estimated directly from those mean fish weights. However, it is also worth noting that $S1_{waa}$ was nearly as accurate as $S1_c$ for the small-pelagic life-history.

Further investigation should include additional configurations (i.e., stage mean fish weights from survey index and growth parameter estimation from them) to explore the ability of these models with the above configuration in logistic or dome-shaped selectivity scenarios. Also worth investigating further is why there is a difference in accuracy of models when growth parameters are estimated differently and why growth parameters estimated from weight-at-age give less accurate estimation in models. This difference may be because the Schnute model population dynamics are underpinned by stage mean fish

weights and estimated growth parameters from those mean weights are a more accurate representation of the assumed dynamics of the stock within the model.

3.4.4 Convergence and identifiability

Also linked to growth/mean weight configuration of Schnute models was the convergence and identifiability of $S1_c$ and $S0_c$ models in logistic and dome-shaped selectivity scenarios. We found that in the majority of cases, $S0_c$ converged but survey catchability q_s was not identifiable, whereas $S1_c$ failed to converge due to some or all of the parameters being unidentifiable. In these instances, stage mean fish weights are skewed by the selectivity pattern, inflating overall mean fish weight and previously exploited biomass mean fish weight. There are two direct consequences to a skewed stage mean fish weight for the “c” configuration: 1) a low ω (fraction of total biomass in a given year due to newly recruited fish) and 2) growth parameters that give a slower linear growth. Evidently, this affects the internal dynamics of the Schnute model to an extent that estimation is not possible. Further investigation to look at different combinations of growth parameter/mean weight configuration would indicate which factor has the larger bearing on convergence.

3.4.5 Limitations and further work

Simulated stocks were set up under certain conditions where natural mortality rate (M) did not vary temporally or across other factors such as size or age. In addition, the true value of natural mortality rate was used a fixed parameter within assessment models, as is common practice (Lee et al., 2011; Maunder and Piner, 2015). These are key assumptions for models and population dynamics of a stock, in all likelihood aiding estimation of other key parameters in assessment models. In reality these assumptions are unlikely to be realistic (Vetter, 1988; Deroba and Schueller, 2013).

Both models implemented here were limited by their assumptions on selectivity. A knife-edged selection is assumed with all fish fully selected, although this assumption can be relaxed within CSA by manually inputting selectivity for the recruit survey. A similar capacity could be implemented within the Schnute model. This presented no issues for the simulated stock as it was set up with a fully selected fishery and survey. However, the white-bellied anglerfish stock selectivity is likely to be more nuanced, as shown by the age-based assessment

model. To improve both stage-based assessment models, it would be useful to investigate whether the models could estimate different selectivity between the stages under certain conditions. In practice, this would mean experimenting with fishing mortality separable by stage as well as survey selectivity.

Modern stock assessment models are often built in a state-space framework, providing flexibility that accounts for uncertainty in both the model processes and the observations (Schnute, 1991; Pella, 1993; Gudmundsson, 1994; Freeman and Kirkwood, 1995; Nielsen and Berg, 2014; Aeberhard et al., 2018). Extending these models into observation and process error state-space would be a logical step to improving the models and their usefulness as assessment models. This may not be possible with CSA as NOAA (2019) states that process errors tend to be difficult to estimate as they cannot be separated from variability in natural mortality rate, survey variance and recruitment. However the Bayesian implementation of an extended CSA accounts for both process and observation error (Li et al., 2019) and we note that the original CSA (Collie and Sissenwine, 1983) implemented a penalized likelihood that could be thought of as a state-space model.

3.5 Conclusions

We aimed to implement the Schnute (1987) delay-difference model and the NOAA (2019) Catch-Survey Analysis (CSA) within the TMB framework and determine the capabilities of these differing approaches to stage-based stock assessment modelling. To the authors' knowledge this was the first time the model described in the theoretical paper by Schnute (1987) has been implemented. Overall, results were poor for both the *Original Schnute Process Error model* and mixed for the *Adapted Schnute Observation Error model*. However, under certain conditions, and with the correct mean weight/growth configuration, the Schnute approach performed well and offers an alternative stage-based assessment model for situations where data are limited. Developing the Schnute models further would be a worthwhile endeavour, perhaps developing a hybrid model with some similar aspects of CSA, as well as identifying the ideal mean weight/growth configuration. More often than not, CSA was not the best performing model in simulated stock scenarios; however, it was the only model to perform consistently (i.e., convergence and generally accurate estimation) across all scenarios. CSA was also the better performing model (in terms of closeness to values from the official assessment model) of

the assessment models on white-bellied anglerfish. Our CSA assessment model implementation from the NOAA Fish and Fisheries Toolbox (NOAA, 2019) with the addition of stand-alone post-recruit indices had more flexibility in the range of survey index inputs, allowing for a more nuanced set up in the model, closer to that of the age-based assessment model.

Stage-based assessment models are a somewhat under-utilised type of fisheries stock assessment model that offer a useful middle ground between aggregated models and more complex compositional models. By incorporating information on the recruitment each year but with relatively simple underlying population dynamics, stage-based assessment models can be important tools in assessing fish populations where estimates of growth are unreliable (e.g. many anglerfish stocks) or not available. This study adds further evidence that stage-based assessment models are an effective tool for stock assessment and accurately assess fish populations with less data requirements and reduced assumptions, compared to more complex models. Looking forward, focussing on case studies would be particularly useful for the development of these models.

3.6 Acknowledgements

This research (Cullen Fellowship: CF/16/03) was carried out with the support of the Marine Institute and is funded under the Marine Research Sub-Programme by the Irish Government. We'd also like to thank Laurence Kell and Iago Mosqueira for developing a useful simulation-testing framework. We are grateful to the associate editor and reviewers for their comprehensive feedback and useful suggestions that greatly improved the manuscript.

Bibliography

- Aeberhard, W. H., J. Mills Flemming, and A. Nielsen (2018). Review of state-space models for fisheries science. *Annual Review of Statistics and Its Application* 5, 215–235.
- Ailloud, L. E., M. W. Smith, A. Y. Then, K. L. Omori, G. M. Ralph, and J. M. Hoenig (2015). Properties of age compositions and mortality estimates derived from cohort slicing of length data. *ICES Journal of Marine Science* 72(1), 44–53.
- Begley, J. and D. Howell (2004). An overview of GADGET, the globally applicable area-disaggregated general ecosystem toolbox. ICES CM 2004/FF:13.
- Cadigan, N. G. (2015). A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. *Canadian Journal of Fisheries and Aquatic Sciences* 73(2), 296–308.
- Cadigan, N. G., E. Wade, and A. Nielsen (2017). A spatiotemporal model for snow crab (*Chionoecetes opilio*) stock size in the southern Gulf of St. Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* 74(11), 1808–1820.
- Cadrin, S. X. (2000). Evaluating two assessment methods for Gulf of Maine northern shrimp based on simulations. *Journal of Northwest Atlantic Fishery Science* 27, 119–132.
- Collie, J. S. and G. H. Kruse (1998). Estimating king crab (*Paralithodes camtschaticus*) abundance from commercial catch and research survey data. In *Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management*, Volume 125, pp. 73–83. NRC Research Press.
- Collie, J. S. and M. P. Sissenwine (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* 40(11), 1871–1879.

- Conser, R. (1995). A modified delury modeling framework for data limited assessment: bridging the gap between surplus production models and age-structured models. Technical report, A working paper for the ICES Working Group on Methods of Fish Stock Assessment. Copenhagen, Denmark.
- Cook, R. M. and M. R. Heath (2018). Population trends of bycatch species reflect improving status of target species. *Fish and Fisheries* 19(3), 455–470.
- Deriso, R. B. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* 37(2), 268–282.
- Deroba, J., D. S. Butterworth, R. Methot, J. De Oliveira, C. Fernandez, A. Nielsen, S. Cadrin, M. Dickey-Collas, C. Legault, J. Ianelli, et al. (2015). Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. *ICES Journal of Marine Science* 72(1), 19–30.
- Deroba, J. J. and A. M. Schueller (2013). Performance of stock assessments with misspecified age-and time-varying natural mortality. *Fisheries Research* 146, 27–40.
- Dichmont, C. M., R. A. Deng, A. E. Punt, J. Brodziak, Y.-J. Chang, J. M. Cope, J. N. Ianelli, C. M. Legault, R. D. Methot, C. E. Porch, et al. (2016). A review of stock assessment packages in the United States. *Fisheries Research* 183, 447–460.
- Fournier, D. A. and I. J. Doonan (1987). A length-based stock assessment method utilizing a generalized delay-difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 44(2), 422–437.
- Fournier, D. A., J. Hampton, and J. R. Sibert (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences* 55(9), 2105–2116.
- Freeman, S. and G. Kirkwood (1995). On a structural time series method for estimating stock biomass and recruitment from catch and effort data. *Fisheries research* 22(1-2), 77–98.

- Gudmundsson, G. (1994). Time series analysis of catch-at-age observations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 43(1), 117–126.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.
- Ibaibarriaga, L., C. Fernández, A. Uriarte, and B. A. Roel (2008). A two-stage biomass dynamic model for Bay of Biscay anchovy: a Bayesian approach. *ICES Journal of Marine Science* 65(2), 191–205.
- ICES (2018). Report of the Benchmark Workshop on Anglerfish Stocks in the ICES Area (WKANGLER), 12–16 February 2018, Copenhagen, Denmark. ICES CM 2018.
- ICES (2019a). Report of the Working Group for the Bay of Biscay and the Iberian Waters Ecoregion (WGBIE).
- ICES (2019b). White anglerfish (*Lophius piscatorius*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2019. ICES Advice 2019, mon.27.78abd.
- Jardim, E., C. P. Millar, I. Mosqueira, F. Scott, G. C. Osio, M. Ferretti, N. Alzorriz, and A. Orió (2014). What if stock assessment is as simple as a linear model? the a4a initiative. *ICES Journal of Marine Science* 72(1), 232–236.
- Johnson, K. F., C. C. Monnahan, C. R. McGilliard, K. A. Vert-pre, S. C. Anderson, C. J. Cunningham, F. Hurtado-Ferro, R. R. Licandeo, M. L. Muradian, K. Ono, et al. (2015). Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. *ICES Journal of Marine Science* 72(1), 137–150.
- Kell, L. (2018). *FLife: Methods for Modelling Life History Traits*. R package version 3.3.0.
- Kell, L. T. and A. Kell (2011). A comparison of age slicing and statistical age estimation for Mediterranean Swordfish (*Xiphias gladius*). *Collect. Vol. Sci. Pap. ICCAT* 66(4), 1522–1534.
- Kell, L. T., I. Mosqueira, and J.-M. Fromentin (2017). FLife: A R Package For Modelling Life History relationships and Population Dynamics. *Collective Volume of Scientific Papers. ICCAT* 73(9), 3009–3024.

- Kienzle, M., A. J. Courtney, and M. F. O'Neill (2014). Environmental and fishing effects on the dynamics of brown tiger prawn (*Penaeus esculentus*) in Moreton Bay (Australia). *Fisheries research* 155, 138–148.
- Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic Differentiation and Laplace approximation. *Journal of Statistical Software* 70(5), 1–21.
- Lee, H.-H., M. N. Maunder, K. R. Piner, and R. D. Methot (2011). Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. *Fisheries Research* 109(1), 89–94.
- Li, Y., L. M. Lee, and J. Rock (2019). Modeling population dynamics and nonstationary processes of difficult-to-age fishery species with a hierarchical Bayesian two-stage model. *Canadian Journal of Fisheries and Aquatic Sciences* 76(12), 2199–2214.
- Maunder, M. N. and K. R. Piner (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science* 72(1), 7–18.
- Maunder, M. N. and A. E. Punt (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research* 142, 61–74.
- Mesnil, B. (2003). The Catch-Survey Analysis (CSA) method of fish stock assessment: an evaluation using simulated data. *Fisheries Research* 63(2), 193–212.
- Methot, R. D. and C. R. Wetzel (2013). Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142, 86–99.
- Mosqueira, I. (2016). stocksims.
- Nielsen, A. and C. W. Berg (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158, 96–101.
- NOAA (2019). NOAA Fish and Fisheries Toolbox: Collie-Sissenwine Analysis (CSA) version 4.3.
- Pedersen, M. W. and C. W. Berg (2017). A stochastic surplus production model in continuous time. *Fish and Fisheries* 18(2), 226–243.

- Pella, J. J. (1993). Utility of structural time series models and the Kalman filter for predicting consequences of fishery actions. In *Proceedings of the international symposium on management strategies for exploited fish populations. Alaska Sea Grant College Program Report*, Number 93-02, pp. 571–593.
- Prager, M. (1992). Aspic-a surplus-production model incorporating covariates. *Collective Volume of Scientific Papers, ICCAT 28*, 218–229.
- Punt, A. E., T. Huang, and M. N. Maunder (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70(1), 16–33.
- Quinn, T. J. and R. B. Deriso (1999). *Quantitative Fish Dynamics*. Oxford University Press.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roel, B. and D. Butterworth (2000). Assessment of the South African chokka squid *Loligo vulgaris reynaudii*: Is disturbance of aggregations by the recent jig fishery having a negative impact on recruitment? *Fisheries Research* 48(3), 213–228.
- Rosenberg, A. A., M. J. Fogarty, A. B. Cooper, M. Dickey-Collas, E. A. Fulton, N. L. Gutiérrez, K. J. Hyde, K. M. Kleisner, T. Kristiansen, C. Longo, et al. (2014). Developing new approaches to global stock status assessment and fishery production potential of the seas. *FAO Fisheries and Aquaculture Circular* (1086), 0_1.
- Schnute, J. (1985). A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* 42(3), 414–429.
- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences* 44(5), 924–940.
- Schnute, J. T. (1991). The importance of noise in fish population models. *Fisheries research* 11(3-4), 197–223.
- Shepherd, J. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science* 56(5), 584–591.

Smith, M. T. and J. T. Addison (2003). Methods for stock assessment of crustacean fisheries. *Fisheries Research* 65(1-3), 231–256.

Thorson, J. (2020). *TMBhelper: Package for basic helper functions that are not worth putting in a specialized contributed package*. R package version 1.3.0.

Vetter, E. (1988). Estimation of natural mortality in fish stocks: a review.

Zhou, S., A. E. Punt, R. Deng, and J. Bishop (2011). Estimating multifleet catchability coefficients and natural mortality from fishery catch and effort data: comparison of bayesian state-space and observation error models. *Canadian Journal of Fisheries and Aquatic Sciences* 68(7), 1171–1181.

Chapter 4

**sbar vignette: An R package for
implementing stage-based
assessment models**

Abstract

The R package ‘sbar’ provides a set of functions to run a number of different stage-based stock assessment models, as well as plot the outputs. Stage-based stock assessment models were: the well-known numbers-based depletion model, CSA, and a biomass-based delay-difference model first described in a theoretical paper by Schnute (1987). An example dataset, catch and survey data for the black-bellied anglerfish stock in the Celtic Seas and Bay of Biscay, is also provided in the package. This vignette introduces the assessment models and demonstrates key aspects for implementation. Details on the different Schnute model versions are given, outlining how the user can adjust relative importance of recruit biomass or previously-exploited biomass, as well run a more classical version of the model. Stock assessment model applications to the the black-bellied anglerfish stock demonstrated the functionality of ‘sbar’ models, as well as providing preliminary results of stock levels and historical exploitation pattern for a stock that is currently unassessed analytically.

4.1 Introduction

Stock assessment models determine the size (relative or absolute) of stocks and the extent to which they are exploited. This information is an important aspect of fisheries management, as it facilitates understanding of population dynamics of fish stocks (Hilborn and Walters, 1992; Cadrin and Dickey-Collas, 2014; Dichmont et al., 2016). Assessment models use a variety of different data sources to estimate key values for fish stocks.

Stock assessment models range in complexity from aggregate models (which require data on total catch over time and an abundance index for the stock) to compositional models (which also require information on the age-composition of the stock). Stage-based models occupy a middle ground in terms of complexity; they generally have simpler population dynamics, more general assumptions and lower data requirements than more complex assessment models, yet they can account for variability in recruitment, which simpler models cannot. Stage-based models are useful in situations where age data are not available but where the cohort of recruiting fish can be identified clearly (e.g., because they form a clear mode in the length frequency distribution). They are particularly useful in situations where the population structure is affected by irregular events of high recruitment.

Considering the large number of stocks without reliable age data for which an assumption of constant recruitment is not appropriate, is it somewhat surprising that stage-based assessment models are not more widely used. This may be in part due to the limited availability of these models in the R environment.

`sbar` is an R package for fitting stage-based fisheries stock assessment models in R. Estimation relies on the **Template Model Builder (TMB)** framework, taking advantage of the automatic differentiation of the likelihood (Kristensen et al., 2016). Two theoretically different stage-based assessment approaches are implemented:

1. **CSA (Catch-Survey Analysis)** - the well-known numbers-based two-stage model
2. Biomass-based delay-difference two-stage models based on the theoretical paper by Schnute (1987)
 - *Original Schnute Process Error model*
 - *Adapted Schnute Observation Error model*

Standard implementations of these two approaches consist of two stages; fish recruited to the exploited population in a given year, and fish that recruited in previous years. Minimum data required for these assessments consists of at least one survey index of relative abundance and a catch time series corresponding to the survey period. There also needs to be some method for splitting the survey data into the two stages (e.g., size or age).

We explore the three models found in the `sbar` R package in this vignette.

4.1.1 General details

- Models require a time series (numbers-based or biomass-based) of catch and at least one survey time series.
- Models are fit by maximising the total log-likelihood of the objective function.
- In both approaches, surveys and catch (where relevant) are assumed to be lognormally distributed.
- Catch is assumed to be fully selected over both stages (i.e. recruitment stage is fully selected)
- Survey catches are assumed to have the same selectivity pattern as the catches (i.e. fully selected), however this assumption can be relaxed for CSA.
- Stages (recruits and post-recruits for CSA, recruit biomass and previously-exploited biomass for Schnute models) can be defined by the user by age, length or any other grouping that offers distinct stages that move from one to the other over a given time period. **NB** post-recruit stage of CSA and previously exploited stage for Schnute models are analogous.
- Assessments models are limited to the time period covered by surveys and cannot estimate stock values outside of this.

4.2 Installation, data and other useful packages

4.2.1 Installation

You can install the `sbar` package from the github repository with the `devtools` package:

```
devtools::install_github("lbatts/sbar")
```

4.2.2 Data

Example stock data loaded with `sbar` are for the black-bellied anglerfish (*Lophius budegassa*) stock in the Celtic Seas and northern Bay of Biscay (ICES Subareas 7, 8a-b and 8d.), which was collated for ICES and processed by Hans Gerritsen (Marine Institute, Ireland).

Data were originally age-structured, so in the interest of simplicity we defined stages (recruits = age 0 and post-recruits/previously exploited = age 1+) according to their age. In reality, users may want to explore other methods to identify the recruit stage in length frequency data (Bhattacharya, 1967, Taylor and Mildenerger (2017); Batts et al., 2019). The `ank78` data object is a list of vectors and dataframes with the necessary data to run the assessment models. It includes: natural mortality (`M`), survey timings (`•_timing`), a dataframe containing catch data, and two dataframes (`ibts` and `ieiams`) containing survey catch-per-unit-effort data for the combined International Bottom Trawl Survey (IBTS), and Irish anglerfish and megrim fisheries survey (IE-IAMS). These dataframes contain time series for total numbers and biomass, post-recruit numbers and biomass, and, recruit numbers and biomass.

4.2.3 Setup

Load the `sbar` and `TMBHelper` packages. `TMBHelper` is a very useful package that should be automatically installed with `sbar` but needs to be loaded if you choose to use it. This package can be used for optimisation of TMB objective functions, as well as further explore convergence and estimability of parameters.

```
library(sbar)
library(TMBHelper)
```

Load the data and define `years` and `no.years` objects. These variables are based on the years that are covered by at least one survey as `sbar` cannot estimate stock levels outside of survey data (i.e. if catch data goes further back).

```
set.seed(1234)
data("ank78")

years <- as.character(2003:2020)
no.years <- length(years)
```

4.3 Introduction to CSA

CSA is a numbers-based depletion model first described in Collie and Sissenwine (1983). The most recent version, which differs somewhat from the original, can be found in the NOAA Fish and Fisheries Toolbox (where a GUI can be downloaded to run the assessment model) (NOAA, 2019). Our implementation of the CSA assessment is very similar to the NOAA version, but with a few minor adjustments to increase the flexibility of the model. These primarily are:

- An option for including stand-alone post recruit survey(s).
- An option to estimate survey specific error.
- General flexibility in estimating or fixing certain parameters in the model (e.g., survey CV, catch CV, natural mortality)

4.3.1 CSA - quick start

Observations needed for a CSA assessment are catch numbers and a matrix of survey indices (catch numbers per unit effort). A survey split into a recruit index and post-recruit index (i.e. one survey, two indices) is the minimum requirement. In this example we use the combined IBTS survey data that has been processed already into numbers-at-ages. This gives a simple way to define recruits (age 0) and post-recruits (age 1+).

```
catch.no <- ank78$catch$total_no

no.ind = 2

obs <- matrix(NA, nrow = no.ind, ncol = no.years)
obs[1, ] <- ank78$ibts$rec_no
obs[2, ] <- ank78$ibts$postrec_no
```

```
obs[obs == 0] <- NA # Make sure missing years are NAs, zeros
will be considered as data
```

Many of the settings for running CSA have defaults but the function requires some user defined values. `indices_att` contains attributes relating to `obs`, indicating if indices (each row) are from the same survey (i.e. same number) and what type of indices they are:

1. recruit index
2. post-recruit index
3. undivided index

Here, we define that `obs` contains two survey indices from the same survey `survey = c(1,1)` and that the first row of the matrix is a recruit index together with the post-recruit index in the second row `type = c(1,2)`. CSA allows stand-alone post-recruit indices (no corresponding recruit index) in addition to at least one split survey (recruit and post-recruit indices). In theory, the user could input a stand-alone recruit index in addition to a split survey, however it is unlikely that the model could estimate a survey catchability parameter.

```
indices_att <- data.frame(survey = c(1,1), type = c(1,2))
timing <- ank78$surv_timing["ibts"] # survey timing
M <- ank78$M #natural mortality
M
#> [1] 0.25
```

Lets run the assessment with default settings. There's warning messages letting you know that defaults are being used for key inputs.

```
obj <- csa(catch_n = catch.no, indices_no = obs,
           indices_att = indices_att, ts = timing, start_nmort = M)
#> Argument 'selrec' missing. Recruits index/indices assumed
fully selected
#> Argument 'start_q' missing. Default start q used for each
survey
#> Argument 'start_surveycv' missing. Default value used for
each survey
```



```
#> Argument 'start_catchcv' missing. Default value used for
each survey
#> Argument 'start_f_calc' missing. Default value used for
each year
#> Argument 'start_prec0' missing. Default value used
#> Argument 'start_rec' missing. Default value used for each
year
```

For those familiar with `TMB`, `csa` is simply a wrapper function that gives the output from `TMB::MakeADFun`, i.e. an objective function with derivatives, Hessian etc. We have kept optimisation and extraction of values from the assessment models separate to allow flexibility with optimisation methods as well as easy access to `TMB` outputs such as the hessian, gradients etc. See `TMB` documentation for details.

Optimising with `nlminb` looks like this, note starting values for parameters are already defined in `obj` by the `csa` function.

```
opt <- nlminb(start=obj$par,objective=obj$fn,gradient=obj$gr)
opt$convergence
#> [1] 0
```

```
opt$par
#>   logitqhat   logphat1   logrhat   logrhat   logrhat   logrhat
#> -10.000918  10.386337   8.931332  11.060665  10.288809  10.106021
#>   logrhat   logrhat   logrhat   logrhat   logrhat   logrhat
#>  10.629008  11.052483   9.120382  10.092034  10.833050  10.498836
#>   logrhat   logrhat   logrhat   logrhat   logrhat   logrhat
#>  11.521685  10.979062  10.513877  10.796414  11.288606  11.025207
#>   logrhat   logrhat   logf_calc   logf_calc   logf_calc   logf_calc
#>  10.950577  12.293373  -1.011232  -1.710150  -1.791739  -1.810962
#>   logf_calc   logf_calc   logf_calc   logf_calc   logf_calc   logf_calc
#>  -2.263982  -1.640474  -1.053673  -1.400797  -1.796318  -1.437401
#>   logf_calc   logf_calc   logf_calc   logf_calc   logf_calc   logf_calc
#>  -1.675537  -1.770628  -2.012250  -2.129092  -2.358514  -2.886793
#>   logf_calc   logf_calc log_surveycv
#>  -2.783537  -3.159775  -1.693288
```

After optimisation, `sdreport` should be used to calculate standard deviations of all model parameters. When summarised (to `obs.srep`) this gives a matrix with parameters estimated from optimisation, as well as standard deviations of those parameters (estimated within `TMB` via the delta-method).

```
obs.srep <- summary(TMB::sdreport(obj))
```

There are also many other estimated or calculated values from the assessment model with associated error estimates (if relevant) given here. For example "phat" is the estimated post-recruit numbers:

```
obs.srep[row.names(obs.srep) == "phat",]
#>      Estimate Std. Error
#> phat  32413.71   5193.317
#> phat  21640.90   4402.343
#> phat  55415.76  10203.765
#> phat  55914.95  10562.577
#> phat  53174.51  10520.097
#> phat  66324.86  12176.978
#> phat  83031.17  12582.110
#> phat  50652.54  10108.767
#> phat  45533.55   9687.902
#> phat  63468.07  12200.225
#> phat  61254.40  12688.090
#> phat 104712.30  20411.281
#> phat 107301.58  22130.937
#> phat  98198.43  21489.277
#> phat 101675.35  23315.569
#> phat 128656.09  27119.992
#> phat 139991.00  27852.914
#> phat 144210.23  28930.612
```

A list of all the outputs that are reported from the assessment models in `sbar` are given in the function help files (e.g., `?csa`).

When optimising with `nlminb` there's also a handy function in `TMBhelper` to do the previous step (and more) that gives practically the same results. See `?TMBhelper::fit_tmb` for details.

```
fit <- fit_tmb(obj = obj, getHessian = T, quiet=T, control =
list(trace=0))
all.equal(fit$par, opt$par)
#> [1] "Mean relative difference: 9.326973e-08"
```

```

all.equal(fit$objective,opt$objective)
#> [1] TRUE
all.equal(summary(fit$SD),obs.srep)
#> [1] "Mean relative difference: 6.740039e-07"

```

4.3.2 Simple plot of CSA results

An easy way to visualise the key outputs from an `sbar` assessment model is to generate an object with `makeasbarclass` and then `plot`. We recommend to follow these with residual diagnostic plots. In the future this feature will be developed for the `sbar` package.

```

survnames<- c("IBTS recruits (CPUE)","IBTS post-recruits
(CPUE)")
x <- makesbarclass(obs.srep,survnames,catch.no,obs,years)

plot(x,out="fit")

```

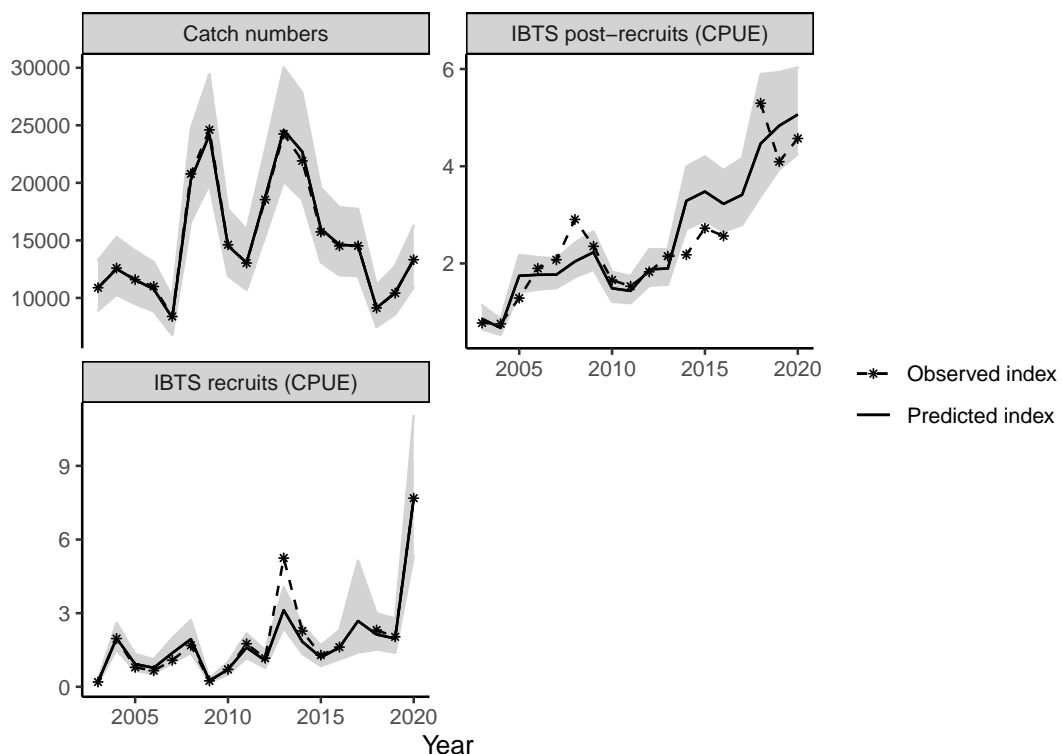


Figure 4.1: Observations and predicted values for catch numbers and survey indices for a CSA assessment model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

```
plot(x, out="stock")
```

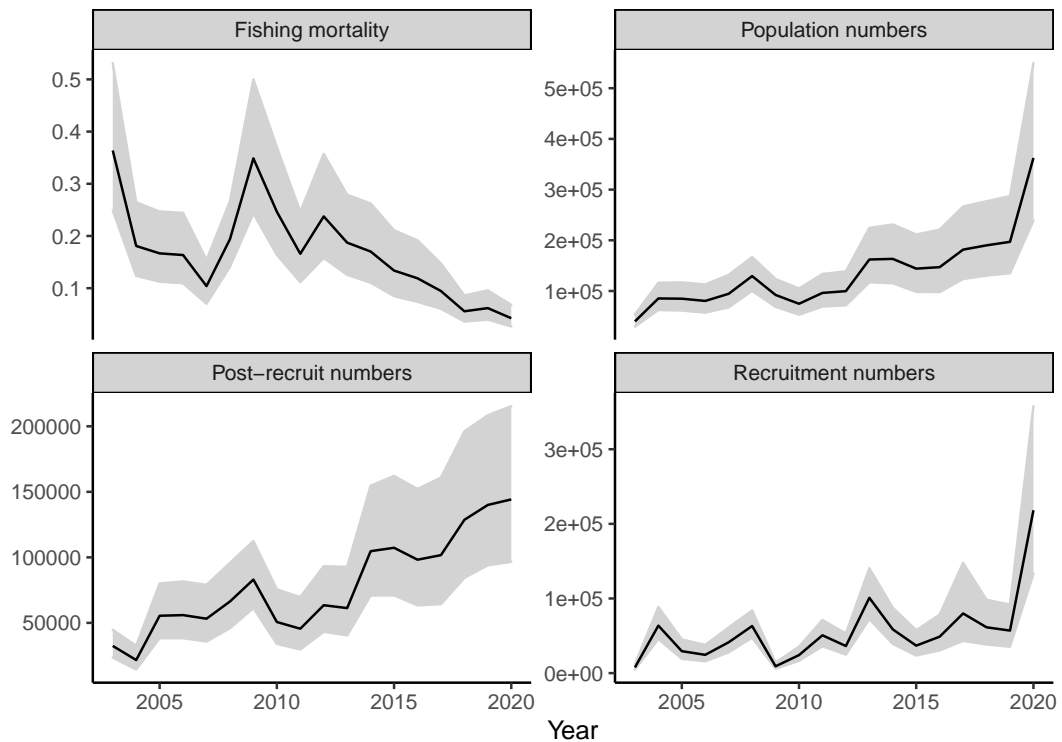


Figure 4.2: Estimated stock numbers and fishing mortality for a CSA assessment model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

4.4 Introduction to Schnute models

Schnute (1987) describes a size-based delay-difference model linking population size structure and mean weights. Growth is assumed to be deterministic and occurs at the end of the year. Mean weights are assumed to not vary throughout the year.

An interesting aspect of this approach is that it offers flexibility in the assumptions regarding the relative importance of recruitment and previously-exploited biomass (i.e. whether the population levels of a fish stock are primarily driven by recruitment or previously-exploited biomass), which underlies the estimation of entire biomass (determined by function argument `version`). We look at this in more detail in the section: *Schnute models in more detail*.

In `sbar` we implement two interpretations of the theoretical model proposed by Schnute (1987). The first which we call the *Original Schnute Process Error*

model (run with `schnute_orig`), which has an autoregressive process error, takes catch biomass as known and assumes no observation error. The other model we call the *Adapted Schnute Observation Error model* (run with `schnute_obserror`), which fits to survey and catch biomass, in addition to being an observation-error-only. Users should note that Batts et al. (2022) found the *Original Schnute Process Error model* was not consistent in self-tests, indicating this model may be structurally biased, particularly under high levels of observation error (Deroba et al., 2015).

4.4.1 Schnute models - quick start

Schnute assumes weight does not change through the year. This maybe a little confusing at first but makes sense under the model assumptions and should be considered when preparing your own data for an `sbar` stock assessment model.

There are three options (1,2 or 3) for the function argument `version`, but for these ‘quick start’ examples we will use `version = 2` (default), which requires a matrix of sampled mean fish weights for the entire stock and for each stage:

- recruit mean weights \bar{Y} (first row)
- previously exploited biomass mean fish weights \bar{Z} (second row)
- entire assessed biomass mean fish weights \bar{X} (third row).

Ideally mean fish weights should come from a sample that is not affected by selectivity (i.e. usually catch has lower selectivity in smaller fish), so in this example we use survey index (combined IBTS survey) to calculate mean fish weights as they’re less likely to be (as) biased.

```
Y <- ank78$ibts$rec_bio/ank78$ibts$rec_no
Z <- ank78$ibts$postrec_bio/ank78$ibts$postrec_no
X <- ank78$ibts$total_bio/ank78$ibts$total_no
```

Then we can populate a matrix with mean fish weights from each stage and the overall mean fish weight.

```
mwts <- matrix(NA, ncol = no.years, nrow = 3)
mwts[1, ] <- Y
```

```
mwts[2, ] <- Z
mwts[3, ] <- X
```

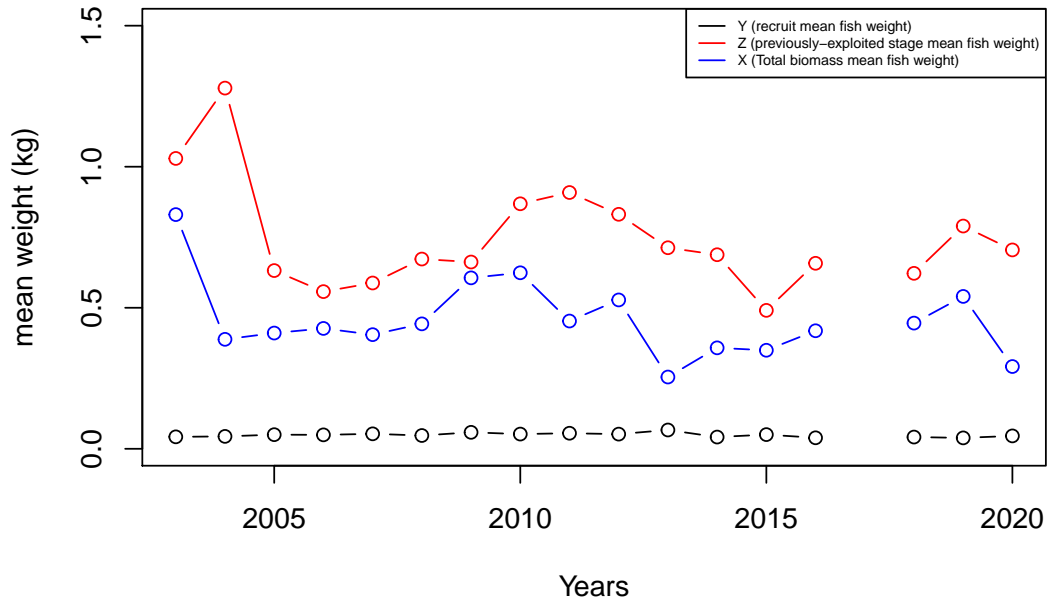


Figure 4.3: Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the combined IBTS survey

There's an issue in 2017 as there's no survey data for this year, so we linearly interpolate between the pre and proceeding points (using the function “approx” in R) for each times series.

```
Y2017<-approx(x = c(2016, 2018), y = mwts[1,c(14,16)], xout =
2017)$y
Z2017<-approx(x = c(2016, 2018), y = mwts[2,c(14,16)], xout =
2017)$y
X2017<-approx(x = c(2016, 2018), y = mwts[3,c(14,16)], xout =
2017)$y

mwts[,15] <- c(Y2017,Z2017,X2017)
```

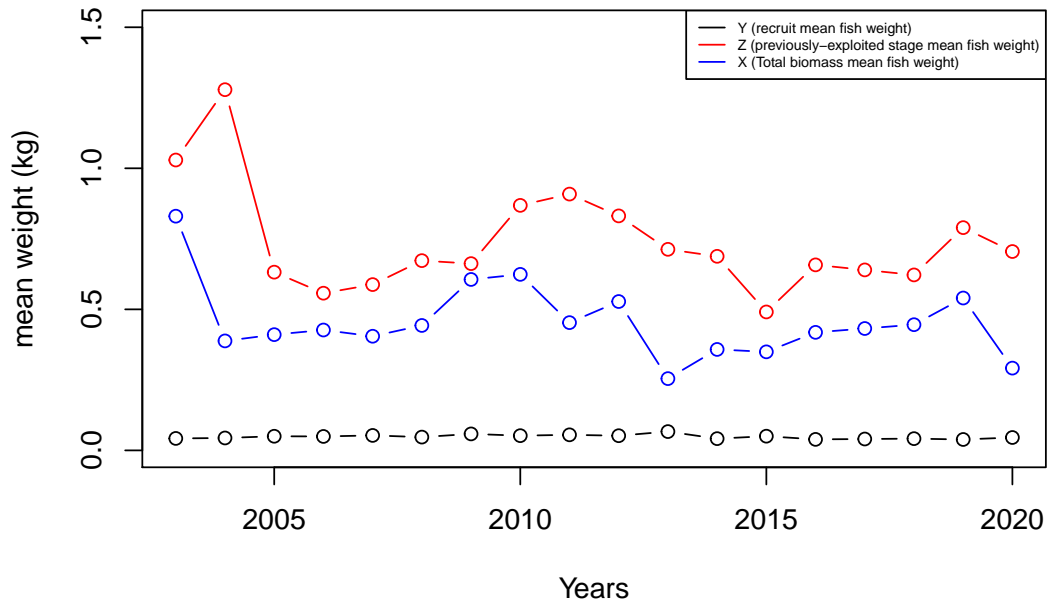


Figure 4.4: Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the combined IBTS survey

Next, we need growth parameters. This aspect is discussed in more detail in *Section 4* but for now just note that we fit a linear model to overall mean fish weights and previously exploited mean fish weights (\bar{Z}_{t+1} vs \bar{X}_t)

```
mod <- lm(mwts[2,2:no.years]~mwts[3,1:no.years-1])
W1 <- coef(mod)[1]
rho1 <- coef(mod)[2]
```

Extract the catch biomass and generate a biomass index from the data.

```
catch_biomass <- ank78$catch$total_bio
index1 <- ank78$ibts$total_bio

obs <- matrix(NA,nrow=1,ncol=no.years)
obs[1,] <- index1 #
obs[obs==0] <- NA

sigma <- exp(-M)
mu <- 0.5
```

`sigma` is survival (i.e. the proportion of the population that survives natural mortality), where $\text{sigma} = e^{-M}$. `mu` is the proportion of the fraction of the

catch removed before natural mortality and is user defined in the *Original Schnute Process Error model* but calculated internally in the *Adapted Schnute Observation Error model*.

4.4.1.1 *Original Schnute Process Error model*

Firstly, we need to use `approx` again to obtain a value for the survey index in the year where we have a missing value. In 2017, there were issues with survey coverage of the IBTS. Due to the internal structure of the process error within the model, the *Original Schnute Process Error model* cannot run with missing survey data.

For demonstration purposes we use `approx`, but ideally other methods should be explored (e.g., Vector-Autoregressive Spatio-Temporal (VAST) (Thorson et al., 2015)).

```
obs_fill<-obs
obs_fill[,15] <- approx(x = c(2016, 2018), y =
index1[c(14,16)]), xout = 2017)$y
```

We can then run the assessment model with some default arguments and then optimise in the same manner as CSA.

```
obj <- schnute_orig(catch_b = catch_biomass, indices_b =
obs_fill, ts = timing, mwts = mwts, rho = rho1, W = W1,
start_sigma = sigma, mu = mu)
#> Argument 'version' missing. Default model version is 2
#> Argument 'ind_l_wt' missing. Default indices likelihood
weighting of 1 used for each survey
#> Argument 'spawn_prop' has length 1 . Given value used for
each year
#> Argument 'start_rec_a' missing. Default value used
#> Argument 'start_rec_b' missing. Default value used
#> Argument 'start_q' missing. Default start q used for each
survey
#> Argument 'start_indexsigma' missing. Default
start_indexsigma used for each survey
fit <- fit_tmb(obj = obj, getHessian = T ,quiet=T,control =
list(trace=0))
```



```

#> Note that `getReportCovariance=FALSE` causes an error in
`TMB::sdreport` when no ADREPORTed variables are present
#> Warning in nlminb(start = startpar, objective = fn,
gradient = gr, control =
#> nlminb.control, : NA/NaN function evaluation

#> Warning in nlminb(start = startpar, objective = fn,
gradient = gr, control =
#> nlminb.control, : NA/NaN function evaluation
obs.srep<-summary(fit$SD)

```

Even though the model converged in `nlminb`, lets do some due diligence and check the estimability with `TMBhelp::check_estimability`.

```

check_estimability(obj)
#> All parameters are estimable

```

All parameters are estimable. However, in other work self-tests for *Original Schnute Process Error model* showed this model was inherently biased and not self-consistent. This indicates structural uncertainty in the model and extra caution is urged when using this model.

```

survnames<- c("IBTS biomass (CPUE)")
x <-
makesbarclass(obs.srep,survnames,catch_biomass,obs_fill,years)

plot(x,out="fit")

```

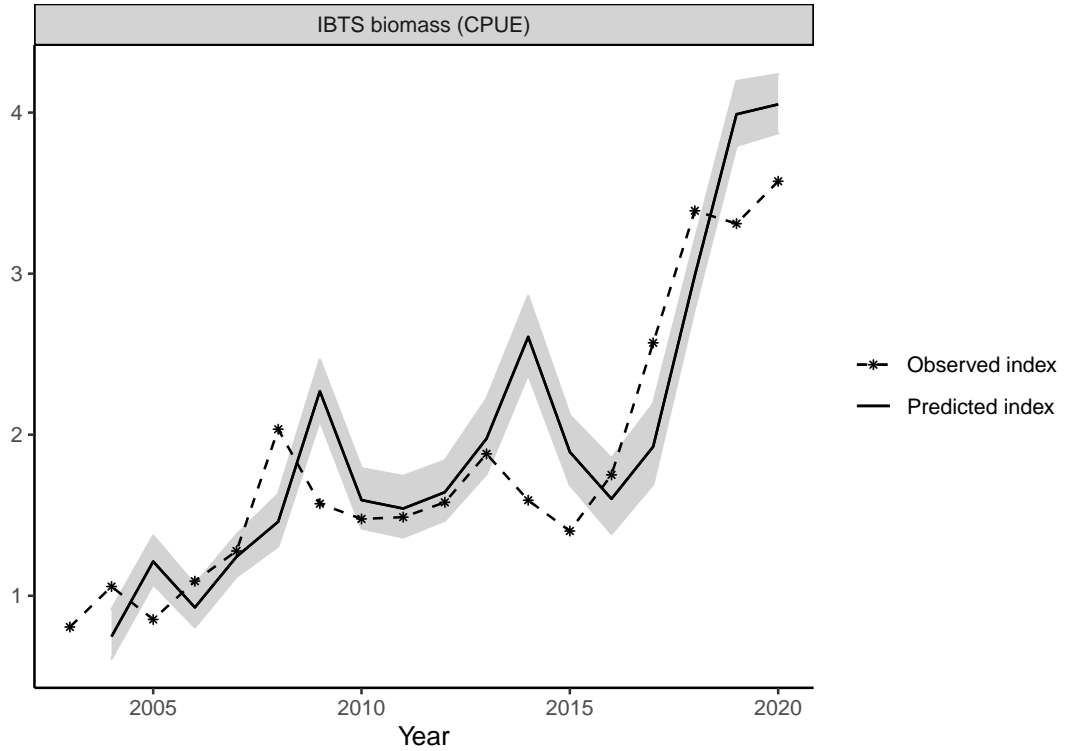


Figure 4.5: Observations and predicted values for the survey index for a Original Schnute Process Error model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

```
plot(x, out="stock")
```

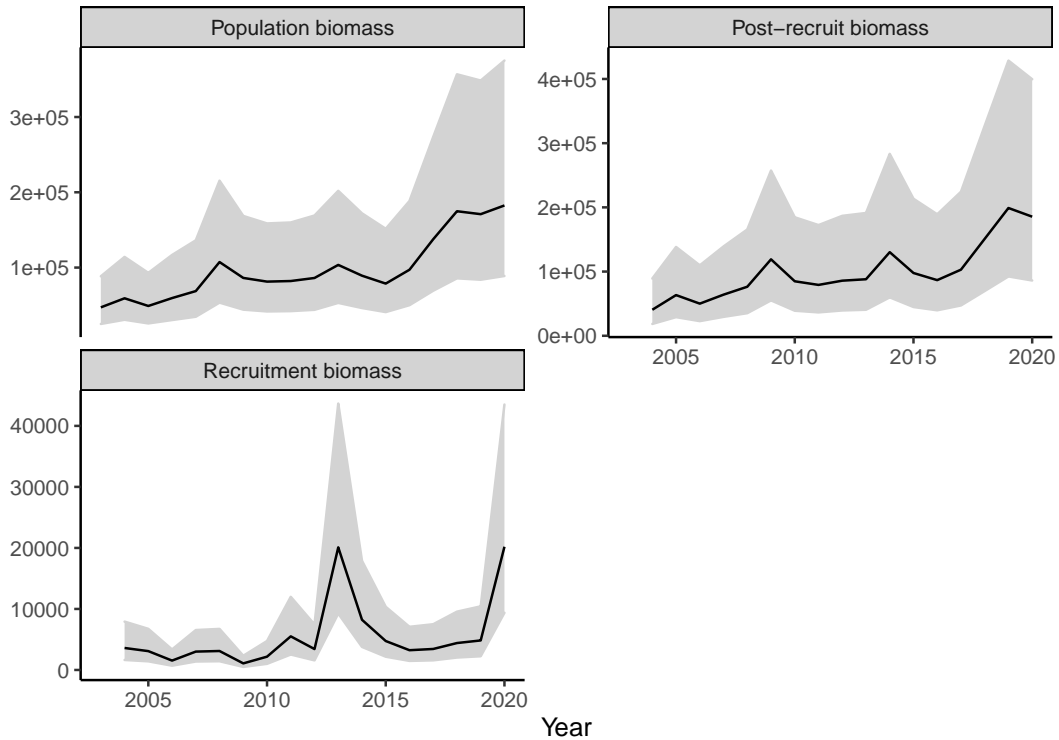


Figure 4.6: Estimated stock biomass, previously-exploited biomass, recruit biomass and fishing mortality for an Original Schnute Process Error model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

Note that this model does not predict the initial year of the survey index and fishing mortality is not estimated as catch is taken in the model without error.

4.4.1.2 *Adapted Schnute Observation Error mode*

We can run the *Adapted Schnute Observation Error model* assessment with the same data. Again, there are a number of defaults which we will leave for this example but will explore in the following section. An important point here is that the assessment can deal with missing survey data so we'll use the original biomass index `obs` rather than `obs_fill`, as predicting without this data may be preferable to interpolating externally.

```
obj <- schnute_obserror(catch_b = catch_biomass, indices_b =
obs, ts = timing, mwts = mwts, rho = rho1, W = W1, start_sigma
= sigma)
#> Argument 'version' missing. Default model version is 2
#> Argument 'ind_l_wt' missing. Default indices likelihood
weighting of 1 used for each survey
```

```

#> Argument 'start_BO' missing. Default value used
#> Argument 'spawn_prop' has length 1 . Given value used for
each year
#> Argument 'start_rec_a' missing. Default value used
#> Argument 'start_rec_b' missing. Default value used
#> Argument 'start_q' missing. Default start q used for each
survey
#> Argument 'start_indexsigma' missing. Default
start_indexsigma used for each survey
#> Argument 'start_catchsigma' missing. Default
start_catchsigma used for each survey
#> Argument 'start_f_calc' missing. Default value used for
each year
#> Argument 'start_f_calc' has only one value which will be
used for each year

fit <- fit_tmb(obj = obj, getHessian = T ,quiet=T,control =
list(trace=0))
#> Note that `getReportCovariance=FALSE` causes an error in
`TMB::sdreport` when no ADREPORTed variables are present

obs.srep <- summary(fit$SD)

check_estimability(obj)
#> All parameters are estimable

```

Check out the fits and stock predictions

```

survnames <- c("IBTS biomass (CPUE)")
x <- makesbarclass(obs.srep,survnames,catch_biomass,obs,years)

plot(x,out="fit")

```

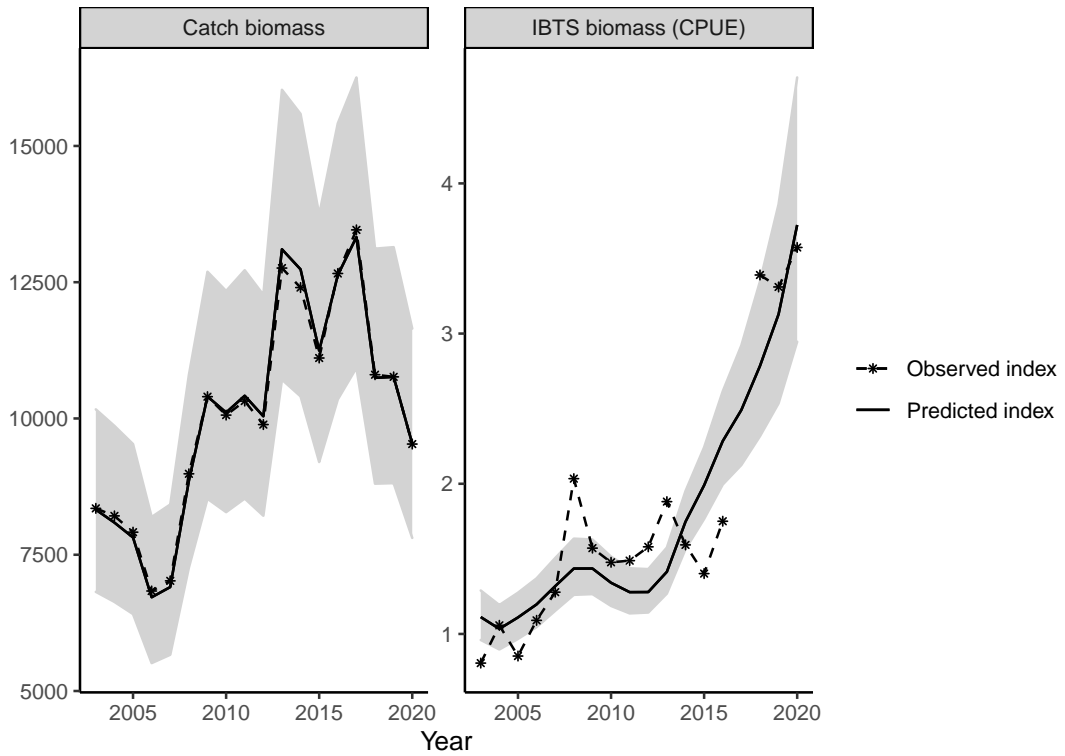


Figure 4.7: Observations and predicted values for the survey index for a Adpated Schnute Observation Error model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

```
plot(x, out="stock")
```

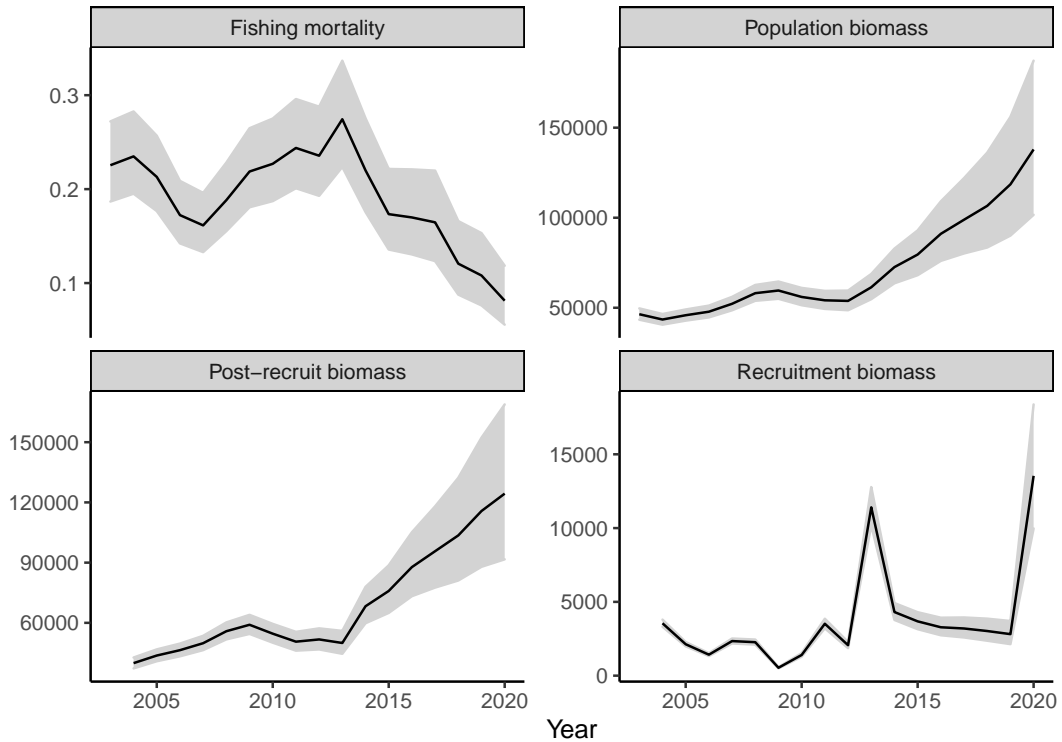


Figure 4.8: Estimated stock biomass, previously-exploited biomass, recruit biomass and fishing mortality for an Adapted Schnute Observation Error model fit on black-bellied anglerfish. Shaded area denotes ± 2 SE on the predicted mean (approximate asymptotic 95% confidence interval).

4.5 Schnute models in more detail

In this section we will go into more detail into certain aspects of the Schnute models implemented within `sbar`. To demonstrate these aspects we'll run through an assessment with the more complex *Adapted Schnute Observation Error model* on the black-bellied anglerfish stock.

In this example we'll also try to fit to an additional survey which is the IE-IAMS monkfish and megrim survey. This survey runs from 2006 to 2020 but has quite a few missing years.

```
head(ank78$ieiams)
#>   year  rec_no postrec_no total_no  rec_bio postrec_bio total_bio
#> 1 2006 9.035172 26.30244 35.33761 0.4459575 13.40772 13.85368
#> 2 2007 3.238904 35.53733 38.77623 0.1717488 18.94035 19.11210
#> 3 2008      NA      NA      NA      NA      NA      NA
#> 4 2009      NA      NA      NA      NA      NA      NA
#> 5 2010      NA      NA      NA      NA      NA      NA
#> 6 2011      NA      NA      NA      NA      NA      NA

index2 <- ank78$ieiams$total_bio
```

We still use the IBTS survey for stage mean fish weights (\bar{Y}' , \bar{Z}' and \bar{X}') as this is the closest to unbiased (i.e. not skewed by selectivity) mean fish weights samples we can get. Catch is likely to be too biased by the selectivity of the fleet and we know the IE-IAMS survey targets larger fish.

Also a reminder that there's no survey data for IBTS in 2017, this isn't an issue for this assessment model with the survey observations but we still need a fully populated mean weight matrix, as discussed above.

Lets visualise the difference selectivity makes on the mean fish weights time series and plot catch mean fish weights with IBTS mean fish weights.

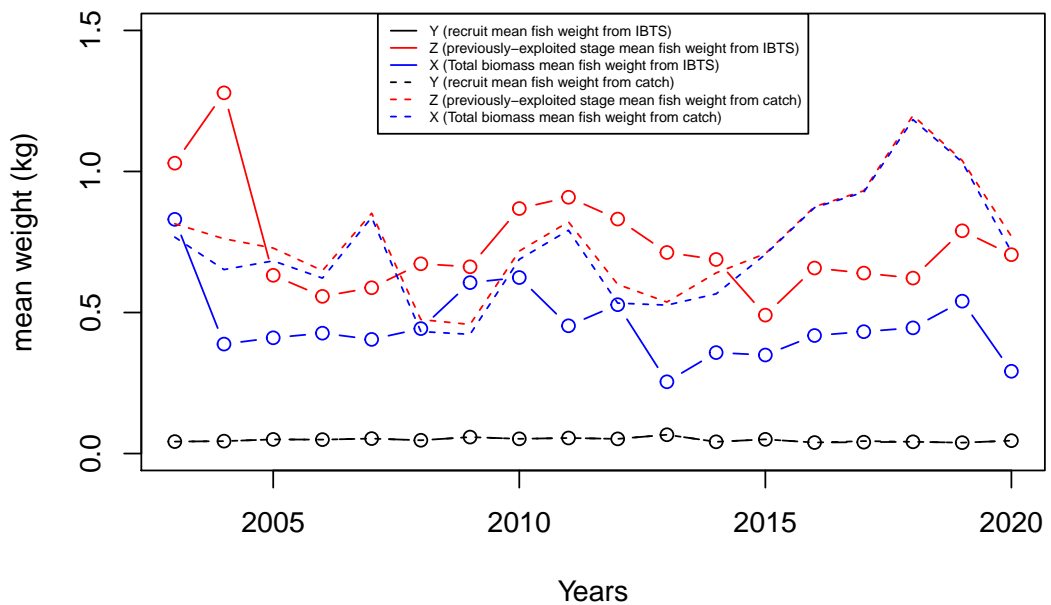


Figure 4.9: Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the combined IBTS survey and catch

4.5.1 Growth and estimating growth parameters

If information on growth is available and weights-at-age are available these can be used (as is common for delay-difference models) to estimate growth parameters with a linear model,

$$\bar{w}_{a+1} = W + \rho\bar{w}_a$$

where \bar{w}_a is the estimated weight-at-age and \bar{w}_{a+1} is the weight-at-age a year older from sampling.

Another option, suggested as a check by Schnute (1987), can be used to

estimate growth parameters through estimation of a linear model on overall mean fish weights and previously-exploited stage mean fish weights from sampling:

$$X'_t = W + \rho\bar{X}_t = \bar{Z}_{t+1}$$

This equation states that the entire population sampled mean fish weight \bar{X}_t , after a year of growth will be \bar{X}'_t , which is the equivalent to the sampled mean fish weight of the previously-exploited population (\bar{Z}) in time $t+1$. This relationship enables the estimation of the parameters W and ρ prior to running an assessment model by fitting a simple linear model.

We encourage users to try both these methods, however simulation-testing indicated that the latter methodology to estimate growth parameters from \bar{Z} and \bar{X} is a better approximation of the deterministic growth assumed within the models Batts et al. (2022). **However, this only the case if sampled mean fish weights are not skewed excessively by selectivity between stages.**

We estimate growth parameters like so,

```
mod <- lm(mwts[2,2:no.years]~mwts[3,1:no.years-1])
W1 <- coef(mod)[1]
rho1 <- coef(mod)[2]
```

Visually this looks like:

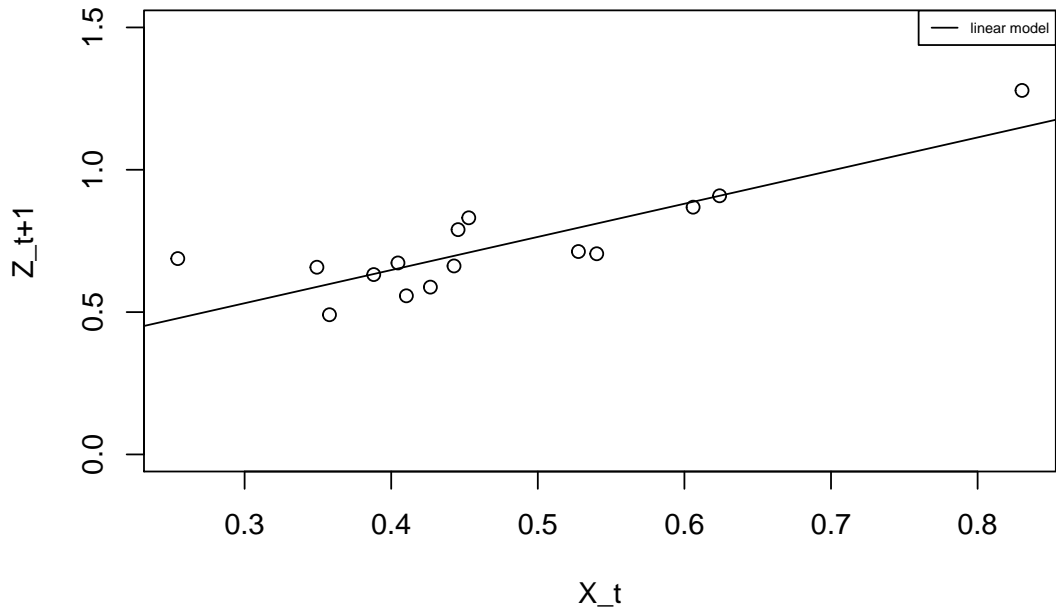


Figure 4.10: Linear relationship between overall mean fish weight and mean fish weight of the previously-exploited biomass.

We can visualise this approximation of linear growth on the mean fish weights time series.

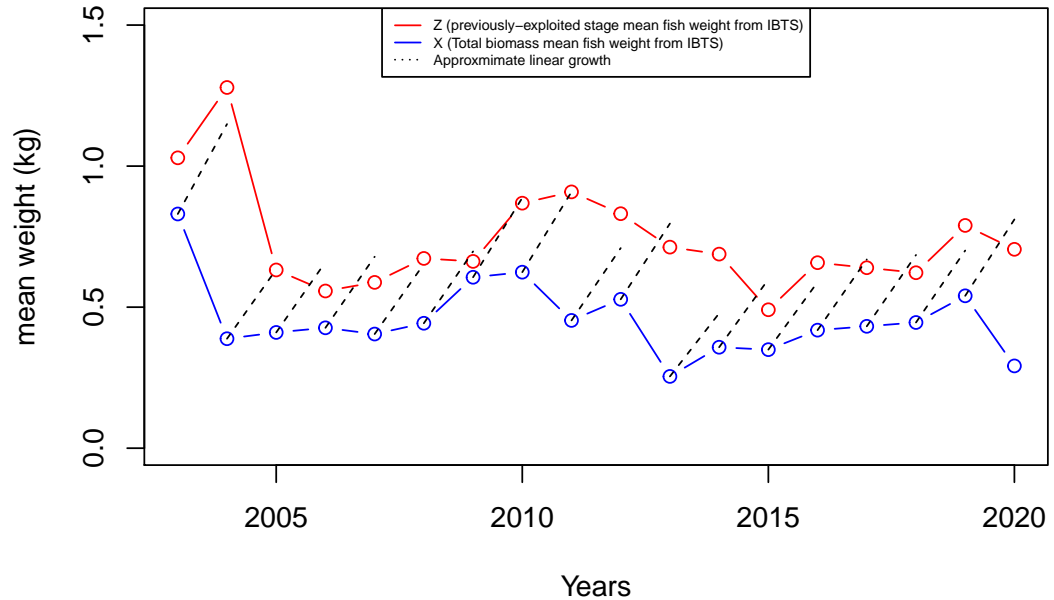


Figure 4.11: Mean fish weights of stages and entire assessed population of black-bellied anglerfish from the combined IBTS survey and approximate linear growth

It is unlikely these surveys do not have any selectivity differences over sizes or ages but these Schnute models assume the same catchability over the entire

assessed population and offer no flexibility in the input.

Note that the `obs` matrix now has two rows, one for each biomass survey index, and `index2` starts at column 4 as this index began in 2006.

```
obs <- matrix(NA,nrow=2,ncol=no.years)
obs[1,]<-index1
obs[2,4:no.years]<-index2 #
obs[obs==0]<-NA
```

At this point in the **quick start** sections we just ran the function with this minimum amount of data and the default arguments. Here, we'll set up a list with data and arguments we want to specify. Note we are setting this list for the *Adapted Schnute Observation Error model*.

```
dat <- list( version = 2,
             catch_b = catch_biomass,
             indices_b = obs,
             ts = ank78$surv_timing,
             mwts = mwts,
             rho = rho1,
             W = W1 ,
             start_q = c(1e-8, 2e-5),
             start_indexsigma = c(0.1, 0.2),
             start_catchsigma = 0.1,
             start_sigma = sigma,
             start_f_calc = 0.5,
             fix_sigma = TRUE,
             fix_indexsigma = TRUE,
             fix_catchsigma = TRUE)
```

Details of arguments can be found in the function documentation but things to note here:

- two survey timings (`ts`)
- two starting survey catchabilities (`start_q`)
- two starting survey standard deviations (`start_indexsigma`)
- starting value for fishing mortality estimates is 0.5 (`start_f_calc`)
- survival (`sigma`) fixed

- survey standard deviations are estimated

and... `version` is set at 2.

4.5.2 Version and the relative importance of recruit or previously-exploited biomass

An intriguing aspect of the model proposed in Schnute (1987) is that there are three model versions, where a predicted total biomass index can be calculated a number of ways. Two of these model versions utilise all three time series of mean fish weights (\bar{X} , \bar{Z} and \bar{Y}) to calculate ω_t . Where, ω_t is defined as the fraction of total biomass in year t due to newly recruited fish:

$$\omega_t = \frac{R_t^*}{N_t^*}$$

where R_t^* is recruitment biomass at time t and N_t^* is population biomass at time t . Schnute (1987) demonstrates that ω_t can be derived from mean fish weights alone:

$$\omega_t = \left(\frac{Y_t}{X_t} \right) \left(\frac{Z_t - X_t}{Z_t - Y_t} \right)$$

With ω_t , N_t^* can be calculated in the population dynamics of the Schnute model with solely either the recruit stage R_t^* (`version = 1`) or the previously-exploited population stage P_t^* (`version = 2`).

version	N_t^* calculation
1	$\frac{R_t^*}{\omega_t}$
2	$\frac{P_t^*}{1-\omega_t}$
3	$R_t^* + P_t^*$

`version = 3` is the more classical model form where estimated biomass in a given year is a combination of recruit biomass and previously exploited biomass.

These versions offer flexibility with the type of model you would like to fit. For example, `version = 1` where the relative importance is shifted to recruit biomass might be useful for a small pelagic stock where recruitment is a big driver of biomass changes. In terms of simplicity `version = 2` would be the

preferred model as no recruitment parameters need to be estimated internally. Versions 1 and 3 fit a Beverton-Holt stock recruitment function internally in the model and these parameters can be difficult to estimate (as we'll see later).

4.5.3 Comparison of Schnute model types

Lets load the data, with the three versions. We use the default recruitment parameters and proportion of biomass mature (`spawn_prop`) (default is 1).

```
ver2<-do.call(schnute_obserror,dat)

dat$version = 1
ver1<-do.call(schnute_obserror,dat)

dat$version = 3
ver3<-do.call(schnute_obserror,dat)
```

Check models evaluate to a finite number with starting parameters

```
ver1$fn(ver1$par)
#> [1] 90079.62

ver2$fn(ver2$par)
#> [1] 101474.3

ver3$fn(ver3$par)
#> [1] 102492.4
```

All three do which is a good start. If they didn't then we would have to play around with starting parameters.

Now let's try and optimise them using `TMBhelper::fit_tmb`, which uses `nlminb`.

```
fit1 <- fit_tmb(obj = ver1, getHessian = T ,quiet=T,control =
list(trace=0))
#> Note that `getReportCovariance=FALSE` causes an error in
`TMB::sdreport` when no ADREPORTed variables are present

fit2 <- fit_tmb(obj = ver2, getHessian = T ,quiet=T,control =
list(trace=0))
```

```

#> Note that `getReportCovariance=FALSE` causes an error in
`TMB::sdreport` when no ADREPORTed variables are present

fit3 <- fit_tmb(obj = ver3, getHessian = T ,quiet=T,control =
list(trace=0))
#> Note that `getReportCovariance=FALSE` causes an error in
`TMB::sdreport` when no ADREPORTed variables are present
#> Warning in fit_tmb(obj = ver3, getHessian = T, quiet = T,
control = list(trace =
#> 0)): Hessian is not positive definite, so standard errors
are not available

```

We can see that the models where `version = 1` and `version = 2` optimised without flagging an issue, whereas the model with `version = 3` could not return a positive definite Hessian.

Remember to follow up convergence with a check on estimability.

```

check_estimability(ver1)
#>          Param          MLE Param_check
#> 1 logrec_param 30.71665          Bad
#> 2 logrec_param 16.78303          Bad
#> 3          logB0 33.12821          Bad
#> 4          logitq -33.12888          Bad
#> 5          logitq -30.61805          Bad
#> 6          logf_calc -23.97572          Bad
#> 7          logf_calc -24.08621          Bad
#> 8          logf_calc -24.68734          Bad
#> 9          logf_calc -25.28017          Bad
#> 10         logf_calc -24.84121          Bad
#> 11         logf_calc -24.73365          Bad
#> 12         logf_calc -26.05905          Bad
#> 13         logf_calc -25.06500          Bad
#> 14         logf_calc -24.08598          Bad
#> 15         logf_calc -24.65532          Bad
#> 16         logf_calc -22.82286          Bad
#> 17         logf_calc -23.99215          Bad

```

```

#> 18    logf_calc -24.35234      Bad
#> 19    logf_calc -24.47183      Bad
#> 20    logf_calc -24.51880      Bad
#> 21    logf_calc -24.86906      Bad
#> 22    logf_calc -25.05115      Bad
#> 23    logf_calc -23.75291      Bad
check_estimability(ver2)
#> All parameters are estimable
check_estimability(ver3)
#>          Param          MLE Param_check
#> 1  logrec_param 11.5357401          OK
#> 2  logrec_param 11.4588316          OK
#> 3      logB0    9.0132474          OK
#> 4      logitq  -7.1149272          OK
#> 5      logitq  -5.9768465          OK
#> 6    logf_calc 18.5382081          Bad
#> 7    logf_calc  1.2268301          OK
#> 8    logf_calc  1.9878611          OK
#> 9    logf_calc  1.0408927          OK
#> 10   logf_calc  0.8391916          OK
#> 11   logf_calc  0.6127227          OK
#> 12   logf_calc  1.0666506          OK
#> 13   logf_calc  1.1370871          OK
#> 14   logf_calc  1.1162249          OK
#> 15   logf_calc  0.9615100          OK
#> 16   logf_calc  0.9402604          OK
#> 17   logf_calc  1.2084133          OK
#> 18   logf_calc  1.2860135          OK
#> 19   logf_calc  1.0672158          OK
#> 20   logf_calc 28.5689851          Bad
#> 21   logf_calc  0.4776739          OK
#> 22   logf_calc  0.5812587          OK
#> 23   logf_calc  0.3451750          OK

```

Models with `version 1` (even though converged with `nlminb` with a reasonable max gradient) and `version = 3` were not able to properly estimate all their parameters. These two models are estimating recruitment parameters, which

is likely the problem, hence the issues flagged with the warning about the Hessian and parameter estimability.

Investigating the `version = 1` and `version = 3` we've found that optimisation with a genetic algorithm to avoid local optima may work for these models. Also worth trying is a stock that is more driven by the recruit stage, e.g., a small pelagic. Stock-Recruit parameters may be easier to estimate in these circumstances.

Bibliography

- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science* 76(7), 2150–2163.
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2022). Numbers or mass? Comparison of theoretically different stage-based stock assessment models and their ability to model simulated and real life stocks. *Canadian Journal of Fisheries and Aquatic Sciences - (Just-IN)*.
- Bhattacharya, C. (1967). A simple method of resolution of a distribution into Gaussian components. *Biometrics*, 115–135.
- Cadrin, S. X. and M. Dickey-Collas (2014). Stock assessment methods for sustainable fisheries. *ICES Journal of Marine Science* 72(1), 1–6.
- Collie, J. S. and M. P. Sissenwine (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* 40(11), 1871–1879.
- Deroba, J., D. S. Butterworth, R. Methot, J. De Oliveira, C. Fernandez, A. Nielsen, S. Cadrin, M. Dickey-Collas, C. Legault, J. Ianelli, et al. (2015). Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. *ICES Journal of Marine Science* 72(1), 19–30.
- Dichmont, C. M., R. A. Deng, A. E. Punt, J. Brodziak, Y.-J. Chang, J. M. Cope, J. N. Ianelli, C. M. Legault, R. D. Methot, C. E. Porch, et al. (2016). A review of stock assessment packages in the United States. *Fisheries Research* 183, 447–460.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.

- Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic Differentiation and Laplace approximation. *Journal of Statistical Software* 70(5), 1–21.
- NOAA (2019). NOAA Fish and Fisheries Toolbox: Collie-Sissenwine Analysis (CSA) version 4.3.
- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences* 44(5), 924–940.
- Taylor, M. and T. K. Mildenerger (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology* 24(4), 330–338.
- Thorson, J. T., A. O. Shelton, E. J. Ward, and H. J. Skaug (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for west coast groundfishes. *ICES Journal of Marine Science* 72(5), 1297–1310.

Chapter 5

Comparative impacts of ageing bias on stage and age-based fisheries stock assessment models

Abstract

Reliable estimation of fish growth is often a key component of a fisheries stock assessment. Growth parameters can be used for age-slicing or cohort-slicing as a means of generating age-composition data for an age-based stock assessment model. A high degree of uncertainty in the growth parameters or misspecification can lead to bias in the age-composition data, potentially altering the perception of the stock and hindering sustainable fisheries management. A management strategy evaluation framework was used to investigate the effect of bias in age-composition data on fisheries stock assessment and management. We test two management procedures (MPs), one based upon an age-based assessment model (a4a) and the other based upon a stage-based stock assessment model (CSA). A biological reference point methodology was also developed for CSA. An underestimated or overestimated growth rate biased age-composition data by shifting the densities of younger and older fish, affecting the different MPs in contrasting ways. For the age-based MP, the perception of the stock was altered directly by the changes in density of the age-composition data, whereas the stage-based MPs were affected indirectly through the management implementation and the use of biased weights-at-age to set catch limits. Overall, the stage-based assessment MP was the more precautionary MP in all scenarios but there was a trade-off with yield. In terms of stock status and management, this study indicates that due consideration should be given to the type of assessment and the potential impact of misspecified growth parameters when providing fisheries management advice.

5.1 Introduction

Effective fisheries management is typically associated with fisheries stock assessment, reflecting the importance of the latter in the advice process for fisheries (Worm et al., 2009; Bianchi et al., 2014; Hilborn et al., 2020). Scientific uncertainty is present in all aspects of assessment and management of fisheries, therefore it is important to account for and quantify this uncertainty (Privitera-Johnson and Punt, 2020). To a large extent, efforts to account for uncertainty and deal with issues of data availability, have lead to progress and development in the field of stock assessment (Patterson et al., 2001; Maunder and Punt, 2013).

There are many forms of stock assessment ranging from surplus-production models (Prager, 1992; Pedersen and Berg, 2017) to age-based models (e.g. Shepherd, 1999; Jardim et al., 2014; Nielsen and Berg, 2014) to ecosystem (e.g. Begley and Howell, 2004) or integrated assessment models (e.g. Bull et al., 2005). However, the majority of contemporary stock assessments are age-based models, when sufficient data is available (Punt et al., 2013).

Age-composition data for both fisheries catch and scientific surveys are essential components of age-based stock assessment models (Hilborn and Walters, 1992). Catch-at-age and age-structured survey data are typically converted from length-composition data using: age estimates of sampled individuals in an age-length key (ALK) (Fridriksson, 1934; Hoenig and Heisey, 1987; Maunder and Piner, 2015); length frequency analysis (Fournier et al., 1990; Taylor and Mildenerger, 2017; Batts et al., 2019); or age-slicing (cohort-slicing) (Hilborn and Walters, 1992; Ailloud et al., 2015).

Uncertainty associated with derived age-composition data is often difficult to account for adequately in traditional statistical catch-at-age stock assessment models (Maunder and Punt, 2013). Age-composition data has the potential to be biased by the conversion process, such as errors in the age-reading method or misspecification or unreliable estimation the of the growth model (Bertignac and De Pontual, 2007; Piner et al., 2005; Maunder and Piner, 2015). Bias at this stage of data collation for an age-based assessment affects other key values of the stock such as catch weight-at-age, stock weight-at-age and maturity-at-age. Cumulatively, these can have a large impact on overall estimates of the fish population when conducting an age-based assessment on potentially biased age-composition data (Bertignac and De Pontual, 2007; Tyszko and

Pritt, 2017).

Integrated stock assessments are increasingly used to account for uncertainty in age-composition data (Maunder and Punt, 2013; Ailloud et al., 2015) such as CASAL for example (Bull et al., 2006; Doonan et al., 2016), Stock Synthesis (Methot and Wetzel, 2013) or MULTIFAN-CL (Fournier et al., 1998). These relatively complex assessment models often allow the user to incorporate many different forms of data (e.g. age-composition data, length-composition data, tagging data, sex or stage data), as well as account for and estimate uncertainty in the data (Maunder and Piner, 2015; Punt et al., 2021). However, these models often require more data than simpler assessment models (Ailloud et al., 2015). In situations where complex compositional models may not be possible, alternative aggregated or stage-based assessment models can be used to potentially reduce (or negate) bias in the age-composition data due to reduced data requirements and more general assumptions within the population dynamics of the models.

Stage-based assessment models bridge the gap between aggregate models and more complex models, as they are based upon relatively simple population dynamics and assumptions. However, they do incorporate information on recruits into the fishery, by modelling the lag between spawning and recruitment (Hilborn and Walters, 1992; Cadrin, 2000; Ailloud et al., 2015; Cook and Heath, 2018; Li et al., 2019). Generally, stage-based assessment models also have lower data requirements than more complex models, and due to the simplified structure of the models they are less likely to be susceptible to issues in ageing uncertainty or bias in age-composition data (Cadrin, 2000; Ailloud et al., 2015).

Catch-Survey Analysis (CSA) is a well-known numbers-based two-stage model (recruit stage and post-recruit stage) with relatively simple population dynamics, assumptions and data requirements. (Collie and Sissenwine, 1983; Mesnil, 2003; NOAA, 2019). Recently CSA has been implemented in the statistical software ‘R’ (Batts et al., 2022a), making use of the automatic differentiation capabilities of the ‘TMB’ framework (Kristensen et al., 2016). Batts et al. (2022b) found that CSA performed relatively well when compared to a statistical catch-at-age assessment for the white-bellied anglerfish in the Celtic Sea and Northern Bay of Biscay.

As part of the stock assessment process, estimated abundance is compared

to reference values for management (Hilborn and Walters, 1992). Biological reference points (BRPs) are key tools for fisheries management, enabling the status of the stock to be assessed and inform advice through harvest control rules (HCR) (Kvamsdal et al., 2016; Silvar-Viladomiu et al., 2021; Zhang and Fong, 2021). Maximum sustainable yield (MSY) is a common and widely accepted approach to fisheries management that is based on targeting estimated reference points of fishing mortality (F_{msy}) and spawning stock biomass (B_{msy}) that will sustain a maximum yield in the long term (Mace, 2001; Mesnil, 2012). The precautionary approach, which is widely advocated and utilised in fisheries management, recognises that there is often marked uncertainty in fisheries systems. Reference points are implemented to account for this uncertainty and minimise the risk of stock collapse (Garcia, 1995; Hilborn et al., 2001; Punt, 2006). The International Council for the Exploration of the Seas (ICES) define target reference points (like F_{msy}) to optimise long-term yield as well as limit reference points (like MSY B_{trigger} and B_{lim}) to minimise the risk of stock collapse (ICES, 2021). Currently there is no protocol to estimate reference points to be used with a CSA assessment under the ICES management framework. In the US, where CSA is used for some fisheries, proxy reference points are more often used (NESFC, 2014).

Various forms of stage-based stock assessment models have been implemented and tested in studies, demonstrating their usefulness in assessing a fish stock with limited data requirements (Cadrin, 2000; Smith and Addison, 2003; Ibaibarriaga et al., 2008; Kienzle et al., 2014; Cook and Heath, 2018; Li et al., 2019). Furthermore, comprehensive simulation testing has shown that stage-based models can perform well under a variety of different stock scenarios (Mesnil, 2003; Deroba et al., 2015; Batts et al., 2022b). Unlike other forms of assessment model, stage based models have yet to be tested for assessment and management capabilities through time, using management strategy evaluation (Carruthers et al., 2014; Fulton et al., 2014; Kell et al., 2014; Walsh et al., 2018).

First developed as a concept within the Scientific Committee of the International Whaling Commission (IWC) (de la Mare, 1986; Kirkwood, 1992; Butterworth, 2007), management strategy evaluation (MSE) is a simulation framework that can be used to gain valuable insights into the assessment and management of stocks (Kell et al., 2005; Butterworth et al., 2010). An MSE consists of two main components: an operating model (OM), and a

management procedure (MP) or management strategy. An OM is a simulated population that can be assessed and managed in a feedback loop via a MP, with uncertainty and stochasticity incorporated into the framework at various points (e.g. observation error, implementation error, recruitment variability) (Smith, 1994; Punt et al., 2014). An MP usually consists of an estimation method (e.g stock assessment) and a harvest control rule (HCR). MSEs can be used in many different contexts to assess the potential risk associated with a range of assessment models, HCRs and utilities. (Dichmont et al., 2006; Holland, 2010; Szuwalski and Punt, 2013; Mildemberger et al., 2021). In other cases MSEs can provide an understanding of how potential biases in the assessment and management process can affect the overall status of the stock over time (Marasco et al., 2007; Kell et al., 2012; Tyszko and Pritt, 2017; Carruthers et al., 2015).

To compare management performance of stage and age-based assessment methods, the goals of this study were to: 1) evaluate the general ability of the CSA MP compared to the age-based MP; 2) develop a procedure for applying bias to generated age-composition data within the MSE; and 3) assess the impact of using biased age-composition data on a fish stock when an age-based assessment MP or a CSA MP is used. Developing a reference points methodology for the CSA assessment MP was also a fundamental component of this study.

5.2 Methods

A short summary of the MSE set up and software is given first, followed by a description of the operating model. The observation error model is detailed next. Descriptions of the management procedures (MPs) are then given and finally performance statistics summarised.

5.2.1 Overview

We simulate a fish stock with similar characteristics to the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d for the OM and then develop a method for implementing bias on age-composition data through age-slicing within the operating error model (OEM) of the MSE. Assessment models are fit to the historical observations from the OEM and relevant reference points are estimated for inclusion in the MPs. MPs are implemented over a 12 year management cycle and their performance evaluated.

5.2.2 General MSE settings

MSEs were implemented within the statistical software R using the FLR framework (Kell et al., 2007). We specifically utilise the ‘mse’ package to run MSEs (Mosqueira and Jardim, 2020) and the “FLife” package (Kell, 2018) to simulate the operating model. Management procedures use either the stage-based assessment model CSA, implemented by Batts et al. (2022b) and available in the ‘sbar’ package (Batts et al., 2022a), or the age-based assessment model a4a from the ‘FLa4a’ package (Jardim et al., 2014). The initial operating model period was 40 years and MSEs were run for a 12 year cycle on top of this. Simulations began with the assessment year in year 40 (as an intermediate year) as was necessary for a management lag and assessment lag of one. Management implementations are first seen in the OM in year 41.

5.2.3 Operating model

The operating model was an age-structured simulated stock created using the “FLife” package in R (Kell, 2018), which uses a combination of given characteristics and life-history theory to generate a stock at equilibrium. Stochasticity in the operating model came from recruitment variability. The stock was created at equilibrium and exploitation began with spawning stock biomass at B_{msy} . Stock-recruit relationship was modelled with a Beverton-Holt model

parameterised with virgin spawning stock biomass and steepness (Equations in *Supplementary Appendix G*). Recruitment variability was assumed lognormal ($\text{sd} = 0.4$) with no autocorrelation in the recruitment process deviations.

Life-history parameters used in the initial set up of the stock were those assumed for the official age-based assessment of the stock ($L_\infty = 171.0$ cm, $K = 0.1075$, $t_0 = -1\text{e-}6$, $a = 3.03\text{e-}5$, $b = 2.82$, $L_{50} = 65.6$ cm and $a_{50} = 4.5$) (ICES, 2019b,a). Steepness of the stock-recruit function was estimated for the genus using the “FishLife” package in R (Thorson, 2019) and set at 0.95. A value for virgin stock biomass was approximated by taking the product of spawners-per-recruit at zero fishing mortality and mean recruitment from the 2019 assessment. Age range was from 0-20 and natural mortality was fixed at 0.25, which is the value used in the age-based assessment (ICES, 2019a). Selectivity of the fishery was a relatively steep logistic curve (age 0 = 0.23, age 1 = 0.85, age 2 = 0.99 and age 3+ = 1), similar to that of the white-bellied anglerfish stock.

Historical exploitation was set over a 40 year period with “roller-coaster” harvest dynamics, which was similar to the fishing mortality pattern estimated for white-bellied anglerfish (ICES, 2019b). In our OM, fishing mortality increased from F_{msy} to $3 \times F_{msy}$ in the first third of the time series, then stayed at this fishing mortality rate for fifteen years. Fishing mortality was then steadily decreased to F_{msy} at the end of the time series (Figure 5.5). The level of overexploitation was exaggerated (high exploitation F is approximately doubled) compared to that estimated in the ICES assessment of the stock; this allowed a clearer investigation into whether management procedures (MPs) could successfully recover or begin to recover the stock if it had been fished well above MSY level for a prolonged period. Estimation of reference points, more specifically the fitting of stock-recruit functions, was also aided by an extended period of high fishing mortality as it provided contrast for the stock-recruit relationship.

5.2.4 Observation Error model

Observations of catch numbers and at least one survey index of relative abundance were required for the two assessment methods used in management procedures (MPs) tested here. For the time period of the initial operating model, a fully-selected CPUE index of numbers-at-age was generated from the OM as the product of stock numbers half way through the year (survey timing

= 0.5), a catchability coefficient of 1e-6 and a random log-normal error (sd = 0.3 for all ages) on numbers-at-age (same error distributions used in Batts et al. (2022b)). A random log-normal error (sd = 0.1) was also applied to catch numbers-at-age from the initial operating model. An additional step in the observation error model was implemented prior to adding random deviances, where catch numbers-at-age and survey CPUE-at-age were age-sliced (cohort-sliced) with or without bias on the growth parameter K (detailed in the following subsection).

As the population was projected forward in time, catch and survey index observations in additional years were generated from the operating model with the same method as above.

5.2.4.1 Bias growth

The OM is age based and generated age-composition data for catch and survey observations. We convert these ages to lengths by creating an age-transition (AT) matrix . This AT matrix is created by applying normal distributions around mean length-at-age from a von Bertalanffy growth function and estimating the probability of a fish in each length class at a given age and vice versa. Simulated length data can then be generated (i.e. the data that is available from sampling the catches and from surveys). We can then apply the transpose of an AT matrix to the length sampling data to convert them into back into age classes for the stock assessment model. We do this using an unbiased AT matrix and AT matrices that are biased towards low and high growth rates. Details of the methodology are given below.

The method of age-slicing catch-at-age and survey CPUE-at-age data was a two stage process using age-length transition matrices (Hilborn and Walters, 1992; Hordyk et al., 2015). An age-transition (AT) matrix is a probability matrix for converting number-at-age to numbers-at-length (and vice versa) where the probability of a fish being at age a and in length class j is given by:

$$P_{a,j} = \begin{cases} \gamma \left(\frac{l_{j+1}^u - L_a}{\sigma_{L_a}} \right) & \text{if } j = 1, \\ \gamma \left(\frac{l_{j+1}^u - L_a}{\sigma_{L_a}} \right) - \left(\frac{l_j^u - L_a}{\sigma_{L_a}} \right) & \text{if } 1 < j \leq J, \\ 1 \left(\frac{l_j^u - L_a}{\sigma_{L_a}} \right) & \text{if } j = J \end{cases} \quad (5.1)$$

where γ is the cumulative normal distribution, L_a is the mean length-at-age, σ_{L_a} is the standard deviation of mean length-at-age and l_j^u is the upper bound of length bin j . Length bins of 0cm (j_{min}) to 205cm (j_{max}) with a width of 1cm were used (columns of AT matrix). Mean lengths at age were estimated using the von Bertalanffy growth equation (rows of AT matrix),

$$L_a = L_\infty \left(1 - e^{-K(a-t_0)}\right); \quad a = a_r + \phi, \dots, a_{max} + \phi \quad (5.2)$$

where L_a is mean length at age, L_∞ is the asymptotic length of the fish, K controls the rate at which the asymptote is approached and t_0 is theoretical age where $L_a = 0$. a_r is the age of recruitment into the fishery, a_{max} is the maximum age and ϕ is the timing of the sampling through the year. Standard deviation of mean length at age was given as:

$$\sigma_{L_a} = \lambda_1 e^{(-1 + \lambda_2 \left(\frac{1 - \rho^{a-1}}{1 - \rho^{A-1}}\right))} \quad (5.3)$$

where λ_1 controls the magnitude of standard deviations, λ_2 controls the trend in σ_{L_a} over ages and $\rho = e^{-K}$ (Fournier et al., 1991; Batts et al., 2019). Numbers at length N_j were established by multiplying the vector of sampled numbers-at-age N_a by the AT matrix \mathbf{P} ,

$$N_j = \sum_{j_{min}}^{j_{max}} \mathbf{P} N_a \quad (5.4)$$

To convert numbers-at-length back into numbers-at-age the AT matrix \mathbf{P} was transposed (\mathbf{P}^T) and standardised ($\dot{\mathbf{P}}^T$) so that the probability of a fish of length j being one of the ages a_r to a_{max} was 1,

$$\dot{P}_{j,a}^T = \frac{P_{j,a}^T}{\sum_{j_{min}}^{j_{max}} P_{j,a}^T} \quad (5.5)$$

where $\dot{\mathbf{P}}^T$ is the standardised and transposed AT matrix. A new vector of numbers-at-age was then given by:

$$N_a = \sum_{a_r}^{a_{max}} \dot{\mathbf{P}}^T N_j \quad (5.6)$$

Three AT matrices were generated: *correct* (unbiased) slicing, *biased high* slicing and *biased low* slicing using the equations above. Minimum and maximum ages a_r (0) and a_{max} (20) were the same as the operating model ages, as were the von Bertalanffy growth parameters used for the *correct* AT matrix ($K^c = 0.1075$, $L_\infty^c = 171\text{cm}$ and $t_0^c = -1\text{e-}6$). A *biased high* AT matrix was generated using a K parameter of $1.2K^c$ (0.129, assumed faster growth) and a *biased low* AT matrix was generated using $0.8K^c$ (0.086, assumed slower growth) to calculate length-at-age 5.1.

Parameters controlling standard deviation at mean length at age were set at appropriate values ($\lambda_1 = 2$ and $\lambda_2 = .4$) that gave relatively small standard deviations ranging from approximately 1.3 - 3.0. Such a small standard deviation at older ages is somewhat unrealistic but deliberate as this facilitated a smooth age-slicing transition from numbers-at-age to numbers at length to numbers-at-age (Figure 5.1). Preliminary testing found that if σ_{L_a} increased to a larger value (e.g 10-15 for age 20 fish) then this meant when performing the *correct* (unbiased) slicing too many fish were being distributed to different ages when converting back to numbers-at-age from the length frequency data, thus skewing catch too far from the original catch and introducing unwanted (and difficult to account for bias). This issue is a general problem with age-slicing and the reason why typically a relatively young plus group is chosen when age-slicing is implemented to generate age-composition data (Kell and Ortiz, 2011).

Table 5.1: Summary of the steps and age-transition matrices used in the observation error model to bias age-slicing. AT' represents the transpose of the AT matrix

Age-slicing observation error	Step	
	Ages to lengths	Lengths to ages
<i>correct</i> (unbiased)	<i>correct</i> AT matrix	<i>correct</i> AT' matrix
<i>biased high</i>	<i>correct</i> AT matrix	AT' matrix with $1.2K^c$
<i>biased low</i>	<i>correct</i> AT matrix	AT' matrix with $0.8K^c$

To bias age-slicing, an AT matrix constructed with biased growth was used in

the second step of the process (Table 5.1). Three observation error models were tested within the MSE framework, all three manipulated observations with the age-slicing methodology, however one of the models used the *correct* AT matrix in the second step (Table 5.1). This accounted for any unknown effects of the age-slicing process. As a further check, un-sliced observations were also included in the MSE framework as an additional operating error model and compared to the *correct* age-slicing results (*Supplementary Appendix G*).

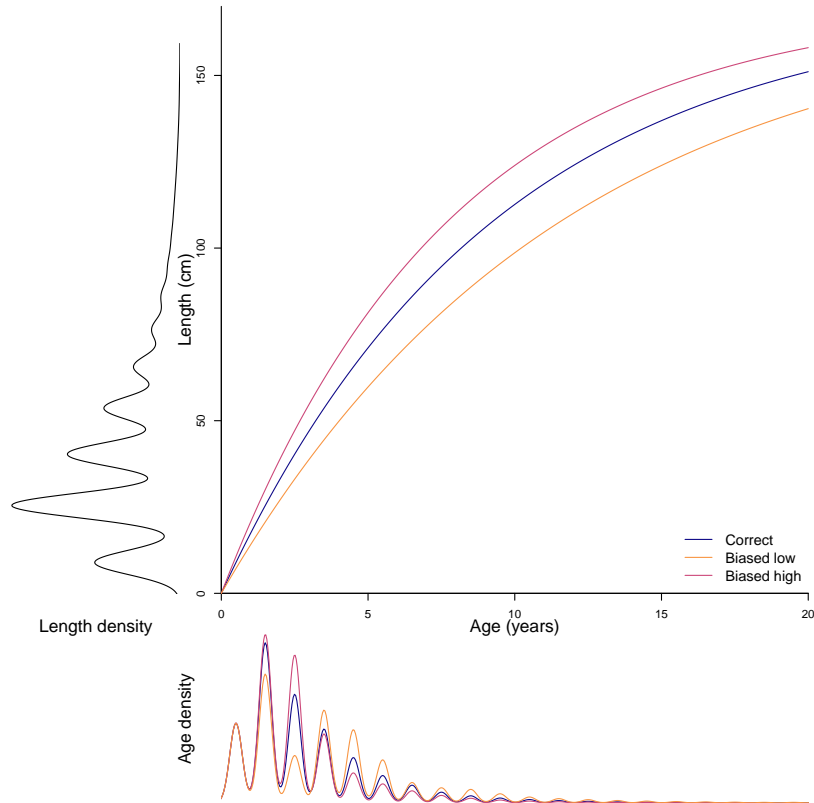


Figure 5.1: Effect of bias on von Bertalanffy growth curves, numbers-at-length density distribution (step 1 of age-slicing) and re-sliced numbers-at-age density distribution (step 2 of age-slicing) for a typical catch-at-age sample (biased high = $1.2K^c$ and *biased low* = $0.8K^c$). Age densities were calculated using a relatively small bandwidth to show peaks in density.

Biased age-slicing changed the age densities of the catch-at-age and age-structured survey data (Figure 5.1). More general shifts in density were seen with increased bandwidth in the density function (Figure 1 in *Supplementary Appendix G*). When a more extreme bias was applied to experiment, differences in density were more pronounced and obvious (Figure 2 in *Supplementary Appendix G*).

Other key values needed for assessments (and subsequent reference points estimation) were also adjusted to account for bias on the K growth parameter. Catch weights-at-age and stock weights-at-age were adjusted by re-calculating weights according to the biased growth curve as well as a and b length-weight parameters (Figure 5.2). Maturity-at-age was also adjusted by shifting the maturity ogive according to the biased ages. An important note here is that with the adjusted weights, total catch weight of observations that have been biased remains almost exactly the same as the *correct* total catch weight. This makes sense as the actual catch weight would not change in real life if bias was present in the ageing/age slicing data collation.

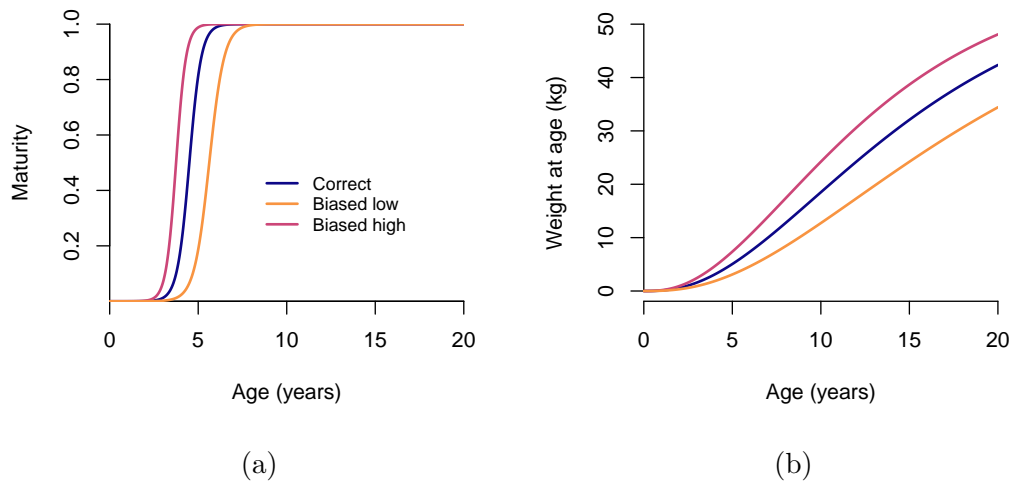


Figure 5.2: Bias adjusted maturity ogives (a) and catch weights-at-age (b) used in the observation error model

5.2.5 Stock assessment methods

Two contrasting estimation methods were used in the construction of MPs tested within the MSE framework. For our study we required estimation methods that produced outputs that can be used to estimate biological reference points, which can then be used in the HCRs to assess stock status and determine future management. Given the age-slicing observation error model affected numbers-at-age in both the survey and catch it was sensible to use an age-based assessment as one of the estimation methods. We used the a4a statistical catch-at-age model (Jardim et al., 2014), which was also the current assessment framework for the real white-bellied anglerfish stock used to condition the OM (ICES, 2019b). To contrast the fully age-based model

we also implemented management procedures that use CSA (Catch-Survey Analysis), which has recently been implemented in the R environment (Batts et al., 2022a), making use of the R package “TMB” taking advantage of the capacity for automatic differentiation (Kristensen et al., 2016; Batts et al., 2022b). CSA is a two-stage numbers-based assessment model that has relatively simple underlying population dynamics compared to a4a and but still incorporates information on recruitment (Collie and Sissenwine, 1983; Batts et al., 2022b). CSA was chosen as it was likely to perform differently to the age-based assessment with regards to age-slicing bias. To our knowledge, CSA has not been MSE tested nor have non-proxy reference points been developed based on the model.

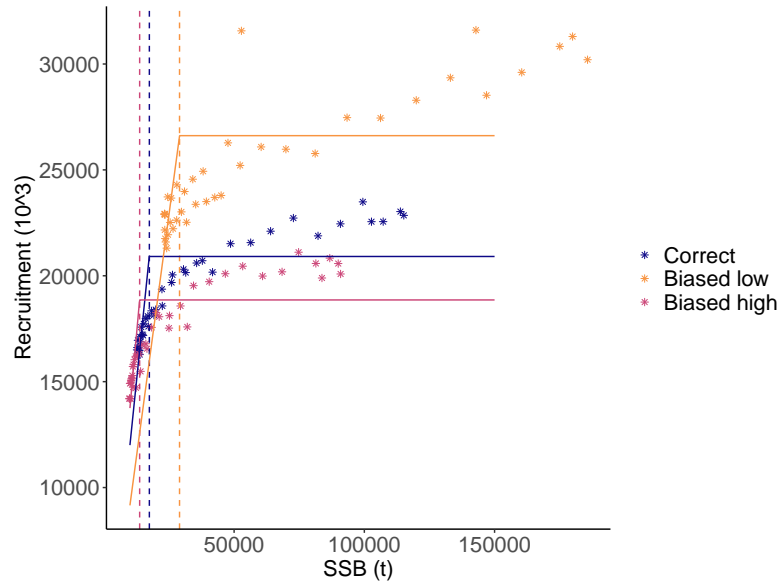
The a4a model requires observations of catch-at-age and a survey CPUE-at-age index, which are both assumed to be normally distributed on the log-scale within the assessment. The a4a assessment framework allows model structure to be controlled by a series of linear models on key aspects of the model (Jardim et al., 2014). Fishing mortality was estimated independently for each age class up to age 5, after which it was assumed to be the same for all older ages. F was assumed to be separable, meaning that only the absolute level of fishing mortality and not the selection pattern (relative F at age) changed over time (same setup as the current assessment for white-bellied anglerfish (ICES, 2019b)). Recruitment was estimated yearly and survey catchability was assumed to be constant across ages and years. Default values and configurations were used for the other aspects of the assessment. Natural mortality was fixed at the value used within the operating model. a4a assessments were run with a plus-group of age 7 and survey timing was set at 0.5.

CSA requires a time series of total catch numbers therefore catch-at-age observations were aggregated. A split survey of two survey indices for recruits and post-recruits was also required for CSA. Survey CPUE-at-age was simplified into a recruit index (we assumed recruits were identified as age = 0 for simplicity) and post-recruit index (ages 1-20 aggregated). CSA assumes both catch numbers and indices are log-normally distributed. Natural mortality was fixed at the same value of the ICES assessment. A key assumption of CSA is that both stages in the catch are fully-selected, which is not the case in these simulations. Key outputs of CSA include recruit and post-recruit numbers in the stock and an estimation of yearly overall fishing mortality.

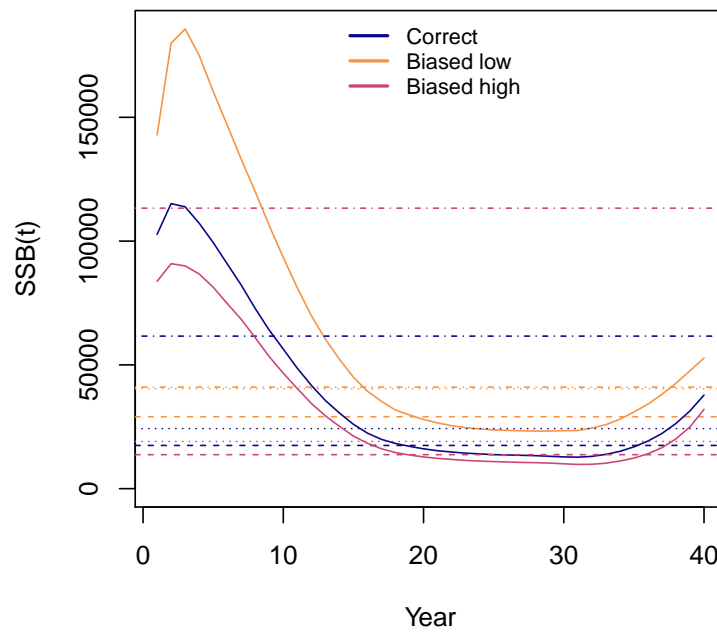
5.2.6 a4a biological reference points

Biological reference points required for the HCR were estimated for each of the observation error models. This involved fitting the *correct*, *biased high* and *biased low* stock observations from the initial operating model time period (across all iterations) with the a4a assessment model. The median values for key outputs from estimated stocks (across the 100 iterations) were then run through the standardised procedure in ICES for estimating biological reference points for the MSY approach (ICES, 2021). This procedure uses “eqsim”, which is stochastic simulation software implemented in the R package “msy” that can be used to estimate and explore reference points (Simmonds et al., 2019; ICES, 2021).

Following ICES protocol B_{lim} was estimated as the breakpoint of a segmented regression as the stock-type was deemed to be type 2 (wide ssb range and evidence that recruitment is or has been impaired) (Figure 5.3). B_{pa} is defined a “stock status reference point above which the stock is considered to have full reproductive capacity, having accounted for estimation uncertainty” (ICES, 2021). $B_{pa} = B_{lim} \times e^{(1.645 \times \sigma)}$ where σ was the default value of 0.2. F_{lim} was estimated with eqsim simulations, as was F_{msy} and MSY $B_{trigger}$. F_{pa} was derived from F_{lim} . It is important to note that because the stocks have not been fished at (estimated) F_{msy} for the last five years of the initial operating model period ICES protocol suggests using B_{pa} as the MSY $B_{trigger}$ value in the advice rule (where MSY $B_{trigger}$ is defined as a biomass reference point that triggers a cautious response within the ICES MSY framework). F_{msy} estimates for the three stocks were all $< F_{pa}$ and $< F_{p.05}$, where $F_{p.05}$ is the fishing mortality that would result in a 95% probability of SSB being above B_{lim} in the long term. However, the ‘*biased low*’ stock F_{msy} was very close to F_{pa} (Table 5.2 and Figure 5.3) (ICES, 2021).



(a) S-R plot



(b) Estimated SSB, B_{lim}, B_{pa} and MSY B_{trigger} values

Figure 5.3: Summary of key outputs across *correct*, *biased high* and *biased low* OEMs from the estimation of a4a reference points including: the stock-recruit data with median estimated segmented regression models and B_{lim} (a); and estimated reference points in relation to estimated SSB. Where dashed lines are B_{lim}, dotted lines are B_{pa} and dash-dot lines are MSY B_{trigger}.

Table 5.2: Summary of biological reference points (biomass BRPs are in tonnes) estimated with a4a (with FBAR =3-6) and eqsim with observations from the initial operating model time period. Observations were from three different observation error models with different age-slicing bias. The analytical F_{msy} value of the simulated stock is shown in parentheses

Observation error model	B_{lim}	B_{pa}	MSY	B_{trigger}	F_{lim}	F_{pa}	F_{msy} (0.148)
<i>Correct</i>	17459	24261		61615	0.41	0.28	0.15
<i>Biased high</i>	13731	19080		113305	0.53	0.36	0.14
<i>Biased low</i>	29045	40360		41014	0.28	0.19	0.19

5.2.7 CSA biological reference points

Biological reference points were also required for the HCR implemented with CSA, however because the assessment is completely numbers-based (no spawning stock biomass estimation) and not age-based, eqsim could not be used. CSA has not been used within the ICES framework before therefore there was no protocol to follow or software developed. We developed our own methodology for estimating reference points for CSA and assessing stock status in an ‘‘ICES style’’ HCR.

Similarly to a4a, CSA was fit to *correct*, *biased high* and *biased low* stock observations from the initial operating model time period. The median estimated stocks (across iterations) were then used in the methodology below. CSA reference points were not estimated stochastically. We developed a methodology for estimating CSA reference points based upon the relatively simple assumptions and outputs of the CSA assessment. Assumptions on growth and weight-at-age were needed in order to calculate a catch biomass at equilibrium.

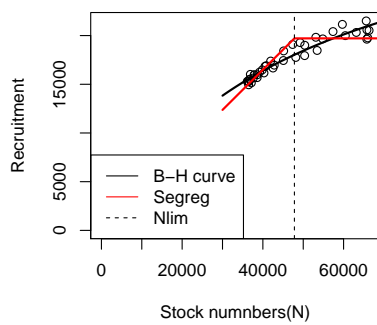
5.2.7.1 F_{msy} and N_{msy}

Fishing mortality that would lead to long-term maximum sustainable yield (F_{msy}) and stock numbers that support F_{msy} (N_{msy}) was found with the following steps using the methodology detailed in the *Supplementary Appendix G*. CSA model-based reference points were derived as follows:

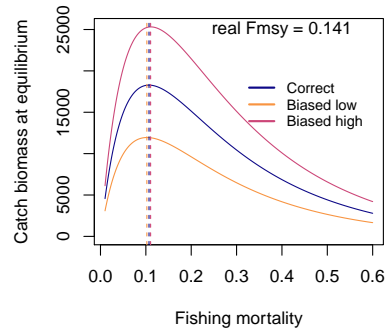
1. Estimate stock-recruit parameters (a and b) by fitting a B-H function to recruit numbers and stocks numbers from a CSA assessment.
2. For a range of F values, calculate stock numbers-at-equilibrium N_n^* and catch-at-equilibrium C^*

- Identify the F value that gives maximum C^* (F_{msy}) and corresponding N_n^* (N_{msy}) (Table 5.3)

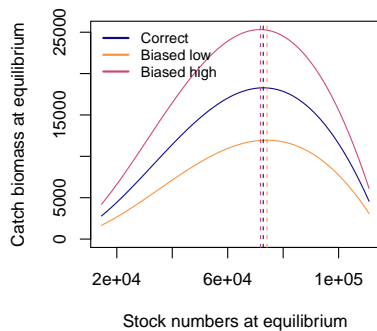
Two key assumptions on growth are introduced in the reference point methodology for CSA: 1) ages of the stock range from 0-20 and 2) that we have growth parameters and length-weight parameters that are used to calculate weight-at-age (biased or not).



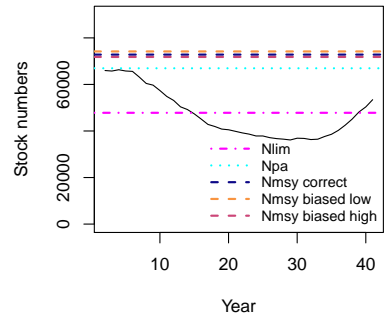
(a) S-R plot



(b) Equilibrium yield vs. Fishing mortality



(c) Equilibrium yield vs. Stock numbers



(d) Operating model with BRPs

Figure 5.4: Summary of key outputs from the estimation of CSA reference points including: the stock-recruit data with fitted models (a); equilibrium yield vs. fishing mortality (b) and stock numbers (c); and Estimated reference points in relation to stock numbers estimated in CSA assessments.

To follow a ICES style Advice rule for the harvest control rule in the CSA MPs additional reference points were needed that were reasonably analogous to those estimated for the age-based assessments for the PA (B_{lim} and B_{pa}). The simplest approach was to develop numbers-based versions of B_{lim} and B_{pa} . Similarly to B_{lim} we define N_{lim} as the estimated breakpoint of a segmented

regression on the stock-recruit curve (rec vs N) (Figure 5.4). ICES protocol states that $B_{pa} = B_{lim} \times e^{(1.645 \times \sigma)}$ but can be calculated as $1.4 \times B_{lim}$ in situations where assessment uncertainty (σ) is unavailable (ICES, 2021). We define $N_{pa} = 1.4 \times N_{lim}$ (Table 5.3). Note that we discuss the limitations of the numbers-based approach in the Discussion.

The assumptions detailed above regarding mean weights-at-age are only relevant to F_{msy} and N_{msy} estimates.

Table 5.3: Summary of biological reference points estimated with CSA and CSA reference point methodology with observations from the initial operating model time period. Observations were from three different observation error models with different age-slicing bias.

Observation error model	N_{lim}	N_{pa}	N_{msy}	F_{msy}
<i>correct</i>	47815	66941	72810	0.11
<i>biased high</i>	47815	66941	71805	0.11
<i>biased low</i>	47815	66941	74181	0.10

5.2.8 Harvest Control Rules and implementation

Management advice was annual and implemented in the year following the assessment. Assessments were run with data from the full time period prior to the assessment year. HCRs were based on the ICES MSY approach for long-lived category 1 and 2 stock (ICES, 2019).

For the age-based assessment management procedures the HCR is simply that from ICES technical guidelines, using B_{lim} , B_{pa} and F_{msy} to determine stock status and fishing mortality (F) for the following year. B_{pa} was used instead of MSY $B_{trigger}$ as the stocks were not estimated to be fished at MSY $B_{trigger}$ for a time period before reference point estimation. HCRs for the age-based MPs follow the ICES Advice Rule:

$$F = \begin{cases} F_{msy} & SSB \geq B_{pa} \text{ (in our case)} \\ F_{msy} \cdot \left(\frac{SSB}{B_{pa}}\right) & B_{lim} < SSB < B_{pa} \\ 0 & SSB \leq B_{lim} \end{cases}$$

In reality, when SSB is found to be below B_{lim} fishing mortality is often not set to 0 but for the purposes of this MSE it was a suitable management decision.

CSA MPs follow a similar procedure to the age-based assessment MPs. We

implemented an “ICES style” advice rule HCR where numbers-based biological reference points (N_{lim} , N_{pa} and F_{msy}) were used to assess stock status (in terms of stock numbers N) and determine fishing mortality for the following year:

$$F = \begin{cases} F_{\text{msy}} & N \geq N_{\text{pa}} \text{ (in our case)} \\ F_{\text{msy}} \cdot \left(\frac{N}{N_{\text{pa}}}\right) & N_{\text{lim}} < N < N_{\text{pa}} \\ 0 & N \leq N_{\text{lim}} \end{cases}$$

Fishing mortality for management was then converted into a total allowable catch (TAC) by estimating stock numbers in the following year given a mean recruitment over the last three years and simulating the management F . For the CSA MPs the methodology for reference point estimation was utilised here. It is important to note that for all MPs TAC was calculated with observation error model weights, thus biased growth assumptions also had an effect here.

5.2.9 Performance statistics

Performance metrics were used to evaluate and compare MPs. Performance focused on two aspects; 1) comparison between MPs that used fundamentally different assessments and reference points (a4a and CSA MPs); and 2) comparison of the performance of each MP when observations were altered by age-slicing bias.

To examine short and medium-term effect of MPs on stock status we calculated F/F_{msy} and B/B_{msy} across replicates half-way through the projected time period of the MSE (six years in) and at the end of of the projected time period (12 years in). Where F and B were fishing mortality ($\bar{f} = 3-6$) and spawning stock biomass of the operating model respectively. F_{msy} and B_{msy} were the analytical values of the these reference points calculated when the operating model was simulated initially. We also examine a Kobe plot of mean and median F/F_{msy} and B/B_{msy} in the final year of simulations.

Risk and sustainability were evaluated by calculating the proportion of replicates for each MP that were at a SSB below 20% of B_0 (risk of stock being fished at low levels, where B_0 is virgin or unexploited spawning stock biomass) (Beddington and Cooke, 1983; Myers et al., 1994) and the proportion of replicates above B_{msy} (sustainability) in each year of simulations. These values are plotted against mean catch biomass in each year to examine trade-offs.

5.3 Results

Results from initial assessments are examined first, followed by the MSE results. Finally we compare across the two different MP estimation methods across the set of performance statistics.

5.3.1 Initial assessments

a4a assessments

a4a assessments on observations from the three OEMs (*correct*, *biased high* and *biased low*) for the initial operating model time period (40 years) gave very different estimations of the stock (Figure 5.5). Estimated biological reference points (using a4a assessment estimates) also differed notably between OEMs (Table 5.2).

When age-slicing was *biased low* the a4a assessment over-estimated spawning stock biomass (SSB) and recruit numbers in comparison to when the *correct* age-slicing was used. Fishing mortality was estimated lower than the assessment using the *correct* OEM observations. For the *biased high* OEM the opposite pattern was seen, with SSB and recruitment estimated lower in comparison to the values estimated with the *correct* OEM observations, along with estimated fishing mortality which was estimated higher (Figure 5.5).

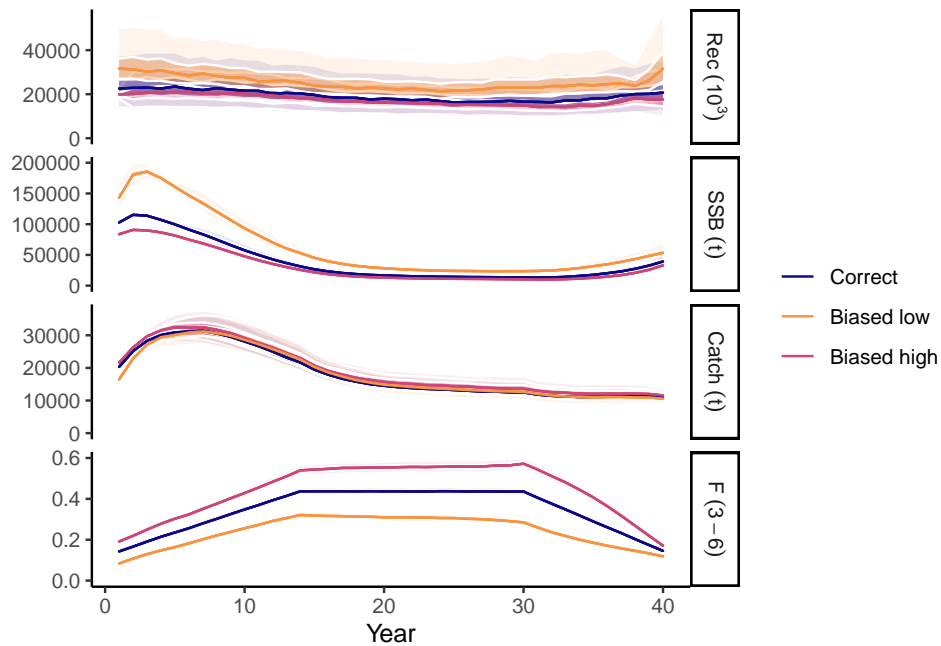


Figure 5.5: a4a stock assessment estimates using observations from three OEMs (*correct*, *biased high* and *biased low*) for initial data years of the operating model. $F(3-6)$ is F_{bar} (mean fishing mortality rate for ages 3-6). Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

Estimated stock numbers-at-age and fishing mortality-at-age showed the same pattern as SSB and mean fishing mortality(f_{bar}), whereas catch-at-age showed a different pattern, similar to that of the OEM observations when biased age-slicing is applied (Figure 5.1) (*Supplementary AppendixG* - Figure 3). Estimated catch numbers at age 0 are very similar, however numbers-at age alternate somewhat over ages and age-slicing bias. Generally *biased low* observations of catch over estimate the number of older fish in the catch and under estimate younger fish compared to the assessment on *correct* age-slicing observations. Assessment results on *biased high* observations show the opposite, estimating higher numbers in younger years (age two mainly) and lower in older years (*Supplementary AppendixG* - Figure 3).

CSA assessments

In contrast to the initial a4a assessments, CSA assessments on observations from the three OEMs gave almost identical stock estimates across differing age-slicing bias (Figure 5.6). There was no difference in the estimation of N_{lim}

and N_{pa} reference points with these assessment results, however F_{msy} and N_{msy} were different but not by a large degree (Table 5.3).

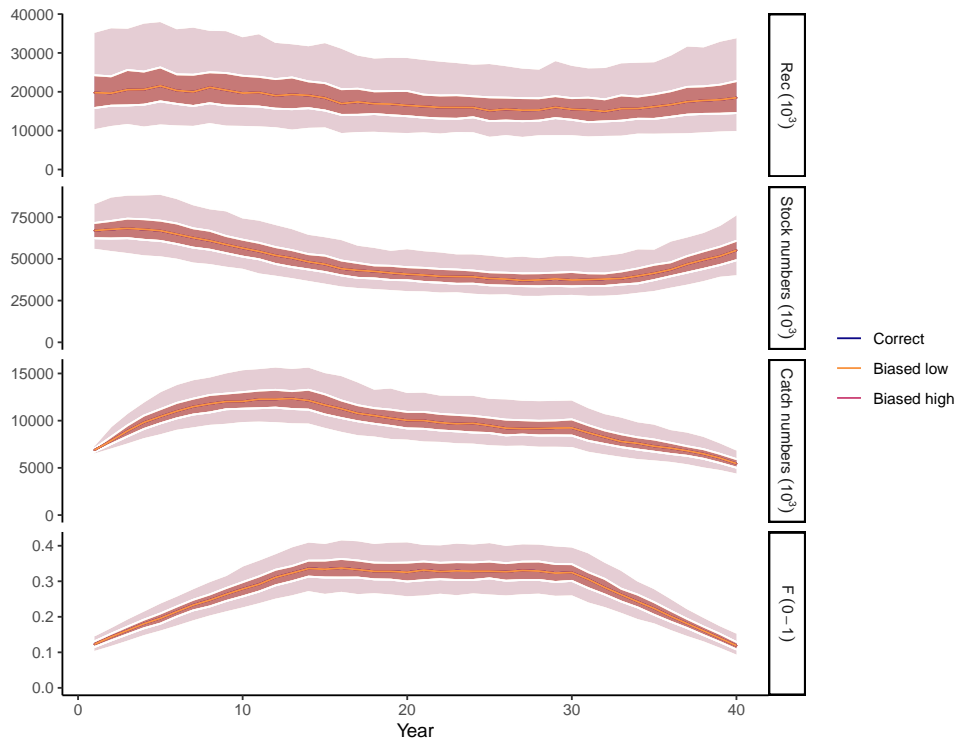


Figure 5.6: CSA stock assessment estimates on initial data years of *correct*, *biased high* and *biased low* simulated observations). Values were extremely similar so are not discernable. $F(0-1)$ is F_{bar} (mean fishing mortality rate for recruit and post-recruit stages). Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

5.3.2 Effect of age-slicing bias in MSE simulations

Age-slicing bias in the OEM on catch-at-age and CPUE-at-age had a large effect on the trajectory of operating models over the MSE projection period for a4a and CSA MPs (Figures 5.7 and 5.9).

a4a MSE results

In simulations where an a4a assessment MP was used on a *biased low* age-slicing OEM fishing mortality increased substantially to a much higher level than with the *correct* OEM and then levelled off. SSB for the *biased low* OEM did not continue the upward trajectory of the initial operating model period and began to level off almost immediately, whereas SSB increased gradually throughout the projection period for the other age-slicing OEMs (Figure 5.7).

A *biased high* OEM resulted in a lower fishing mortality and lower catch being implemented on the stock, relative to the other OEMs. Coupled with this was a steeper rise in SSB over the projection period and a higher SSB at the end of the projection period (Figure 5.7).

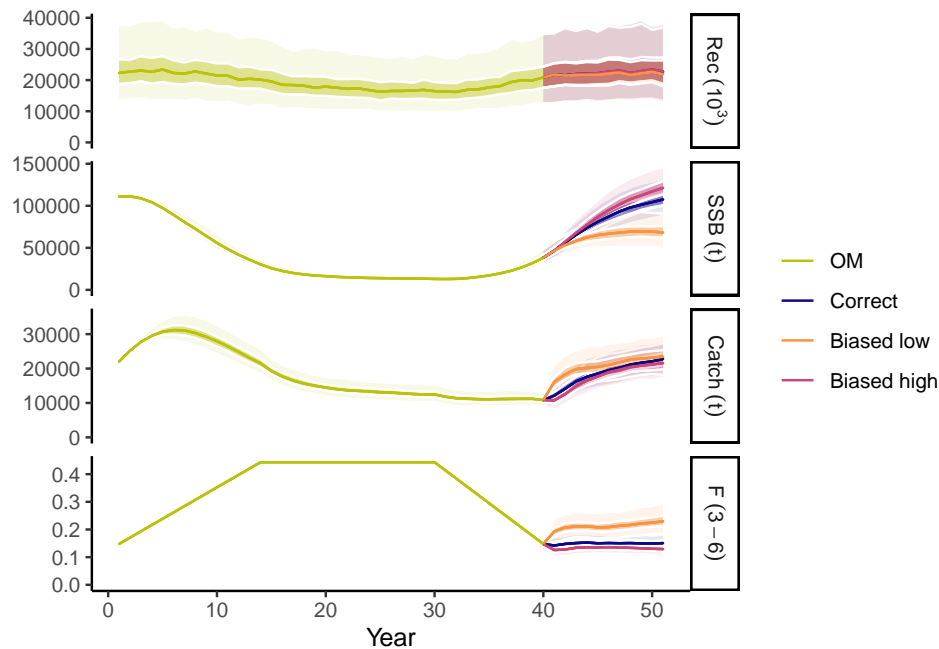


Figure 5.7: MSE simulations for a4a assessment MPs. Showing operating model projections across three OEMs (age-slicing bias: *correct*, *biased low* and *biased high*) displaying recruitment, ssb, catch and fishing mortality (fbar 3-6). Initial operating model period (OM) is also shown. Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

For age-based assessment MPs, when age-slicing bias was *biased low* or *biased high* in the OEM there was a large amount of assessment bias in the annual assessments of the stocks. A *biased low* OEM resulted in the a4a assessments consistently over estimating SSB and under estimating fishing mortality in comparison to the operating model. In contrast, the *biased high* OEM caused the a4a assessment to under-estimate SSB and over-estimate fishing mortality relative to the operating model. Assessment bias was not notable for the a4a assessment with the *correct* age-slicing OEM, however, SSB was marginally over-estimated relative to the operating model (Figure 5.8).

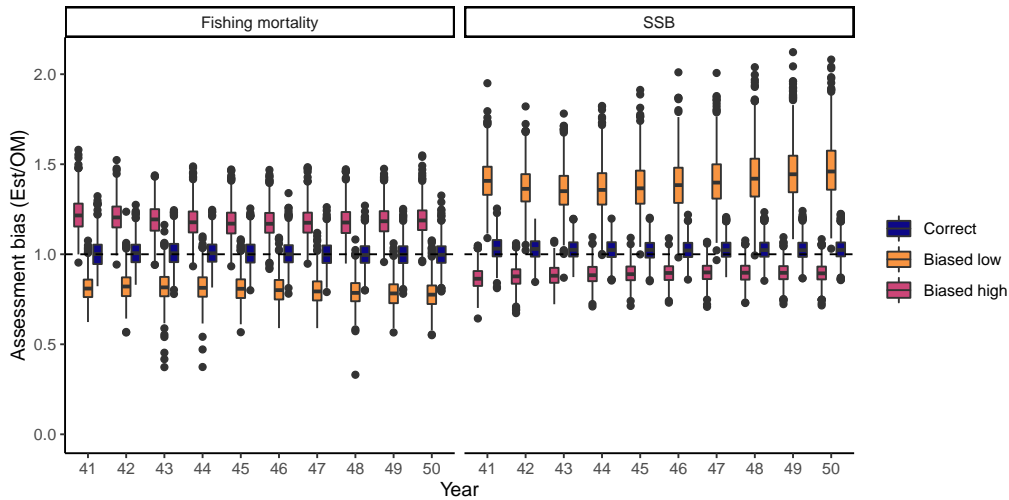


Figure 5.8: a4a assessment bias of spawning stock biomass (SSB) and fishing mortality (\bar{f} = ages 3 to 6) across three OEMs (age-slicing bias: *correct*, *biased low* and *biased high*) for MSE simulations with a4a assessment MPs. Where assessment bias = estimate/OM value. X axis refers to the year an assessment was run (e.g. assessment year= 41, data years 1:40, OM comparison = end of year 40).

CSA MSE results

Simulations with CSA assessment MPs showed the opposite effect of age-slicing bias compared to a4a MP results. SSB in the operating models showed a relatively steep increase over the projection period with the *biased low* OEM and a steady increase for the *correct* OEM. SSB for the *biased high* OEM appears to level off (Figure 5.9). Fishing mortality is highest in the operating model where the *biased high* OEM is used and lowest with the *biased low* OEM. Trajectories appear to be levelling off but still decreasing somewhat over the simulation period (Figure 5.9).

In the CSA MPs, catch and fishing mortality show a large degree of variability in the early years of the projection for all OEMs. Notably, a number of replicates have zero fishing mortality and catch in those early years. Catch and fishing mortality for the *biased high* OEM shows the aforementioned variability throughout the time period of the simulation (Figures 5.9 and 5.11).

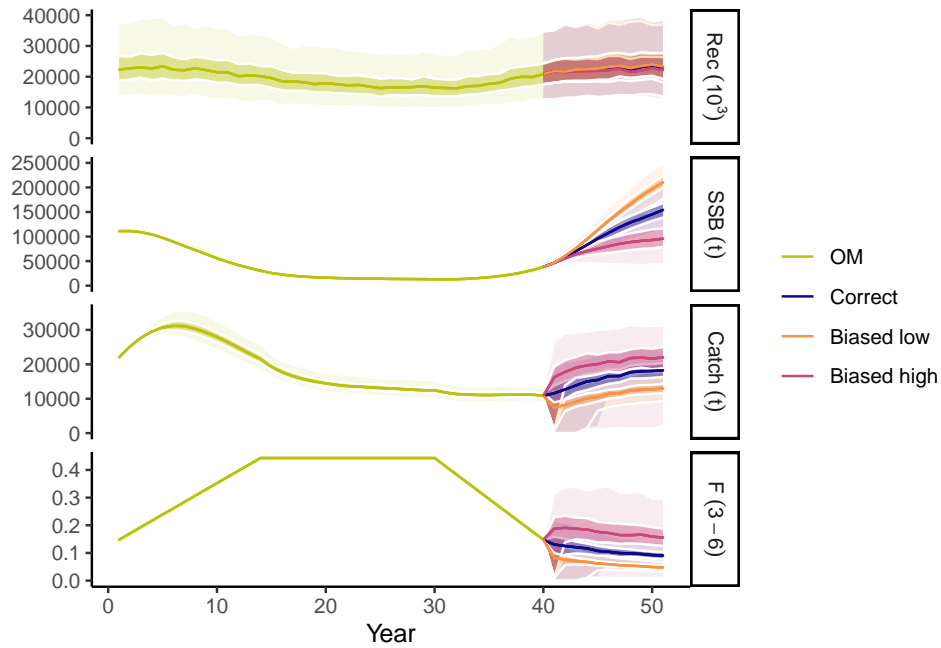


Figure 5.9: MSE simulations for CSA assessment MPs. Showing operating model projections across three OEMs (age-slicing bias: *correct*, *biased low* and *biased high*) displaying recruitment, ssb, catch and fishing mortality (fbar 3-6). Initial operating model period (OM) is also shown. Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

CSA assessment MPs showed assessment bias, demonstrating that CSA underestimated both stock numbers and fishing mortality across OEMs. There was no discernable difference in stock numbers assessment bias in the annual assessments across OEMs. There was also little difference across OEMs for fishing mortality assessment bias, however a slight difference did develop through the simulations as the operating models diverged under differing OEMs and MPS (Figure 5.10).

There were notable differences in the application of the CSA harvest control rule (given the different F_{msy} across OEMs), where fishing mortality for the following year differed between age-slicing bias in the OEMs. These differences were apparent and amplified somewhat in the implementation step converting the given fishing mortality into a total allowable catch (TAC). HCR and TAC implementation steps for the *biased high* OEM resulted in a much higher TAC on average than the *correct* age-slicing OEM. Conversely, the *biased low* age-slicing OEM implemented a much lower TAC on average than the *correct* age-slicing OEM (Figure 5.11).

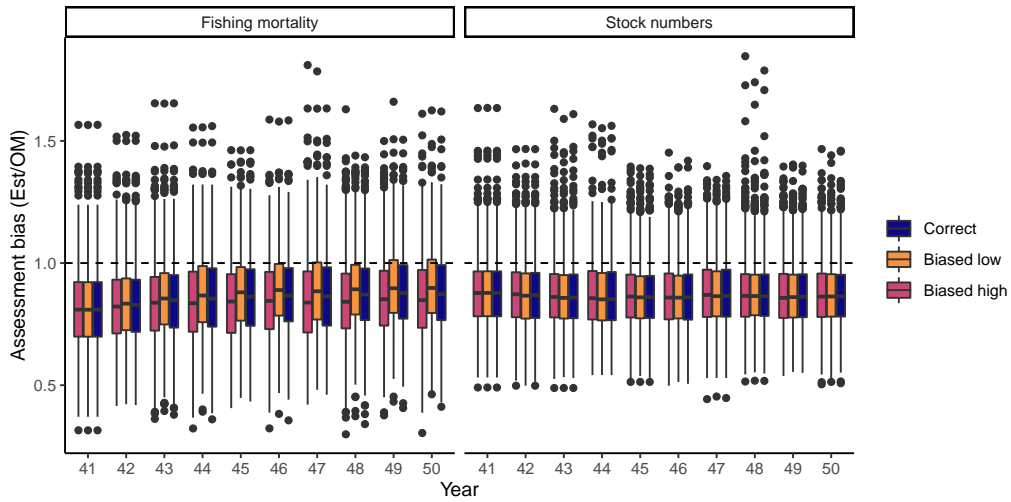


Figure 5.10: CSA assessment bias of total stock numbers and fishing mortality ($\bar{f} = 3-6$) across three OEMs (age-slicing bias: *correct*, *biased low* and *biased high*) for MSE simulations with CSA assessment MPs. Where assessment bias = estimate/OM value.

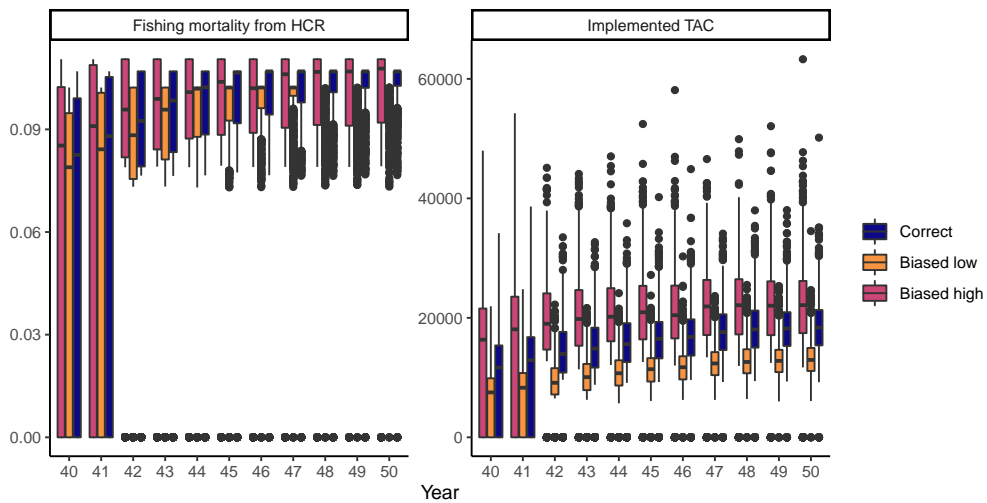


Figure 5.11: CSA MP outputs of fishing mortality from the harvest control rule (HCR) and implemented total allowable catch (TAC) across three OEMs.

5.3.3 Performance statistics comparison

The *biased high* CSA combination showed the greatest variability in stock status (B/B_{msy} and F/F_{msy}) in both the short-term and mid-term of management simulations. The a4a MP with *biased low* OEM generally maintained $F/F_{msy} > 1$ and $B/B_{msy} < 1$ over the short term and mid term with no indication of movement towards a more favourable stock status (Figures 5.12 and 5.13).

CSA MPs with *correct* and *biased low* OEMs reached and maintained F/F_{msy} generally <1 (i.e. fished below F_{msy}) in the short-term and mid-term over simulations, as did the a4a MP with *biased high* OEM. a4a MP with *correct* OEM reached maintained a distribution of F/F_{msy} values clustered around one for the short-term and mid-term (Figures 5.12 and 5.13).

CSA MPs with *correct* and *biased low* OEMs, as well as the a4a MP with *biased high* OEM had distributions of $B/B_{msy} <1$ in the short-term but by mid-term (end of simulations) B/B_{msy} values for replicates were generally >1 (i.e. SSB $>$ than B_{msy}) and actually well above in the case of CSA MPs. a4a MP with *correct* OEM had a distribution of B/B_{msy} values clustered just below one by mid-term (Figures 5.12 and 5.13).

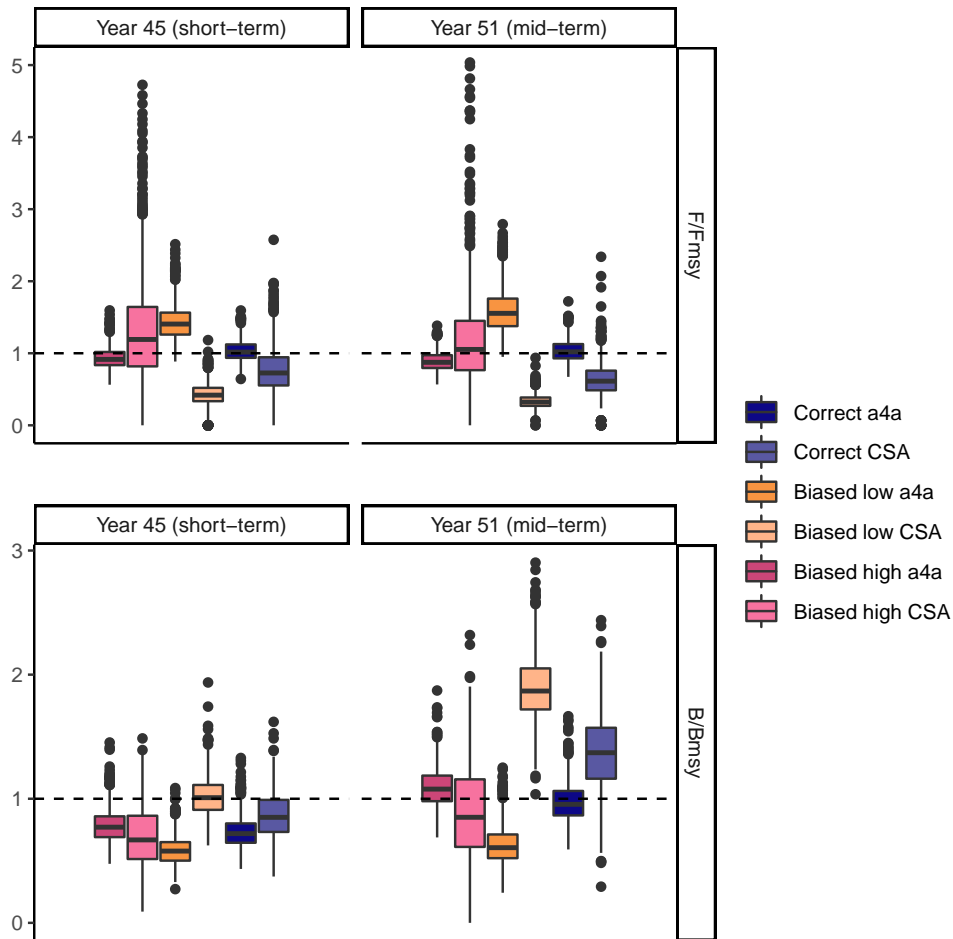


Figure 5.12: Distribution of stock status (F/F_{msy} and B/B_{msy}) in the 6th year (short-term) and final (12th) year (mid-term) of MSE simulations. Boxplots summarise replicates across management procedures (a4a and CSA) and observation error models (*correct*, *biased high* and *biased low*). A small number of F/F_{msy} values > 5 were omitted for plot clarity.

The a4a MPs with *high biased* and *correct* OEMs were closest to optimising exploitation (i.e. middle of Kobe plot) by the end of the simulation period as their mean/median B/B_{msy} and F/F_{msy} were close to the centre of the Kobe plot (Figure 5.13). CSA MPs with *biased low* and *correct* OEMs were comfortably in the under-fished section of the Kobe plot. Average B/B_{msy} and F/F_{msy} values placed CSA MP with *biased high* OEM and a4a MP with *biased low* OEM in the area of the Kobe plot where stocks are likely to be overexploited. The a4a MP with *biased low* OEM was by far the worst in terms of its position in the Kobe plot given median values of B/B_{msy} and F/F_{msy} at the end of the simulation period (Figure 5.13).

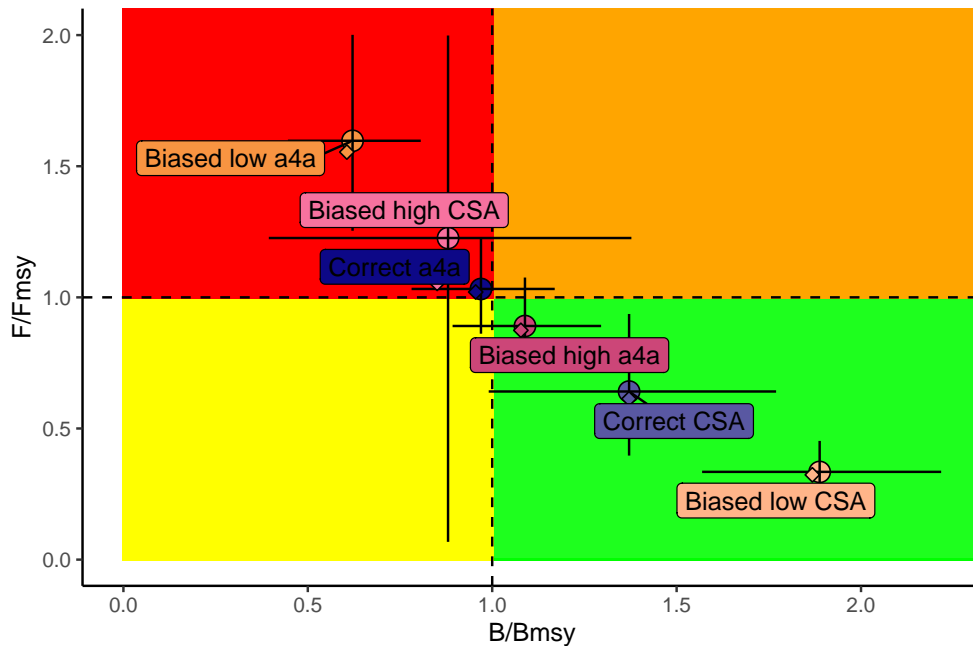


Figure 5.13: Kobe plots (B/B_{msy} vs. F/F_{msy}) for the final (12th) year of MSE simulations across management procedures (a4a and CSA) and observation error models (*correct*, *biased high* and *biased low*). Circles are means and diamonds are median values across replicates. Solid black lines show 10% - 90% quantiles across replicates for B/B_{msy} (horizontal lines) and F/F_{msy} (vertical lines).

Biased low OEM with a4a MP was the only scenario where there was no reasonable increase in the proportion of replicates above B_{msy} or below 20% of B_0 , however it did sustain the highest catches over the simulation period. Overall, for the other scenarios, mean catch increased over time with increases in the proportion of replicates above B_{msy} and decreases in the proportion below 20% of B_0 . Each MP and OEM combination did so at similar rates but

CSA MPs and *biased low* and *correct* OEMS reached very high proportions of replicates $> B_{msy}$ by the end of simulations. Only *biased low* with a CSA MP and *biased high* with a4a MP did not reach a high proportion of replicates $< 20\%$ of B_0 in over the time period (Figure 5.14).

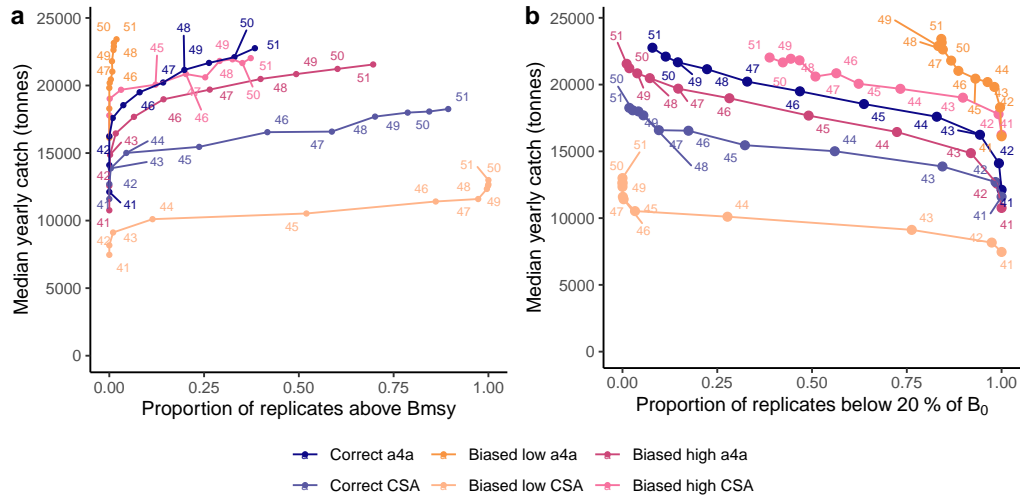


Figure 5.14: Yearly trade-offs between risk or sustainability vs mean catch biomass from MSE simulations across management procedures (a4a and CSA) and observation error models (*correct*, *biased high* and *biased low*). Labels denote year of simulations in plots.

5.4 Discussion

Age-slicing or cohort-slicing is a commonly used method for obtaining age-composition data for stock assessments (Ailloud et al., 2015). This study offers an insight into the impact that age-slicing with a biased growth model can have on the assessment and management of a stock over time with an age-based or stage-based assessment. In addition, we have also evaluated the suitability of CSA as an alternative assessment/reference point framework.

Overall, the simulation analysis demonstrated that both age-based and CSA MPs were considerably affected by bias in age-slicing, however the point where bias altered the outcomes of the MPs clearly differed. The age-based assessment was directly affected by the shifts in density of the biased age-composition data, whereas stage-based assessment were affected indirectly by the mean weights-at-age used in reference points and catch limits. Biased growth assumptions also had contrasting effects on the ability of MPs to assess the stock and implement advice. As a result, stock status and trajectory differed across MP and age-slicing combinations. *Biased low* age-slicing (assumed slower growth) with an age-based MP had a negative impact on the trajectory of the stock compared to *correct* growth, whereas with a CSA MP management was conservative in comparison to the *correct* simulations and resulted in more positive trajectory for the stock. The opposite was true for *biased high* age-slicing and the two MPs. In general, the CSA MP performed well in the *correct* and *biased low* scenarios, having a positive effect on stock status. However the CSA MP appeared to be somewhat more precautionary than the age-based MP in the *correct* scenario.

5.4.1 Effect of age-slicing bias

Biased age slicing clearly introduces a large degree of bias in the assessment and perception of the stock with an age-based MP. In the scenario where age-slicing was *biased low* the perception of stock biomass was approximately one and a half that of the real value. For the *biased high* scenario the opposite was true. If we look at the methodology we can see that these results generally reflect the changes in age densities caused by biased slicing. Whilst not a simple pattern, we can say that overall *biased high* slicing results in catch-at-age and index CPUE-at-age to be higher than the *correct* numbers in young fish and lower than the *correct* numbers in older fish (there is more nuance than this if looking at specific ages) and vice-versa for *biased low* age

slicing. These shifts in age densities are reflected in the assessment, where for example the a4a MP with *biased high* observations underestimated SSB and over-estimated fishing mortality. These differences in perception of the stock resulted in estimated reference points being biased as well, which may somewhat compensate assessment bias. Other studies also found that bias in age-composition data can alter estimated levels of the stock and also suggest an assumed faster growth could lead to greater sustained SSB and increased landings in the future or vice-versa for a slower growth rate (Lai and Gunderson, 1987; Tyler et al., 1989; Bertignac and De Pontual, 2007).

In contrast, CSA showed no obvious effect of bias age-slicing in initial assessments or assessment bias in MSE simulations. The structure of the stage-based model is likely to be the main reason why the perception of the stocks by CSA was not influenced by biased age-slicing (Collie and Sissenwine, 1983; Mesnil, 2003). Observations of recruit and post-recruit CPUE in the survey indices showed no discernable differences across age-slicing scenarios (as age 0 was scarcely influenced by the biasing process and post-recruits were aggregated) and catch numbers were aggregated, leading to extremely similar data inputs across age-slicing scenarios.

Nevertheless, CSA MPs were affected by biased age-slicing. CSA MP outputs were altered by bias age-slicing in the HCR and implementation steps of the MP, whereas age-based MPs were mainly altered by the perception of the assessment. Decisions controlled by the HCR differ somewhat across age-slicing scenarios as F_{msy} was estimated at marginally different levels across scenarios as assumptions on growth and weight-at-age were needed to calculate yield for F_{msy} and N_{msy} estimation. Weights-at age (which were biased in the biased scenarios) also influenced the implementation step as they were used to convert the fishing mortality for the next year into a TAC.

The effect of age-slicing bias indicates that certain MPs are more precautionary depending on the age-slicing scenario. Our findings indicate that, for a given stock, if there was some uncertainty in growth parameters and a possibility that they are biased the user should carefully consider what assessment should be used (Lai and Gunderson, 1987; Tyler et al., 1989; Chrysafi et al., 2019). If fish growth was suspected of being biased high, both a4a and CSA provide for good stock status with compromised yields (in the case of a4a) or higher yields (in the case of CSA). For suspected biased low growth CSA provides for good stock status but compromised yields, but a4a would lead to overexploitation

of the stock.

Interestingly, the CSA MP appears to slowly be recovering the stock even in the *biased high* scenario, indicating if uncertainty was high in growth parameters CSA would be the more likely MP to recover a stock regardless of any bias in age-composition data, albeit with reduced yield. However, the range in stock status (B/B_{msy} and F/F_{msy}) and mean catch for replicates for this MP and age-slicing scenario also demonstrate that the HCR and TAC implementation did not perform well at stabilising the stock in the time period of simulations.

5.4.2 General effectiveness of CSA

Overall, a combination of CSA assessment, CSA derived reference points and ‘ICES style’ HCR had a positive effect on stock status and stock trajectory given a *correct* scenario. However, CSA MPs were generally more precautionary than the age-based MPs, probably due to a combination of underestimating the stock and reference point estimation method. With this risk averse approach, there is a trade off with yield, as is common in fisheries management (Thorpe and De Oliveira, 2019). Nevertheless, from a management perspective this is a much preferred outcome than an over-estimation and unsustainable high yields (Wetzel and Punt, 2011; Chrysafi et al., 2019).

CSA underestimated stock numbers and fishing mortality within the MSE simulations, which is somewhat contrary to the results of Batts et al. (2022b) study where CSA gives reasonably close estimates of stock number for simulated stocks with logistic selection. The key difference with this study is the timing of survey indices is half way through the year, therefore the survey index (whilst catchability is constant across ages) was affected by the selectivity pattern of the stock. Interestingly, further investigations indicated that this issue manifests itself within the assessment by underestimating post-recruit numbers. Further exploration of this issue would be worthwhile.

Broadly speaking, MSE simulations demonstrated the reference point estimation method and subsequent HCR developed for this study could be used to recover and sustainably manage a stock under *correct* or *biased low* scenarios. However, there was a considerable number of replicates that were classified by the HCR as below N_{lim} in the early years of simulations, resulting in a zero TAC being implemented for those replicates. Further development of the N_{lim} limit reference point may be needed for management of real stocks, as it is

not often realistic to stop a fishery completely (Hilborn, 2007). Furthermore, expanding the reference point estimation method for CSA into a stochastic simulation analogous to eqsim would be a useful development (Simmonds et al., 2019; ICES, 2021).

Another issue worth considering, is whether using numbers for biological reference points is viable at all. In our study, these values were successful in managing the stock tested but in others there may be issues arising from the mismatch of numbers and biomass. In certain situations, where fishing mortality is high, a stock may have high stock numbers (driven by young fish) but low biomass (because high proportion of young fish). How this scenario would be managed under the current CSA reference point procedure developed here is difficult to ascertain without further extensive simulation testing of the procedure. We suggest that this should be the next step in bringing CSA assessment and management forward. In addition, due consideration should be given to alternative approaches to developing non-proxy biological reference points for CSA, as well as stage-based models in general.

5.4.3 Conclusions

Adequately accounting for uncertainty and bias in age-composition can be difficult in traditional statistical catch-at-age assessments (Maunder and Punt, 2013; Ailloud et al., 2015). In this study we aimed to assess the impact that biased age-composition data could have on a stock when age-based or stage-based assessment and management were implemented. We also investigated the suitability of a stage-based assessment model as an alternative assessment and management framework, developing a reference point estimation method in the process. MSEs offered a useful simulation framework to fulfill these objectives. Age-slicing, leading to biased age-composition data, was found to affect both the age-based management procedures and the stage-based management procedures, although in very different ways. A key finding here was that age-composition data generated with a faster growth rate than the true growth rate, has a more precautionary impact on management advice for a stock in comparison to a biased slow growth rate. Our study also established CSA could be a potentially useful tool for stock assessment and management in Europe under ICES guidelines in the future.

Bibliography

- Ailloud, L. E., M. W. Smith, A. Y. Then, K. L. Omori, G. M. Ralph, and J. M. Hoenig (2015). Properties of age compositions and mortality estimates derived from cohort slicing of length data. *ICES Journal of Marine Science* 72(1), 44–53.
- Batts, L., C. Minto, and H. Gerritsen (2022a). sbar: Stage-based assessments in r. R package version 0.0.0.9000.
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science* 76(7), 2150–2163.
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2022b). Numbers or mass? Comparison of theoretically different stage-based stock assessment models and their ability to model simulated and real life stocks. *Canadian Journal of Fisheries and Aquatic Sciences - (Just-IN)*.
- Beddington, J. R. and J. G. Cooke (1983). The potential yield of fish stocks. *FAO Fisheries Technical Paper* 242, 47.
- Begley, J. and D. Howell (2004). An overview of GADGET, the globally applicable area-disaggregated general ecosystem toolbox. ICES CM 2004/FF:13.
- Bertignac, M. and H. De Pontual (2007). Consequences of bias in age estimation on assessment of the northern stock of European hake (*Merluccius merluccius*) and on management advice. *ICES Journal of Marine Science* 64(5), 981–988.
- Bianchi, M., F. Chopin, T. Farme, N. Franz, C. Fuentevilla, L. Garibaldi, and A. Laurenti (2014). FAO: The State of World Fisheries and Aquaculture. *Food and Agriculture Organization of the United Nations: Rome, Italy*.

- Bull, B., R. Francis, A. Dunn, A. McKenzie, D. Gilbert, M. Smith, R. Bian, and D. Fu (2005). CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.
- Bull, B., A. D. R.I.C.C. Francis, D. G. A. McKenzie, R. B. M.H. Smith, and D. Fu (2006). CASAL (C++ Algorithmic Stock Assessment Laboratory). *CASAL User manual v2.30-2012/03/21. NIWA Technical Report 127*.
- Butterworth, D. S. (2007). Why a management procedure approach? some positives and negatives. *ICES Journal of Marine Science* 64(4), 613–617.
- Butterworth, D. S., N. Bentley, J. A. De Oliveira, G. P. Donovan, L. T. Kell, A. M. Parma, A. E. Punt, K. J. Sainsbury, A. D. Smith, and T. K. Stokes (2010). Purported flaws in management strategy evaluation: basic problems or misinterpretations? *ICES Journal of Marine Science* 67(3), 567–574.
- Cadrin, S. X. (2000). Evaluating two assessment methods for Gulf of Maine northern shrimp based on simulations. *Journal of Northwest Atlantic Fishery Science* 27, 119–132.
- Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick, and J. Cope (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research* 153, 48–68.
- Carruthers, T. R., J. F. Walter, M. K. McAllister, and M. D. Bryan (2015). Modelling age-dependent movement: an application to red and gag groupers in the Gulf of Mexico. *Canadian Journal of Fisheries and Aquatic Sciences* 72(8), 1159–1176.
- Chrysafi, A., J. M. Cope, and A. Kuparinen (2019). Eliciting expert knowledge to inform stock status for data-limited stock assessments. *Marine Policy* 101, 167–176.
- Collie, J. S. and M. P. Sissenwine (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* 40(11), 1871–1879.
- Cook, R. M. and M. R. Heath (2018). Population trends of bycatch species reflect improving status of target species. *Fish and Fisheries* 19(3), 455–470.
- de la Mare, W. K. (1986). Simulation studies on management procedures. Technical report, Reports of the International Whaling Commission, Volume 36, pg. 429–450.

- Deroba, J., D. S. Butterworth, R. Methot, J. De Oliveira, C. Fernandez, A. Nielsen, S. Cadrin, M. Dickey-Collas, C. Legault, J. Ianelli, et al. (2015). Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. *ICES Journal of Marine Science* 72(1), 19–30.
- Dichmont, C. M., A. R. Deng, A. E. Punt, W. Venables, and M. Haddon (2006). Management strategies for short lived species: the case of Australia’s Northern Prawn Fishery: 2. Choosing appropriate management strategies using input controls. *Fisheries Research* 82(1-3), 221–234.
- Doonan, I., K. Large, A. Dunn, S. Rasmussen, C. Marsh, and S. Mormede (2016). Casal2: New Zealand’s integrated population modelling tool. *Fisheries Research* 183, 498–505.
- Fournier, D. A., J. Hampton, and J. R. Sibert (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences* 55(9), 2105–2116.
- Fournier, D. A., J. R. Sibert, J. Majkowski, and J. Hampton (1990). MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian journal of fisheries and aquatic sciences* 47(2), 301–317.
- Fournier, D. A., J. R. Sibert, and M. Terceiro (1991). Analysis of length frequency samples with relative abundance data for the Gulf of Maine northern shrimp (*Pandalus borealis*) by the MULTIFAN method. *Canadian Journal of Fisheries and Aquatic Sciences* 48(4), 591–598.
- Fridriksson, Á. (1934). On the calculation of age-distribution within a stock of cod by means of relatively few age determinations as a key to measurements on a large scale. *Rapports et Procbs-Verbaux des Reunions, Conseil Permanent Intemational pour l’Exploration de la Mer*. 86(2).
- Fulton, E. A., A. D. Smith, D. C. Smith, and P. Johnson (2014). An integrated approach is needed for ecosystem based fisheries management: insights from ecosystem-level management strategy evaluation. *PloS one* 9(1), e84242.

- Garcia, S. (1995). The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review. *FAO Technical Paper 350*.
- Hilborn, R. (2007). Managing fisheries is managing people: what has been learned? *Fish and Fisheries* 8(4), 285–296.
- Hilborn, R., R. O. Amoroso, C. M. Anderson, J. K. Baum, T. A. Branch, C. Costello, C. L. de Moor, A. Faraj, D. Hively, O. P. Jensen, H. Kurota, L. R. Little, P. Mace, T. McClanahan, M. C. Melnychuk, C. Minto, G. C. Osio, A. M. Parma, M. Pons, S. Segurado, C. S. Szuwalski, J. R. Wilson, and Y. Ye (2020). Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences* 117(4), 2218–2224.
- Hilborn, R., J.-J. Maguire, A. M. Parma, and A. A. Rosenberg (2001). The precautionary approach and risk management: can they increase the probability of successes in fishery management? *Canadian Journal of Fisheries and Aquatic Sciences* 58(1), 99–107.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.
- Hoenig, J. M. and D. M. Heisey (1987). Use of a log-linear model with the em algorithm to correct estimates of stock composition and to convert length to age. *Transactions of the American Fisheries Society* 116(2), 232–243.
- Holland, D. S. (2010). Management strategy evaluation and management procedures: tools for rebuilding and sustaining fisheries. *OECD Food, Agriculture and Fisheries Papers, No. 25, OECD Publishing, Paris*.
- Hordyk, A., K. Ono, S. Valencia, N. Loneragan, and J. Prince (2015). A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science* 72(1), 217–231.
- Ibaibarriaga, L., C. Fernández, A. Uriarte, and B. A. Roel (2008). A two-stage biomass dynamic model for Bay of Biscay anchovy: a Bayesian approach. *ICES Journal of Marine Science* 65(2), 191–205.
- ICES (2019). Advice basis. *In Report of the ICES Advisory Committee, 2019. ICES Advice 2019, section 1.2.*

- ICES (2019a). Report of the Working Group for the Bay of Biscay and the Iberian Waters Ecoregion (WGBIE).
- ICES (2019b). White anglerfish (*Lophius piscatorius*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2019. ICES Advice 2019, mon.27.78abd.
- ICES (2021). ICES fisheries management reference points for category 1 and 2 stocks; technical guidelines. *In Report of the ICES Advisory Committee, 2021. ICES Advice 2021, Section 16.4.3.1.*
- Jardim, E., C. P. Millar, I. Mosqueira, F. Scott, G. C. Osio, M. Ferretti, N. Alzorriz, and A. Orío (2014). What if stock assessment is as simple as a linear model? the a4a initiative. *ICES Journal of Marine Science* 72(1), 232–236.
- Kell, L. (2018). *FLife: Methods for Modelling Life History Traits*. R package version 3.3.0.
- Kell, L., G. Pilling, G. Kirkwood, M. Pastoors, B. Mesnil, K. Korsbrekke, P. Abaunza, R. Aps, A. Biseau, P. Kunzlik, et al. (2005). An evaluation of the implicit management procedure used for some ICES roundfish stocks. *ICES Journal of Marine Science* 62(4), 750–759.
- Kell, L. T., J.-M. Fromentin, and S. Bonhommeau (2012). An evaluation of the implications of population structure on the current bluefin tuna advice framework. *Collective Volume of Scientific Papers, ICCAT* 68(1), 370–378.
- Kell, L. T., G. Merino, P. De Bruyn, J. O. de Urbina, H. Arrizabalaga, and J. S. H. Murua (2014). An example management strategy evaluation of a harvest control rule. *Collective Volume of Scientific Papers, ICCAT* 70(5), 2096–2110.
- Kell, L. T., I. Mosqueira, P. Grosjean, J.-M. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. Pastoors, J. Poos, et al. (2007). FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science* 64(4), 640–646.
- Kell, L. T. and M. Ortiz (2011). A comparison of statistical age estimation and age slicing for atlantic bluefin tuna (*thunnus thynnus*). *Collect. Vol. Sci. Pap. ICCAT* 66(2), 948–955.

- Kienzle, M., A. J. Courtney, and M. F. O'Neill (2014). Environmental and fishing effects on the dynamics of brown tiger prawn (*Penaeus esculentus*) in Moreton Bay (Australia). *Fisheries research* 155, 138–148.
- Kirkwood, G. (1992). Background to the development of revised management procedures. *Report of the International Whaling Commission* 42, 236–243.
- Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic Differentiation and Laplace approximation. *Journal of Statistical Software* 70(5), 1–21.
- Kvamsdal, S. F., A. Eide, N.-A. Ekerhovd, K. Enberg, A. Gudmundsdottir, A. H. Hoel, K. E. Mills, F. J. Mueter, L. Ravn-Jonsen, L. K. Sandal, et al. (2016). Harvest control rules in modern fisheries management. *Elementa: Science of the Anthropocene* 4.
- Lai, H. L. and D. R. Gunderson (1987). Effects of ageing errors on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). *Fisheries Research* 5(2-3), 287–302.
- Li, Y., L. M. Lee, and J. Rock (2019). Modeling population dynamics and nonstationary processes of difficult-to-age fishery species with a hierarchical Bayesian two-stage model. *Canadian Journal of Fisheries and Aquatic Sciences* 76(12), 2199–2214.
- Mace, P. M. (2001). A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. *Fish and fisheries* 2(1), 2–32.
- Marasco, R. J., D. Goodman, C. B. Grimes, P. W. Lawson, A. E. Punt, and T. J. Quinn II (2007). Ecosystem-based fisheries management: some practical suggestions. *Canadian Journal of Fisheries and Aquatic Sciences* 64(6), 928–939.
- Maunder, M. N. and K. R. Piner (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science* 72(1), 7–18.
- Maunder, M. N. and A. E. Punt (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research* 142, 61–74.

- Mesnil, B. (2003). The Catch-Survey Analysis (CSA) method of fish stock assessment: an evaluation using simulated data. *Fisheries Research* 63(2), 193–212.
- Mesnil, B. (2012). The hesitant emergence of maximum sustainable yield (MSY) in fisheries policies in Europe. *Marine Policy* 36(2), 473–480.
- Methot, R. D. and C. R. Wetzel (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142, 86–99.
- Mildenberger, T. K., C. W. Berg, A. Kokkalis, A. R. Hordyk, C. Wetzel, N. S. Jacobsen, A. E. Punt, and J. R. Nielsen (2021). Implementing the precautionary approach into fisheries management: Biomass reference points and uncertainty buffers. *Fish and Fisheries*.
- Mosqueira, I. and E. Jardim (2020). The mse package. *Joyournal of Open Source Software* X, XX.
- Myers, R., A. Rosenberg, P. Mace, N. Barrowman, and V. Restrepo (1994). In search of thresholds for recruitment overfishing. *ICES Journal of Marine Science* 51(2), 191–205.
- NESFC (2014). 58th Northeast Regional Stock Assessment Workshop (58th SAW) Assessment Summary Report. US Department of Commerce. NEFSC (Northeast Fisheries Science Center). *Northeast Fisheries Science Center Ref. Doc. 14-03, 44 pp.*
- Nielsen, A. and C. W. Berg (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158, 96–101.
- NOAA (2019). NOAA Fish and Fisheries Toolbox: Collie-Sissenwine Analysis (CSA) version 4.3.
- Patterson, K., R. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D. W. Skagen, et al. (2001). Estimating uncertainty in fish stock assessment and forecasting. *Fish and fisheries* 2(2), 125–157.
- Pedersen, M. W. and C. W. Berg (2017). A stochastic surplus production model in continuous time. *Fish and Fisheries* 18(2), 226–243.
- Piner, K. R., O. S. Hamel, J. L. Menkel, J. R. Wallace, and C. E. Hutchinson (2005). Age validation of canary rockfish (*Sebastes pinniger*) from off the

- Oregon coast (USA) using the bomb radiocarbon method. *Canadian Journal of Fisheries and Aquatic Sciences* 62(5), 1060–1066.
- Prager, M. (1992). Aspic-a surplus-production model incorporating covariates. *Collective Volume of Scientific Papers, ICCAT 28*, 218–229.
- Privitera-Johnson, K. M. and A. E. Punt (2020). A review of approaches to quantifying uncertainty in fisheries stock assessments. *Fisheries Research* 226, 105503.
- Punt, A. E. (2006). The FAO precautionary approach after almost 10 years: have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management? *Natural Resource Modelling* 19(4), 441–464.
- Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. De Oliveira, and M. Haddon (2014). Management strategy evaluation: best practices. *Fish and Fisheries* 17(2), 303–334.
- Punt, A. E., T. Huang, and M. N. Maunder (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70(1), 16–33.
- Punt, A. E., G. N. Tuck, J. Day, P. Burch, R. B. Thomson, and P. Bessell-Browne (2021). The impact of alternative age-length sampling schemes on the performance of stock assessment methods. *Fisheries Research* 238, 105904.
- Shepherd, J. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science* 56(5), 584–591.
- Silvar-Viladomiu, P., C. Minto, G. Halouani, L. Batts, D. Brophy, C. Lordan, and D. G. Reid (2021). Moving reference point goalposts and implications for fisheries sustainability. *Fish and Fisheries*.
- Simmonds, J., E. Hjørleifsson, and C. Millar (2019). *msy: Estimation of Equilibrium Reference Points for Fisheries*. R package version 0.1.19.
- Smith, A. (1994). Management strategy evaluation: the light on the hill. *Population dynamics for fisheries management*, 249–253.

- Smith, M. T. and J. T. Addison (2003). Methods for stock assessment of crustacean fisheries. *Fisheries Research* 65(1-3), 231–256.
- Szuwalski, C. S. and A. E. Punt (2013). Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. *ICES Journal of Marine Science* 70(5), 955–967.
- Taylor, M. and T. K. Mildenerger (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology* 24(4), 330–338.
- Thorpe, R. B. and J. A. De Oliveira (2019). Comparing conceptual frameworks for a fish community MSY (FCMSY) using management strategy evaluation—an example from the North Sea. *ICES Journal of Marine Science* 76(4), 813–823.
- Thorson, J. (2019). *FishLife: Predict Life History Parameters For Any Fish*. R package version 2.0.0.
- Tyler, A., R. Beamish, and G. McFarlane (1989). Implications of age determination errors to yield estimates. *Canadian Special Publication of Fisheries and Aquatic Sciences* 108, 27–35.
- Tyszko, S. M. and J. J. Pritt (2017). Comparing otoliths and scales as structures used to estimate ages of Largemouth Bass: consequences of biased age estimates. *North American Journal of Fisheries Management* 37(5), 1075–1082.
- Walsh, J. C., C. Minto, E. Jardim, S. C. Anderson, O. P. Jensen, J. Afflerbach, M. Dickey-Collas, K. M. Kleisner, C. Longo, G. C. Osio, et al. (2018). Trade-offs for data-limited fisheries when using harvest strategies based on catch-only models. *Fish and Fisheries* 19(6), 1130–1146.
- Wetzel, C. R. and A. E. Punt (2011). Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. *Fisheries Research* 110(2), 342–355.
- Worm, B., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings, et al. (2009). Rebuilding global fisheries. *science* 325(5940), 578–585.
- Zhang, Z. and K. Fong (2021). Estimation of key population parameters and msy-based reference points for sidestripe shrimp (*Pandalopsis dispar*) in the Fraser River Delta, British Columbia. *Fisheries Research* 238, 105893.

Chapter 6

General Discussion

Tracking cohorts is key to many different stock assessment models. Multiple factors can influence our ability to track cohorts, such as: available data, type of stock assessment, exploitation pattern, recruitment deviations, and life-history characteristics of a fish stock. The main aims of my thesis were to: develop methods to address cohort uncertainty in fisheries stock assessment and management (Chapters 2-4), as well as developing a framework for testing how uncertainty in the form of biased age-composition data may influence stock status and management goals over time (Chapter 5). Cohort uncertainty in the form of age-estimation and growth was specifically focussed upon. Anglerfish *Lophius piscatorius* and *Lophius budegassa* stocks in the Celtic Seas and Bay of Biscay (ICES subareas 7, 8.a-b and 8.d) were used throughout as case study stocks.

Chapters 2-5 addressed these objectives. Here, the results are placed in overall context of current fisheries stock assessment and management, as well as specifically for anglerfish, with a view to further improvements on the basis of the understanding developed.

6.1 Summary of main thesis results

Chapter 2 developed a novel method of length frequency analysis called LFEM, or Length Frequency analysis with the Expectation-Maximisation algorithm. This method is a maximum likelihood procedure that uses mixture models and the Expectation-Maximisation (EM) algorithm to estimate von Bertalanffy growth parameters from length frequency data from survey data. The basic version of LFEM fits mixture models over all years and multiple surveys to give a single set of estimates for von Bertalanffy growth parameters. The basic model was extended into a hierarchical framework, where growth variability could be modelled through bivariate random effects on cohort specific growth parameters. Testing on haddock and anglerfish determined that the basic model gave reasonable estimates of growth parameters and their associated uncertainty. The hierarchical extension successfully modelled growth variability in cohorts. Overall this method proved to be a robust and accurate method of length frequency analysis, providing a useful alternative to other widely-used analyses. Along with being published in a peer-reviewed paper (Batts et al., 2019), LFEM was presented at the ICES stock assessment benchmark workshop on anglerfish in ICES areas and contributed to the development of an ICES assessment for white-bellied anglerfish in the Celtic Sea and Bay of

Biscay (WKANGLER) (ICES, 2018) (*Supplementary Appendix C*).

Chapter 3 implemented and compared the performance of two theoretically different stage-based stock assessment models: the well-known numbers-based depletion model, CSA, first developed by Collie and Sissenwine (1983) and currently available in the NOAA Fish and Fisheries Integrated Toolbox (NOAA, 2019); and a biomass-based delay-difference model first described in a theoretical paper by Schnute (1987). Model performance was compared across an extensive simulation-testing framework, as well as the white-bellied anglerfish in the Celtic Sea and Bay of Biscay. Results from the simulation-testing demonstrated that the biomass-based delay-difference (Schnute) models could estimate close to the true values of stocks in certain scenarios, however performance and convergence were sensitive to selectivity assumptions and growth/mean weight configurations. Overall, CSA was more robust to selectivity assumptions, converging consistently with reasonable accuracy, but had inherently higher uncertainty than Schnute models. CSA estimates of stock numbers for the real anglerfish stock were remarkably close to the estimates from the official age-based assessment. This work was presented as a working document at the ICES Working Group for Bay of Biscay and Iberian Waters Ecoregion (WGBIE) (*Supplementary Appendix E*) and has been published (Batts et al., 2022).

An R package (‘sbar’) was developed to enable easy application of the stage-based assessment models implemented in Chapter 3. The ‘sbar’ package is available on github and the manual can be found in *Appendix F*. Chapter 4 is a vignette that documents key functions and demonstrates the different models to users, including the various settings and variants of the models that could be implemented. Details on the Schnute model versions that were not tested in Chapter 3 are given, outlining how the user can adjust the relative importance of recruits or previously-exploited biomass, as well run a more classical version of the model. The case study stock for this vignette was the black-bellied anglerfish stock for the Celtic Seas and Bay of Biscay, providing preliminary results on estimated stock size and fishing mortality trends for this stock, which has not been analytically assessed before.

Chapter 5 uses a management strategy evaluation (MSE) framework to investigate the impact of biased age-slicing of length-based data on fisheries stock assessment and management over time. The operating model was a life-history based simulated stock with similar characteristics to the white-bellied

anglerfish stock in the Celtic Seas and Bay of Biscay. A biased age-slicing procedure was applied to both survey and catch observations, generating biased age-composition data in the operating error model (OEM). Two different management procedures (MPs) were tested within the simulation framework. An MP based upon a statistical catch-at-age model (a4a) and an MP based upon a stage-based model (CSA). A procedure for estimating non-proxy reference points for CSA was also developed in order to implement a harvest control rule based on the output of the CSA assessment. Results demonstrated that biased age-composition data impacted the contrasting MPs in opposite ways. Age-based assessment models were affected directly, with biased age-composition data altering the perception and in turn the management of the stock. The stage-based assessment model was robust to biased age-composition, with perception of the stock similar across age-slicing scenarios. However, bias did enter indirectly into the CSA MP through the use of biased mean weights-at-age in the reference point estimation and the implementation of the harvest control rule.

In the following sections, the methods developed in this thesis are discussed in the context of their application to future fisheries stock assessments, as well as the wider implications of the findings on stock assessment and management over time.

6.2 Applications to fisheries stock assessment

6.2.1 Acquiring reliable growth information from length frequency data

Often it is the reliability and availability of data that determines the choice of stock assessment model for fisheries (Hilborn and Walters, 1992). In many fisheries, due to the extensive use of age-based stock assessment models (Punt et al., 2013), reliable age-composition data is an essential component (Hoggarth, 2006; Maunder et al., 2016). Typically, age-composition data is generated by converting length-structured data into age-structured data through an age-length key (ALK) (Fridriksson, 1934; Hoenig and Heisey, 1987; Maunder and Piner, 2015). Age-estimates are needed to build an ALK, however there are often situations in fisheries where age-estimates are not reliable (e.g. unclear incremental growth rings) or collection is not feasible (e.g. lack of resources or time) (Hilborn and Walters, 1992). If an age-structured stock

assessment model is the desired approach, an alternative method for estimating growth and generating age-composition data is needed. Length frequency analysis is one well-known method for estimating fish growth and processing numbers-at-length data prior to running an assessment model (Hasselblad, 1966; Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017; Batts et al., 2019). Growth variability in fish is also an extensive area of research (Quinn and Deriso, 1999; Wang, 1999; Eveson et al., 2015; Morrongiello and Thresher, 2015; Cadigan et al., 2016), however there have been few studies that have modelled growth variability from length frequency data (Fournier et al., 1998; Roa-Ureta, 2010; Lloyd-Jones et al., 2016).

There was much scope for development of a new method of length frequency analysis for fisheries survey data as part of this thesis. Chapter 2 introduced an alternative approach to estimating growth parameters from length frequency data, in addition to modelling variability in fish growth. Whilst widely-used analyses, such as MULTIFAN (Fournier et al., 1990) and ELEFAN (Taylor and Mildenerger, 2017) are reliable and useful procedures in their own right, there could be circumstances where the LFEM method may be preferred. For example, in situations where the user wanted more flexibility in the set up of the model, LFEM may be useful, as the method does not require the fixing of L_{∞} or binning of fish lengths prior to running the procedure.

Another situation where LFEM could prove effective would be where there are notable differences in the structure of periodic length frequency data, possibly caused by growth variability between fish cohorts or years. Taking advantage of the functionality of the automatic differentiation framework TMB (Kristensen et al., 2016), the basic LFEM model was extended into a hierarchical framework, which estimated growth variability with bivariate random effects. The LFEM hierarchical model inherently tracks cohorts, accounting for growth variability, without explicitly modelling the changes in mixing proportions of the mixture models. This approach differs somewhat to that used in the integrated assessment MULTIFAN-CL to account for growth variability, where density dependence is incorporated into the model to track cohorts (Fournier et al., 1998). This is a useful attribute of MULTIFAN-CL, however because it is an integrated assessment it is difficult to disentangle the fitting of length frequency data from the population dynamics. On the contrary, the LFEM approach is useful because it doesn't include explicit cohort strength tracking. Numbers-at-age can be examined after length frequency analysis to

determine good cohort tracking.

Accounting for variability in cohort growth was an important aspect of LFEM and is particularly useful in the wider context of fisheries stock assessment, where cohort specific growth curves could be used to improve the quality of age-composition data by tailoring the age-composition data to differences in the growth of cohorts (Whitten et al., 2013). In other situations, estimated growth parameters from the slowest and fastest growing cohorts could be a useful guide for bounds of growth parameters estimated in assessment models or tested within a simulation framework such as an MSE. It may also be possible to expand the methodology into an assessment method and this is discussed in section 6.4.

Although LFEM could be considered less subjective than other length frequency conversion methods, consideration needs to be given to the age of the first component ('age0' in the model), which is needed to run the model. This input is used internally by the model to convert the estimated parameters into the classical von Bertalanffy growth parameters and, as such, it has no bearing on the estimated means of the components of the mixture model. However, care should be taken (supported by expert knowledge) in situations where the user is not certain of age0, as estimated parameters would differ depending on the value defined by the user. While age-estimation might be difficult for a species, this issue could be addressed through tag-recapture studies (Hamel et al., 2014) or daily increment analysis (Hislop et al., 2001; Brophy et al., 2021) to validate age0. It is worth noting here, the value of age0 may not be that important in the main class of assessment models investigated in Chapters 3-5.

A notable limitation of LFEM was that the iterative procedure is relatively slow and the automatic differentiation of TMB was used to improve the maximisation. LFEM was also somewhat sensitive to starting parameter values, occasionally estimating growth parameters that were biologically unrealistic. The EM algorithm is a well-known procedure for fitting mixture models and obtaining latent variables (i.e. latent variable in this case being the component of the length-frequency data that a fish of a given length is likely to belong to) (McLachlan and Peel, 2004). Convergence of the EM algorithm is generally slow but stable (Varadhan and Roland, 2008) and is also known to be sensitive to the choice of starting parameters (Biernacki et al., 2003), which could explain the sensitivity of LFEM. Correlation of growth parameters could also

be the cause, as this issue arises in other length frequency models and is typically dealt with by initially fixing L_∞ (Fournier et al., 1990; Taylor and Mildenerger, 2017). To counteract this sensitivity for LFEM, the procedure outlined in Chapter 2 was recommended, where a set of sensitivity runs were used in combination with the model selection criteria (Batts et al., 2019).

Finally, it may be the case that LFEM is somewhat dependent on reasonably defined components in the length frequency data. Lack of clear components in larger length classes may lead to higher uncertainty in the growth estimates. In these cases, LFEM can still be useful, as the variability (i.e. mean of the normal distribution) of the first component in length frequency data, which is usually well-defined, can still be modelled with the hierarchical extension. This analytical method for identifying the recruitment stage in yearly survey data could provide improved stage-composition data, which in turn could be used in stage-based assessment models such as those developed in Chapter 3 (Batts et al., 2022).

6.2.2 Stage-based assessment models: a diverse and useful middle ground

Many different stage-based assessment models have been developed in various different forms (biomass-based and numbers-based) and varying levels of complexity (Deriso, 1980; Collie and Sissenwine, 1983; Schnute, 1985, 1987; Ibaibarriaga et al., 2008; Zhou et al., 2011; Kienzle et al., 2014; Cook and Heath, 2018; Li et al., 2019). Many of these models are tailored to particular stocks or only tested on single stocks, but these studies show that stage-based models, whilst comparatively under-utilised, can be reliable stock assessment models. In addition to demonstrating the ability of stage-based models to perform well on the real stock of white-bellied anglerfish in Celtic Sea and Bay of Biscay, Chapter 3 has also given further insight into how well stage-based models perform more generally. The comprehensive simulation-testing framework in Chapter 3 (Batts et al., 2022) demonstrated that implementations of CSA and the *Adapted Schnute Observation Error model* could estimate stock size to a good degree of accuracy and fishing mortality to a lesser extent across a range of simulated stock scenarios.

Chapter 3 gave an opportunity to compare two theoretically different stage-based assessment models (numbers-based or biomass-based), however, it was difficult say which model was the ‘best’ overall. The answer was not simple,

as the *Adapted Schnute Observation Error model* was the min-max solution in more scenarios than CSA, whereas CSA was much more consistent in its convergence and accuracy. As is common in the field of fisheries stock assessment, the preferred assessment model is dependent on the attributes of the stock. As CSA was more robust, it lends itself to situations where little tailoring is needed in the implementation of the assessment model. The *Adapted Schnute Observation Error model* requires more thought into the validity of the growth/mean weight configuration, but the reward is lower uncertainty in estimates and no estimation of yearly recruitment parameters. Notwithstanding, the closeness of the CSA stock estimates to the estimates from the official age-based assessment for white-bellied anglerfish stock adds extra weight to the argument for CSA.

The combination of traditional age-based assessment models with methods such as LFEM is one approach to assessing fisheries where age-estimates are unreliable (e.g. biased) or not possible, however, other forms of stock assessment model with different data requirements are often preferred instead (Smith and Addison, 2003; Punt et al., 2013). Stage-based models offer an alternative middle ground between aggregate and compositional models, incorporating information on recruitment but with simpler population dynamics, suiting situations where data is lacking or subject to a high degree of uncertainty. They are also likely to be more robust to assumption violations in comparison to more complex models. Validation of CSA, specifically to its robustness to cohort uncertainty in the form of biased age-composition data, was demonstrated in Chapter 5. Here, fitting of the CSA assessment model to either true or biased age-composition data resulted in almost exactly the same estimates of the stock. The CSA assessment model was not affected by changes in age-composition due to the simplified internal structure of the model.

In general, findings in Chapter 3 indicated that the stage-based models could be used for assessments on real stocks in the future. Not only on anglerfish stocks, but in situations where age-composition data is not reliable and where there is not sufficient data or expertise to apply more complex integrated assessments such as Stock Synthesis (Methot and Wetzel, 2013). The stage-based assessment models tested here are a form of relatively simple model that are likely to be robust to assumption violations and issues with uncertainty or bias in the data. In addition, they are comparatively easy to run and check diagnostics. Stage-based assessment models also make good candidates

for secondary assessment models, which can be used to validate a primary assessment that is likely to have more complex population dynamics and assumptions (Mesnil, 2003). However, conducting an extra assessment in the often time-constrained fisheries stock assessment and advice process may be a luxury that is not always feasible. There are also likely to be some potential barriers to the immediate adoption of these models for stock assessment: the lack of development and testing of biological reference points (although this is somewhat addressed for CSA in Chapter 5), the fact these models are not state-space models, and the assumption that the fishery is fully selected. If these points could be addressed in the further development of this class of assessment models then there would likely be a wider take-up.

To the author's knowledge this is the first time the approach found in the theoretical paper by Schnute (1987) has been implemented. The population dynamics of this model were intriguing as this approach allowed accurate modelling of the stock without estimating recruitment parameters or assumptions on the spawning proportion of the stock. However, a caveat for this type of population dynamics is it is heavily dependent on the stage mean fish weights and their reliability. Evidence of this impact could be seen with the effect the growth/mean weight configuration had on both the convergence and accuracy of the *Adapted Schnute Observation Error model*. Whilst further investigation into growth/mean weight configuration is needed, the recommended configuration is growth parameters estimated from a linear regression on stage mean fish weights paired with sampled stage mean fish weights that are unlikely to be heavily skewed by selectivity pattern (e.g. mean weights from a survey).

An important point to note was that self-tests indicated that the implementation of the *Original Schnute Process Error model* was not a consistent or reliable assessment model and was intrinsically biased (Deroba et al., 2015). This was unfortunate and cause for concern, particularly because the model showed signs of divergence even when there was no observation error on the survey index or catch biomass. As recommended in Deroba et al. (2015), the first step was to thoroughly check the code, although no issues were found. This evidence suggests that the model (or more specifically this interpretation and implementation of the theoretical model of Schnute (1987)) simply may not be an adequate assessment model due to underlying structural uncertainty (Deroba et al., 2015). The importance of self-tests is highlighted here, as the *Original Schnute Process Error model* performed reasonably well

in some instances across the simulation-testing framework and on the real anglerfish stock, but due to the evidence from the self-tests, the model cannot be recommended for use in its current form.

In Chapter 4, an R package and vignette for implementing the stage-based assessment models from Chapter 3 were developed. Both the vignette and package are a key development, as they enable ease of use of the stage-based assessment models that were developed for Chapter 3 and also encourage other users to test the models with different case study fish stocks (Lortie et al., 2020). Assessment models would be more likely to be adopted for real stocks if there is accessible documentation and examples of the assessment models in use.

The vignette demonstrates the stage-based assessment models on the black-bellied anglerfish stock in the Celtic Seas and Bay of Biscay, which is an important stock currently managed with survey trends (ICES, 2018). These preliminary implementations, in Chapter 4, demonstrate that these assessment models could be used in the future to assess the size and trends of the stock. Comparing the historical exploitation pattern and stock abundance trends between the white and black-bellied anglerfish stocks in the Celtic Seas and Bay of Biscay indicates that for both species the fishing pressure has been decreasing in recent years and this has led to a gradual increase in stock levels (ICES, 2021a,b; Batts et al., 2022).

Furthermore, the vignette offered a convenient space to explore the Schnute model and its structure in more detail. In Chapter 3 (Batts et al., 2022), one version of the theoretical model proposed by Schnute (1987) was focused upon: that which derives future biomass from surviving biomass using population dynamics based on stage mean fish weights (Schnute, 1987). Two other versions proposed in Schnute (1987) were also implemented, which are outlined and discussed in the vignette. Testing found that the other two versions were more difficult to optimise with standard optimisers and suggest using a genetic algorithm for optimisation, such as that found in the ‘GA’ R package (Scrucca, 2017), may resolve this issue.

6.2.3 Applications to anglerfish stock assessment

This thesis has addressed key challenges facing stock assessment of anglerfish fisheries that arise from uncertainties in data observation and processing

(Perez et al., 2005; Farina et al., 2008; Maguire et al., 2008; Richards, 2016; ICES, 2018). LFEM is a useful tool in stock assessment for estimating and understanding growth of anglerfish, as demonstrated by the use of the model in the generation of age-composition data for the white-bellied anglerfish stock in the Celtic Seas and Bay of Biscay (ICES, 2018). For other anglerfish stocks where there is uncertainty in growth parameters LFEM could be used in a similar fashion, perhaps in conjunction with other length frequency analysis such as ELEFAN (Mildenberger et al., 2017; Taylor and Mildenberger, 2017) to validate the results. Furthermore, as demonstrated here, anglerfish survey data typically have a very defined recruit component (Batts et al., 2019), which would suit a hierarchical LFEM model to objectively define the recruit stage in yearly length frequency data, before using the stage-composition data in a stage-based assessment model.

The ability of CSA to fit the white-bellied anglerfish stock in Chapter 3, estimating both stock numbers and fishing mortality very similar to the official age-based assessment, demonstrates that CSA is a suitable candidate assessment model for *Lophius* species despite the limitations (e.g. assumptions on selectivity). Chapter 4 further supports this, indicating that stocks such as the black-bellied anglerfish stock in the Celtic Seas and Bay of Biscay, which is currently not managed with an assessment model (ICES, 2021a), can be assessed with the relatively simple stage-based assessment models implemented in the ‘sbar’ R package.

6.3 Applications to fisheries stock assessment and management over time

Age-slicing is a common method for obtaining age-composition data for stock assessments (Hilborn and Walters, 1992; Ailloud et al., 2015). In addition to accounting for uncertainty and bias in age-estimates or age-composition data with the methods discussed in Chapters 2-4, it was also important to understand the impact that age-slicing with a biased growth curve could have on fisheries stock assessment and management over time. Specifically, Chapter 5 investigated this impact with respect to either an age-based or stage-based assessment management procedure (MP). Management strategy evaluation (MSE) is a full feedback system evaluation framework in which these impacts could be assessed.

Other knowledge gaps were also addressed in Chapter 5, where CSA and age-based assessment MPs were generally tested for performance in an MSE framework and an OM with similar characteristics to white-bellied anglerfish was used. Here, the development of non-proxy reference points for CSA was also an important aspect of this research, providing for future stock assessment and management with CSA. Comparison of the performance statistics outlining risk and trade-offs of the two contrasting MPs was particularly informative in highlighting the differences in each MP's impact on management of a stock.

6.3.1 Implications of biased age-composition data for management advice and performance

Age-based and stage-based assessment and management were affected by biased age-composition data in different ways. The perception of the stock by the age-based assessment model was directly impacted by the shifting density of age-composition in observations when the growth curve used for age-slicing was biased. Changes in the ratio between younger and older fish in observations resulted in the age-based assessment overestimating stock size and underestimating fishing mortality when the growth rate was biased low (i.e. more older fish and less younger fish than in reality), and vice versa for the biased high growth model. Other studies have also found that assumed faster growth could lead to the larger SSB (and landings) into the future of a fishery, due the underestimation of the stock (Lai and Gunderson, 1987; Tyler et al., 1989; Bertignac and De Pontual, 2007). On the other hand CSA was impacted indirectly by its use of mean weights-at-age in the implementation of management, moreover, it was impacted in the opposite way to the age-based MPs. These findings have considerable implications on the sustainability and potential overexploitation of a stock, suggesting due consideration should be given to whether an assumed growth model is potentially biased and what stock assessment model is more suitable.

6.3.2 Performance of stage-based management procedures

To the author's knowledge, stage-based assessment models have not been tested within an MSE framework. In addition, development of non-proxy biological reference points for CSA was a fundamental challenge achieved in the research in Chapter 5. Results showed that overall the stage-based

assessment model/reference points combination performed well, if somewhat more precautionary than the age-based approach (possibly due to general underestimation of the stock). The challenge was to develop reference points for a model with numbers-based outputs but still incorporate shifts in catch biomass-at-equilibrium with changing fishing mortality. In the approach developed, assumptions on growth, age range and length-weight parameters were needed, as using catch numbers-at-equilibrium was heavily influenced by recruits numbers and skewed the estimated F_{msy} . Here, CSA reference points avoid recruitment overfishing (Myers et al., 1994) but cannot, in isolation, say anything about growth overfishing. This is a reversal of the historical approach where yield-per-recruit was used to avoid growth overfishing, whereas most developed countries now construct their fisheries management process around preventing recruitment overfishing (Hilborn and Walters, 1992; Ben-Hasan et al., 2021).

The findings outlined above further advance the prospect of using CSA for stock assessment and fisheries management. However, despite using mean-weights-at-age to calculate catch biomass-at-equilibrium in the procedure outlined in Chapter 5, it is important to consider whether numbers-based biological reference points are a valid and robust approach. Numbers-based reference points may be susceptible to a mis-match between biomass and stock numbers at high levels of fishing mortality (low biomass but high numbers). Also worth mentioning here is the development of reference points for the Schnute model may be simpler, as it is biomass-based and assumptions of growth and mean fish weights are implicit in the model already.

6.3.3 Performance of age-based management procedures

Management strategy evaluation essentially links monitoring, advice and policy, and a key factor of a true MSE is stakeholder engagement (Butterworth, 2007; Punt et al., 2014). Generally, engagement facilitates open dialogue and aids identifying key management goals important to the industry (Rademeyer et al., 2007). Representatives of the Irish fishing industry that exploits the anglerfish stocks in the Celtic Seas and Bay of Biscay were engaged to inform them about the research and gain their perspectives on management for the stocks. Their interest in the project lay in an initial assessment of the status quo (i.e. whether the ICES advice rule and the MSY approach was a reasonable way

to manage the stock). Although the operating model used in the MSE was a simulated population with characteristics based upon white-bellied anglerfish, the findings indicate that an age-based assessment and the ICES Advice rule as a combination of assessment and management was reasonable in terms of yield and sustainability in comparison to another approach (CSA MP) in non-biased age-composition scenarios. Unfortunately, external circumstances meant that further planned meetings with stakeholders during the development of the MSE did not occur, however it is important that the presentation of these results is followed up in the near future. Continuing the dialogue with the industry is important to keep the industry informed and maintain a good relationship between fisheries scientists and the fishing industry. A meeting of the Irish Fisheries Science Research Partnership (IFSRP) would be an appropriate setting for this.

6.4 Future directions

There are multiple areas that would be worthwhile pursuing with the LFEM model. LFEM source code and brief directions for use are readily available, however development of a LFEM R package for easy implementation of the model would encourage wider use and testing. Further development of the LFEM procedure could include extending the model into an integrated assessment model, similar to MULTIFAN-CL (Fournier et al., 1998), where LFEM would be integrated with an age-structured population model. The hierarchical version of such an assessment would be particularly interesting to develop. In this regard, researchers at Memorial University in Newfoundland have been in contact to say they are currently developing something along these lines. Another worthwhile way to extend the hierarchical LFEM model would be to generally focus on modelling the mixing proportions of the mixture models and possibly incorporating density-dependence, in order to model the abundance of a cohort through time. Finally, length frequency distribution derived cohort growth parameter estimates could also be used to investigate environmental covariates such as sea temperature (Baudron et al., 2014; Barrow et al., 2018).

State-space assessment models are common in modern stock assessment, providing a flexible framework that accounts for both process error and observation error in the model (Schnute, 1991; Pella, 1993; Gudmundsson, 1994; Freeman and Kirkwood, 1995; Nielsen and Berg, 2014; Aeberhard et al., 2018). Ex-

tending the stage-based assessment models implemented in Chapter 3 and 4 would be a sensible progression in the development of the models. Providing both model process and observation uncertainty can be accounted for in the extended models, the flexibility would likely improve the models overall, as well as increasing the likelihood of the stage-based models being used as primary assessments. Biomass-based (Meyer and Millar, 1999) and numbers-based (Li et al., 2019) stage-based assessment models have previously been implemented with a state-space framework within a Bayesian framework. It is also worth noting that the original implementation of CSA by Collie and Sissenwine (1983) used a penalised likelihood that could be thought of as an early state-space model. On the contrary, NOAA (2019) state that their CSA model does not allow the estimation of process errors as they can be difficult to estimate given the structure of the model. Exploration of whether this is the case for the CSA implementation for this thesis would be constructive.

Following the implementation and testing of the Schnute model in this thesis, there are many possibilities for further extensions, developments and testing for the model. Priority should be given to additional examination of the growth/mean weight configuration and whether the recommended configuration (see above) can improve the model estimation under logistic or dome-shaped selectivity scenarios. The good performance of the *Adapted Schnute Observation Error model* coupled with the population dynamics where the relative importance of the biomass of each stage can be changed suggests that developing a hybrid model, with some aspects of CSA, may be worthwhile. For example, if the option for stage-specific indices were incorporated into the model, an interesting question would be whether entire biomass of a stock could be estimated from solely a recruit index. Such questions could further stock assessment research hypotheses and testing.

Obtaining suitable reference points is often key for successful fisheries assessment and management (Hilborn and Walters, 1992; Punt, 2006; Kvamsdal et al., 2016; Hilborn et al., 2020; Silvar-Viladomiu et al., 2021; Zhang and Fong, 2021). Linking the work of Chapters 3-5 was the stage-based stock assessment model CSA. In Chapter 5, a procedure for estimating biological reference points was developed, which could be used in combination with an ICES style harvest control rule to manage a stock with CSA. Further development and comprehensive testing of this procedure, as well as consideration of other options (such as using theory on stage mean fish weights from Schnute (1987))

would be a worthwhile pursuit, particularly considering the potential issue in high exploitation scenarios.

The MSE framework used to investigate bias in age-slicing was a comprehensive full-feedback system evaluation based on life history theory, fishery understanding, and incorporated widespread uncertainty. This is an important development for this particular fishery and the framework can be easily extended to improve our knowledge of anglerfish assessment and management more generally. Three such extensions that should be prioritised are: 1) implementation of an alternative OM that is based upon the ICES assessment model fit for white-bellied anglerfish; 2) investigate other key uncertainties of anglerfish, such as separation of species in the catch; and 3) further engagement with the industry. Development of multiple OMs is common practice in MSEs (Punt et al., 2014) and including one based on the real anglerfish assessment is a logical step that would provide extra insights into the performance of the management procedures under different stock dynamics. Regarding suggested extension 2, as described in the Introduction of the thesis, anglerfish stock assessment and management are somewhat impeded by a number of issues and not just age estimation/growth uncertainty. It is important to address these other uncertainties in an MSE framework, in order to better understand the impact of those uncertainties and potentially prioritise future research. The frameworks developed within this thesis could contribute to that effort.

Bibliography

- Aeberhard, W. H., J. Mills Flemming, and A. Nielsen (2018). Review of state-space models for fisheries science. *Annual Review of Statistics and Its Application* 5, 215–235.
- Ailloud, L. E., M. W. Smith, A. Y. Then, K. L. Omori, G. M. Ralph, and J. M. Hoenig (2015). Properties of age compositions and mortality estimates derived from cohort slicing of length data. *ICES Journal of Marine Science* 72(1), 44–53.
- Barrow, J., J. Ford, R. Day, and J. Morrongiello (2018). Environmental drivers of growth and predicted effects of climate change on a commercially important fish, *Platycephalus laevigatus*. *Marine Ecology Progress Series* 598, 201–212.
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science* 76(7), 2150–2163.
- Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2022). Numbers or mass? Comparison of theoretically different stage-based stock assessment models and their ability to model simulated and real life stocks. *Canadian Journal of Fisheries and Aquatic Sciences - (Just-IN)*.
- Baudron, A. R., C. L. Needle, A. D. Rijnsdorp, and C. Tara Marshall (2014). Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes. *Global change biology* 20(4), 1023–1031.
- Ben-Hasan, A., C. Walters, A. Hordyk, V. Christensen, and M. Al-Husaini (2021). Alleviating growth and recruitment overfishing through simple management changes: Insights from an overexploited long-lived fish. *Marine and Coastal Fisheries* 13(2), 87–98.

- Bertignac, M. and H. De Pontual (2007). Consequences of bias in age estimation on assessment of the northern stock of European hake (*Merluccius merluccius*) and on management advice. *ICES Journal of Marine Science* 64(5), 981–988.
- Bhattacharya, C. (1967). A simple method of resolution of a distribution into gaussian components. *Biometrics*, 115–135.
- Biernacki, C., G. Celeux, and G. Govaert (2003). Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis* 41(3-4), 561–575.
- Brophy, D., S. Pérez-Mayol, R. Duncan, K. Hüsey, A. J. Geffen, H. D. Gerritsen, M. C. Villanueva, and B. Morales-Nin (2021). Elemental composition of illicia and otoliths and their potential application to age validation in white anglerfish (*Lophius piscatorius* linnaeus, 1758). *Estuarine, Coastal and Shelf Science* 261, 107557.
- Butterworth, D. S. (2007). Why a management procedure approach? some positives and negatives. *ICES Journal of Marine Science* 64(4), 613–617.
- Cadigan, N. G., S. E. Campana, and H. editor: Jan Jaap Poos (2016). Hierarchical model-based estimation of population growth curves for redfish (*Sebastes mentella* and *sebastes fasciatus*) off the Eastern coast of Canada. *ICES Journal of Marine Science* 74(3), 687–697.
- Collie, J. S. and M. P. Sissenwine (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* 40(11), 1871–1879.
- Deriso, R. B. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* 37(2), 268–282.
- Deroba, J., D. S. Butterworth, R. Methot, J. De Oliveira, C. Fernandez, A. Nielsen, S. Cadrin, M. Dickey-Collas, C. Legault, J. Ianelli, et al. (2015). Simulation testing the robustness of stock assessment models to error: some results from the ices strategic initiative on stock assessment methods. *ICES Journal of Marine Science* 72(1), 19–30.

- Eveson, J. P., J. Million, F. Sardenne, and G. Le Croizier (2015). Estimating growth of tropical tunas in the Indian ocean using tag-recapture data and otolith-based age estimates. *Fisheries research* 163, 58–68.
- Farina, A., M. Azevedo, J. Landa, R. Duarte, P. Sampedro, G. Costas, M. Torres, and L. Cañas (2008). *Lophius* in the world: a synthesis on the common features and life strategies. *ICES Journal of Marine Science: Journal du Conseil* 65(7), 1272–1280.
- Fournier, D. A., J. Hampton, and J. R. Sibert (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences* 55(9), 2105–2116.
- Fournier, D. A., J. R. Sibert, J. Majkowski, and J. Hampton (1990). MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian journal of fisheries and aquatic sciences* 47(2), 301–317.
- Freeman, S. and G. Kirkwood (1995). On a structural time series method for estimating stock biomass and recruitment from catch and effort data. *Fisheries research* 22(1-2), 77–98.
- Fridriksson, Á. (1934). On the calculation of age-distribution within a stock of cod by means of relatively few age determinations as a key to measurements on a large scale. *Rapports et Procès-Verbaux des Reunions, Conseil Permanent International pour l'Exploration de la Mer*. 86(2).
- Gudmundsson, G. (1994). Time series analysis of catch-at-age observations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 43(1), 117–126.
- Hamel, M. J., J. D. Koch, K. D. Steffensen, M. A. Pegg, J. J. Hammen, and M. L. Rugg (2014). Using mark-recapture information to validate and assess age and growth of long-lived fish species. *Canadian Journal of Fisheries and Aquatic Sciences* 71(4), 559–566.
- Hasselblad, V. (1966). Estimation of parameters for a mixture of normal distributions. *Technometrics* 8(3), 431–444.

- Hilborn, R., R. O. Amoroso, C. M. Anderson, J. K. Baum, T. A. Branch, C. Costello, C. L. de Moor, A. Faraj, D. Hively, O. P. Jensen, H. Kurota, L. R. Little, P. Mace, T. McClanahan, M. C. Melnychuk, C. Minto, G. C. Osio, A. M. Parma, M. Pons, S. Segurado, C. S. Szuwalski, J. R. Wilson, and Y. Ye (2020). Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences* 117(4), 2218–2224.
- Hilborn, R. and C. J. Walters (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media.
- Hislop, J. R., A. Gallego, M. R. Heath, F. M. Kennedy, S. A. Reeves, and P. J. Wright (2001). A synthesis of the early life history of the anglerfish, *Lophius piscatorius* (Linnaeus, 1758) in northern British waters. *ICES Journal of Marine Science* 58(1), 70–86.
- Hoenig, J. M. and D. M. Heisey (1987). Use of a log-linear model with the em algorithm to correct estimates of stock composition and to convert length to age. *Transactions of the American Fisheries Society* 116(2), 232–243.
- Hoggarth, D. D. (2006). *Stock assessment for fishery management: A framework guide to the stock assessment tools of the fisheries management and science programme*. Number 487. Food & Agriculture Org.
- Ibaibarriaga, L., C. Fernández, A. Uriarte, and B. A. Roel (2008). A two-stage biomass dynamic model for Bay of Biscay anchovy: a Bayesian approach. *ICES Journal of Marine Science* 65(2), 191–205.
- ICES (2018). Report of the Benchmark Workshop on Anglerfish Stocks in the ICES Area (WKANGLER), 12–16 February 2018, Copenhagen, Denmark. ICES CM 2018.
- ICES (2021a). Black-bellied anglerfish (*Lophius budegassa*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, ank.27.78abd.
- ICES (2021b). White anglerfish (*Lophius piscatorius*) in Subarea 7 and divisions 8.a–b and 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2021. ICES Advice 2021, mon.27.78abd.

- Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic Differentiation and Laplace approximation. *Journal of Statistical Software* 70(5), 1–21.
- Kvamsdal, S. F., A. Eide, N.-A. Ekerhovd, K. Enberg, A. Gudmundsdottir, A. H. Hoel, K. E. Mills, F. J. Mueter, L. Ravn-Jonsen, L. K. Sandal, et al. (2016). Harvest control rules in modern fisheries management. *Elementa: Science of the Anthropocene* 4.
- Lai, H. L. and D. R. Gunderson (1987). Effects of ageing errors on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). *Fisheries Research* 5(2-3), 287–302.
- Li, Y., L. M. Lee, and J. Rock (2019). Modeling population dynamics and nonstationary processes of difficult-to-age fishery species with a hierarchical bayesian two-stage model. *Canadian Journal of Fisheries and Aquatic Sciences* 76(12), 2199–2214.
- Lloyd-Jones, L. R., H. D. Nguyen, G. J. McLachlan, W. Sumpton, and Y.-G. Wang (2016). Mixture of time-dependent growth models with an application to blue swimmer crab length-frequency data. *Biometrics* 72(4), 1255–1265.
- Lortie, C. J., J. Braun, A. Filazzola, and F. Miguel (2020). A checklist for choosing between r packages in ecology and evolution. *Ecology and evolution* 10(3), 1098–1105.
- Maguire, J.-J., P. Pereda, R. Duarte, H. Dobby, and M. Azevedo (2008). Monkfish/anglerfish across the world; common problems and common solutions: an introduction to papers presented at the ices theme session in september 2007. *ICES Journal of Marine Science* 65(7), 1270–1271.
- Maunder, M. N., P. R. Crone, A. E. Punt, J. L. Valero, and B. X. Semmens (2016). Growth: Theory, estimation, and application in fishery stock assessment models. *Fisheries Research* (180), 1–3.
- Maunder, M. N. and K. R. Piner (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science* 72(1), 7–18.
- McLachlan, G. and D. Peel (2004). *Finite mixture models*. John Wiley & Sons.

- Mesnil, B. (2003). The Catch-Survey Analysis (CSA) method of fish stock assessment: an evaluation using simulated data. *Fisheries Research* 63(2), 193–212.
- Methot, R. D. and C. R. Wetzel (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142, 86–99.
- Meyer, R. and R. B. Millar (1999). Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 56(1), 37–52.
- Mildenberger, T. K., M. H. Taylor, and M. Wolff (2017). Tropfishr: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution* 8(11), 1520–1527.
- Morrongiello, J. R. and R. E. Thresher (2015). A statistical framework to explore ontogenetic growth variation among individuals and populations: a marine fish example. *Ecological Monographs* 85(1), 93–115.
- Myers, R., A. Rosenberg, P. Mace, N. Barrowman, and V. Restrepo (1994). In search of thresholds for recruitment overfishing. *ICES Journal of Marine Science* 51(2), 191–205.
- Nielsen, A. and C. W. Berg (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158, 96–101.
- NOAA (2019). NOAA Fish and Fisheries Toolbox: Collie-Sissenwine Analysis (CSA) version 4.3.
- Pella, J. J. (1993). Utility of structural time series models and the Kalman filter for predicting consequences of fishery actions. In *Proceedings of the international symposium on management strategies for exploited fish populations. Alaska Sea Grant College Program Report*, Number 93-02, pp. 571–593.
- Perez, J. A. A., P. R. Pezzuto, and H. A. de Andrade (2005). Biomass assessment of the monkfish *lophius gastrophysus* stock exploited by a new deep-water fishery in southern brazil. *Fisheries research* 72(2-3), 149–162.
- Punt, A. E. (2006). The FAO precautionary approach after almost 10 years: have we progressed towards implementing simulation-tested feedback-control

- management systems for fisheries management? *Natural Resource Modelling* 19(4), 441–464.
- Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. De Oliveira, and M. Haddon (2014). Management strategy evaluation: best practices. *Fish and Fisheries* 17(2), 303–334.
- Punt, A. E., T. Huang, and M. N. Maunder (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70(1), 16–33.
- Quinn, T. J. and R. B. Deriso (1999). *Quantitative Fish Dynamics*. Oxford University Press.
- Rademeyer, R. A., É. E. Plagányi, and D. S. Butterworth (2007). Tips and tricks in designing management procedures. *ICES Journal of Marine Science* 64(4), 618–625.
- Richards, R. (2016). Monkfish operational assessment. NEFSC (Northeast Fisheries Science Center). *Northeast Fisheries Science Center Ref Doc Ref Doc. 16-09; 109 p.*.
- Roa-Ureta, R. H. (2010). A likelihood-based model of fish growth with multiple length frequency data. *Journal of agricultural, biological, and environmental statistics* 15(3), 416–429.
- Schnute, J. (1985). A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* 42(3), 414–429.
- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences* 44(5), 924–940.
- Schnute, J. T. (1991). The importance of noise in fish population models. *Fisheries research* 11(3-4), 197–223.
- Scrucca, L. (2017). On some extensions to GA package: hybrid optimisation, parallelisation and islands evolution. *The R Journal* 9(1), 187–206.
- Silvar-Viladomiu, P., C. Minto, G. Halouani, L. Batts, D. Brophy, C. Lordan, and D. G. Reid (2021). Moving reference point goalposts and implications for fisheries sustainability. *Fish and Fisheries*.

- Smith, M. T. and J. T. Addison (2003). Methods for stock assessment of crustacean fisheries. *Fisheries Research* 65(1-3), 231–256.
- Taylor, M. and T. K. Mildenerger (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology* 24(4), 330–338.
- Tyler, A., R. Beamish, and G. McFarlane (1989). Implications of age determination errors to yield estimates. *Canadian Special Publication of Fisheries and Aquatic Sciences* 108, 27–35.
- Varadhan, R. and C. Roland (2008). Simple and globally convergent methods for accelerating the convergence of any EM algorithm. *Scandinavian Journal of Statistics* 35(2), 335–353.
- Wang, Y.-G. (1999). Estimating equations for parameters in stochastic growth models from tag–recapture data. *Biometrics* 55(3), 900–903.
- Whitten, A. R., N. L. Klaer, G. N. Tuck, and R. W. Day (2013). Accounting for cohort-specific variable growth in fisheries stock assessments: a case study from south-eastern Australia. *Fisheries Research* 142, 27–36.
- Zhang, Z. and K. Fong (2021). Estimation of key population parameters and MSY-based reference points for sidestripe shrimp (*Pandalopsis dispar*) in the Fraser River Delta, British Columbia. *Fisheries Research* 238, 105893.
- Zhou, S., A. E. Punt, R. Deng, and J. Bishop (2011). Estimating multifleet catchability coefficients and natural mortality from fishery catch and effort data: comparison of bayesian state–space and observation error models. *Canadian Journal of Fisheries and Aquatic Sciences* 68(7), 1171–1181.

Appendix A

Published manuscript I

Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models

Luke Batts ^{1,2,*}, Cólín Minto¹, Hans Gerritsen², and Deirdre Brophy¹

¹Marine and Freshwater Research Centre, Galway-Mayo Institute of Technology, Galway, H91 T8NW, Ireland

²Marine Institute, Rinville, Oranmore, Galway, H91 R673, Ireland

*Corresponding author: tel: +353 91 742525; e-mail: luke.batts@research.gmit.ie.

Batts, L., Minto, C., Gerritsen, H., and Brophy, D. Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsz103.

Received 1 November 2018; revised 30 April 2019; accepted 8 May 2019.

Analysis of length frequency distributions from surveys is one well-known method for obtaining growth parameter estimates where direct age estimates are not available. We present a likelihood-based procedure that uses mixture models and the expectation–maximization algorithm to estimate growth parameters from length frequency data (LFEM). A basic LFEM model estimates a single set of growth parameters that produce one set of component means and standard deviations that best fits length frequency distributions over all years and surveys. The hierarchical extension incorporates bivariate random effects into the model. A hierarchical framework enables inter-annual or inter-cohort variation in some of the growth parameters to be modelled, thereby accommodating some of the natural variation that occurs in fish growth. Testing on two fish species, haddock (*Melanogrammus aeglefinus*) and white-bellied anglerfish (*Lophius piscatorius*), we were able to obtain reasonable estimates of growth parameters, as well as successfully model growth variability. Estimated growth parameters showed some sensitivity to the starting values and occasionally failed to converge on biologically realistic values. This was dealt with through model selection and was partly addressed by the addition of the hierarchical extension.

Keywords: anglerfish *Lophius piscatorius*, bivariate random effects, EM algorithm, haddock *Melanogrammus aeglefinus*, LFEM, von Bertalanffy growth

Introduction

Fish growth is a widely studied aspect of fish biology (Pardo *et al.*, 2013), reflecting its importance for understanding life histories (Denney *et al.*, 2002), the effects of changing environmental conditions (Jobling, 2002; Baudron *et al.*, 2014) as well as ecosystem and population dynamics (Quinn and Deriso, 1999; Travers *et al.*, 2007; Smith *et al.*, 2015). Growth parameters are a fundamental component of fisheries stock assessments (Quinn and Deriso, 1999) and their accurate estimation can be key to ensuring that assessments are reliable (Hilborn and Walters, 1992). Estimation of other important population dynamics parameters such as fishing mortality and selectivity depend on the accurate estimation of growth (Hoggarth, 2006; Maunder *et al.*, 2016).

For teleost fish, growth information is usually obtained by estimating age from incremental growth layers laid down within the otoliths (fish ear bones), scales or fin rays. When age estimation

using calcified structures is too costly to implement, not practical or is suspected to give unreliable estimates then other methods such as tagging or length frequency analysis provide an alternative means of obtaining growth estimates (Hilborn and Walters, 1992). Clear, unambiguous growth increments are often lacking in the calcified structures of tropical or deep sea species that are not exposed to marked seasonal temperature changes, while direct ageing of crustaceans using hard parts still requires further validation (Kilada and Driscoll, 2017). As a consequence, assessments of fisheries for these species often rely on alternative methods for estimating age (Sparre and Venema, 1998; Chang *et al.*, 2012).

The most widely used fish growth model is the von Bertalanffy growth function (VBGF), which relates length to age using three parameters; L_{∞} (asymptotic mean length), K (rate of approach to asymptote), and t_0 (the theoretical time at which fish length

equals zero) (von Bertalanffy, 1938). Various modifications of the traditional VBGF appear in the literature. For example, Schnute and Fournier (1980) re-parameterized the VBGF to include more biologically intuitive parameters (i.e. the mean lengths of the first and last ages of the sampled population and a parameter that controls growth between ages). Another example is the incorporation of seasonal oscillation into the VBGF (Pitcher and MacDonald, 1973; Lloyd-Jones et al., 2016). Other growth functions that have been used to model fish growth include the Gompertz (Winsor, 1932), allometric (Quinn and Deriso, 1999), logistic (Cormon et al., 2016), and bi-phasic models (Minte-Vera et al., 2016); the most appropriate choice of model depends on the species (Katsanevakis and Maravelias, 2008). Here, the re-parameterized VBGF was chosen as the underlying model because of the widespread use of the VBGF and its appropriateness for length frequency analysis (Schnute and Fournier, 1980).

Growth can often vary spatially, temporally, between cohorts, between individuals, and within individuals (Quinn and Deriso, 1999; Wang, 1999; Morrongiello and Thresher, 2015). Cohort effects can vary to a lesser or greater extent depending on the species and have been linked to environmental changes (Baudron et al., 2014; Morrongiello and Thresher, 2015). Accounting for cohort growth variability can also improve stock assessment models (Whitten et al., 2013). Methods for accounting for individual growth variation in length-at-age data have also been explored, for example Pilling et al. (2002) used a non-linear random effects model to model growth parameter variability between individuals length-at-age data. Cadigan and Campana (2016) developed a hierarchical mixed-effects model to account for between stock and between individual variability in length-at-age data. Growth models incorporating random effects have also been applied to analysis of tagging data, where between-individual growth variability is often considerable (Eveson et al., 2015).

In situations where fish cannot be aged directly, length frequency analysis offers an alternative means of estimating fish growth parameters, provided there is a distinct spawning season (Hasselblad, 1966; Bhattacharya, 1967; Fournier et al., 1990; Taylor and Mildenerger, 2017). Perhaps the most prominent length frequency method in the literature is the robust maximum likelihood-based estimation procedure MULTIFAN, which is now incorporated into a length-based statistical catch at age stock assessment model (Fournier et al., 1990, 1998). The non-parametric ELEFAN procedure (Pauly, 1987) is also widely used and has recently been extended into the R environment where modern optimization algorithms are used to improve its performance (Mildenerger et al., 2017; Taylor and Mildenerger, 2017).

Currently available approaches to length frequency analysis present some limitations. Both MULTIFAN and ELEFAN require that length bins are defined prior to the analysis, potentially introducing subjectivity to the identification of cohorts. For ELEFAN it is recommended that L_∞ be fixed in the initial analysis and is then re-calculated post-optimization. Both models are also largely constrained to a single growth curve and do not incorporate inter-annual or inter-cohort variability in growth (although this is addressed to some extent in the development of MULTIFAN-CL) (Fournier et al., 1998). Current limitations could be addressed by developing a model that uses an expectation-maximization (EM) algorithm to estimate growth parameters (Dempster et al., 1977). This approach would remove the

need to predefine length classes and offers the flexibility to include variability in growth.

Whilst the EM algorithm is widely used to model length frequency distributions, it has rarely been used to estimate mixture model parameters with an incorporated von Bertalanffy growth structure. However, the methodology of MULTIFAN is similar to the EM algorithm in its application of maximum likelihood theory (Fournier et al., 1990). The “mixdist” package in the R environment enables the user to fit a mixture model with component means constrained to a growth curve and this utilizes an EM algorithm (Macdonald and Du, 2011). In addition, a recent paper by Lloyd-Jones et al. (2016) used a minorization-maximization (MM) algorithm (the EM algorithm is a form of MM algorithm) to estimate seasonally oscillating von Bertalanffy growth over a year on monthly length frequency data of a crab species (Lloyd-Jones et al., 2016), where L_∞ is considered a random variable, although it is not modelled explicitly. These examples leave much scope for expanding this area of research further.

This study aims to develop a new method, based on the EM algorithm, to model fish growth from length frequency data whilst incorporating von Bertalanffy growth structure and accounting for variability in growth between cohorts or years. Gaussian mixture models are developed and fit to length frequency distributions, where means of components within a cohort (represented by normal distributions in the mixture models) are restricted to von Bertalanffy growth. First, a basic version (analogous to MULTIFAN) of the methodology is introduced and tested using length frequency data for two fish species; haddock *Melanogrammus aeglefinus* and white-bellied anglerfish *Lophius piscatorius*. Second, hierarchical models that introduce various bivariate random effects, are introduced and tested with the intention of modelling some of the variation in the recruitment and growth of the fish species. The overall aim of the study is to introduce an alternative approach to obtaining a range of credible estimates of growth parameters from length frequency data and account for variability in growth by cohort or year.

Methods

A description of the growth function used within the model is first introduced, followed by a description of the length frequency expectation-maximization (LFEM) model itself. A summary of model parameters, inputs, and outputs is then given, as well as a description of the model selection criteria and software used for development. Finally, the two data sources that the LFEM model is applied to are described.

Growth function

Length frequency distributions were modelled using mixture models, where the means of the cohorts (normal distributions within the mixture models) follow a VBGF. Schnute and Fournier (1980) re-parameterized the classical von Bertalanffy growth (CVBG) equation into one more appropriate to length frequency analysis, represented by

$$\mu_i = l + (L - l) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M, \quad (1)$$

where μ_i is the mean length of fish at age a_i ; l and L are the first and final component or cohort mean lengths, i.e. μ_1 and μ_M , where M is the maximum age class; k is equivalent to $\exp(-K)$

and controls the distance between two successive component means [Equation (1)]. The CVBG parameters can then be calculated from these parameters using

$$L_\infty = \frac{L - lk^{M-1}}{1 - k^{M-1}}, \quad (2)$$

$$K = -\ln(k), \quad (3)$$

$$t_0 = a_1 - \frac{1}{\ln(k)} \ln \left\{ \frac{L - l}{L - lk^{M-1}} \right\}, \quad (4)$$

where L_∞ is the asymptotic length of the fish, K controls the rate at which the asymptote is approached, and t_0 is the theoretical time when an individual would have length zero. a_1 is the assumed age of the l component. Standard deviation (SD) of the normal distribution of the component (σ_i) is either estimated as constant (CSD) or a linear (LSD) function of component means (Schnute and Fournier, 1980). The equation for LSD is

$$\sigma_i = s + (S - s) \frac{\mu_i - l}{L - l}; \quad i = 1, \dots, M, \quad (5)$$

where s is the SD of the first component and S is the SD of the final component.

LFEM model and algorithm

This methodology summarizes the basic model developed in this study where a single set of growth parameters that produce one set of component means and SD s are estimated (i.e. if l , L , and k are identical for every cohort then component means in each annual mixture for a particular survey model are identical).

Based on the observed data log-likelihood for a single mixture model (McLachlan and Peel, 2004), the observed data log-likelihood for a set of finite mixture models on length frequency distributions of fish where cohort component means are restricted to the re-parameterized VBGF (RVBG) is

$$\ln L(\Psi) = \sum_{v=1}^V \sum_{r=1}^R \sum_{j=1}^{n_{vr}} \ln \left\{ \sum_{i=1}^M \lambda_{vri} f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) \right\}, \quad (6)$$

$$f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) = N(y_{jvr}; \mu_{vri}, \sigma_{vri}, a_{1v}), \quad (7)$$

$$\mu_{vri} = l_v + (L_v - l_v) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V, \quad (8)$$

where the components of the mixture models are normally distributed [Equation (7)] and their means, μ_{vri} are given by Equation (8). Component SD s are either assumed constant or constrained by Equation (5). Indices denote survey (v), year (r), component (i), and observation (j). Ψ is a vector that contains all unknown parameters in the model, V is the number of surveys, R is the number of years, M is the number of components, and n_{vr} is the total number of fish in any given survey and year. λ_{vri} is the mixing proportion for the i th component in the v th survey in the r th year, y_{jvr} is the length of fish j in the v th survey in the r th year, and θ_{vri} is the parameters of the i th component in the v th survey in the r th year. a_{1v} determines the assumed age of first component of the v th survey and is used to incorporate surveys of differing timings by back or forward projecting component means using Equations (2)–(4).

The complete data log-likelihood can be given as

$$\log L_c(\Psi) = \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \sum_{j=1}^{n_{vr}} z_{jvri} \{ \log \lambda_{vri} + \log f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) \}, \quad (9)$$

where Equations (7) and (8) apply. z_{jvri} is a data label vector distributed according to a multinomial distribution of length of M with $z_{jvri} = (z_{jvri})_i = 1$ or 0, according to whether that particular fish or y_j belongs to the i th component in the v th survey in the r th year (McLachlan and Peel, 2004). However, as we do not know what component each fish belongs to the EM algorithm was used to treat z_{jvri} as missing data and estimate Ψ iteratively. Direct estimation of the observed log-likelihood is possible but not practical in this case and the EM algorithm offered a simple framework to implement the model in Minto *et al.* (2018). For details on the EM algorithm see Supplementary Appendix S1.

Hierarchical extensions

Hierarchical models were developed, incorporating bivariate random effects and allowing inter-annual or inter-cohort variation to be modelled. In contrast to the basic model that applies a single set of component means to the data, these models allow cohorts to be modelled through years and surveys. For the hierarchical models both ‘‘CSD and LSD’’ models were implemented. To avoid over-parameterization these hierarchical models were not extended beyond bivariate random effects and in each model either the key parameter for the asymptotic length (L) or the growth parameter (k) is random, not both. The logic behind this is also discussed by Eveson *et al.* (2007) and Lloyd-Jones *et al.* (2016), suggesting that a random effect on either K or L_∞ incorporates sufficient variability and has less bias than if both parameters were random.

Hierarchical models were fitted using the same number of components as the best fitting basic model (based on the model selection criteria described below). Three different hierarchical model structures were tested: (i) bivariate random effects on cohort specific l and L ; (ii) bivariate random effects on cohort specific l and k ; (iii) bivariate random effects on cohort specific l and yearly k . Correlation between the bivariate random effects in each scenario is also modelled. A random effect on l was specifically chosen to account for variability in the mean length of the first component in each cohort. Other random effects were chosen to investigate and incorporate growth variability. For details on the model formulation see Supplementary Appendix S1.

Model parameters, inputs, and outputs

A wide range of starting values for parameters of the basic model were tested (Table 1, see Table 2 for parameter descriptions), amounting to over 1000 model runs for each species and each SD type. To test the suitability of the model selection criteria for identifying the number of identifiable age classes in the length frequency data, the number of components tested ranged from 3 to 14 for haddock and 6 to 14 for white-bellied anglerfish. SD s for the random effects of l and k for the haddock hierarchical model were fixed at $\exp(-5)$ to give the model stability but still retain the random effect. More information on model stability in the haddock hierarchical models and general model inputs can found in Supplementary Appendix S1.

Table 1. Summary of the starting values tested for the model parameters of the basic model.

Species	Parameter	Starting values tested		
		From	To	By
Haddock	l	10	10	NA
	L	40	70	10
	k	0.6	0.9	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	3	14	1
White-bellied anglerfish	λ_i	1/No. components	NA	NA
	l	16	16	NA
	L	100	130	10
	k	0.7	0.99	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	6	14	1
	λ_i	1/No. components	NA	NA

Model selection criteria and standard errors

Through exploratory simulation analysis it was found that Akaike information criterion (AIC) performed best from a range of model selection criteria. Model selection for this study was performed by selecting the model with highest observed log-likelihood value for each “number of components” tested. A variation on AIC we call “sub-AIC” was used to select the best fitting model. “sub-AIC” is similar to AIC, however the observed log-likelihood used is only contributed to by a subset of the length frequency data where reasonable numbers of fish were observed (i.e. haddock ≤ 40 cm & anglerfish ≤ 100 cm). This method ensures model selection is based on the majority of the data and not confounded by low fish numbers at high lengths. A similar method where model deviance is calculated from a subset of length classes when numbers in some length classes are low was used in model selection for fish selectivity curves (Millar and Fryer, 1999).

The EM algorithm does not directly provide estimates of uncertainty of parameters (McLachlan and Peel, 2004). Standard errors were approximated by first fitting the model to convergence with the EM algorithm, then using the maximum likelihood parameter estimates to optimize the observed data log-likelihood function [Equation (6)] (Minto et al., 2018). Standard errors were approximated via the delta method with the R package “TMB” (Kristensen et al., 2016).

Validation for haddock

Age data on the North Sea International Bottom Trawl Survey (NS-IBTS) are collected on a length-stratified basis. To obtain unbiased length-at-age data for haddock the length-stratified age data was used to construct annual age length keys (ALKs). Annual length frequency distributions (i.e. the same used in the length frequency analysis) were applied to their respective ALKs using the method described by Isermann and Knight (2005), resulting in an assigned age for each individual fish of the length frequency distribution. To enable comparison to the growth parameters estimated from the length frequency data alone, a VBGF was fit to this length-at-age data using maximum likelihood and mean length-at-age was also calculated.

Table 2. Summary of model parameters, and other relevant inputs/ outputs of models.

Notation	Type	Description
Basic model		
L	Parameter	Mean of final component
l	Parameter	Mean of first component
k	Parameter	Controls rate of growth
s	Parameter	SD of first component
S	Parameter	SD of final component
j	Index	Individual fish length
i	Index	Component number
ν	Index	Survey number
r	Index	Number of year
$\lambda_{\nu ri}$	Parameter(s)	Mixing proportion of i th component in survey(ν) and year(r)
μ_i	Output	Mean of i th component
σ_i	Output	SD i th component
L_∞	Output	CVBG parameter
K	Output	CVBG parameter
t_0	Output	CVBG parameter
$a_{1\nu}$	Input	Used to calculate t_0
M	Input	Number of components/ages assumed present in length frequency data
$n_{\nu r}$	Input	Number of observations(fish) in a given survey(ν) and year(r)
Hierarchical models		
L_c	Parameter	Mean of final component of a cohort
l_c	Parameter	Mean of first component of a cohort
k_c	Parameter	Mean of random effect that controls rate of growth for a cohort
k_r	Parameter	Mean of random effect that controls rate of growth for a particular year
σ_L	Parameter	SD of random effect on L
σ_l	Parameter	SD of random effect on l
σ_k	Parameter	SD of random effect on k
ρ_{lk}	Parameter	Correlation of bivariate random effects on l, k for example
$\mu_{\nu ri}$	Output	Mean of i th component in survey(ν) and year(r)

Not all hierarchical model parameters are used in each model.

Software

We develop a hybrid EM algorithm that uses the R package “TMB” within the “maximization” step (Kristensen et al., 2016). Specifically, complete data log-likelihood functions were written in C++ using the TMB library class, compiled and automatically differentiated (Kristensen et al., 2016). These objective functions, including gradients, were then optimized at each iteration with the statistical software R using the `nminb` function within base R (R Core Team, 2018). The remaining computations of the algorithm and other analysis were conducted with R (R Core Team, 2018). The “FSA” R package (Ogle, 2018) was used to apply ALKs to the un-aged length frequency data, using the semi-random method from Isermann and Knight (2005).

Data

Haddock was chosen to validate the method as otolith age readings and growth parameters derived from lengths at age are reliable. White-bellied anglerfish was chosen because age readings are considered unreliable (ICES, 2011) and there is need for a cohort analysis to build an assessment model on. Both species are characterized by relatively fast growth and a discrete recruitment season, leading to

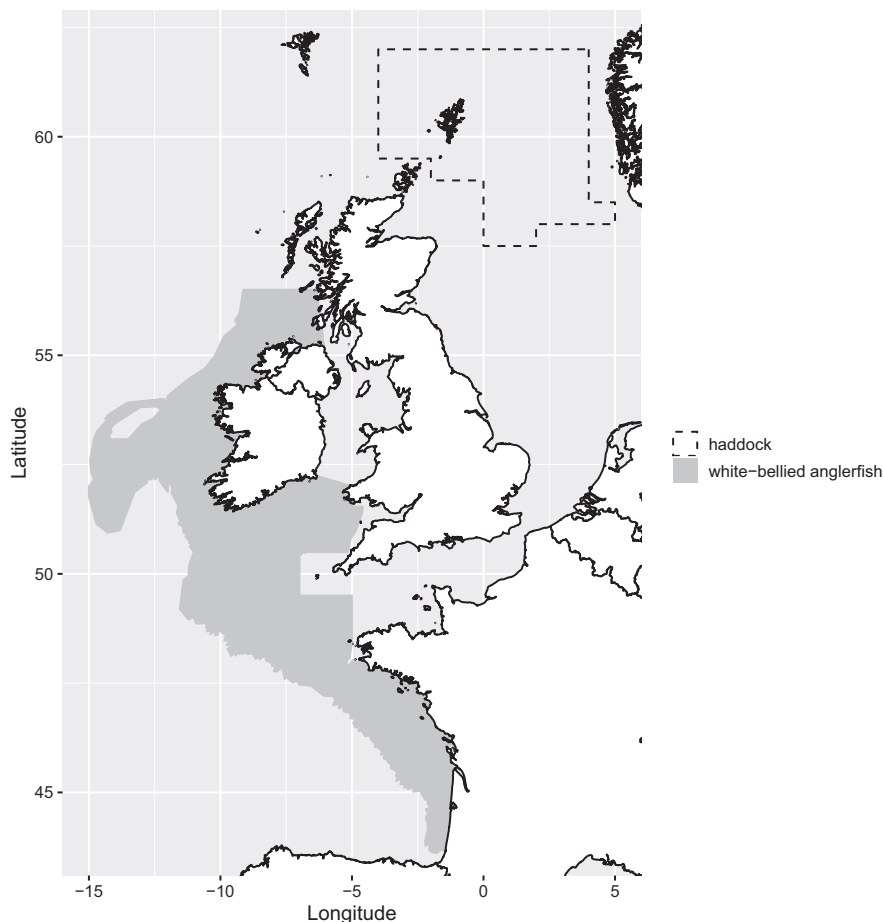


Figure 1. Areas of survey coverage from where length frequency data were obtained for haddock (standard roundfish area one of the NS-IBTS) and white-bellied anglerfish (EVHOE, IE-IGFS, and SP-PORC).

distinct length cohorts that can be tracked over time. Combined sex data was used for both species as is convention for haddock and white-bellied anglerfish in the most recent ICES assessment of the stock (ICES, 2018c). Survey catch data were obtained from the ICES DATRAS database (ICES, 2018b). Haddock length frequency data were from standard roundfish area one of the NS-IBTS (Figure 1). Only data collected in the third quarter of the year (Q3) were used. The corresponding age-length key for this area and time of year was also extracted from DATRAS. Length frequency data for white-bellied anglerfish were obtained from three surveys that cover the anglerfish stock in ICES areas 7.b-k, 8.a-b, and 8.d; the French EVHOE groundfish survey, the Irish groundfish survey (IE-IGFS) and the Spanish Porcupine groundfish survey (SP-PORC) (Figure 1). The French and Irish surveys were conducted mainly in the fourth quarter of the year and the Spanish survey in September. A Grande Ouverture Verticale fishing gear was used on all surveys with the exception of the Spanish Porcupine bank survey which used “porcupine baca” fishing gear (ICES, 2018a).

Results

Haddock

Haddock basic LFEM model

The basic model (single set of growth parameters estimated) was sensitive to starting parameters. Some of the model runs tended to

converge on k parameter values very close to 1. Whilst these gave good or even the best model fits according to the sub-AIC model selection criteria, the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). Some model runs were discarded (406 of 1488 CSD models and 42 of 1488 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for “biologically unreasonable” L_∞ estimates. To further filter out any spurious models the results were further refined by identifying and removing any models that had a second component mean ≤ 15 cm. An example of why refining is important when a wide range of starting parameters are tested can be found in Supplementary Appendix S2.

A nine component model was the best fitting (lowest sub-AIC) for CSD and no models above nine components had reasonable fits for this SD type. LSD models gave generally lower sub-AIC values than CSD models and the best fitting model overall had 12 components and LSD (Table 3). Within both SD types, but particularly in LSD models, parameter estimates, and component means were consistent across the range of number of components tested (Tables 3 and 4). When compared to the VBGF estimated from length-at-age data ($\hat{K} = 0.499$, $\hat{L}_\infty = 39.5$, and $\hat{t}_0 \approx 0$), the VBGF from the best fitting basic LSD model is similar, although it does slightly underestimate early age mean lengths (Figure 2). The best fitting basic CSD model follows a similar growth pattern until the fifth

Table 3. von Bertalanffy growth parameters from (lowest sub-AIC) model runs on haddock length frequency data after refining valid model runs.

SD type	No. components	RVBG parameters			CVBG parameters			sub-AIC	
		\hat{l} (cm)	\hat{L} (cm)	\hat{k}	\hat{L}_∞ (cm)	\hat{K}	\hat{t}_0		
CSD	3	10.5 (0.008)	32.4 (0.007)	0.816 (0.0012)	76	0.20	-0.10	6 167 506	
	4	10.3 (0.006)	37.1 (0.011)	0.780 (0.0007)	61	0.25	-0.12	6 031 048	
	5	10.3 (0.005)	39.7 (0.014)	0.754 (0.0006)	53	0.28	-0.13	5 989 126	
	6	10.3 (0.005)	41.8 (0.018)	0.750 (0.0005)	51	0.29	-0.15	5 974 158	
	7	10.3 (0.005)	43.0 (0.021)	0.752 (0.0005)	50	0.29	-0.18	5 957 923	
	8	10.4 (0.004)	44.0 (0.022)	0.761 (0.0004)	49	0.27	-0.23	5 946 292	
	9	10.4 (0.004)	45.0 (0.026)	0.773 (0.0004)	50	0.26	-0.27	5 941 101	
	LSD	3	10.1 (0.003)	31.3 (0.008)	0.854 (0.0013)	88	0.16	-0.14	5 982 480
		4	10.1 (0.003)	34.9 (0.012)	0.720 (0.0007)	49	0.33	-0.07	5 918 972
5		10.1 (0.003)	37.0 (0.015)	0.689 (0.0006)	44	0.37	-0.06	5 901 698	
6		10.1 (0.003)	37.5 (0.016)	0.665 (0.0005)	41	0.41	-0.06	5 889 986	
7		10.1 (0.003)	38.4 (0.018)	0.661 (0.0005)	40	0.41	-0.06	5 886 624	
8		10.1 (0.003)	39.0 (0.020)	0.658 (0.0004)	40	0.42	-0.06	5 885 749	
9		10.1 (0.003)	39.5 (0.021)	0.658 (0.0004)	40	0.42	-0.06	5 885 240	
10		10.1 (0.003)	40.0 (0.023)	0.661 (0.0004)	40	0.41	-0.06	5 885 145	
11		10.1 (0.003)	40.4 (0.024)	0.663 (0.0004)	40	0.41	-0.07	5 885 031	
12		10.1 (0.003)	40.6 (0.025)	0.665 (0.0004)	40	0.41	-0.07	5 885 004	
13		10.1 (0.003)	40.8 (0.026)	0.666 (0.0004)	41	0.41	-0.07	5 885 005	
14		10.1 (0.003)	40.9 (0.027)	0.667 (0.0004)	41	0.41	-0.07	5 885 023	

Number of components is the number of ages assumed to be present in the data. Rows highlighted in grey are the models with the lowest sub-AIC model selection criteria values. Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameter estimates are given to 3 s.f. and CVBG to 2 s.f.

Table 4. Component/cohort means and their SD parameters for haddock, corresponding to the basic models presented in Table 3.

SD type	No. components	Component											\hat{s}	\hat{S}				
		1	2	3	4	5	6	7	8	9	10	11						
CSD	3	10.5	22.6	32.4													3.66	
	4	10.3	21.5	30.2	37.1												2.89	
	5	10.3	21.0	29.0	35.1	39.7											2.65	
	6	10.3	20.6	28.4	34.2	38.5	41.8										2.50	
	7	10.3	20.2	27.7	33.3	37.5	40.7	43.0									2.39	
	8	10.3	19.8	27.0	32.4	36.6	39.8	42.2	44.0								2.29	
	9	10.4	19.4	26.4	31.9	36.1	39.3	41.8	43.8	45.3							2.21	
	LSD	3	10.0	21.5	31.3													1.63
		4	10.1	21.1	29.1	34.9												1.65
5		10.1	20.9	28.3	33.5	37.0											1.66	
6		10.1	20.7	27.7	32.4	35.5	37.5										1.67	
7		10.1	20.6	27.5	32.1	35.1	37.1	38.4									1.67	
8		10.1	20.6	27.4	32.0	34.9	36.9	38.2	39.0								1.68	
9		10.1	20.5	27.4	31.9	34.9	36.8	38.1	39.0	39.5							1.68	
10		10.1	20.5	27.4	31.9	34.9	36.9	38.2	39.1	39.6	40.0						1.68	
11		10.1	20.5	27.3	31.9	34.9	36.9	38.3	39.1	39.7	40.1	40.4					1.68	
12		10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.4	40.6				1.68	
13		10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.5	40.7	40.8			1.68	
14		10.1	20.4	27.3	31.9	35.0	37.0	38.3	39.3	39.9	40.3	40.5	40.7	40.8	40.9		1.69	

In the case of CSD models $S = s$ but values are only given in the final (S) column.

component, the following component means then appeared to overestimate the mean length of fish in comparison to the LSD model and the length-at-age estimated model fit (Figure 2). Estimated SD was very similar within SD types (Table 4).

Haddock hierarchical LFEM model with cohort specific l and k

Overall CVBG parameters for the hierarchical CSD model ($\hat{K} = 0.241$, $\hat{L}_\infty = 52.9$ cm, and $\hat{t}_0 = -0.31$) were similar to those given

by the basic CSD model and followed a similar trajectory. Hierarchical LSD model parameters ($\hat{K} = 0.394$, $\hat{L}_\infty = 42.1$ cm and $\hat{t}_0 = -0.06$) were also similar to estimates from its corresponding basic model (Table 3 and Figure 2).

The performance of the hierarchical l and k LSD model (corresponding to the best fitting basic model) was examined in more detail. Ranges of the cohort specific L_∞ estimates were 41.8–42.5 cm for \hat{L}_∞ and 0.334–0.463 for \hat{K} . Correlation

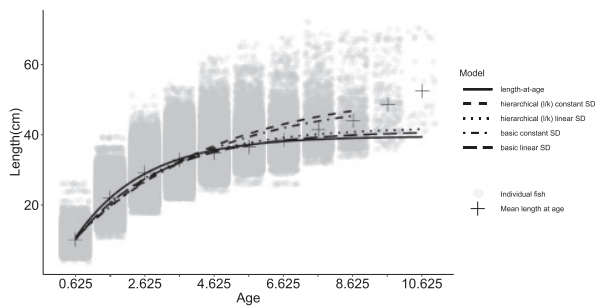


Figure 2. Length-at-age data and estimated von Bertalanffy growth curves for haddock. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data, growth curves from best fitting basic length frequency models, and overall growth curves (i.e. means of RE parameters) from hierarchical l/k models. Ages are jittered for visualization.

parameter $\hat{\rho}$ for the random effects l and k was estimated as -0.23 . Cohort-specific random effects gave the model flexibility in its estimated growth trajectories for each cohort. When examining the cohorts that were observed at least until the ninth component over the period analysed (2000–2009) the estimated cohort growth curves from this study’s hierarchical model appear to fit reasonably well in some cohorts (e.g. cohorts 2002, 2003, and 2004) and very well in others (e.g. cohorts 2000, 2005, and 2009) (Figure 3). When cross-referenced with Figure 4 the hierarchical model cohort curves that are very similar to length-at-age estimated cohort curves are those cohorts where a strong signal can be seen through a number of years (i.e. cohort 2000) (Figures 3 and 4). Deviations (difference between LFEM model length-at-age and ML estimated length-at-age) of these cohorts showed that in the early ages of some cohorts the hierarchical LFEM model underestimated the mean length-at-age in comparison to the ML fit on the length-at-age data (Supplementary Appendix S2). In the later years of the length frequency distributions the hierarchical model begins to severely underestimate growth (Figure 4). Also of note is the flexibility the random effect on the mean of the first component (l) gives when modelling the shifting position of the first component over the years, such as in 2006 or 2013 (Figure 4).

White-bellied anglerfish

White-bellied anglerfish basic LFEM model

As with haddock, the basic LFEM model was sensitive to starting parameters, particularly k and some of the model runs tended to converge on k parameter values very close to 1. These models often gave good or even the best model fits according to the sub-AIC model selection criteria, however the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). A large number of model runs was discarded (720 of 1080 CSD models and 740 of 1080 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for “biologically unreasonable” L_∞ estimates.

Overall, the nine component model for LSD was best fitting (lowest sub-AIC) and LSD models consistently had lower sub-AIC values than their counterpart CSD models. For LSD models as the number of components assumed was increased the final

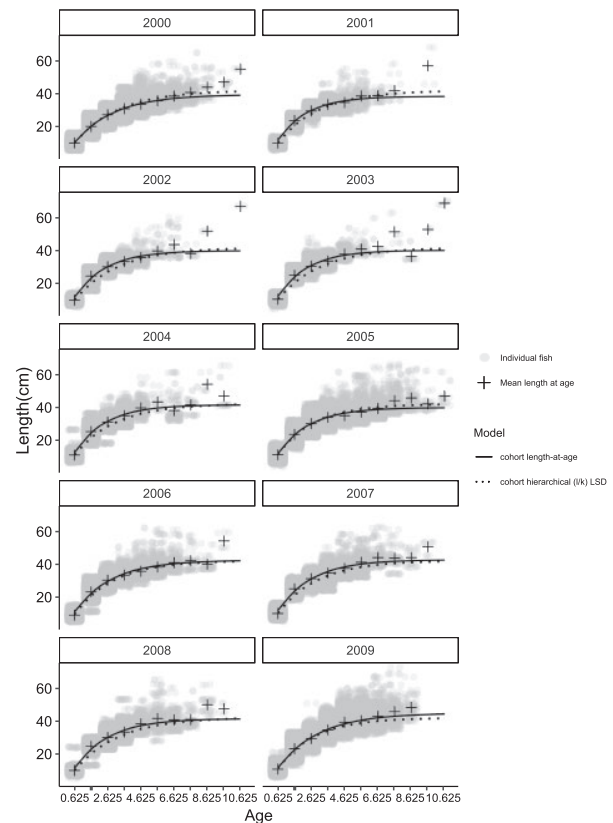


Figure 3. Length-at-age data and estimated von Bertalanffy growth curves for cohorts of haddock observed up to at least their 9th component. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data of the cohort and cohort-specific growth curves estimated within the LSD hierarchical model that varies l and k . Ages are jittered for visualization.

component means (L) increased, as did the CVBG parameter K , while L_∞ decreased. CSD models showed no such pattern (Tables 5 and 6).

Component means were consistent between and within SD types up to the fourth component particular for models with an assumed number of components of 11 or above. LSD model component means did not change with increasing number of components assumed and the best fitting model across all models (LSD, nine components) had very similar component means to the best fitting CSD model (Table 6). When plotted on the raw length frequency data the best fitting basic model appears to fit the data reasonably well in some years but not in others and cohort progression can be seen up until the seventh component in some cohorts (Figure 6a).

The growth curve of the parameters estimated from the best fitting basic LFEM model differs somewhat from the growth curves estimated in a previous study (Landa *et al.*, 2013). Growth curves were reasonably similar up to age six, although this study’s VBGF shows marginally faster growth. After age six the growth curves begin to diverge more, with this study estimates of length-at-age markedly higher (Figure 5a).

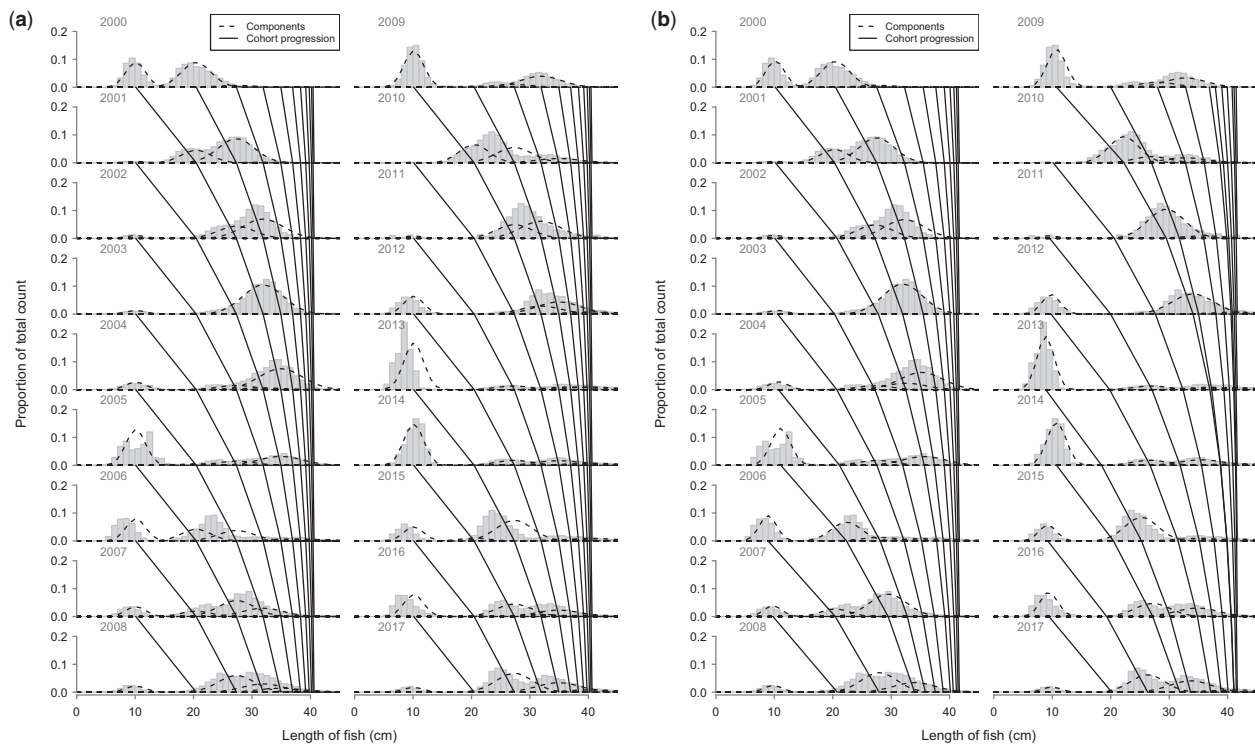


Figure 4. Length frequency distributions of haddock with fitted distributions from: (a) the lowest sub-AIC basic model (12 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and k . Also shown is the progression of the cohort mean values. Note that the maximum length of fish was 72 cm but only fish up to 45 cm were plotted here as there was little data above this length, all estimated cohorts are displayed.

White-bellied anglerfish hierarchical LFEM model with cohort specific l and L

Overall CVBG parameters for the hierarchical LSD model with cohort specific l and L corresponding to the best fitting basic model ($\hat{K} = 0.089$, $\hat{L}_\infty = 215.2$ cm, and $\hat{t}_0 = -0.14$) differed from those given by the basic model. \hat{K} was increased and \hat{L}_∞ reduced dramatically, \hat{t}_0 was also slightly reduced. The resulting VBGF curve was different from the basic model curve as well as the growth curves from a previous study (Landa et al., 2013). The hierarchical model showed faster growth than the other curves and estimated component means somewhat different from the best fitting basic LFEM model (Figure 5a). Cohort-specific growth curves from the hierarchical LFEM model showed a range of trajectories and cohort \hat{L}_∞ estimates ranged from 184.9 to 246.6 cm (Figure 5b). The slowest growing cohort with the lowest estimated L_∞ (2014 cohort) followed a similar trajectory to the basic LFEM model (Figure 5).

There was a negative correlation estimated ($\hat{\rho} = -0.69$) between the bivariate random effects l and L of each cohort (Figure 7). The bivariate cohort-specific random effects show a moderate degree of variation over a few centimetres for \hat{l} and ~ 30 cm for \hat{L} . There is also a negative trend in L (Figure 7b). There is a high degree of uncertainty around estimates of L for the last few cohorts (Figure 7b).

Other white-bellied anglerfish hierarchical models

Hierarchical models with a cohort specific k or yearly varying k were also fit to the anglerfish data. In both cases the RVBG

parameter K varies to a large extent between cohorts, however L_∞ also varies and the resulting cohort curves differ but not dramatically. Parameter estimates can be found in [Supplementary Appendix S3](#).

Discussion

Overview of LFEM

This study aimed to improve upon currently available approaches for growth estimation of fish through length frequency analysis. A basic method utilizing the EM algorithm and analogous to MULTIFAN (Fournier et al., 1990) was produced. The methodology was then extended further to incorporate bivariate random effects to capture some of the growth variation in length frequency data.

Few studies have modelled growth variation in length frequency data (Fournier et al., 1998; Roa-Ureta, 2010; Lloyd-Jones et al., 2016). The approach of using a VBGF to constrain mixture model component means through a time series (i.e. along a cohort) whilst explicitly estimating cohort or year specific random effects on two growth parameters is a worthwhile extension to this area of research [see Lloyd-Jones et al. (2016) for an alternative]. The aim of this integrated approach was to aid identification of growth parameters in variable length frequency distributions obtained from multiple annual surveys. Furthermore, the associated uncertainty of the parameter estimates is also approximated. Overall both versions of the methodology produced reasonable estimates of von Bertalanffy growth parameters and their variability between cohorts/years. However,

Table 5. von Bertalanffy growth parameters from LFEM model runs that gave the lowest sub-AIC value on white-bellied anglerfish length frequency data.

SD type	No. components	RVBG parameters			CVBG parameters			sub-AIC
		\hat{l} (cm)	\hat{L} (cm)	\hat{k}	\hat{L}_∞ (cm)	\hat{K}	\hat{t}_0	
CSD	8	17.1 (0.090)	114.4 (0.572)	0.966 (0.003)	469	0.035	-0.20	99 430
	9	17.0 (0.087)	118.6 (0.621)	0.945 (0.003)	294	0.057	-0.17	99 397
	10	17.0 (0.085)	119.9 (0.621)	0.930 (0.003)	232	0.072	-0.17	99 419
	11	16.6 (0.082)	125.3 (0.653)	0.974 (0.002)	489	0.026	-0.45	99 341
	12	16.6 (0.081)	128.4 (0.699)	0.969 (0.002)	401	0.031	-0.48	99 367
	13	16.6 (0.080)	130.1 (0.734)	0.951 (0.002)	267	0.050	-0.40	99 419
	14	16.5 (0.078)	126.8 (0.653)	0.940 (0.002)	215	0.062	-0.41	99 444
LSD	8	16.5 (0.076)	107.3 (0.919)	0.941 (0.005)	276	0.061	-0.13	99 112
	9	16.3 (0.075)	108.7 (0.997)	0.967 (0.005)	412	0.033	-0.34	99 102
	10	16.3 (0.076)	113.1 (0.999)	0.974 (0.004)	471	0.027	-0.45	99 110
	11	16.4 (0.076)	114.9 (1.017)	0.951 (0.004)	264	0.051	-0.39	99 173
	12	16.4 (0.076)	116.3 (0.992)	0.938 (0.003)	213	0.064	-0.37	99 237
	13	16.4 (0.076)	120.1 (1.022)	0.930 (0.003)	195	0.072	-0.34	99 312
	14	16.4 (0.076)	122.1 (1.077)	0.921 (0.003)	177	0.082	-0.31	99 393

Number of components is the number of ages assumed to be present in the data. Parameters l and L are those estimated for the surveys that occurred in the fourth quarter of the year (EVHOE and IE-IGFS). Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameters estimates are given to 3 s.f. and CVBG to 2 s.f. (with the exception of L_∞). The Spanish survey estimates of can be found in [Supplementary Appendix S3](#).

Table 6. Component/cohort means and their SD parameters for white-bellied anglerfish, corresponding to the basic models presented in [Table 5](#).

SD type	No. components	Component														\hat{s}	\hat{S}		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14				
CSD	8	17.1	32.5	47.3	61.7	75.5	88.9	101.9	114.4										4.91
	9	17.0	32.4	46.9	60.6	73.6	85.8	97.4	108.3	118.6									4.75
	10	16.9	32.0	45.9	58.9	71.0	82.3	92.7	102.5	111.5	119.9								4.62
	11	16.6	28.8	40.7	52.3	63.5	74.5	85.2	95.6	105.8	115.7	125.3							4.34
	12	16.6	28.4	39.8	50.9	61.7	72.1	82.3	92.1	101.6	110.8	119.7	128.4						4.23
	13	16.6	28.9	40.6	51.7	62.2	72.3	81.9	90.9	99.6	107.8	115.6	123.0	130.1					4.20
	14	16.5	28.5	39.8	50.4	60.3	69.7	78.5	86.7	94.5	101.8	108.6	115.1	121.1	126.8				4.12
LSD	8	16.5	32.0	46.5	60.2	73.1	85.2	96.6	107.3										3.96 11.15
	9	16.3	29.3	41.8	53.9	65.6	76.9	87.9	98.5	108.7									3.83 12.03
	10	16.3	28.2	39.9	51.2	62.2	72.9	83.4	93.5	103.5	113.1								3.85 9.90
	11	16.3	28.6	40.2	51.3	61.8	71.8	81.4	90.4	99.0	107.2	114.9							3.86 9.28
	12	16.4	28.6	40.1	50.9	61.1	70.5	79.4	87.8	95.6	103.0	109.9	116.3						3.87 8.55
	13	16.4	28.9	40.5	51.3	61.3	70.6	79.3	87.4	94.9	101.9	108.4	114.5	120.1					3.90 7.72
	14	16.4	29.1	40.8	51.6	61.5	70.6	79.1	86.8	94.0	100.5	106.6	112.2	117.3	122.1				3.91 7.60

In the case of CSD models $S = s$ but values are only given in the final (S) column.

there were problems with convergence to biologically unreasonable estimates of CVBG parameters particularly in the basic models, as well as issues with model stability in the hierarchical models where information was intermittently sparse (i.e. weak or low abundance cohorts).

Haddock

Basic and overall hierarchical parameter estimates for haddock were within or close to previously reported ranges ($K \approx 0.1-0.5$ and $L_\infty \approx 45-55$ cm) (Baudron *et al.*, 2011). Parameter estimates from the best fitting basic model (12 components and LSD) and overall parameter estimates from the lk hierarchical LSD model gave very similar growth curves to that of a maximum likelihood fit on length-at-age data for fish from the same region and time

of year (Figure 2). CSD models overestimated growth somewhat in older fish indicating the importance of identifying the correct SD. When applying their method to data for pike and abalone Schnute and Fournier (1980) also found that SD treatment was important when deriving growth from observed length frequencies. The number of components identified by the best fitting basic LFEM model (12 components) is close to the number of ages present in the ALK (11 ages). The model selection criterion was also particularly useful for identifying the more suitable SD type. The effect of misidentifying component number is also negated somewhat by the consistency of the component means with differing number of components (Table 4).

From the hierarchical model (random effects on l and k) we can see a similar variation in cohort-specific growth parameters as Baudron *et al.* (2011) found in their study. Hierarchical

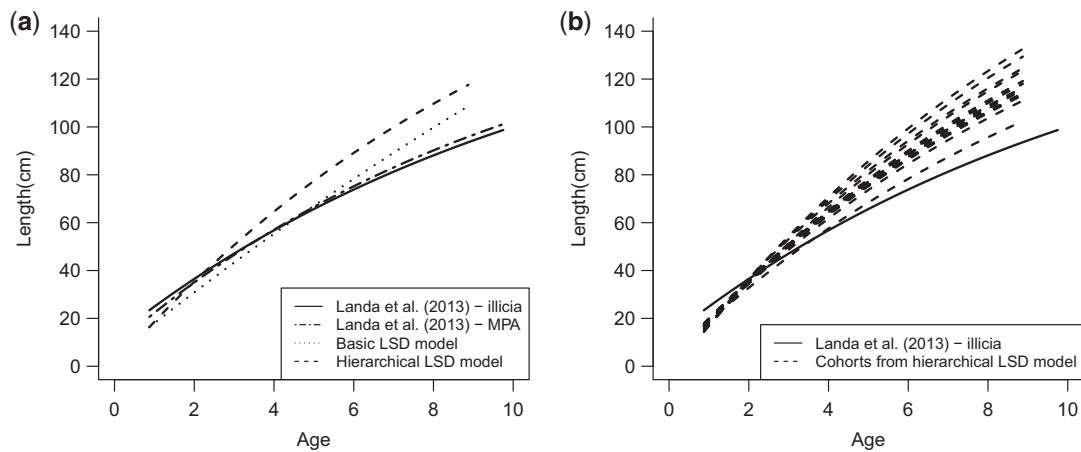


Figure 5. von Bertalanffy growth curves estimated for white-bellied anglerfish from: (a) the best fitting basic model and overall parameters from the corresponding hierarchical model with a cohort-specific random effect on l and L , as well as growth curves [estimated from illicia and modal progression analysis (MPA)] from a previous study (Landa et al., 2013), (b) estimated cohort growth curves from the hierarchical LSD model with cohort-specific random effects on l and L .

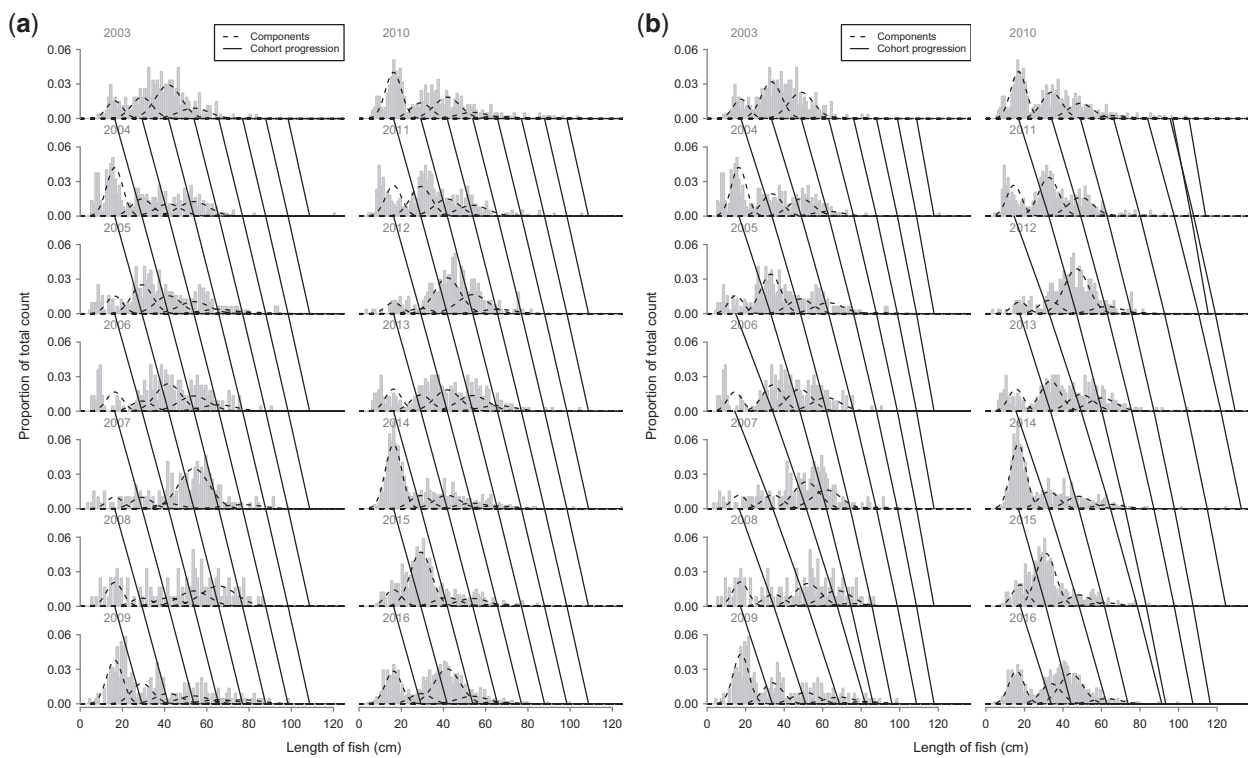


Figure 6. Length frequency distributions of white-bellied anglerfish from the quarter four IE-IGFS with fitted distributions from: (a) the lowest sub-AIC basic LFEM model (9 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and L . Also shown is the progression of the cohort mean values (black solid lines). Similar plots for the Spanish Porcupine Bank survey (SP-PORC) on the French groundfish survey (EVHOE) can be found in [Supplementary Appendix S3](#).

model and length-at-age estimated cohort growth curves were similar in cohorts observed up until the ninth component. However outside of these cohorts growth was often underestimated, indicating the effect of the lack of information in later

ages. Cohort growth curves that showed the most similar fits to the data between the two methods also tended to be for cohorts where strong signals were present in sequential years (Figure 3 and 4).

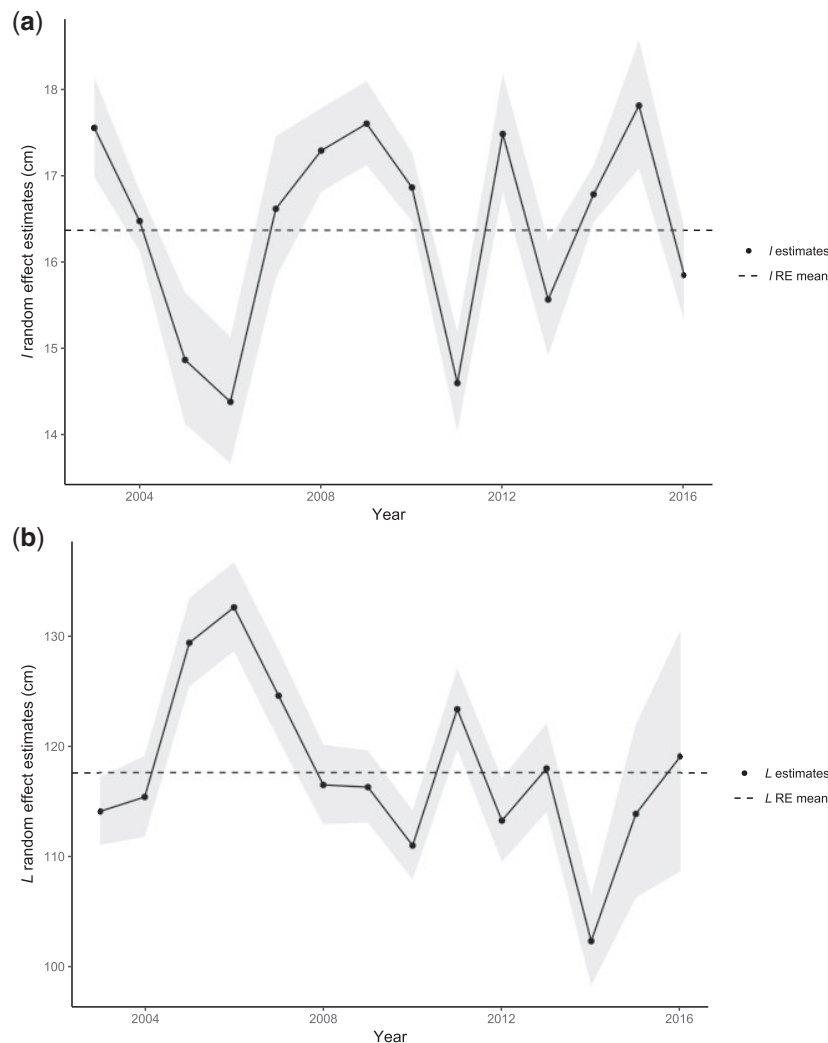


Figure 7. Estimated random effects on first (*I*) and last (*L*) component means of cohorts from a nine component hierarchical LSD model on length frequency distributions of white-bellied anglerfish. Shaded area denotes 2*standard error of the random effect estimates and dashed lines denote the mean of the random effect. Parameter estimates are for quarter four surveys (EVHOF and IE-IGFS).

White-bellied anglerfish

Parameter estimates for white-bellied anglerfish (both sexes combined) were broadly similar to those in the literature, although differences are more pronounced in the later ages and estimates of L_{∞} from this study were consistently higher than previously reported (Farina *et al.*, 2008; Landa *et al.*, 2013).

The basic and hierarchical models produced different growth curves, despite having the same starting parameters. The 2014 cohort from the hierarchical model follows a similar trajectory to the best fitting basic LFEM model growth curve, indicating that 2014–2016 (the years 2014 cohort is observed) may be dominating in the estimation of the basic model. Flexibility within the hierarchical model estimated a slower growth for the 2014 cohort, whilst estimating a slightly faster growth for other cohorts. Further investigations showed that when years 2014–2016 (these years had a higher number of fish than other years) were removed from the data the basic LFEM model component means were more alike to the hierarchical model estimates for the majority of years. Subjectively, fitted distributions of the hierarchical model

seem to fit better than the basic LFEM model (Figure 6). Estimated growth curves predicted faster growth particularly in the older fish compared to anglerfish age validation studies (Figure 5) (Farina *et al.*, 2008; Landa *et al.*, 2013). Faster growth in the later ages was less pronounced in models that assumed a higher number of components, however these were not selected as the best fitting model (Table 5).

Model selection based on sub-AIC indicated that a nine component model with LSD provided the best fit to the data. Although *L. piscatorius* is thought to live frequently to over ten years of age (Farina *et al.*, 2008), fish of this age are not sufficiently well represented in the length frequency data to support the selection of a large number of components. The nine component model produced a higher estimate of L_{∞} and a lower estimate of K compared to other studies (Farina *et al.*, 2008; Landa *et al.*, 2013), however, the estimates of components means and SDs fit the length frequency distributions reasonably well (Figure 6).

As with haddock SD treatment was important. Including a linear increasing component SD produced components with very

wide length distributions and mixing proportions appeared to be biased towards a single component in some cases. CSD allowed for more overlap between components and more balanced mixing proportions but not necessarily a better fitting model (Table 5).

Overall, discrepancies between parameter estimates from this study and those from studies that use direct ageing methods may reflect modelling challenges with the data or the difficulties of obtaining reliable age estimates for anglerfish (ICES, 2011). Landa *et al.* (2013) used length frequency analysis techniques to validate direct ageing of illicia and found good agreement between the parameter estimates produced by both methods. However, that study was based on survey data from the Porcupine Bank while this study also included data from two other surveys as model inputs. Further difficulty in estimating growth may have been introduced here by combining data for males and females as white-bellied anglerfish show sexual dimorphism in growth (Farina *et al.*, 2008).

Caveats

The EM algorithm is an accepted method for estimating latent variables but is sensitive to the choice of starting parameters (Biernacki *et al.*, 2003), as was evident from the model outputs presented here. This was accounted for by conducting sensitivity runs on some of the main parameters (k , L and number of components), producing a wide range of values for each parameter estimate. Some other starting parameters were fixed, such as the mean of the first component (l) and the starting mixing proportions, which was arbitrarily set as $1/\text{No. components}$. It is likely that the starting mixing proportions in conjunction with a k growth rate parameter drove the variation in parameter estimates at convergence. Lloyd-Jones *et al.* (2016) modelled length frequency distributions using a methodology similar to the one employed here, but used a MM algorithm to estimate parameters.

The number of sensitivity runs was large for this study (>1000 for each SD type) and a wide range of starting parameter values was used. Many k parameter starting values were very different from known estimates, which may partially account for the number of model runs that were excluded before further analysis. This study's approach to parameter sensitivity by multiple model runs followed by exclusion of models with "biologically unreasonable" parameter estimates is a robust method for dealing with the sensitivity, although caution should be used when setting a threshold (e.g. this study used a L_∞ of 500 cm, so as not to introduce subjectivity into model selection). Haddock required a further refining of models by identifying and removing any models with a second component mean ≤ 15 cm. Second component means ≤ 15 cm were caused by certain (slow growth) starting parameters and the second component mixing proportion tending to zero.

In this study, the sub-AIC model selection criteria appeared to work reasonably well (details of sub-AIC are given in "Model selection criteria and standard errors" section). sub-AIC was chosen as the model selection criteria in part due to the performance of AIC in exploratory simulation analysis with the LFEM model. AIC was able to identify the correct number of components when the SD of components was small (~ 1), as did Bayesian information criterion (BIC). With increasing SD of length frequency data AIC was more robust, still under-fitting the data somewhat but much less severely than other model selection criteria. However, model selection is often problematic in studies such as this, hence

should be conducted cautiously using expert knowledge of the species in question (Schnute and Fournier, 1980; Fournier *et al.*, 1990; Taylor and Mildenerger, 2017).

Correlation between parameters could contribute to the observed variation in parameter estimates. CVBG parameters (K and L_∞) have been shown to covary strongly (Gallucci and Quinn, 1979; Pilling *et al.*, 2002). Here, the re-parameterized von Bertalanffy growth parameters appeared strongly correlated, with many combinations of k and L parameters resulting in very similar model fits. The same issue arises with other length frequency analysis methods such as MULTIFAN and ELEFAN and is usually dealt with by fixing either K or L_∞ in the initial stages of the analysis (Fournier *et al.*, 1990; Taylor and Mildenerger, 2017). Variation in estimated maximum likelihood parameters could also be driven by the substantial variability that exists in the input data.

Hierarchical models appeared to be less sensitive to the choice of starting parameters, although due to computational demands these models were not subjected to the same level of sensitivity testing. Model testing using the haddock data showed that when length frequency data indicates very variable recruitment (i.e. in some years the first component is almost not present) the hierarchical model initially struggled to estimate these components, shifting the first component to the second component position and thus confounding estimates of cohort progression. It seemed that there was not sufficient information within the estimated distribution of the random effect to prevent this. This issue is a form of the label switching problem that is well known in mixture models (Yao, 2015). Fixing the SD of the bivariate random effects at a suitably low value [i.e. $\exp(-5)$] as detailed in "Model selection criteria and standard errors" section, retained enough flexibility in the model to estimate variable random effects but sufficient constraint to keep l and k random effects estimates within reasonable bounds, thus preventing label switching.

Further work

A comprehensive comparison of the basic and hierarchical models with other methodologies such as MULTIFAN, the updated ELEFAN procedure, and the more subjective Bhattacharya method would be a useful avenue for further research (Bhattacharya, 1967; Fournier *et al.*, 1990; Taylor and Mildenerger, 2017). In MULTIFAN-CL density dependence can be included in the structure of the model as abundance of a cohort can effect growth rate in some fish species (Fournier *et al.*, 1998). The hierarchical model presented in this study is currently designed to model growth variability between cohorts independently of cohort abundance but could be modified to include cohort abundance as a covariate, increasing complexity and modelling the connection between mixing probabilities of a cohort. Laslett *et al.* (2004) and Roa-Ureta (2010) devised stage-based approaches to model fitting. A comparison of these stage-based approaches with hierarchical model presented here would be useful to assess the effectiveness of integrating constraints and random effects into the mixture model parameter estimation.

Currently, the components of the mixture models are assumed to be normally distributed. Further development of the method could enable the user to specify alternative distributions such as log normal, gamma, or other, thus allowing non-normality and skewness to be modelled (Macdonald and Pitcher, 1979; Lloyd-Jones *et al.*, 2016).

Conclusions

Use of length frequency analysis to obtain growth parameters is a well-developed area of research (Hasselblad, 1966; Fournier *et al.*, 1990, 1998; Taylor and Mildenerger, 2017). These techniques provide growth information when age of individuals cannot be determined otherwise. Parameter estimates, their associated uncertainty and model fits presented here demonstrate the usefulness of our method for obtaining growth information from length frequency data.

Growth parameters are estimated whilst simultaneously fitting finite mixture models using the EM algorithm. The use of the EM algorithm removes the need to define length bins, allowing a more objective set up of the modelling procedure. Models can take inputs from multiple surveys from different times of year and offer an alternative framework to other well-known approaches for estimating growth parameters from length frequency data. Furthermore, within the hierarchical framework we explicitly model bivariate random effects on growth parameters, successfully modelling growth variability and allowing trends to be identified in the estimates.

Where age data are not readily available modelling cohort growth variability through this study's methodology can improve information available for stock assessments (i.e. cohort-specific growth curves for age slicing). In addition, estimates from these length frequency models give useful credible estimates of the boundaries of growth parameters that could be tested for stock assessments in a simulation context, such as in a management strategy evaluation (Punt *et al.*, 2013). Finally, length frequency distribution derived cohort growth parameter estimates could also be used to investigate environmental covariates such as sea temperature (Baudron *et al.*, 2014; Barrow *et al.*, 2018).

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

We'd also like to thank ICES, the reviewers and editor for their comments and those that provided and collated the data that was used in the study.

Funding

We thank the Cullen fellowship programme of the Marine Institute, Ireland for the funding for this postgraduate research Reference No. CF/16/03.

References

- Barrow, J., Ford, J., Day, R., and Morrongiello, J. 2018. Environmental drivers of growth and predicted effects of climate change on a commercially important fish, *Platycephalus laevis*. *Marine Ecology Progress Series*, 598: 201–212.
- Baudron, A. R., Needle, C. L., and Marshall, C. 2011. Implications of a warming North Sea for the growth of haddock *Melanogrammus aeglefinus*. *Journal of Fish Biology*, 78: 1874–1889.
- Baudron, A. R., Needle, C. L., Rijnsdorp, A. D., and Tara Marshall, C. 2014. Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes. *Global Change Biology*, 20: 1023–1031.
- Bhattacharya, C. 1967. A simple method of resolution of a distribution into Gaussian components. *Biometrics*, 23: 115–135.
- Biernacki, C., Celeux, G., and Govaert, G. 2003. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis*, 41: 561–575.
- Cadigan, N. G., and Campana, S. E. 2016. Hierarchical model-based estimation of population growth curves for redbfish (*Sebastes mentella* and *Sebastes fasciatus*) off the eastern coast of Canada. *ICES Journal of Marine Science*, 74: 687–697.
- Chang, Y.-J., Sun, C.-L., Chen, Y., and Yeh, S.-Z. 2012. Modelling the growth of crustacean species. *Reviews in Fish Biology and Fisheries*, 22: 157–187.
- Cormon, X., Ernande, B., Kempf, A., Vermard, Y., and Marchal, P. 2016. North Sea saithe *Pollachius virens* growth in relation to food availability, density dependence and temperature. *Marine Ecology Progress Series*, 542: 141–151.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39: 1–38.
- Denney, N. H., Jennings, S., and Reynolds, J. D. 2002. Life-history correlates of maximum population growth rates in marine fishes. *Proceedings of the Royal Society of London B: Biological Sciences*, 269: 2229–2237.
- Eveson, J. P., Million, J., Sardenne, F., and Le Croizier, G. 2015. Estimating growth of tropical tunas in the Indian Ocean using tag-recapture data and otolith-based age estimates. *Fisheries Research*, 163: 58–68.
- Eveson, J. P., Polacheck, T., and Laslett, G. M. 2007. Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. *Canadian Journal of Fisheries and Aquatic Sciences*, 64: 602–617.
- Farina, A., Azevedo, M., Landa, J., Duarte, R., Sampedro, P., Costas, G., Torres, M. *et al.* 2008. Lophius in the world: a synthesis on the common features and life strategies. *ICES Journal of Marine Science*, 65: 1272–1280.
- Fournier, D. A., Hampton, J., and Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 2105–2116.
- Fournier, D. A., Sibert, J. R., Majkowski, J., and Hampton, J. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian Journal of Fisheries and Aquatic Sciences*, 47: 301–317.
- Gallucci, V. F., and Quinn, T. J. 1979. Reparameterizing, fitting, and testing a simple growth model. *Transactions of the American Fisheries Society*, 108: 14–25.
- Hasselblad, V. 1966. Estimation of parameters for a mixture of normal distributions. *Technometrics*, 8: 431–444.
- Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Kluwer Academic Publishers, Boston/Dordrecht/London.
- Hoggarth, D. D. 2006. Stock assessment for fishery management: a framework guide to the stock assessment tools of the fisheries management and science programme. FAO Fisheries Technical Paper. No. 487. Food & Agriculture Organisation, Rome.
- ICES. 2011. Report of the anglerfish (*Lophius piscatorius*) illicia and otoliths exchange 2011, ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS).
- ICES. 2018a. International Bottom Trawl Survey (2000–2017). <http://datras.ices.dk> (last accessed 9 May 2018).
- ICES. 2018b. DATRAS Survey Descriptions. <http://datras.ices.dk> (last accessed 9 May 2018).
- ICES. 2018c. ICES Advice on Fishing Opportunities, Catch, and Effort Bay of Biscay and the Iberian Coast, Celtic Seas, Greater North Sea, and Oceanic Northeast Atlantic Ecoregions—mon.27.78abd.

- Isermann, D. A., and Knight, C. T. 2005. A computer program for age-length keys incorporating age assignment to individual fish. *North American Journal of Fisheries Management*, 25: 1153–1160.
- Jobling, M. 2002. Environmental factors and rates of development and growth. *Handbook of fish biology and fisheries*, 1: 97–122.
- Katsanevakis, S., and Maravelias, C. D. 2008. Modelling fish growth: multi-model inference as a better alternative to a priori using von Bertalanffy equation. *Fish and Fisheries*, 9: 178–187.
- Kilada, R., and Driscoll, J. G. 2017. Age determination in crustaceans: a review. *Hydrobiologia*, 799: 21–36.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. 2016. TMB: automatic differentiation and Laplace approximation. *Journal of Statistical Software*, 70: 1–21.
- Landa, J., Barrado, J., and Velasco, F. 2013. Age and growth of angler-fish (*Lophius piscatorius*) on the porcupine bank (west of Ireland) based on illicia age estimation. *Fisheries Research*, 137: 30–40.
- Laslett, G. M., Eveson, J. P., and Polacheck, T. 2004. Fitting growth models to length frequency data. *ICES Journal of Marine Science*, 61: 218–230.
- Lloyd-Jones, L. R., Nguyen, H. D., McLachlan, G. J., Sumpton, W., and Wang, Y.-G. 2016. Mixture of time-dependent growth models with an application to blue swimmer crab length-frequency data. *Biometrics*, 72: 1255–1265.
- Macdonald, P., and Du, J. 2011. *mixdist: Finite Mixture Distribution Models*. R Package Version, 0.5-4. <https://CRAN.R-project.org/package=mixdist>.
- Macdonald, P., and Pitcher, T. 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. *Journal of the Fisheries Board of Canada*, 36: 987–1001.
- Maunder, M. N., Crone, P. R., Punt, A. E., Valero, J. L., and Semmens, B. X. 2016. Growth: theory, estimation, and application in fishery stock assessment models. *Fisheries Research*, 180: 1–3.
- McLachlan, G., and Peel, D. 2004. *Finite Mixture Models*. John Wiley & Sons, New York.
- Mildenberger, T. K., Taylor, M. H., and Wolff, M. 2017. TropFishR: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution*, 8: 1520–1527.
- Millar, R. B., and Fryer, R. J. 1999. Estimating the size-selection curves of towed gears, traps, nets and hooks. *Reviews in Fish Biology and Fisheries*, 9: 89–116.
- Minte-Vera, C. V., Maunder, M. N., Casselman, J. M., and Campana, S. E. 2016. Growth functions that incorporate the cost of reproduction. *Fisheries Research*, 180: 31–44.
- Minto, C., Hinde, J., and Coelho, R. 2018. Including unsexed individuals in sex-specific growth models. *Canadian Journal of Fisheries and Aquatic Sciences*, 75: 282–292.
- Morrongiello, J. R., and Thresher, R. E. 2015. A statistical framework to explore ontogenetic growth variation among individuals and populations: a marine fish example. *Ecological Monographs*, 85: 93–115.
- Ogle, D. H. 2018. *FSA: Fisheries Stock Analysis*. R Package Version 0.8.20. <https://CRAN.R-project.org/package=FSA>.
- Pardo, S. A., Cooper, A. B., and Dulvy, N. K. 2013. Avoiding fishy growth curves. *Methods in Ecology and Evolution*, 4: 353–360.
- Pauly, D. 1987. A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. *In* Length-based Methods in Fisheries Research. Ed. by D. Pauly and G. R. Morgan. ICLARM Conf Proc, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait, pp. 7–34.
- Pilling, G. M., Kirkwood, G. P., and Walker, S. G. 2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 424–432.
- Pitcher, T., and MacDonald, P. 1973. Two models for seasonal growth in fishes. *Journal of Applied Ecology*, 10: 599–606.
- Punt, A. E., Trinnie, F., Walker, T. I., McGarvey, R., Feenstra, J., Linnane, A., and Hartmann, K. 2013. The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face of non-stationary dynamics. *Fisheries Research*, 137: 116–128.
- Quinn, T. J., and Deriso, R. B. 1999. *Quantitative Fish Dynamics*. Oxford University Press, New York/Oxford.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Roa-Ureta, R. H. 2010. A likelihood-based model of fish growth with multiple length frequency data. *Journal of Agricultural, Biological, and Environmental Statistics*, 15: 416–429.
- Schnute, J., and Fournier, D. 1980. A new approach to length-frequency analysis: growth structure. *Canadian Journal of Fisheries and Aquatic Sciences*, 37: 1337–1351.
- Smith, M., Fulton, E., Day, R., Shannon, L., and Shin, Y.-J. 2015. Ecosystem modelling in the southern Benguela: comparisons of Atlantis, Ecosim, and OSMOSE under fishing scenarios. *African Journal of Marine Science*, 37: 65–78.
- Sparre, P., and Venema, S. C. 1998. *Introduction to Tropical Fish Stock Assessment—Part 1: Manual*. FAO, Rome.
- Taylor, M., and Mildenberger, T. K. 2017. Extending electronic length frequency analysis in R. *Fisheries Management and Ecology*, 24: 330–338.
- Travers, M., Shin, Y.-J., Jennings, S., and Cury, P. 2007. Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. *Progress in Oceanography*, 75: 751–770.
- von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on growth laws. ii). *Human Biology*, 10: 181–213.
- Wang, Y.-G. 1999. Estimating equations for parameters in stochastic growth models from tag-recapture data. *Biometrics*, 55: 900–903.
- Whitten, A. R., Klaer, N. L., Tuck, G. N., and Day, R. W. 2013. Accounting for cohort-specific variable growth in fisheries stock assessments: a case study from south-eastern Australia. *Fisheries Research*, 142: 27–36.
- Winsor, C. P. 1932. The Gompertz curve as a growth curve. *Proceedings of the National Academy of Sciences*, 18: 1–8.
- Yao, W. 2015. Label switching and its solutions for frequentist mixture models. *Journal of Statistical Computation and Simulation*, 85: 1000–1012.

Handling editor: Jan Jaap Poos

Appendix B

Supplementary material for chapter 2

Supplementary Appendix

MS: Estimating fish growth parameters and growth variability from length frequency data using hierarchical mixture models

Luke Batts, Hans Gerritsen, Deirdre Brophy and C oil n Minto

June 4, 2019

Code and examples of model runs can be found at <https://github.com/lbatts/LFEM>

1 S1 - Additional methodology

1.1 Model and algorithm

The EM algorithm is a commonly used iterative method for finding the maximum likelihood parameters of a mixture model, when the model depends on latent variables (Dempster *et al.*, 1977). In this case the EM algorithm can be used to fit the finite mixture models that are used to model length frequency distributions of fish. Latent variables are the data labels associated with each fish length that determine which component of the mixture model the fish belongs to (i.e. a proxy for age group).

An EM algorithm is made up of two steps; the expectation (E) step and the maximisation (M) step. The E step requires the calculation of the conditional expectation given the current estimate of the parameters at the k th iteration (in the first step this is according to some reasonable starting parameters).

$$Q(\Psi; \Psi^{(k)}) = E_{\Psi^{(k)}} \{\log L_c(\Psi) | y\} \quad (1.1.1)$$

Computation of the expected value or posterior probability that each data point (y_{jvr}) belongs to each component of the mixture model given the observed data (y) and current estimate of parameters ($\Psi^{(k)}$) is needed.

$$E_{\Psi^{(k)}}(Z_{jvri} | y) = pr_{\Psi^{(k)}} \{Z_{jvri} = 1 | y\} \quad (1.1.2)$$

$$= \tau_{vri}(y_{jvr}; \Psi^{(k)}) \quad (1.1.3)$$

Where

$$\tau_{vri}(y_{jvr}; \Psi^{(k)}) = \lambda_{vri}^{(k)} f_{vri}(y_{jvr}; \theta_{vri}^{(k)}) / \sum_{h=1}^M \lambda_{vrh}^{(k)} f_{vrh}(y_{jvr}; \theta_{vrh}^{(k)}) \quad (1.1.4)$$

Using these posterior probabilities in the M step, new estimates of the parameters are found by maximising the conditional expectation in the following form

$$Q(\Psi; \Psi^{(k)}) = \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \sum_{j=1}^{n_{vr}} \tau_{vri}(y_{jvr}; \Psi^{(k)}) \{ \log \lambda_{vri} + \log f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) \} \quad (1.1.5)$$

where.

$$f_{vri}(y_j; \theta_{vri}) = N(y_{jvr}; \mu_{vri}, \sigma_{vri}, a_{1v}) \quad (1.1.6)$$

$$\mu_{vri} = l_v + (L_v - l_v) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V \quad (1.1.7)$$

New parameter estimates are then used in the successive E step and so on until the difference between $L(\Psi^{(k+1)})$ and $L(\Psi^{(k)})$ is suitably small (See (Dempster *et al.*, 1977) for details on the EM algorithm).

1.2 Hierarchical models

Note that random effects are only estimated on the timing of the first survey ($v=1$), l_{vr} and L_{vr} (if present in model) where $v > 1$ are re-calculated from equations in section 2.5 of the article and a_{1v} of the survey.

1.2.1 Cohort specific l and L

The observed data log-likelihood for this model for the k th iteration of the algorithm is given by

$$\log L(\Psi^{(k)}) = \sum_{j=1}^{n_{vr}} \int_{l_c} \int_{L_c} \log \left\{ \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \lambda_{vri} f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) + f(l_c; u) + f(L_c; U) dl_c dL_c \right\} \quad (1.2.1)$$

where

$$f(l_c; u) = N(\mu_l, \sigma_l) \quad (1.2.2)$$

$$f(L_c; U) = N(\mu_L, \sigma_L) \quad (1.2.3)$$

$$\mu_{vri} = l_{vc} + (L_{vc} - l_{vc}) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V; r = 1, \dots, R \quad (1.2.4)$$

and c denotes the cohort, given by

$$c = r - (i - 1) \quad (1.2.5)$$

Random effects are modelled on the first (l) and final (L) component means for each observed cohort. Cohorts of fish increase in length every year according to the cohort specific random effects on the means of the first and final components and an overall growth parameter (k). Correlation between cohort specific random effects are accounted for with a correlation parameter.

1.2.2 Cohort specific l and k

The observed data log likelihood for this model for the k th iteration of the algorithm is given by

$$\log L(\Psi^{(k)}) = \sum_{j=1}^{n_{vr}} \int_{l_c} \int_{k_c} \log \left\{ \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \lambda_{vri} f_{vri}(y_{jvr}; \theta_{vri}, a_{1v} + f(l_c; u) + f(k_c; U) dl_c dk_c \right\} \quad (1.2.6)$$

where

$$f(l_c; u) = N(\mu_l, \sigma_l) \quad (1.2.7)$$

$$f(k_c; U) = N(\mu_k, \sigma_k) \quad (1.2.8)$$

$$\mu_{vri} = l_{vc} + (L_v - l_{vc}) \frac{1 - k_c^{i-1}}{1 - k_c^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V; r = 1, \dots, R \quad (1.2.9)$$

where k_c denotes the cohort specific k random effect. Random effects are modelled on the first (l) component mean and growth parameter (k) for each observed cohort. Cohorts of fish increase in length every year according to the cohort specific random effects on the mean of the first component and and growth parameter (k). Correlation between cohort specific random effects are accounted for with a correlation parameter.

1.2.3 Cohort specific l and yearly k

The observed data log likelihood for this model for the k th iteration of the algorithm is given by

$$\log L(\Psi^{(k)}) = \sum_{j=1}^{n_{vr}} \int_{l_c} \int_{k_r} \log \left\{ \sum_{v=1}^V \sum_{r=1}^R \sum_{i=1}^M \lambda_{vri} f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) + f(l_c; u) + f(k_r; U) \right\} dl_c dk_r \quad (1.2.10)$$

where

$$f(l_c; u) = N(\mu_l, \sigma_l) \quad (1.2.11)$$

$$f(k_r; U) = N(\mu_k, \sigma_k) \quad (1.2.12)$$

$$\mu_{vri} = l_{vc} + (L_v - l_{vc}) \frac{1 - k_r^{i-1}}{1 - k_r^{M-1}}; \quad i = 1, \dots, M; v = 1, \dots, V; r = 1, \dots, R \quad (1.2.13)$$

Random effects are modelled on the first (l) component for each observed cohort and a yearly growth parameter (k). Cohorts of fish increase in length every year according to the cohort specific random effects on the mean of the first component and and growth parameter (k). Correlation between cohort specific random effects are accounted for with a correlation parameter.

1.3 Hierarchical model stability

In some years of the haddock data there is very little information for the first component of the mixture model. This created issues with the hierarchical models as the random effects for these years were able to vary too much, sometimes confounding the fitting of many of the mixture models around these years. Like with any length frequency analysis method, data or lack of it can be key. Fixing the SD for the random effects to a suitably small value still allowed the flexibility of the random effect but restricted it enough that we do not obtain confounding results. Another way to improve the stability of the hierarchical models for haddock would be to include more data, either from another survey or other area for example.

1.4 Model inputs etc.

Both versions of the model allow surveys from different periods of the year to be introduced through the a_{1v} argument, where a_{1v} is the assumed age of the first component. This argument allows for surveys from different periods of the year to be entered into the models, as long as a reasonable a_{1v} is assigned to it. For example, assuming a birth date of January 1st, haddock from the North Sea International Bottom Trawl Survey, which takes place in the third quarter of the year have an a_{1v} of 0.625 (i.e. the proportion of the year that has passed by the middle of Quarter 3). Similarly, a_{1v} inputs for white-bellied anglerfish for each survey were as follows: EVHOE (Q4) = 0.875 (middle of Q4); IE-IGFS (Q4) = 0.875; SP-PORC (Q4) = 0.73 (middle

of September).

2 S2 - haddock additional results

2.1 Model selection and inference

Due the wide array of starting parameters that were tested within the sensitivity analysis some models converged on “biologically unreasonable” L_∞ values and occasionally strange models where the second component mean μ_2 was in between the first actual components and converged to a very very small mixing proportion. This was dealt with by subsetting the models out of the final model selection.

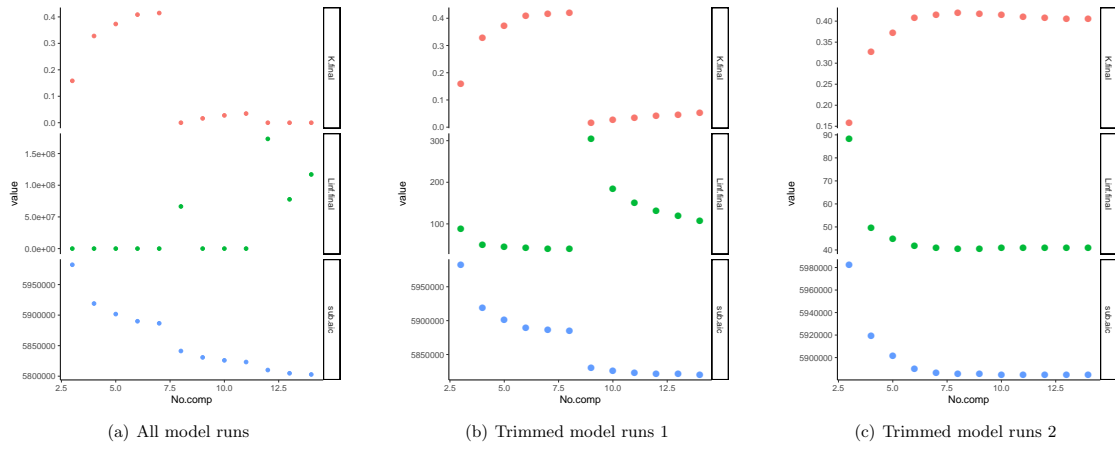


Figure 1: Final von Bertalanffy parameter estimates and sub-AIC values for best fitting (i.e. lowest sub-AIC values) basic LSD models on haddock from: (a) the full set of model runs; (b) the trimmed set of model runs with models $L_\infty < 500$ removed and (c) a trimmed set of model runs with models $L_\infty < 500$ and $\mu_2 \leq 15$ removed.

2.2 Deviations

Deviations (i.e. cohort specific means estimated with hierarchical LFEM model - cohort specific means estimated by ML on length -age-data) indicate some under-estimation in the early years in some of the observed cohorts.

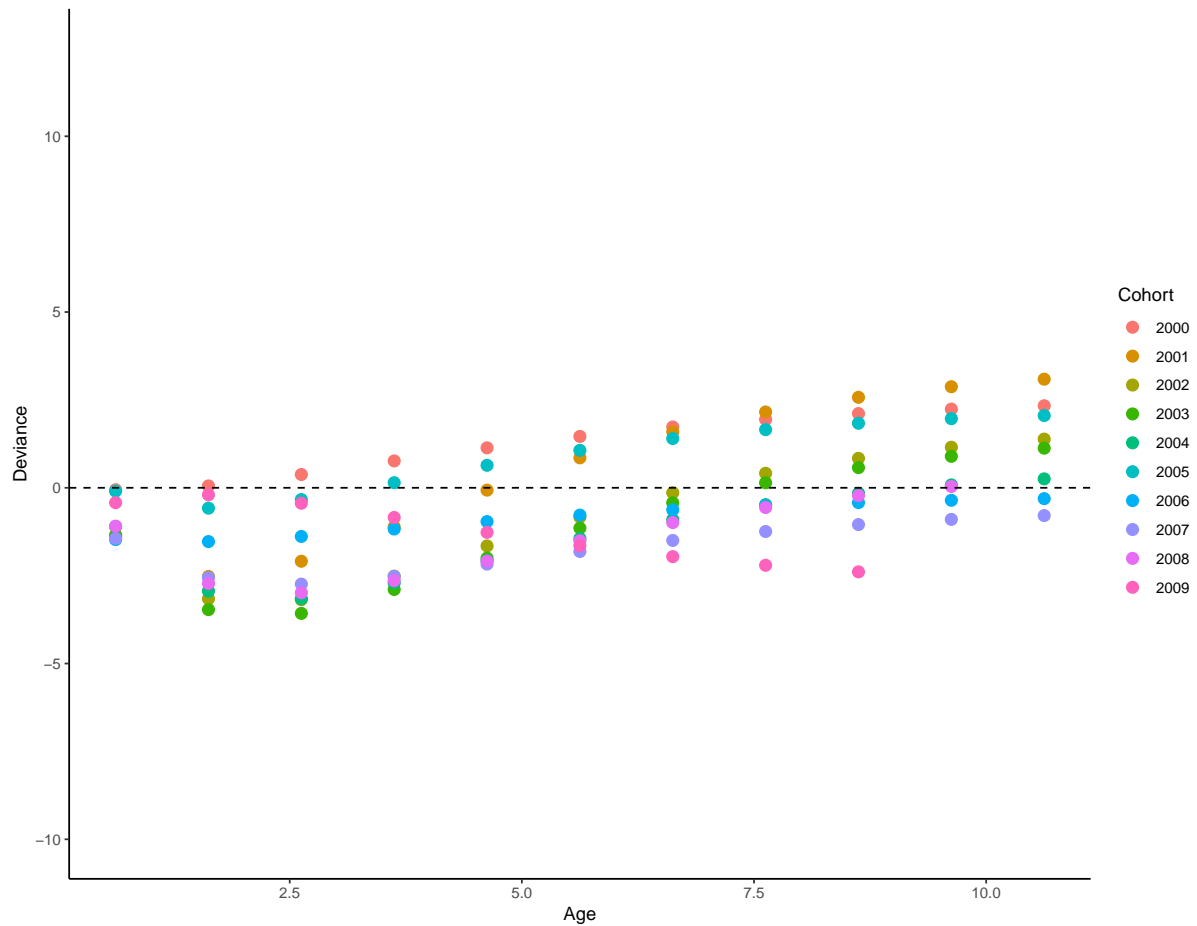


Figure 2: Deviations for cohorts of haddock observed at least up to the ninth component. Showing the difference between mean length-at-age from the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data of the cohort and cohort specific mean lengths estimated within the linear SD hierarchical LFEM model that varies l and k .

3 S3 - white-bellied anglerfish additional results

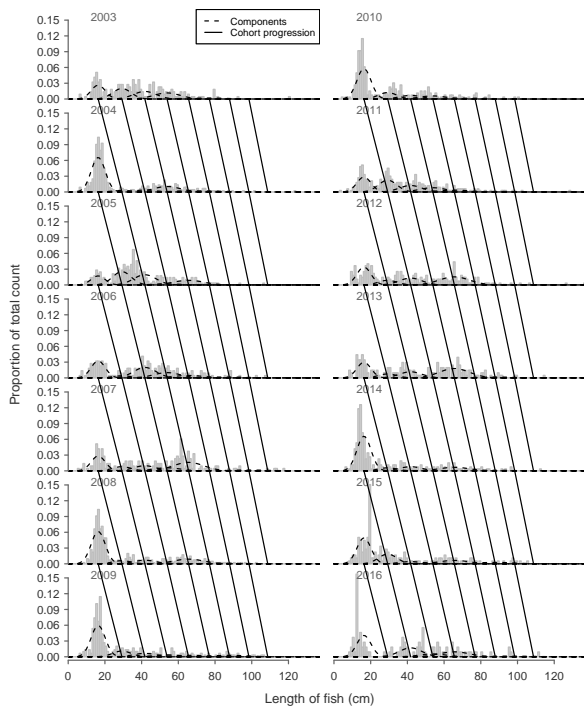
3.1 Spanish Porcupine Bank survey for best fitting basic model

Results presented here are for the Spanish Porcupine Bank survey (SP-PORC) which is conducted at a slightly different time of year to the other surveys, hence why component means differ.

Table 1: Component means and their standard deviation for white-bellied anglerfish for the Spanish Porcupine Bank survey, corresponding to the model runs presented in the main article

SD type	No. components	Component														<i>s</i>	<i>S</i>	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14			
CSD	8	14.81	30.27	45.20	59.63	73.57	87.03	100.04	112.60									4.91
	9	14.71	30.21	44.85	58.68	71.74	84.08	95.73	106.74	117.13								4.75
	10	14.68	29.86	43.98	57.12	69.34	80.70	91.27	101.11	110.25	118.76							4.62
	11	14.84	27.08	38.99	50.60	61.91	72.93	83.67	94.13	104.31	114.24	123.91						4.34
	12	14.81	26.69	38.20	49.36	60.17	70.65	80.81	90.66	100.20	109.45	118.42	127.11					4.23
	13	14.76	27.14	38.91	50.10	60.75	70.87	80.50	89.65	98.35	106.63	114.50	121.99	129.11				4.20
	14	14.73	26.83	38.20	48.88	58.92	68.36	77.22	85.55	93.38	100.74	107.65	114.15	120.26	125.99			4.12
LSD	8	14.20	29.80	44.48	58.28	71.26	83.48	94.96	105.77									3.96
	9	14.42	27.41	39.99	52.15	63.91	75.29	86.30	96.95	107.25								3.83
	10	14.55	26.53	38.19	49.55	60.61	71.38	81.87	92.09	102.03	111.72							3.85
	11	14.52	26.86	38.59	49.74	60.34	70.42	80.00	89.11	97.77	106.01	113.83						3.86
	12	14.52	26.91	38.52	49.41	59.63	69.21	78.19	86.62	94.52	101.93	108.88	115.40					3.87
	13	14.53	27.12	38.84	49.74	59.88	69.32	78.09	86.26	93.86	100.93	107.51	113.63	119.32				3.90
	14	14.51	27.35	39.18	50.07	60.11	69.36	77.88	85.72	92.96	99.62	105.75	111.41	116.61	121.41			3.91

3.2 EVHOE fitted distributions for best fitting basic model and corresponding hierarchical l/L model



(a) basic LFEM fit on EVHOE

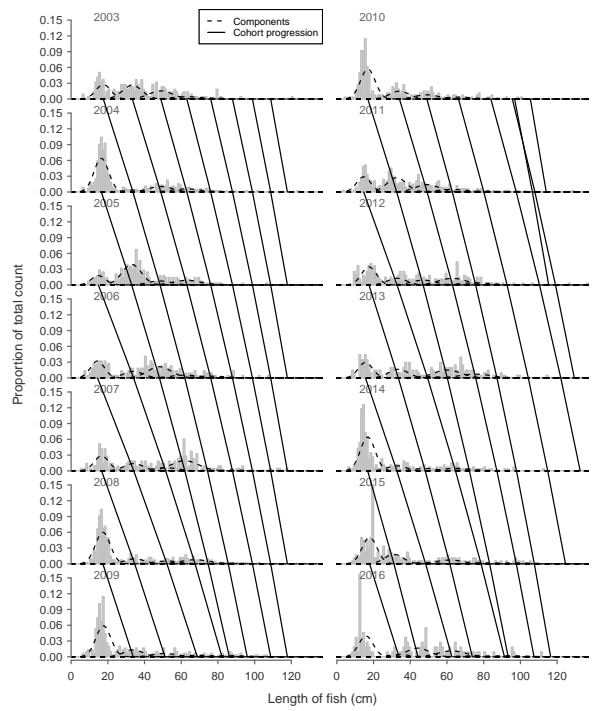
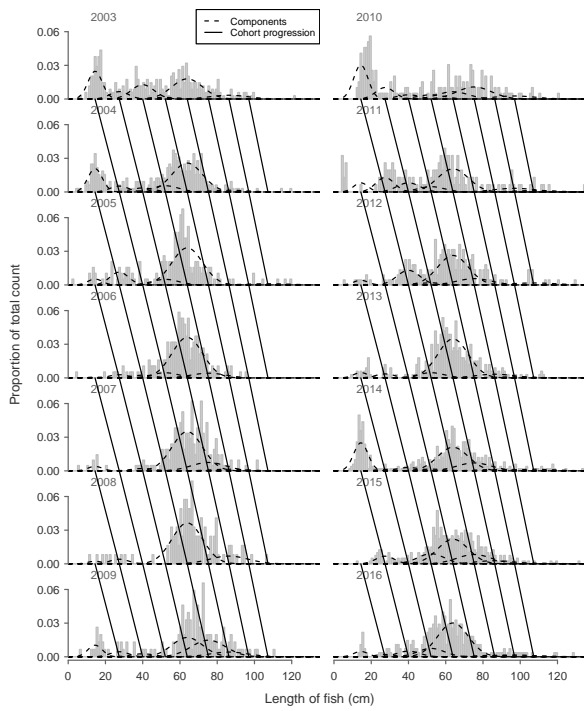
(b) corresponding hierarchical LFEM model fit to (a) with bivariate random effects on l and L on EVHOE

Figure 3: Length frequency distributions of white-bellied anglerfish from the quarter four French groundfish survey (EVHOE) with fitted distributions from: (a) the lowest sub-AIC basic LFEM model (9 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and L . Also shown is the progression of the cohort mean values (black solid lines). A similar plot for the Irish groundfish survey (IE-IGFS) can be found in section 3.2 of the main manuscript.

3.3 SP-PORC fitted distributions for best fitting basic model and corresponding hierarchical l/L model



(a) basic LFEM fit on SP-PORC

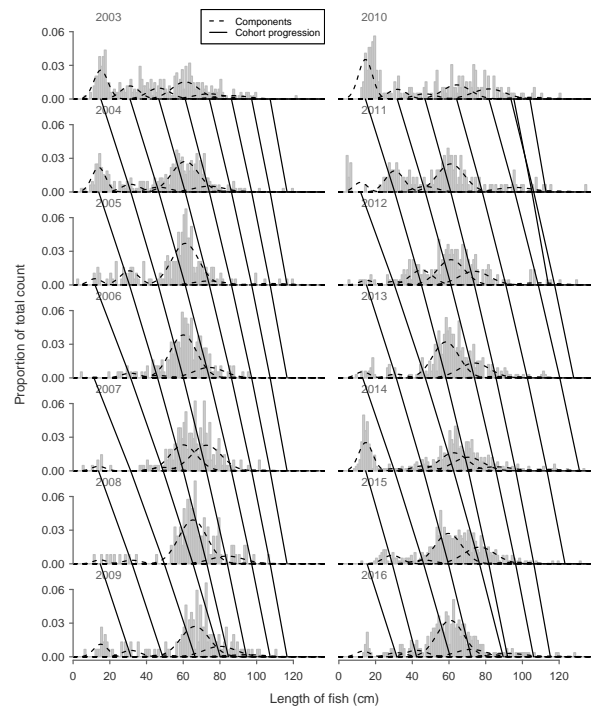
(b) corresponding hierarchical LFEM model fit to (a) with bivariate random effects on l and L on SP-PORC

Figure 4: Length frequency distributions of white-bellied anglerfish from the quarter three Spanish Porcupine Bank survey (SP-PORC) with fitted distributions from: (a) the lowest sub-AIC basic LFEM model (9 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and L . Also shown is the progression of the cohort mean values (black solid lines). A similar plot for the Irish groundfish survey (IE-IGFS) can be found in section 3.2 of the main manuscript.

3.3.1 White-bellied anglerfish hierarchical model with cohort specific l and k

Overall CVBG parameters for the hierarchical LSD model with cohort specific l and k corresponding to the best fitting basic model ($\hat{K} = 0.093$, $\hat{L}_\infty = 207.7$ cm and $\hat{t}_0 = -0.006$) showed a similar pattern to the previous hierarchical model in that parameter estimates differed to those given by the basic model. The resulting growth curve was very similar to the l/L hierarchical model but differed somewhat from the best fitting basic LFEM model. There was also a correlation estimated for cohort l and k random effects ($\hat{\rho} = 0.66$). Cohort specific \hat{L}_∞ estimates ranged from 156.9 - 362.7 cm and \hat{K} estimates ranged from = 0.04 - 0.16. However, as \hat{L}_∞ and \hat{K} strongly covary these parameters only equate to a couple of centimetres difference in growth each year between cohorts. Patterns in the cohort specific random effect on k showed marked variation between cohorts (Figure 5).

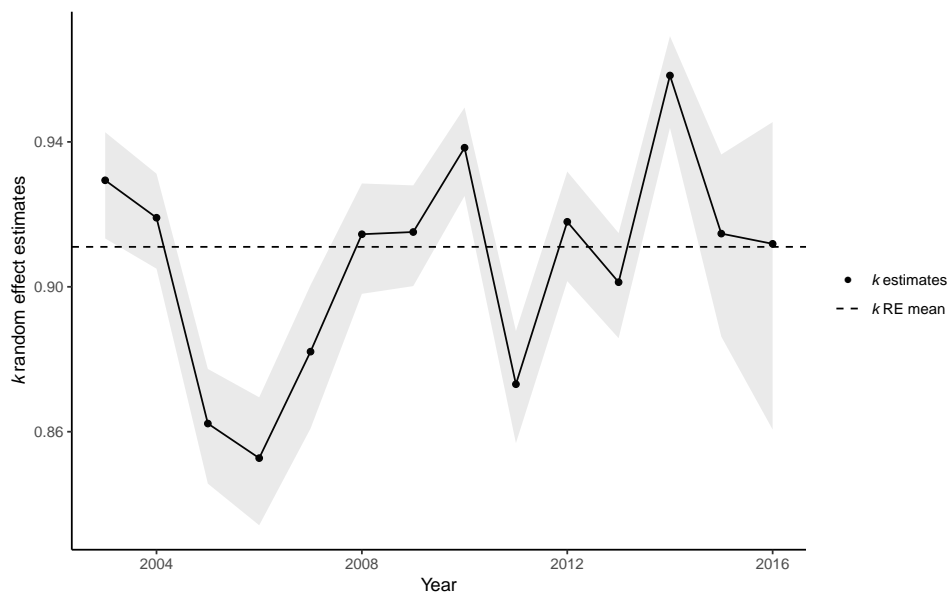


Figure 5: Estimated random effects on cohort specific growth rate parameter (k) from a nine component hierarchical LSD model on length frequency distributions of white-bellied anglerfish. Shaded area denotes 2*standard error of the random effect estimates and the dashed line denotes the mean of the random effect. Parameter estimates are for quarter four surveys (EVHOE and IE-IGFS).

3.3.2 White-bellied anglerfish hierarchical model with cohort specific l and yearly k

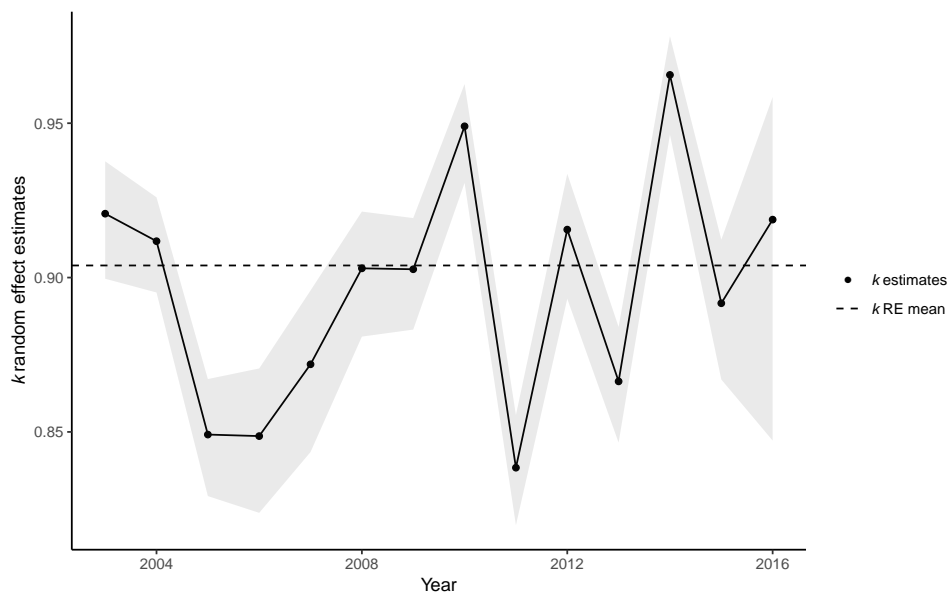


Figure 6: Estimated random effects on year specific growth rate parameter (k) from a nine component hierarchical LSD model on length frequency distributions of white-bellied anglerfish. Shaded area denotes 2*standard error of the random effect estimates and the dashed line denotes the mean of the random effect. Parameter estimates are for quarter four surveys (EVHOE and IE-IGFS)

Overall CVBG parameters for the hierarchical LSD model with cohort specific l and yearly k corresponding to the best fitting basic model ($\hat{K} = 0.10$, $\hat{L}_\infty = 194.1$ cm and $\hat{t}_0 \approx 0$) were different from those estimated from the basic model but similar to other hierarchical models. There was a correlation estimated for cohort l and yearly k random effects ($\hat{\rho} = 0.58$). \hat{L}_∞ estimates ranged from 187.0 - 204.6 cm and yearly \hat{K} estimates ranged from = 0.03 - 0.18. Variation in the year specific random effect on \hat{k} were similar to those of the cohort specific \hat{k} (Figure 6).

References

Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38.

Appendix C

**Working document to ICES
Benchmark Workshop on
Anglerfish Stocks in the ICES
Area (WKANGLER), 2018**

An alternative method for length frequency analysis and its application to monkfish

Luke Batts, C oil n Minto, Hans Gerritsen and Deidre Brophy

October 3, 2017

1 Introduction

Much work has been conducted on European anglerfish (*Lophius piscatorius* and *Lophius budegassa*) life history traits over the years and much of this has focused on growth patterns [Farina et al., 2008]. These studies have predominantly used calcified structures with annual rings to age fish and produce growth estimates, however there has been well documented difficulties with ageing anglerfish this way [Woodroffe et al., 2003, Farina et al., 2008]. A notable exception in recent years where age and growth estimation from illicia has been further validated with various length frequency analyses is Landa *et al.s* (2013) study on *L. piscatorius* on the Porcupine Bank. Using modal progression analysis on length frequency distributions over a number of years Landa *et al.* (2013) were able to track a cohort of *L. piscatorius* through eight successive years of the Spanish Porcupine Bank survey. Thus providing support for the aging by illicia that had also been conducted in the study.

Fisheries surveys are an important aspect of fisheries research and offer fishery independent estimates of abundance and structure of fish populations. Length frequency distributions from surveys have been used across many marine species to produce growth estimates. With this in mind this work has looked at developing an alternative approach to mixture models to estimate modes of cohorts across years and surveys. The intention is to both estimate credible growth parameters, as well as explore the differences between surveys and the possibility of combining them.

2 Methods

2.1 Data

Length frequency data of both *Lophius* species was used from the three surveys that cover the anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d; the French EVHOE survey, the Spanish Porcupine Bank survey and the Irish Groundfish survey. Preliminary results shown here were conducted on mixed sex distributions of each species.

2.2 Model and Algorithm

The length frequency distributions were modelled using mixture models, where the means of the components (distributions within the mixture model) were constrained to follow a von Bertalanffy growth function, commonly represented by

$$L(t) = L_{\infty}\{1 - \exp[-K(t - T_0)]\}. \quad (2.2.1)$$

Where L_{∞} is the asymptotic length of the fish, K controls the curvature or growth rate, t is the time step and T_0 is the theoretical time when an individual would have length zero. In terms of length frequency distributions of fish and the components of a mixture model this can be better represented by

$$\mu_i = L_{\infty}\{1 - \exp[-K(a_i - T_0)]\}. \quad (2.2.2)$$

Where μ_i is the mean length of fish at age a_i and $i = 1, \dots, \text{maximum age class}(M)$. Schnute and Fournier (1980) re-parametrised this equation into one more appropriate to length frequency analysis and this is represented by

$$\mu_i = l + (L - l) \frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \dots, M. \quad (2.2.3)$$

Where l and L are the first and final mean lengths of observed fish cohorts, i.e. μ_1 and μ_M . k is equivalent to $\exp(-K)$ and is a fraction that controls the distance between two successive mean lengths. The classical von Bertalanffy parameters can then be calculated from these parameters (see [Schnute and Fournier, 1980] for details).

The parameters of a mixture model (i.e. the mixing proportions of the components as well as their means and standard deviations) can easily be estimated with maximum likelihood if the specific component that a measurement belongs to is known. However, this is rarely the case when analysing length frequency distributions and this study has used the EM algorithm as the basis for estimating the parameters. The EM algorithm is a commonly used iterative method for finding the maximum likelihood parameters of a mixture model, when the model depends on latent variables [Dempster et al., 1977]. The EM algorithm is made up of two steps; the expectation (E) step and the maximisation (M) step. The E step essentially calculates the posterior probability for each data point in each component (in the first step some reasonable starting values for the components are needed in order for these can be calculated). Using this posterior probability in the M step, we can use maximum likelihood to gain new estimates of the parameters by maximising the conditional expectation. This study uses an adapted M step, where the means of the mixture model's components are constrained to the re-parametrised von Bertalanffy growth function within the conditional expectation. This process is iterated many times over, calculating new posterior probabilities from the new parameter estimates, then re-estimating those parameters and so on until the algorithm converges.

3 Results and discussion

These results are preliminary and for the moment the model only allows a constant standard deviation for the components and no yearly effects to be modelled. Multiple surveys and years are incorporated into the model, here we have assumed that the data from the three surveys catches nine cohorts/components of the monkfish species. Also of note is that for the calculation of T_0 from l, L, k and number of components/cohorts captured in the survey data, it is assumed the age of the first cohort is 0. However this can be easily changed in the R function, under the argument "age0".

Table 3.1: von Bertalanffy growth parameters for monkfish species

Species	re-parameterised VBG parameters			classical VBG parameters		
	$l(\text{cm})$	$L(\text{cm})$	k	$L_\infty(\text{cm})$	K	T_0
<i>L.piscatorius</i>	16.78	117.08	0.929912	244.31	0.072665	-0.979066
<i>L.budegassa</i>	11.07	77.74	0.888131	119.84	0.118635	-0.816986

Table 3.2: Component/cohort means and their standard deviation for monkfish species, assuming surveys capture nine cohorts

Species	Component/cohort									SD
	1	2	3	4	5	6	7	8	9	
<i>L.piscatorius</i>	16.78	32.72	47.55	61.34	74.17	86.09	97.18	107.49	117.08	4.796
<i>L.budegassa</i>	11.07	23.24	34.04	43.64	52.17	59.74	66.46	72.43	77.76	3.165

Figures 1-3 show length frequency distributions for *L.piscatorius* through the years of the surveys, with the components/cohorts modelled by the mixture model overlaid. Figures 4-6 show length frequency distributions for *L.budegassa* through the years of the surveys, with the components/cohorts modelled by the mixture model overlaid.

The main advantage of this method has over other length frequency analyses is that it is completely objective. Issues such as selecting the correct bin size, although may alter how the data is plotted, has no bearing on the computation of the parameters as the algorithm uses the individual lengths of each fish. As mentioned previously the model is still being developed and it is hoped that a variable standard deviation can be incorporated as well as some process for allowing component means to vary somewhat between years. This may then allow the model to deal with slight differences in spawning and growth, in addition to "anomalous" length frequency distributions such as that in Fig. 5 in the year 2014 or peaks in the distribution before the first modelled component/cohort. The sensitivity of the algorithm to different numbers of components will also be tested.

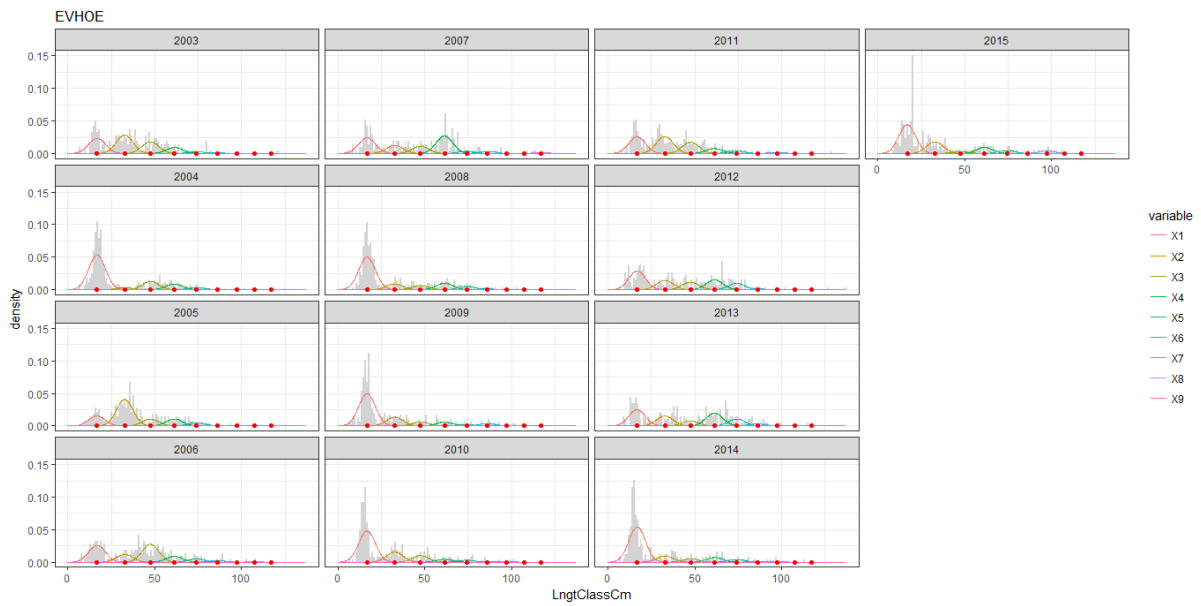


Figure 1: Length frequency distributions of *L. piscatorius* across 13 years of survey data from the French survey (EVHOE), with the fitted normally distributed components of the mixture model

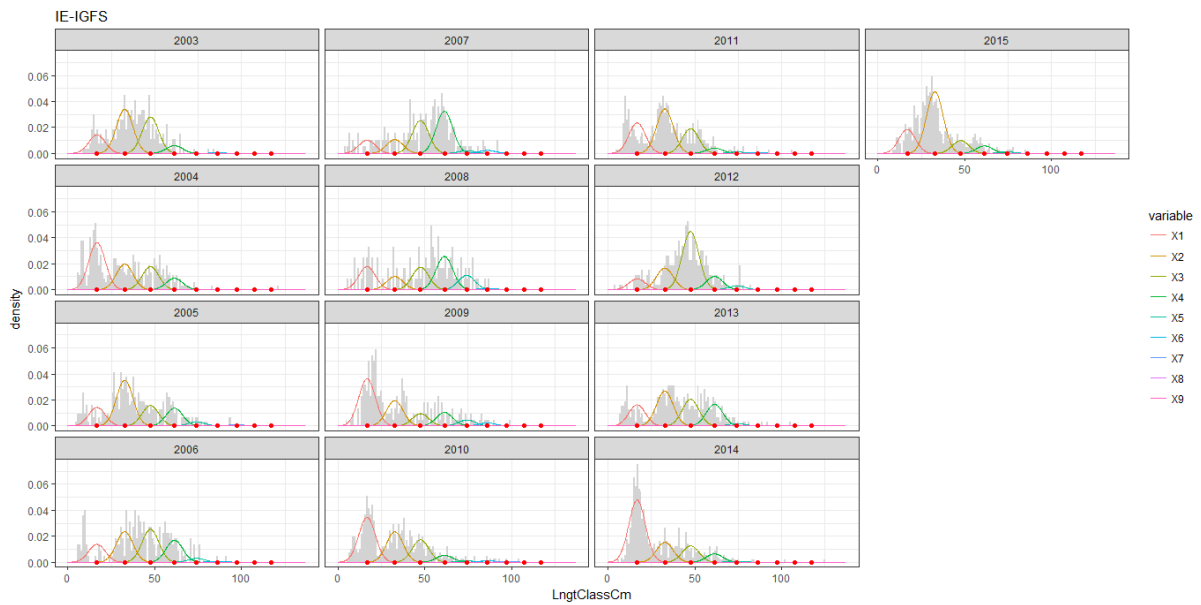


Figure 2: Length frequency distributions of *L. piscatorius* across 13 years of survey data from the Irish Groundfish survey (IE-IGFS), with the fitted normally distributed components of the mixture model

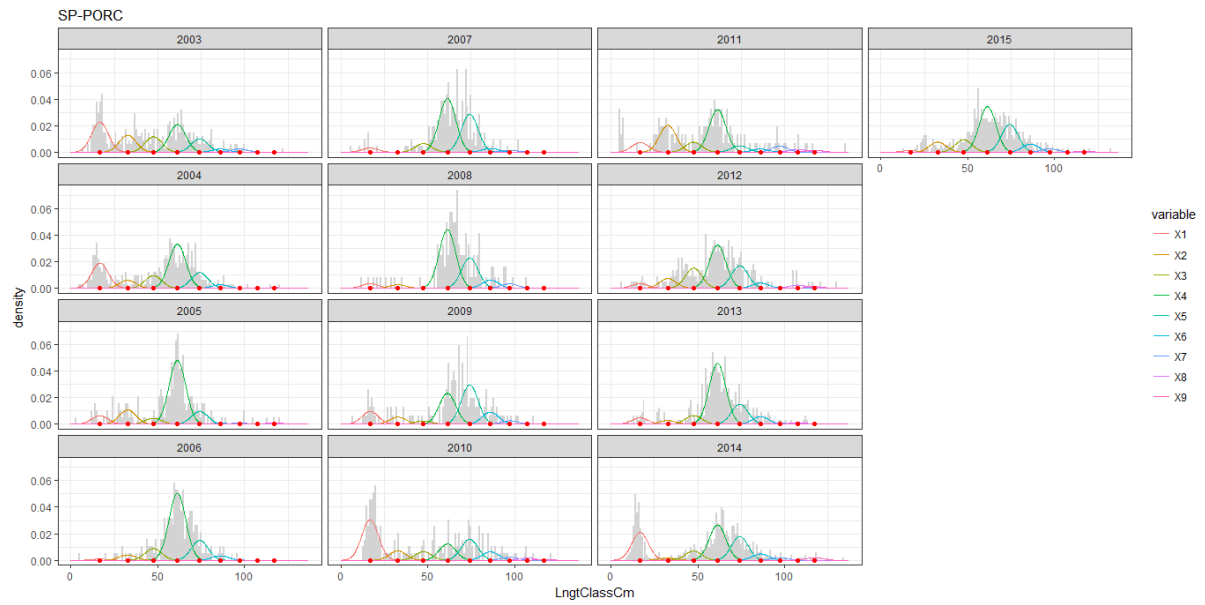


Figure 3: Length frequency distributions of *L. piscatorius* across 13 years of survey data from the Spanish Porcupine Bank survey (SP-PORC), with the fitted normally distributed components of the mixture model

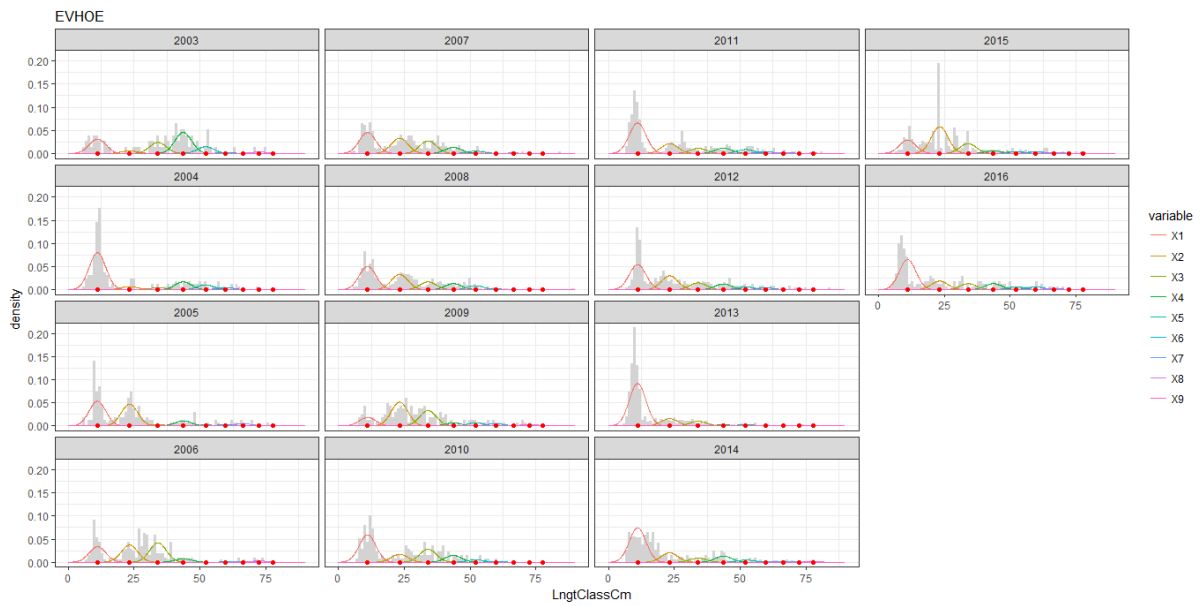


Figure 4: Length frequency distributions of *L.budegassa* across 14 years of survey data from the French survey (EVHOE), with the fitted normally distributed components of the mixture model

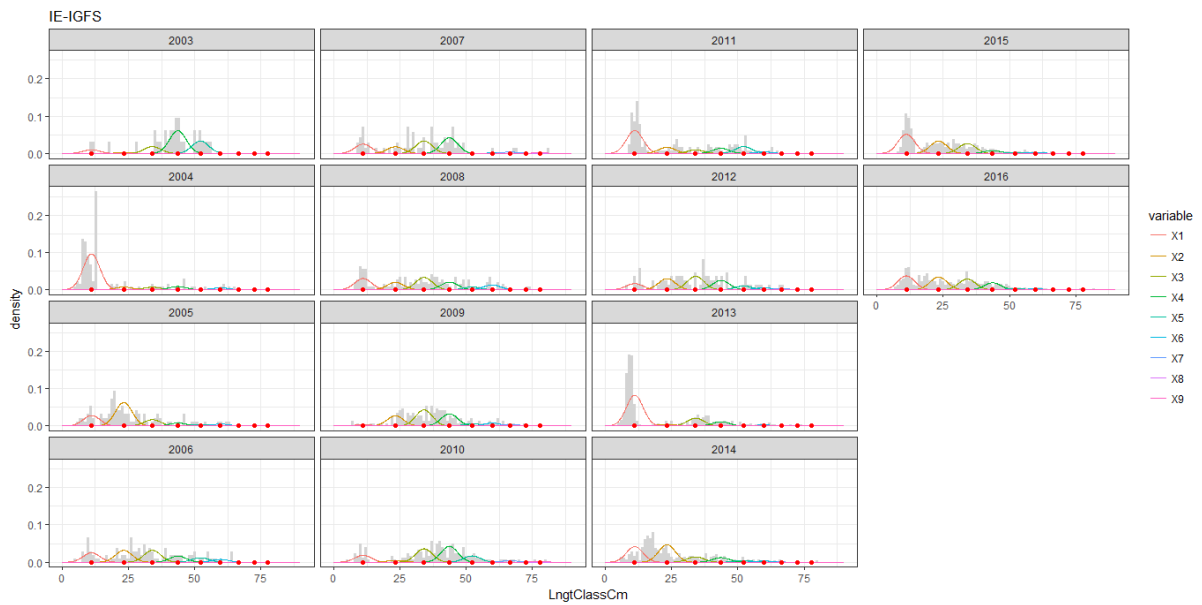


Figure 5: Length frequency distributions of *L.budegassa* across 14 years of survey data from the Irish Groundfish survey (IE-IGFS), with the fitted normally distributed components of the mixture model

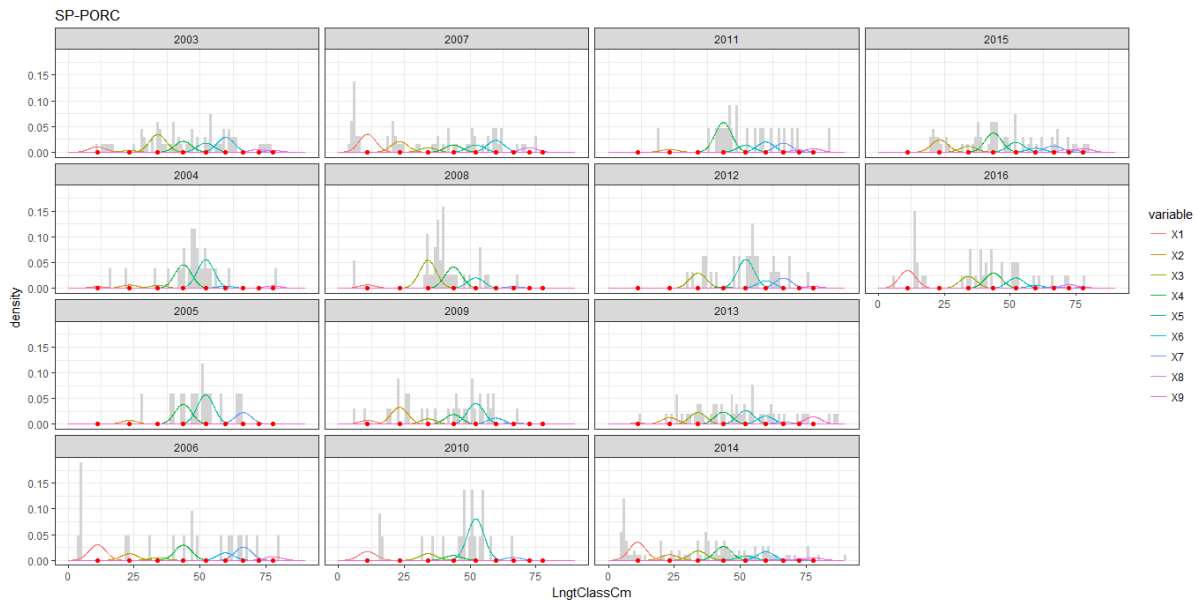


Figure 6: Length frequency distributions of *L. budegassa* across 14 years of survey data from the Spanish Porcupine Bank survey (SP-PORC), with the fitted normally distributed components of the mixture model

References

- [Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.
- [Farina et al., 2008] Farina, A., Azevedo, M., Landa, J., Duarte, R., Sampedro, P., Costas, G., Torres, M., and Cañas, L. (2008). Lophius in the world: a synthesis on the common features and life strategies. *ICES Journal of Marine Science: Journal du Conseil*, 65(7):1272–1280.
- [Schnute and Fournier, 1980] Schnute, J. and Fournier, D. (1980). A new approach to length–frequency analysis: growth structure. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(9):1337–1351.
- [Woodroffe et al., 2003] Woodroffe, D., Wright, P., and Gordon, J. (2003). Verification of annual increment formation in the white anglerfish, *lophius piscatorius* using the illicia and sagitta otoliths. *Fisheries Research*, 60(2):345–356.

Appendix D

Supplementary material for chapter 3

Supplementary Appendices

MS: Numbers or mass? Comparison of two theoretically different stage-based stock assessment models and their ability to model simulated and real- life stocks

Luke Batts, C oil n Minto Hans Gerritsen and Deirdre Brophy

February 4, 2022

Code and examples of model runs can be found at https://github.com/lbatts/sim_sbar. An R package was also developed and can be found at: <https://github.com/lbatts/sbar>

1 Additional methods

1.1 Schnute

1.1.1 Population dynamics

In this weight structured population model the number of fish in a population can be described as a density (as weight is continuous), the integrals for population number (N_t) and biomass (N_t^*) respectively are

$$(1) \quad N_t = \int_V^{V^\infty} N(w, t)dw$$

$$(2) \quad N_t^* = \int_V^{V^\infty} wN(w, t)dw$$

Given the entire fully selected population is comprised of two stages (as described in the main text), recruits (numbers R_t and biomass R_t^*) and previously exploited population (numbers P_t and biomass P_t^*), these stages can also be represented as such with the specific subscripts and superscripts of the stage's integral

$$(3) \quad R_t = \int_V^{V'} R(w, t)dw$$

$$(4) \quad R_t^* = \int_V^{V'} wR(w, t)dw$$

$$(5) \quad P_t = \int_{V'}^{V^\infty} P(w, t)dw$$

$$(6) \quad P_t^* = \int_{V'}^{V^\infty} wP(w, t)dw$$

19 It is important to note here that as time is not continuous in this model, there is an
 20 implicit assumption that growth occurs at the end of the year. Total survival of the
 21 population from one year to the next is also assumed to be weight-independent (analogous
 22 to age-structured delay-difference models (Schnute, 1985)). These fundamentals lead to
 23 the key dynamic equation for a weight structured population model described in Schnute
 24 (1987) as

$$(7) \quad \int_{w'_1}^{w'_2} N(w, t) dw = \tau_t \int_{w_1}^{w_2} N(w, t) dw$$

25 where τ_t is the fraction of the population that survives total mortality (natural and
 26 fishing) during year t . This equation states that any fish in the arbitrary weight interval
 27 $[w_1, w_2]$ that survives to the next year will reach the weight interval $[w'_1, w'_2]$.

28 1.1.2 Deterministic equations

29 The basic delay-difference equation can be simply described as predicting biomass at
 30 time $t+1$ from three basic components: recruitment biomass at $t+1$, surviving biomass
 31 from time t and growth of the surviving fish (Deriso, 1980; Schnute, 1985; Hilborn and
 32 Walters, 1992). Schnute (1987) represents this process with the equation

$$(8) \quad N_{t+1}^* - R_{t+1}^* = \tau_t \left(\frac{X'_t}{X_t} \right) N_t^* = P_{t+1}^*$$

33 which states that the biomass in year $t + 1$ minus recruitment, or previously exploited
 34 biomass P_{t+1}^* , is derived from the surviving biomass from year t multiplied by a growth
 35 factor of $\frac{X'_t}{X_t}$ as defined in the main text. This implies that the growth of biomass from
 36 one year to the next depends solely on average weights (growth parameters applied to
 37 calculate X'_t are estimated from average weights).

38 Schnute (1987) introduces catch $C(w, t)$, spawning stock $S(w, t)$ and the population index
 39 $I(w, t)$ as similar density functions to equation 1. Three timing parameters were also
 40 introduced, in addition to multiplying the density functions by w and integrating each
 41 from V to V_∞ , to give the key biomass equations

$$(9) \quad \tau_t N_t^* = \sigma [N_t^* - \mu C_t^*] - (1 - \mu) C_t^*$$

$$(10) \quad S_t^* = (1 - \nu) N_t^* + \nu \tau_t N_t^*$$

$$(11) \quad I_t^* = q [(1 - \theta) N_t^* + \theta \tau_t N_t^*]$$

42 where σ is the survival from natural mortality, q is the index catchability and the timing
 43 parameters are defined as μ fraction of catch removed before natural mortality, ν timing
 44 of spawning or fraction of total mortality prior to spawning and θ timing of the index
 45 or fraction of total mortality prior to indexing the population (Schnute, 1987). A key

46 assumption here is that catchability is constant for the survey and is equivalent to the
 47 fully selected catch.

48 Following on from the biomass equations above, τ_t can be removed from equations 10 and
 49 11 using 9

$$(12) \quad S_t^* = [1 - \nu(1 - \sigma)]N_t^* + \nu[1 - \mu(1 - \sigma)]C_t^*$$

$$(13) \quad I_t^* = q \left([1 - \theta(1 - \sigma)]N_t^* - \theta[(1 - \sigma)]C_t^* \right)$$

50 Equation 13 predicts a total biomass index from N_t^* , which is comprised of R_t^* and P_t^* .
 51 Both R_t^* and P_t^* are calculated from the biomass in the previous year N_{t-1}^* . Schnute (1987)
 52 also proposes two alternative equations for predicting the population index, making use
 53 of all three time series of mean weights (X , Z and Y). ω_t is defined as the fraction of
 54 total biomass in year t due to newly recruited fish and its derivation is detailed in the
 55 main text.

56 ω_t values can then used to compute estimates of N_t^* using only either the recruit stage
 57 R_t^* or the previously exploited population stage P_t^* .

$$(14) \quad N_t^* = \frac{R_t^*}{\omega_t} = \frac{P_t^*}{1 - \omega_t}$$

58 We implement this model version that estimates N_t^* from P_t^* (P_t^* being calculated from
 59 N_{t-1}^*) at each time step.

60 1.1.3 Recruitment

61 When using a size-structured model, recruits reaching the critical size V are potentially
 62 different ages depending on the number of years it takes for fish to reach V (Schnute,
 63 1987). Assumptions within the model hold because regardless of the variable ages within
 64 the recruits size class $[V, V')$, all fish within this size class will in theory grow on to be
 65 part of previously exploited biomass in the following year, providing the recruit stage has
 66 been correctly defined.

67 1.1.4 Estimation

68 If information on growth is available and weights-at-age are available these can be used
 69 (as is common for delay-difference models) to estimate growth parameters with a linear
 70 model. This was the unbiased mean weight/ growth configuration,

$$(15) \quad \bar{w}_{a+1} = W + \rho\bar{w}_a$$

71 where \bar{w}_a is the estimated weight-at-age and \bar{w}_{a+1} is the weight-at-age a year older from
 72 sampling.

73 Another option, the (potentially) skewed mean weight/ growth configuration, suggested
74 as a check by Schnute (1987), can be used to estimate growth parameters through esti-
75 mation of a linear model on overall mean weights and previously-exploited stage mean
76 weights from catch sampling:

$$(16) \quad X'_t = W + \rho \bar{X}_t = \bar{Z}_{t+1}$$

77 This equation states that the entire population sampled mean weight (\bar{X}) in time t , after a
78 year of growth, will be equivalent to the sampled mean weight of the previously-exploited
79 population (\bar{Z}) in time $t+1$. This relationship enables the estimation of the parameters
80 W and ρ prior to assessment model by fitting a simple linear model where \bar{X}_t and \bar{Z}_{t+1}
81 are generally calculated from the chosen weight intervals applied to the catch data. When
82 fitting these linear models prior to running the assessments, residuals were assumed to
83 be normally distributed.

Table 1: Summary of the starting values for key parameters in the Schnute (1987) assessment model. A hat denotes estimated parameters in the model.

Notation	Starting value	
	simulated (large demersal/small pelagic)	large demersal
σ	$e^{-0.25}/e^{-0.7}$ (fixed)	$e^{-0.25}$ (fixed)
\hat{q}_s	1E-08	1E-08, 1E-05, 2E-08
$\hat{\delta}_s$	0.1	0.1
λ	0.1 (fixed)	0.2 (fixed)
\hat{N}_1^*	5 x maximum C_t^*	age-based assessment value + 1E+04
\hat{F}_t	0.5	0.3
ν	0 (fixed)	0 (fixed)
θ_s	0 (fixed)	0.875, 1, 0.875 (fixed)

84 1.2 CSA Estimation

Table 2: Summary of the starting values for key parameters in CSA assessment model. Hat operator denotes estimated parameters in the model.

Notation	Starting value	
	simulated (large demersal/small pelagic)	large demersal
M	0.25/0.7 (fixed)	0.25 (fixed)
\hat{q}_s	1E-08	1E-06, 1E-05, 2E-06
$\hat{\tau}_s$	0.1	0.1
ω	0.1 (fixed)	0.2 (fixed)
θ_s	0 (fixed)	0.875, 1, 0.875 (fixed)
\widehat{P}_1	4 x maximum C_t	age-based estimated value + 1E+04
\widehat{R}_t	2 x maximum C_t	age-based estimated value + 1E+03
\widehat{F}_t	0.5	0.3

1.3 Simulated stocks data generation

Initial small pelagic fish parameters (von Bertalanffy growth parameters and weight-length parameters, $weight = aLength^b$) were means taken across all entries for *Clupea harengus harengus* from the online Fishbase database (Froese et al., 2021) ($L_\infty = 30.8$ cm, $k = 0.35$, $t_0 = -0.87$, $a = 6.27E-06$, $b = 3.09473$). Values for length and age where 50% of the population are mature were taken from a spring-spawning herring stock assessment for Skagerak, Kattegat and western Baltic ($L^{50} = 21.4$ cm and $a^{50} = 2.5$) (ICES, 2019c). Steepness of the stock-recruit function was estimated for the genus using the “Fishlife” package in R (Thorson, 2019) and set at 0.73. An value for virgin stock biomass was approximated from the spring-spawning herring stock assessment for Skagerak, Kattegat and western Baltic and set at approximately the maximum spawning stock biomass the assessment estimated (ICES, 2019a). Age range was from 0-8 and natural mortality $M = 0.7$. Natural mortality was constant, estimated through life-history theory (Griffiths and Harrod, 2007).

Initial large demersal fish parameters were those used in the official assessment of the *Lophius piscatorius* stock in ICES areas 7.b-k, 8.a-b and 8.d ($L_\infty = 171.0$ cm, $k = 0.1075$, $t_0 = -1E-06$, $a = 3.03E-05$, $b = 2.82$, $L^{50} = 65.6$ cm and $a^{50} = 4.5$) (ICES, 2019d,b). Steepness of the stock-recruit function was estimated for the genus using the “Fishlife” package in R (Thorson, 2019) and set at .95. A value for virgin stock biomass was approximated by taking the product of spawners-per-recruit at fishing mortality = 0 and mean recruitment from the 2019 assessment. Age range was from 0-20 and natural mortality $M = 0.25$. Natural mortality was constant and set at the value used in the age-based assessment (ICES, 2019b).

Fishing mortality was set at $0.5 \times F_{msy}$ for the entire time series for stocks set up with constant harvest dynamics. Fishing mortality for stocks set up with the one-way trip harvest dynamics had a steadily increasing fishing mortality from $0.5 \times F_{msy}$ towards $0.8 \times F_{crash}$ over the time period. The rate of the increase depended on the length of the time series. Finally, stocks set up with the roller-coaster harvest dynamics had an increasing fishing mortality from $0.5 \times F_{msy}$ to $0.8 \times F_{crash}$ for the first third of the time series, then stayed at this fishing mortality rate for five years. Fishing mortality was then steadily decreased to F_{msy} at the end of the time series. Here, F_{crash} is the level of fishing mortality that will drive the stock to extinction and F_{msy} is the level of fishing mortality that provides maximum sustainable yield. Values for F_{msy} and F_{crash} are calculated according to the deterministic equations of the simulation model used within FLR (Kell, 2018).

Stock-recruitment model for the simulated stocks was a Beverton-Holt function:

$$(17) \quad R_{t+1} = \frac{aS_t^*}{b + S_t^*}$$

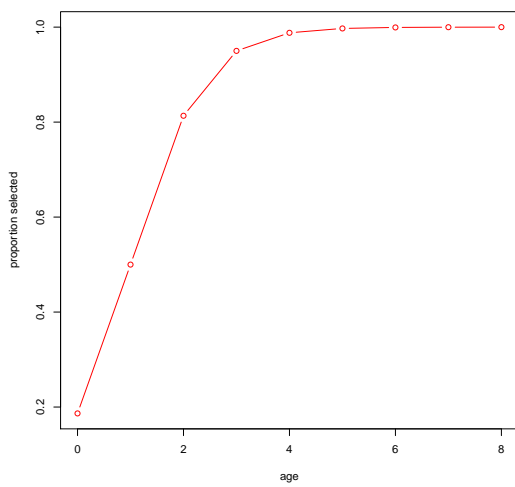
where S_t^* is spawning stock biomass and R_t is recruit numbers. a and b are the parameters of the stock-recruitment curve, equivalent to the maximum or asymptotic number of recruits (a) and the spawning stock biomass needed to produce recruitment equal to $a/2$ on average (Beverton and Holt, 1957). The S-R function was parameterised for virgin biomass and steepness, which were defined for each life-history strategy

$$(18) \quad a = \frac{4 \cdot vb \cdot h}{spr0 (5 \cdot h - 1.0)}$$

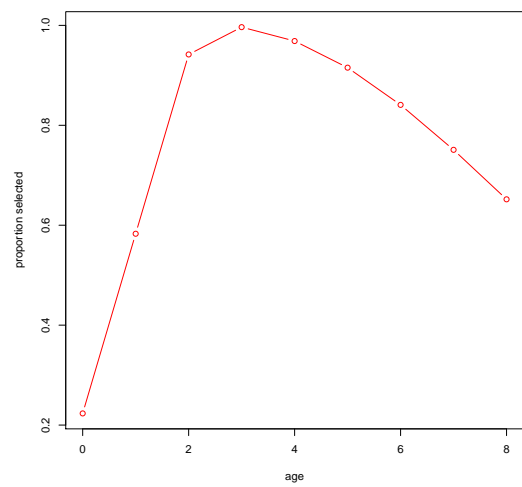
$$(19) \quad b = \frac{vb (1.0 - h)}{5 \cdot h - 1.0}$$

126 where vb is virgin biomass, h is steepness and $spr0$ is the spawners-per-recruit when
 127 fishing mortality is zero.

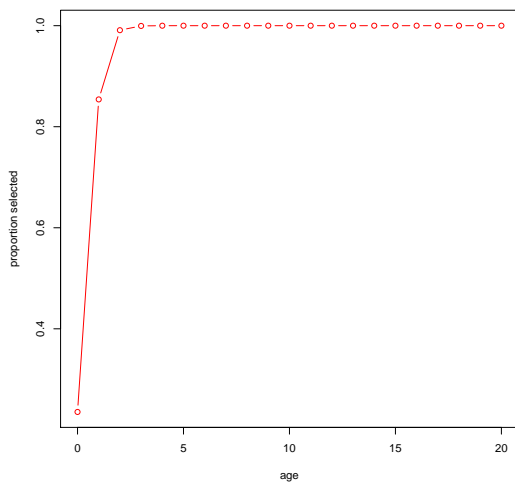
128 1.3.1 Selectivity patterns for simulated stocks



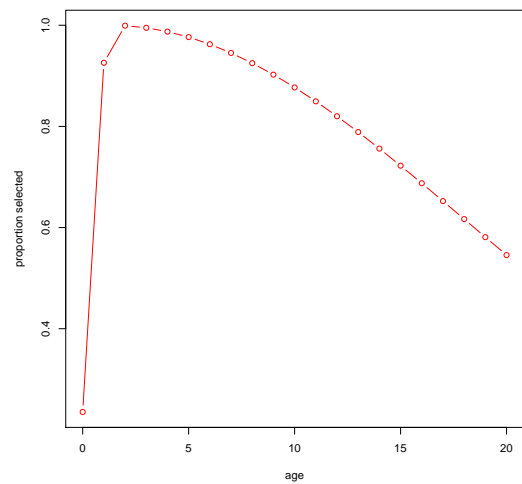
(a) small pelagic logistic selectivity



(b) small pelagic dome-shaped selectivity



(c) large demersal logistic selectivity



(d) large demersal dome-shaped selectivity

Figure 1: Selectivity patterns for logistic and dome-shaped scenarios for large demersal and small pelagic life-history strategies

129 **2 Additional results**

130 **2.1 Simulated stock convergence**

131 All factors of the simulated stock scenarios, except selectivity pattern, did not affect the ability of assessment models to converge with identifiable parameters. This is
 132 indicated by the approximately equal percentages across factor levels, for each factor, for
 133 each assessment model (Table 3).
 134

Table 3: Percentage (to two decimal points) Convergence of assessment runs across the five assessments and all factors from the simulation framework except selectivity.

Assessment	Classification	HD			LH		TS		SR		AR	
		c	ow	rc	large demersal	small pelagic	short	long	recsd0.1	recsd0.4	nocor	0.6rho
CSA	All parameters are estimable	99.85	100.00	100.00	99.99	99.92	99.90	100.00	99.96	99.94	99.97	99.93
	Converged but not all parameters identifiable	0.06				0.04	0.04		0.01	0.03	0.01	0.03
	Did not converge	0.08			0.01		0.06		0.03	0.03	0.01	0.04
S1 _c	All parameters are estimable	33.33	33.33	36.92	35.72	33.33	35.72	33.33	33.69	35.36	34.60	34.46
	Converged but not all parameters identifiable	8.15	6.81	9.40	8.89	7.35	5.76	10.47	7.58	8.65	7.75	8.49
	Did not converge	58.52	59.85	53.69	55.39	59.32	58.51	56.19	58.72	55.99	57.65	57.06
S1 _{waa}	All parameters are estimable	100.00	100.00	99.60	99.74	100.00	99.74	100.00	99.83	99.90	99.88	99.86
	Converged but not all parameters identifiable											
	Did not converge			0.40	0.26		0.26		0.17	0.10	0.12	0.14
S0 _c	All parameters are estimable	33.33	33.33	33.65	33.54	33.33	33.54	33.33	33.33	33.54	33.38	33.50
	Converged but not all parameters identifiable	66.52	65.90	66.35	65.85	66.67	65.85	66.67	66.51	66.00	66.44	66.07
	Did not converge	0.15	0.77		0.61		0.61		0.15	0.46	0.18	0.43
S0 _{waa}	All parameters are estimable	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Converged but not all parameters identifiable											
	Did not converge											

135 **2.2 Time series plots for 20 year time series**

136 Generally, recruitment error, recruitment autocorrelation and time-series length scenarios
 137 had little effect on the ability of each model to estimate population trends and abundance.
 138 However, CSA was somewhat affected by differing recruitment scenarios, which was par-
 139 ticularly evident in stocks set up with constant fishing mortality (HD: c) (Figures 2, 3,
 140 4, 5, 6, 7).

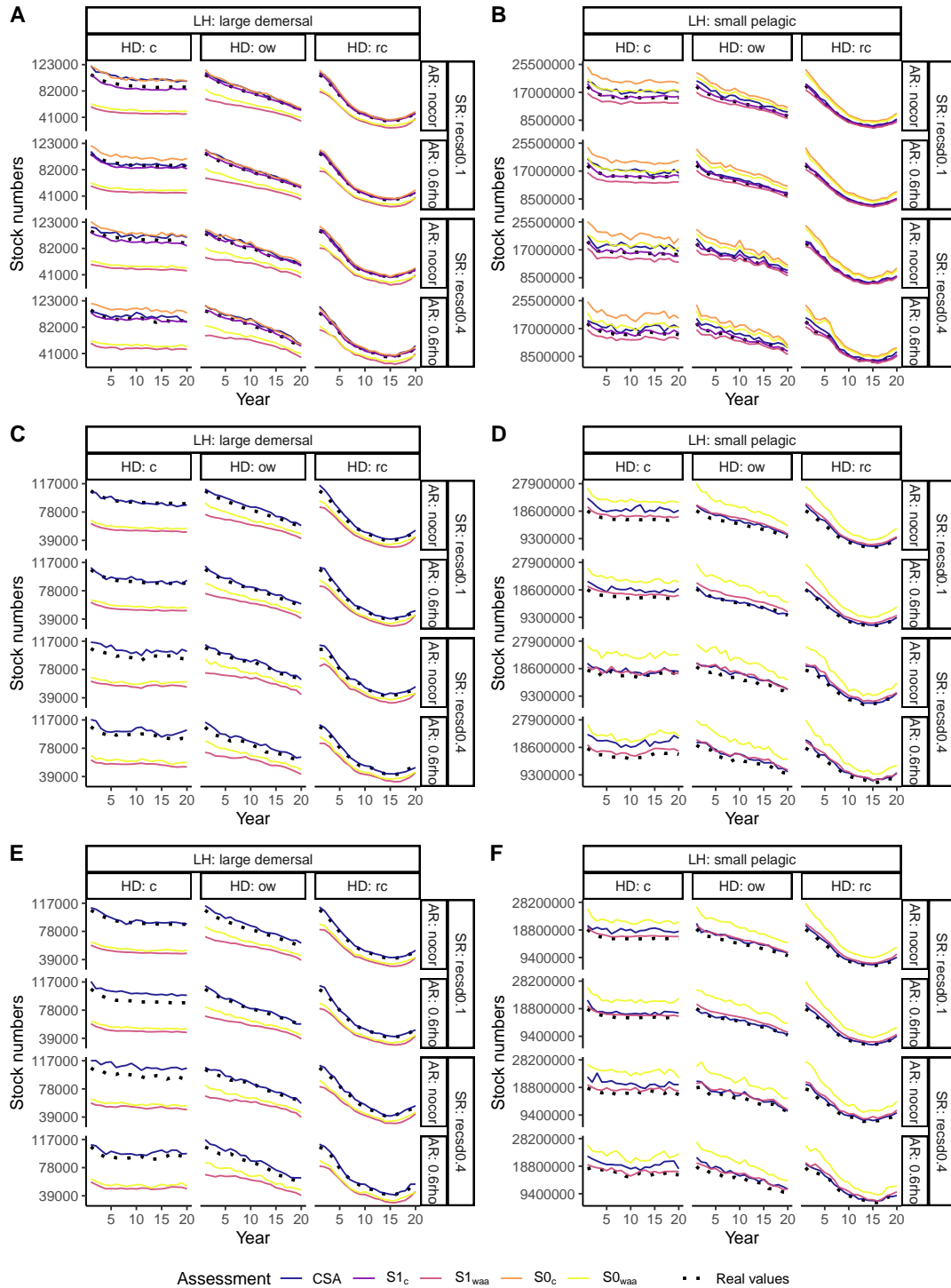


Figure 2: Median estimated stock numbers, as well the corresponding real values for large demersal life history strategy (A, C, E) and small pelagic life history strategy (B, D, F) life history strategies for simulated stock scenarios over a short time series (20 years) for knife-edged (A, B), logistic (C, D) and dome-shaped (E, F) selection patterns. Estimates from five different assessment models (CSA, $S1_c$, $S1_{waa}$, $S0_c$ and $S0_{waa}$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ and $S0_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

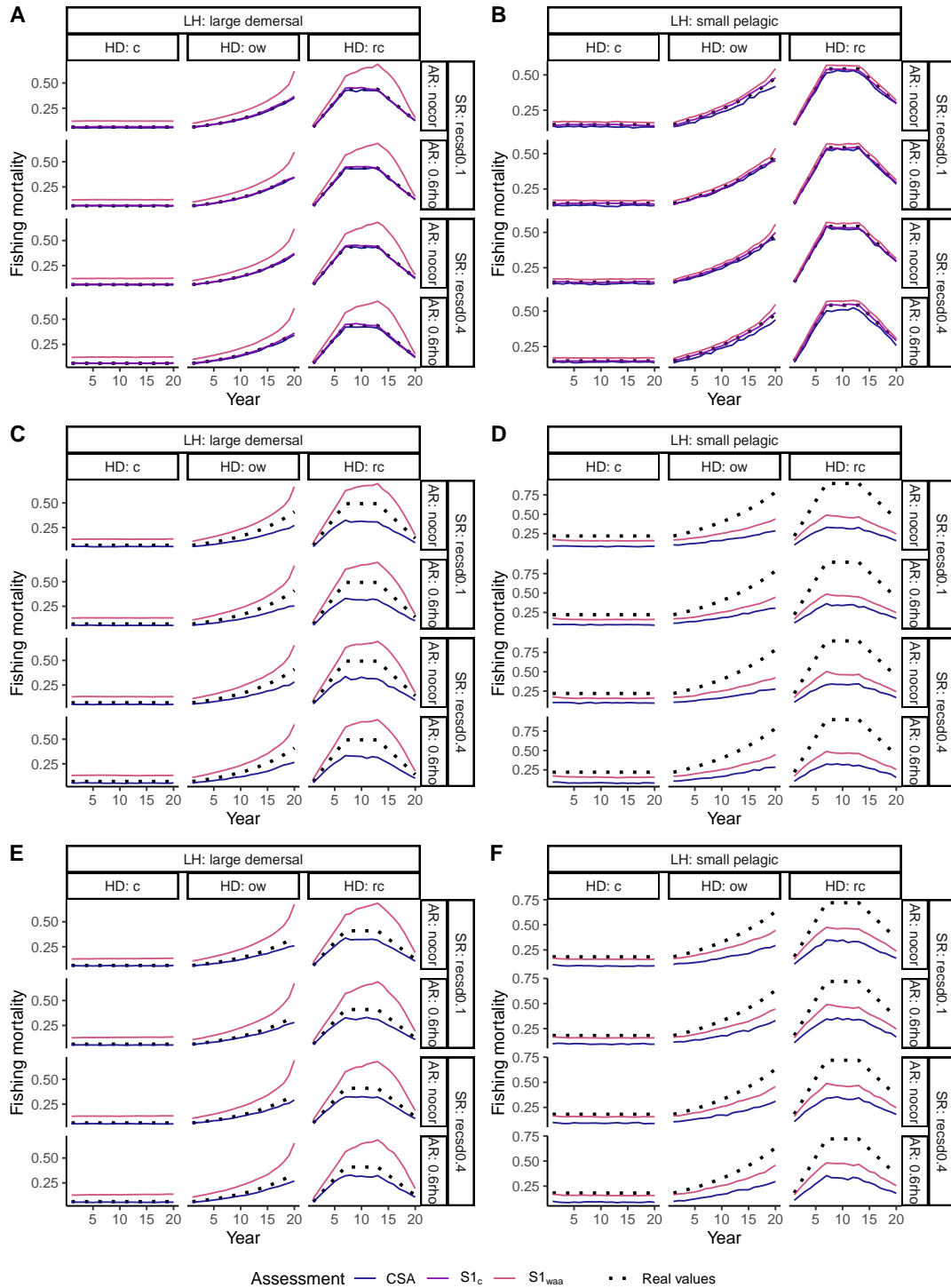


Figure 3: Median estimated fishing mortality, as well the corresponding real values for large demersal (A, C, E) and small pelagic (B, D, F) life history strategies for simulated stock scenarios over a short time series (20 years) for knife-edged (A, B), logistic (C, D) and dome-shaped (E, F) selection patterns. Estimates from three different assessment models (CSA, $S1_c$ and $S1_{waa}$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

¹⁴¹ **2.3** Time series plots for 40 year time series

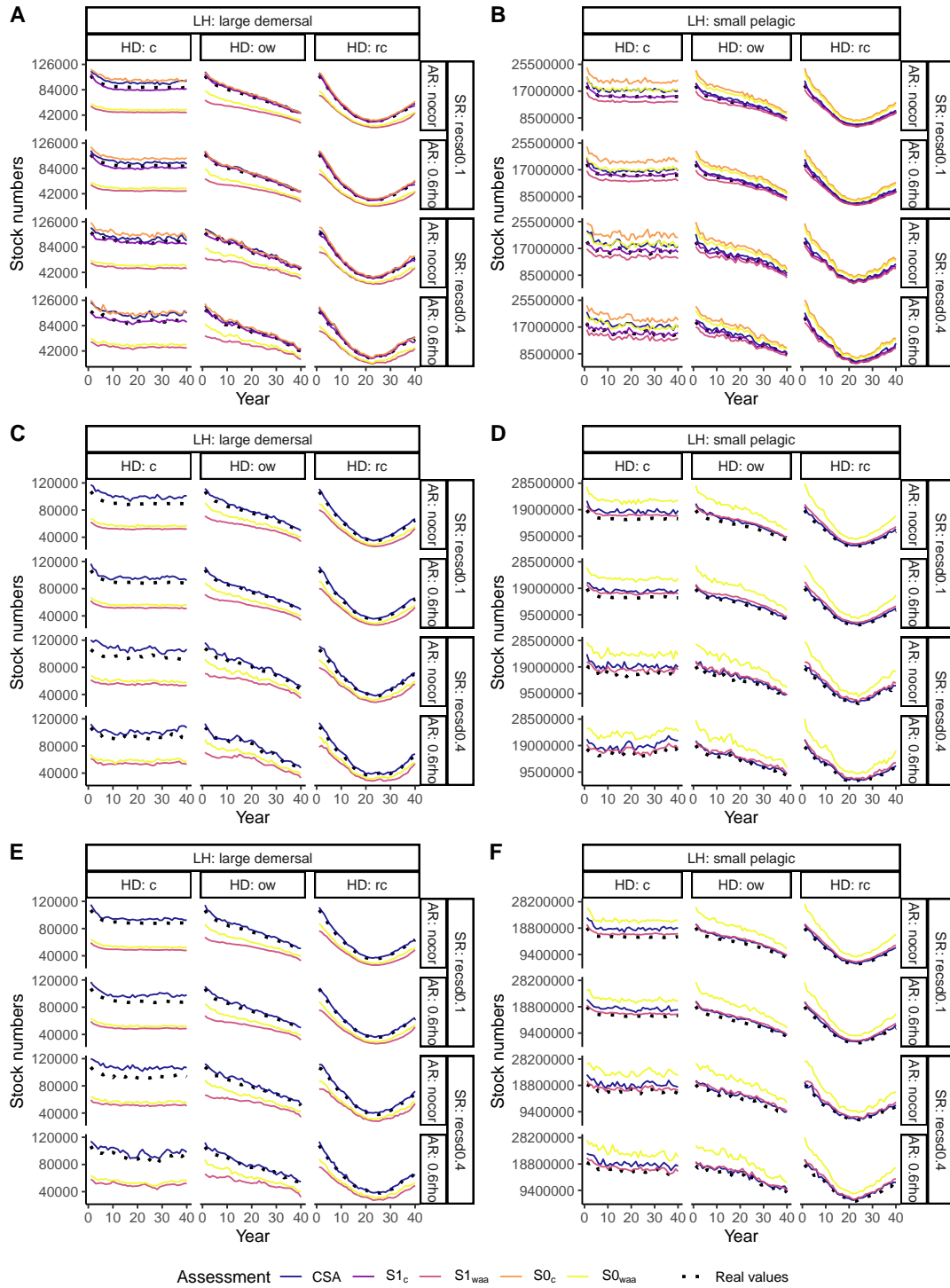


Figure 4: Median estimated stock numbers, as well the corresponding real values for large demersal (A, C, E) and small pelagic (B, D, F) life history strategies for simulated stock scenarios over a long time series (40 years) for knife-edged (A, B), logistic (C, D) and dome-shaped (E, F) selection patterns. Estimates from five different assessment models (CSA , $S1_c$, $S1_w$, $S0_c$ and $S0_w$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ and $S0_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

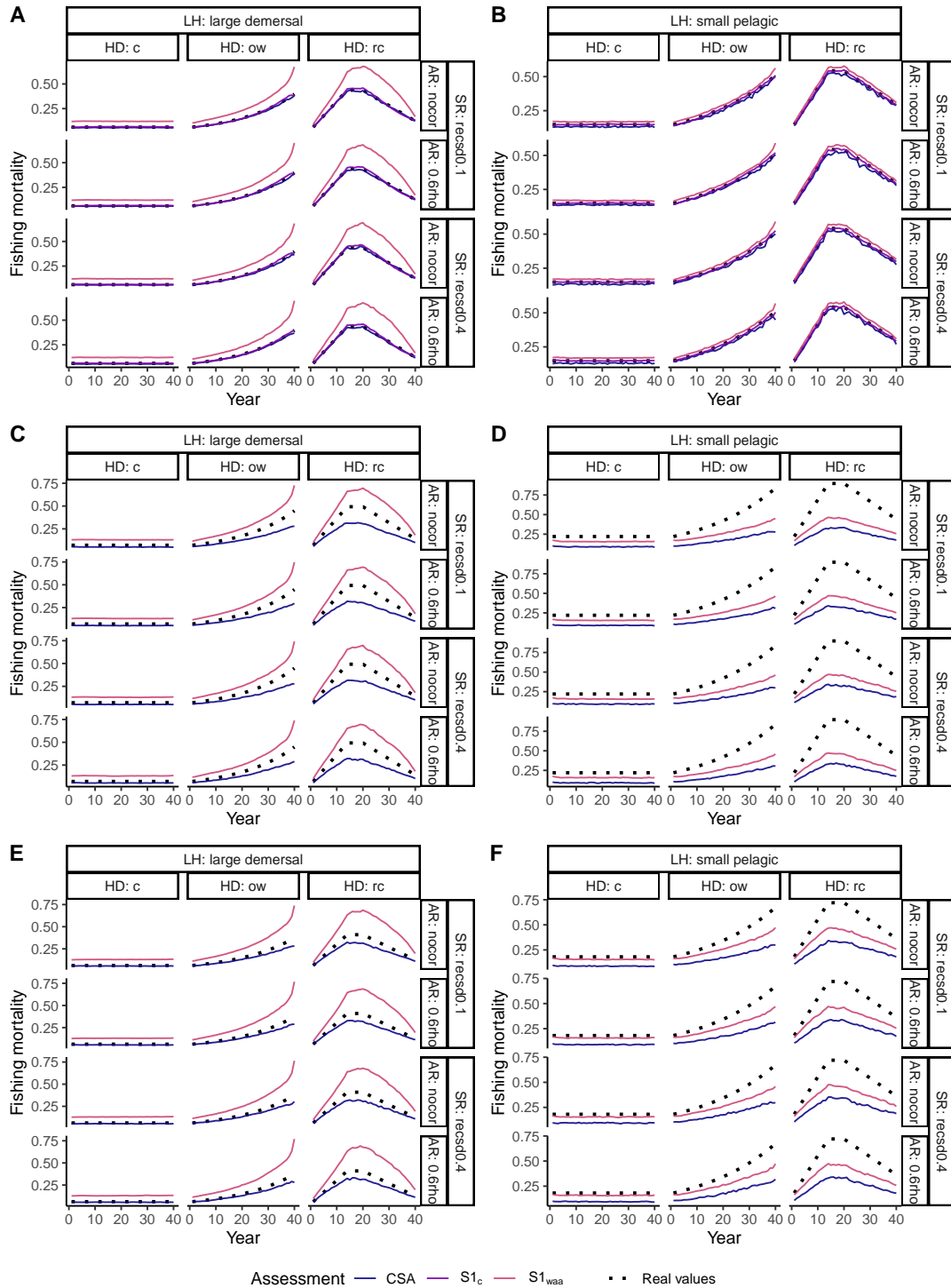


Figure 5: Median estimated fishing mortality, as well the corresponding real values for large demersal (A, C, E) and small pelagic (B, D, F) life history strategies for simulated stock scenarios over a long time series (40 years) for knife-edged (A, B), logistic (C, D) and dome-shaped (E, F) selection patterns. Estimates from three different assessment models (CSA , $S1_c$ and $S1_w$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

142 **2.4 Violin plots for all selectivity patterns, SR and AR scenar-**
143 **ios**

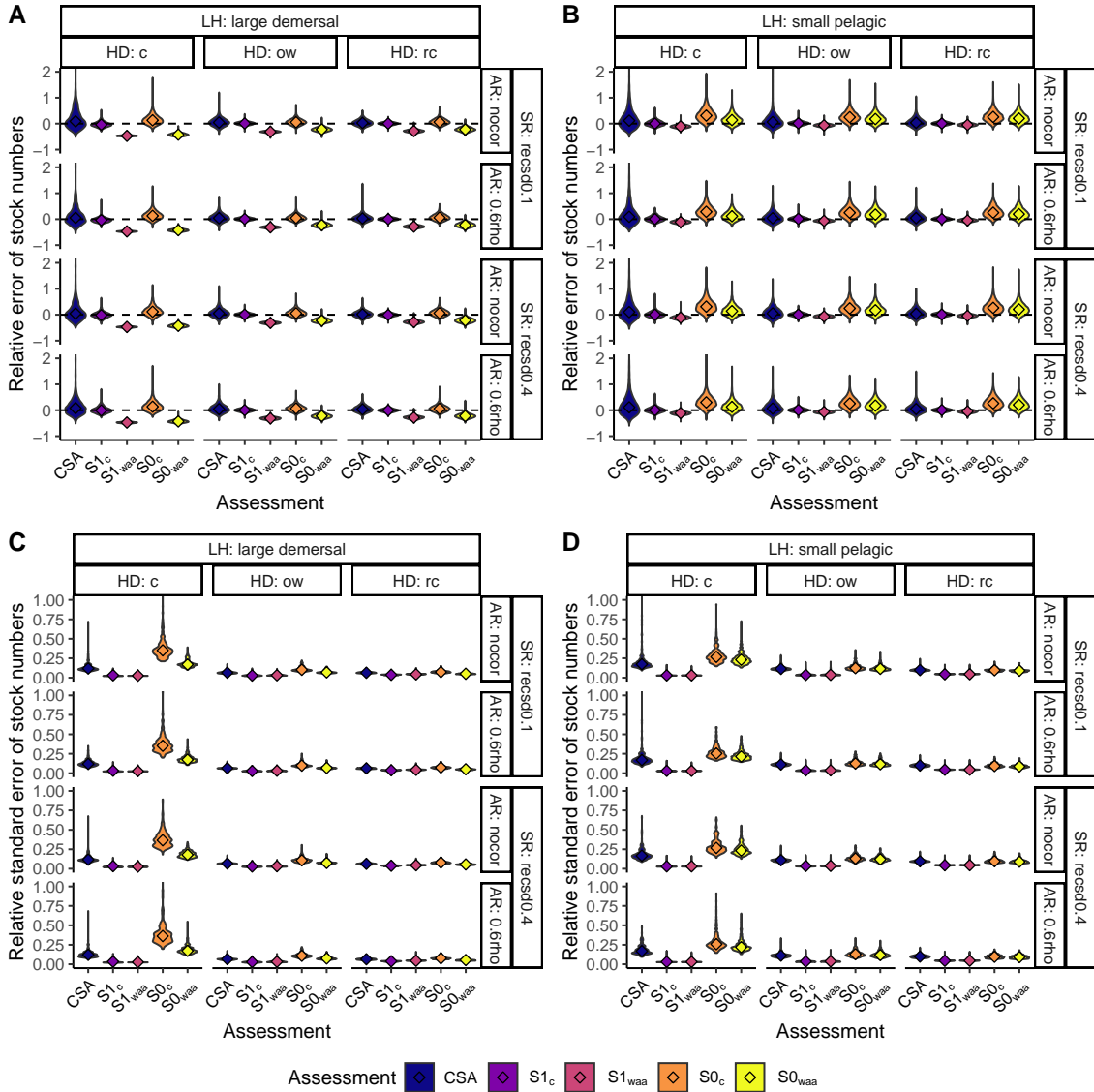


Figure 6: Distribution of relative error (A, B) and relative standard error (C, D) of the stock numbers in the current year for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a knife-edged selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from five different assessment models (CSA, $S1_c$, $S1_{waa}$, $S0_c$ and $S0_{waa}$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ and $S0_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

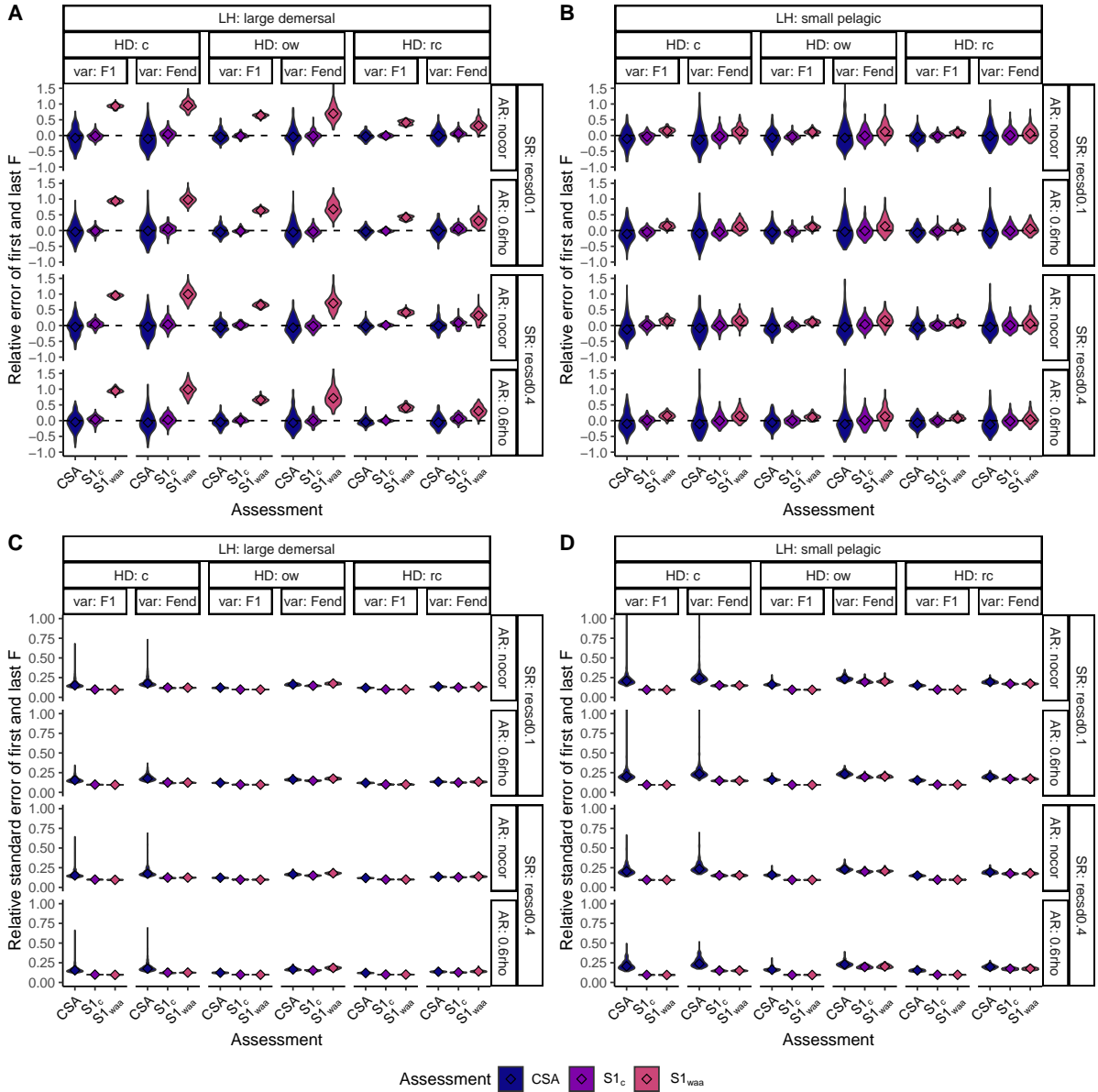


Figure 7: Distribution of relative error (A, B) and relative standard error (C, D) of fishing mortality at the start (F1) and end (Fend) of a time series for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a knife-edged selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from three different assessment models (CSA, $S1_c$ and $S1_{waa}$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (SR)) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

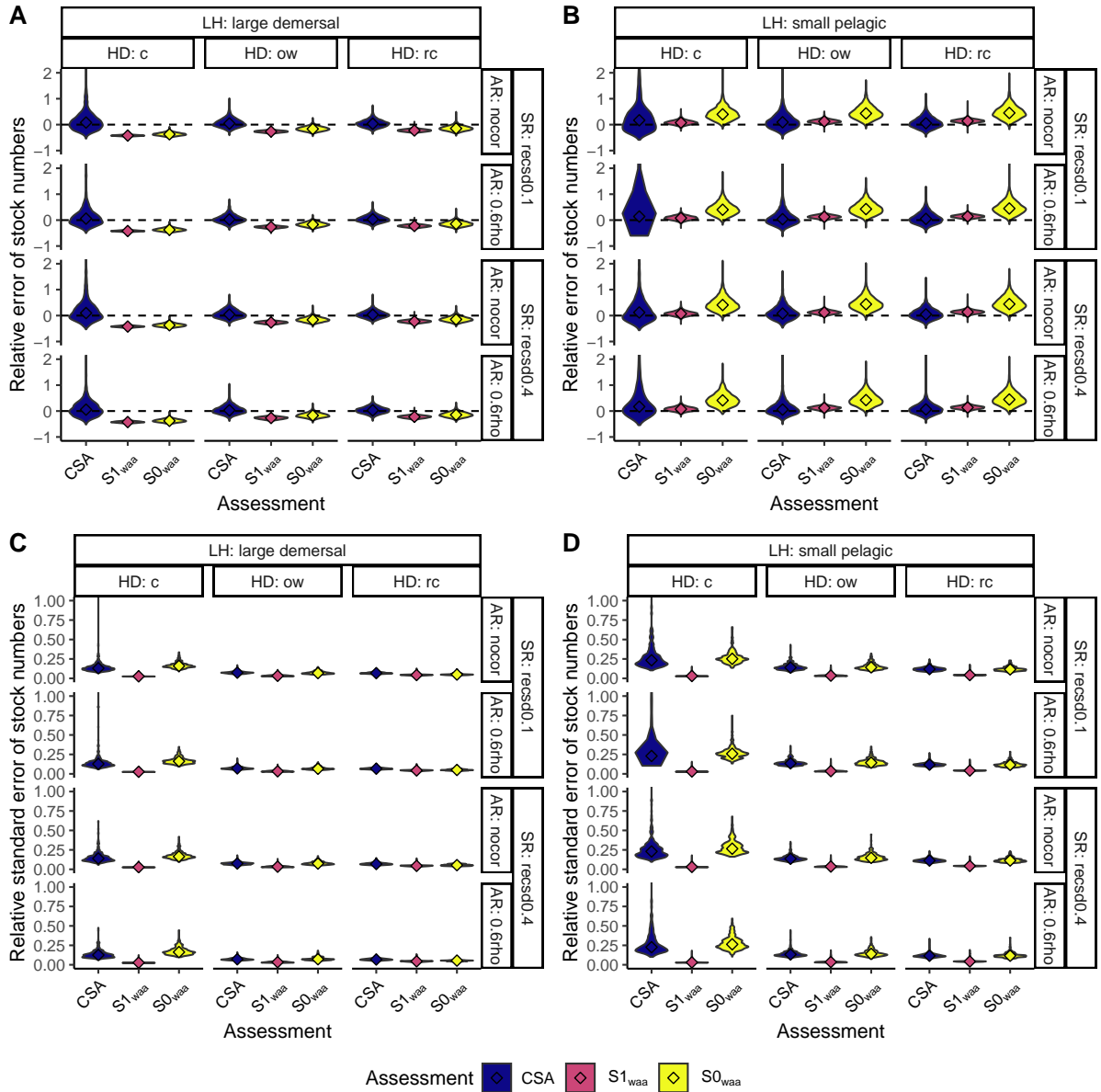


Figure 8: Distribution of relative error (A, B) and relative standard error (C, D) of the stock numbers in the current year for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a logistic selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from three different assessment models (CSA , $S1_u$ and $S0_u$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (SR)) are given. As the majority of $S1_c$ and $S0_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

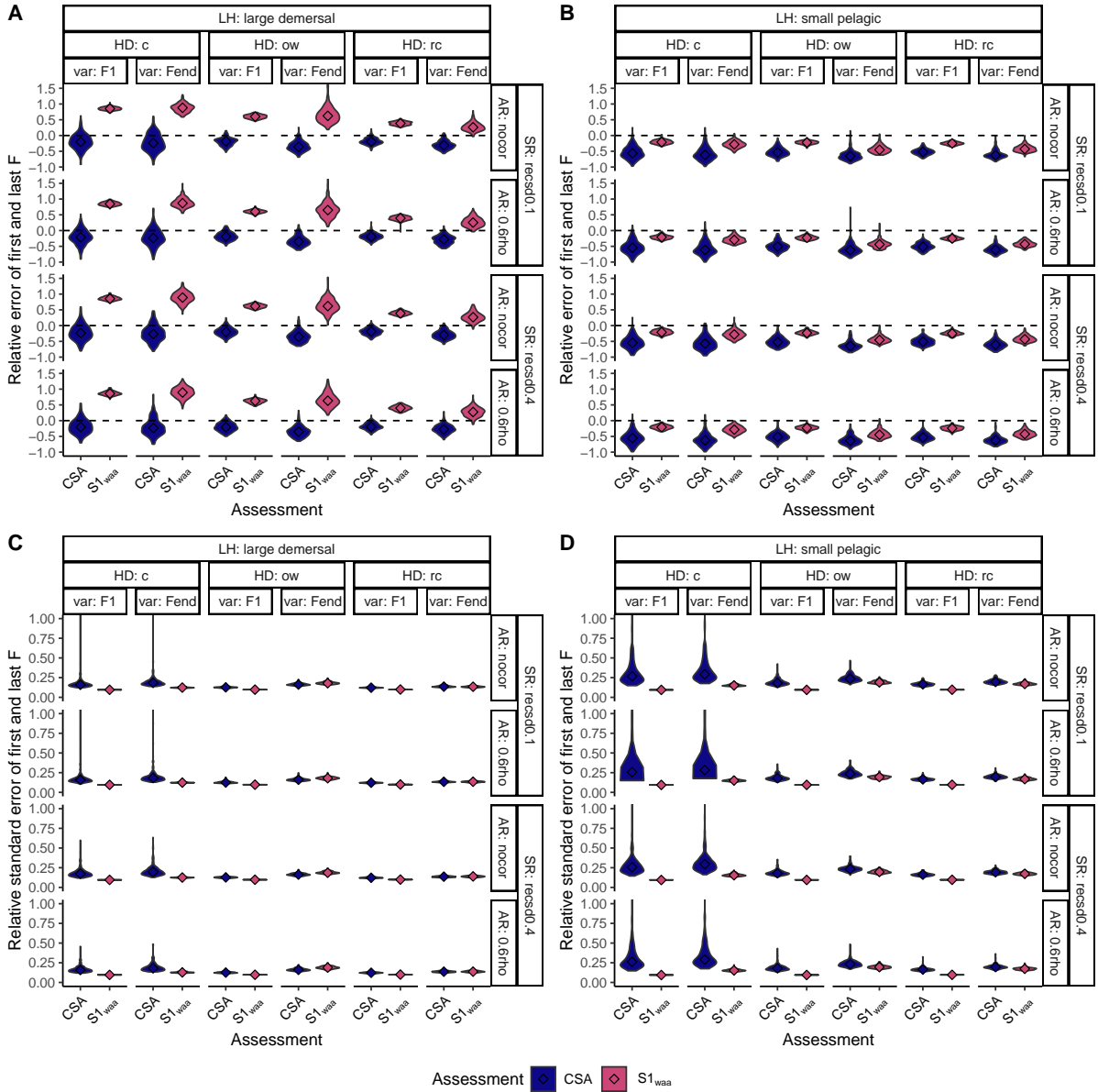


Figure 9: Distribution of relative error (A, B) and relative standard error (C, D) of fishing mortality at the start (F1) and end (Fend) of a time series for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a logistic selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from assessment models CSA and $S1_u$ over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

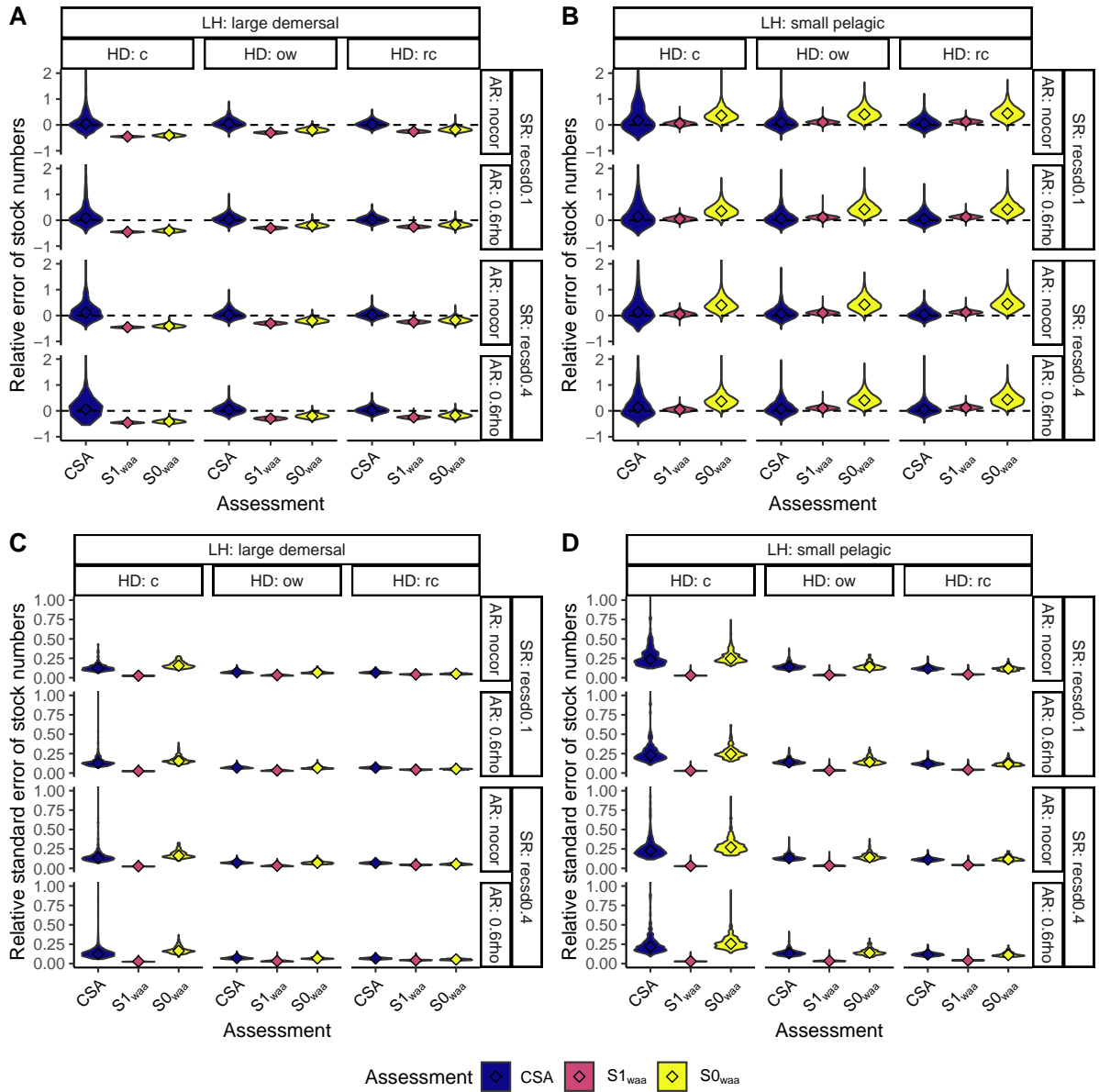


Figure 10: Distribution of relative error (A, B) and relative standard error (C, D) of the stock numbers in the current year for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a dome-shaped selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from three different assessment models (CSA , $S1_u$ and $S0_u$) over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. A very small number of runs categorised as "realistic estimates" for $S1_c$ in dome-shaped selection scenarios were omitted as estimated values were very large and prevented comparison of other models.

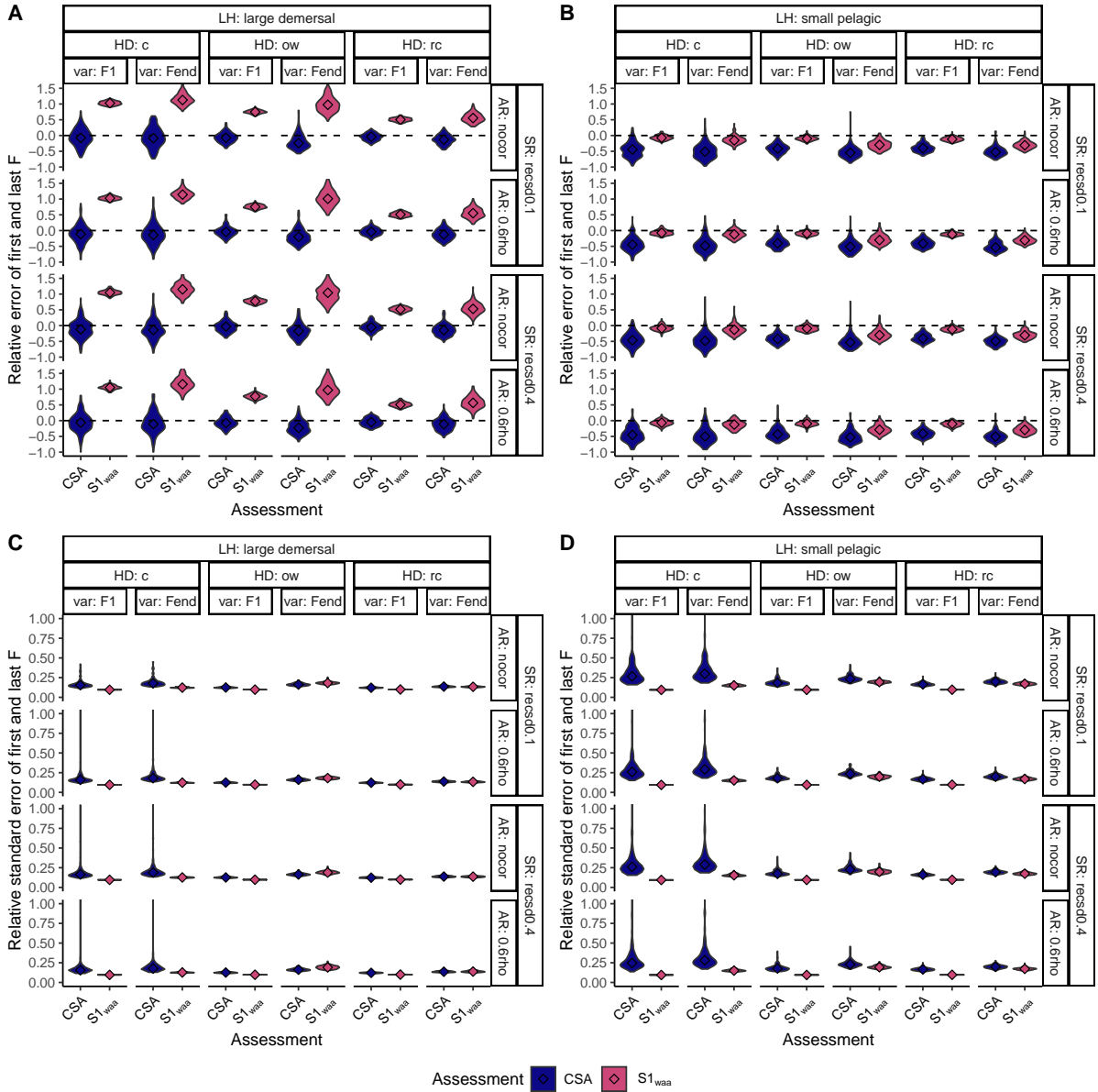


Figure 11: Distribution of relative error (A, B) and relative standard error (C, D) of fishing mortality at the start (F1) and end (Fend) of a time series for large demersal (A, C) and small pelagic (B, D) life history strategies for simulated stock scenarios with a dome-shaped selectivity pattern. Median relative error and median relative standard error is also shown by the diamond points. Shape of violin plots represent the kernel probability density of the data. Estimates from assessment models CSA and $S1_u$ over a number of factors from the simulation framework (i.e. harvest dynamics (HD), recruitment variability (SR) and recruitment autocorrelation (AR)) are given. As the majority of $S1_c$ did not converge or could not identify all parameters in logistic or dome-shaped selection scenarios these models are not shown for these scenarios. A small amount of runs that did identify parameters for these models were omitted, as estimated values were very large and prevented comparison of other models.

144 2.5 Self-tests for white-bellied anglerfish case study

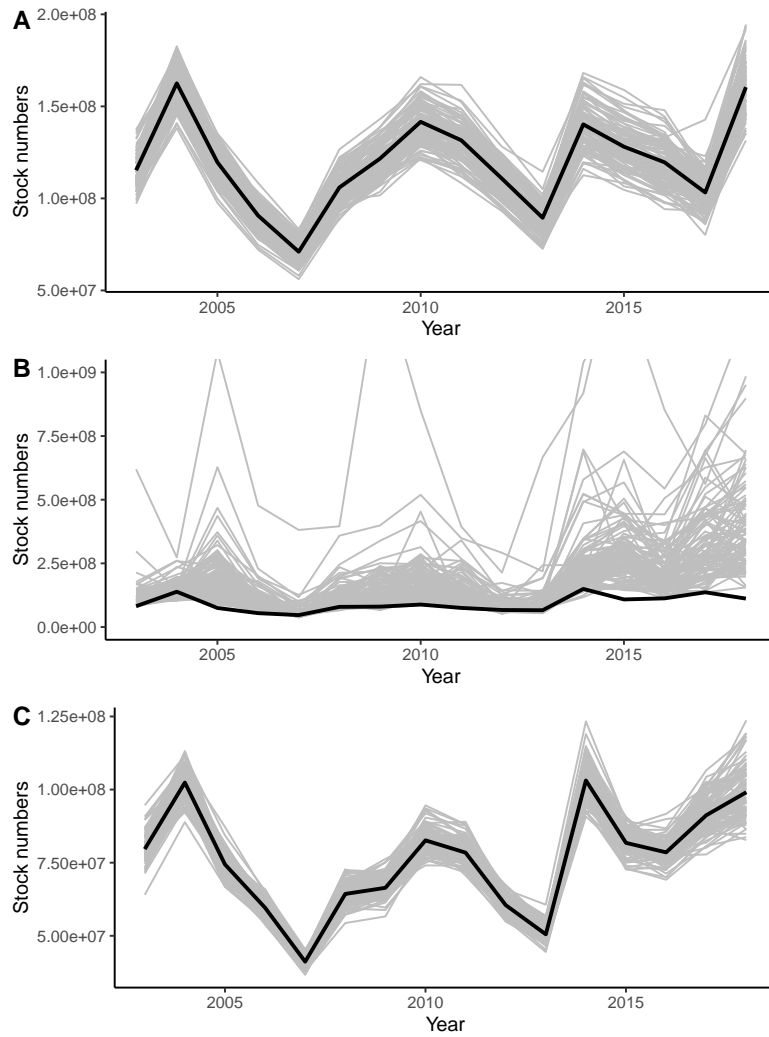


Figure 12: Estimated population numbers (\hat{N}) and "self-test" replicates for three assessment models; CSA, (A), $S0_{waa}$ (B) and $S1_{waa}$ (C) on the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d. Estimate from actual data (bold black line) and re-estimated stock numbers from conditional simulations (thin grey lines).

References

- 145
146 Beverton, R.J. and Holt, S.J. 1957. On the dynamics of exploited fish populations.
147 Fisheries Investigation Series 2, volume 19, UK Ministry of Agriculture. Fisheries, and
148 Food, London, UK .
- 149 Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured
150 model. *Canadian Journal of Fisheries and Aquatic Sciences* **37**(2): 268–282.
- 151 Froese, R., Pauly, D. et al. 2021. FishBase.
- 152 Griffiths, D. and Harrod, C. 2007. Natural mortality, growth parameters, and environ-
153 mental temperature in fishes revisited. *Canadian Journal of Fisheries and Aquatic*
154 *Sciences* **64**(2): 249–255.
- 155 Hilborn, R. and Walters, C.J. 1992. Quantitative Fisheries Stock Assessment: Choice,
156 Dynamics and Uncertainty. Springer Science & Business Media.
- 157 ICES 2019a. Herring *Clupea harengus* in subdivisions 20-24, spring-spawners (Skagerrak,
158 Kattegat and western Baltic). in Report of the ICES Advisory Committee, 2019. ICES
159 Advice 2019, her.27.20-24.
- 160 ICES 2019b. Report of the Working Group for the Bay of Biscay and the Iberian Waters
161 Ecoregion (WGBIE) .
- 162 ICES 2019c. Stock annex: Herring *Clupea harengus* in subdivisions 20-24, spring-
163 spawners (Skagerrak, Kattegat and western Baltic).
- 164 ICES 2019d. White anglerfish (*Lophius piscatorius*) in Subarea 7 and divisions 8.a–b and
165 8.d (Celtic seas, Bay of Biscay). in Report of the ICES Advisory Committee, 2019.
166 ICES Advice 2019, mon.27.78abd.
- 167 Kell, L. 2018. FLife: Methods for Modelling Life History Traits. R package version 3.3.0.
- 168 Schnute, J. 1985. A general theory for analysis of catch and effort data. *Canadian Journal*
169 *of Fisheries and Aquatic Sciences* **42**(3): 414–429.
- 170 Schnute, J. 1987. A general fishery model for a size-structured fish population. *Canadian*
171 *Journal of Fisheries and Aquatic Sciences* **44**(5): 924–940.
- 172 Thorson, J. 2019. FishLife: Predict Life History Parameters For Any Fish. R package
173 version 2.0.0.

Appendix E

Working document to ICES
working group for the Bay of
Biscay and The Iberian Waters
Ecoregion (WGBIE), 2020

WD for ICES WGBIE 2020

Stage-based assessment models for black and white anglerfish in areas
7b-k, 8abd

Luke Batts, C oil n Minto, Hans Gerritsen and Deirdre Brophy

May 5, 2020

1 Introduction

Stage-based assessment models offer an alternative middle ground between aggregate and compositional models (Hilborn and Walters, 1992; Li *et al.*, 2019). Two well known but theoretically different approaches to stage-based assessment models are: biomass-based delay-difference models, first described by Deriso (1980) and developed by Schnute (1985, 1987); Fournier and Doonan (1987); and numbers-based depletion models, the most well-known of which is the Catch-Survey Analysis (CSA) (Collie and Sissenwine, 1983; Smith and Addison, 2003). Here we compare parameter estimation of these distinct approaches.

Schnute (1987) describes a size-based delay-difference model linking population size structure and mean weights. In the most simple form this consists of two stages of biomass (recruits and previously exploited biomass) and assumed deterministic growth of all individuals in the exploited stock. The model also offers flexibility for alternative processes to estimating total biomass from different assumptions regarding the relative importance of recruitment and previously exploited biomass.

CSA is a relatively simple two-stage model (numbers of recruits and post-recruits), which has changed somewhat from the original model first described in Collie and Sissenwine (1983). The most recent version can be found in the NOAA Fisheries Integrated Toolbox (FIT), where population dynamics centre on Baranov’s catch equation and estimation is through maximum likelihood.

We implemented both Schnute (1987) and CSA within the “TMB” framework in R, taking advantage of automatic differentiation of the likelihood for optimisation. The assessments’ capabilities in accurately modelling both *Lophius piscatorius* and *Lophius budegassa* stocks in the Celtic Sea and Northern Bay of Biscay were investigated.

2 Methods

2.1 Data

Data for the white-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d was collated from the 2019 ICES stock assessment. This consisted of catch data and three survey indices that cover the anglerfish stock; a combined index of the French EVHOE survey (Q4) and Irish Groundfish survey (IGFS, Q4) spanning 2003-2018, Spanish Porcupine Groundfish survey (SP-PORC, Q3/Q4) spanning 2001-2018, Irish Anglerfish and Megrin Survey (IAMS, Q1) in years 2007-2008 and 2016-2018 .

Data for the black-bellied anglerfish stock in ICES areas 7.b-k, 8.a-b and 8.d was collated from the 2020 ICES stock assessment. This consisted of catch data and two survey indices that cover the anglerfish stock; a combined index of the French EVHOE survey (Q4) and Irish Groundfish survey (IGFS, Q4) spanning 2003-2019 and the Irish Anglerfish and Megrin Survey (IAMS, Q1) in years 2007-2008 and 2016-2019 .

2.2 key assumptions and things of note

- Both models assume all stages are fully selected, although this can be relaxed in CSA within the input if external data on selectivity is available (but was not done so here).
- For both species of anglerfish the number of ages to which each survey indices is trimmed to is quite specific for the a4a assessments, i.e. ages 0-2 for IBTS combined survey, 1-5 for IAMS and 2-6 for SP-PORC. Schnute and CSA do not allow for this subtlety, although CSA does allow indices to be classified as “recruit”, “post-recruit” or “whole” population indices. Schnute inputs are only entire biomass indices.
- For both assessments and species the natural mortality was not estimated and was fixed at 0.25, as is the case for the age-based assessments. There is ongoing work on whether these assessments can estimate natural mortality under certain conditions.
- Recruits and previously exploited biomass/post-recruits were crudely split by length; white-bellied anglerfish recruits were ≤ 25 cm, black-bellied recruits were ≤ 15 cm (Figure 1).

2.2.1 Schnute

- By using of a time series of mean weights within the model Schnute (1987) provides a model framework where the entire biomass of the population can be calculated either solely from the recruitment biomass, solely from the previously-exploited population biomass or both. All runs presented here were with Schnute version 2, where biomass in a given year is computed from the previously exploited biomass. This version has the advantage that no estimation of stock-recruitment parameters is necessary.
- Recruitment biomass at each time step is calculated through a relationship between mean weights and biomass of each stage.

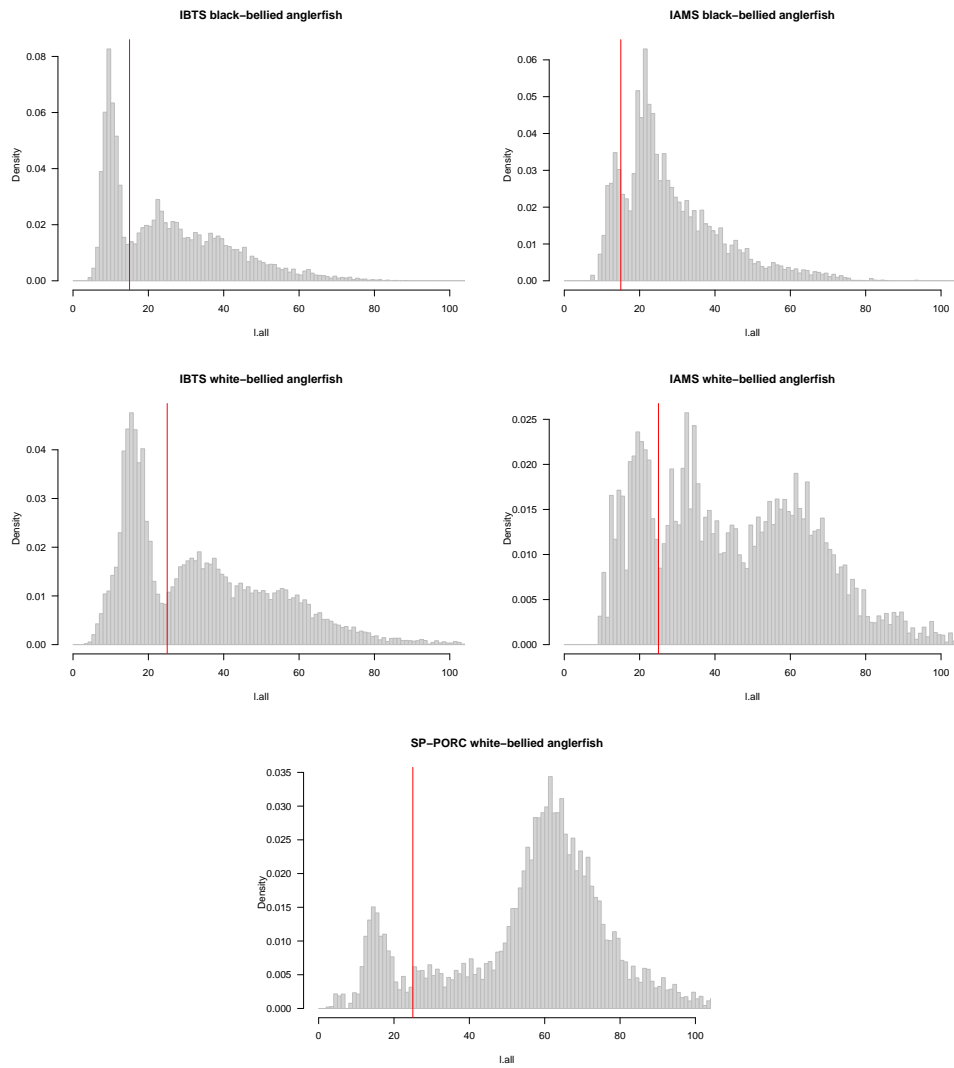


Figure 1: Length distributions of survey data aggregated over years. Red line indicates where cut off length was decided upon for each species

2.2.2 CSA

- In order to mimic the conditions of the a4a assessments as much as possible IAMS and SP-PORC (for white-bellied) were only input as post-recruit indices. CSA requires at least one survey to be split into recruits and post-recruits, IBTS combined survey provided these two indices.
- Recruitment numbers at each time step are latent states/parameters estimated within the model.

3 Results

3.1 white-bellied anglerfish

3.1.1 Schnute model fit

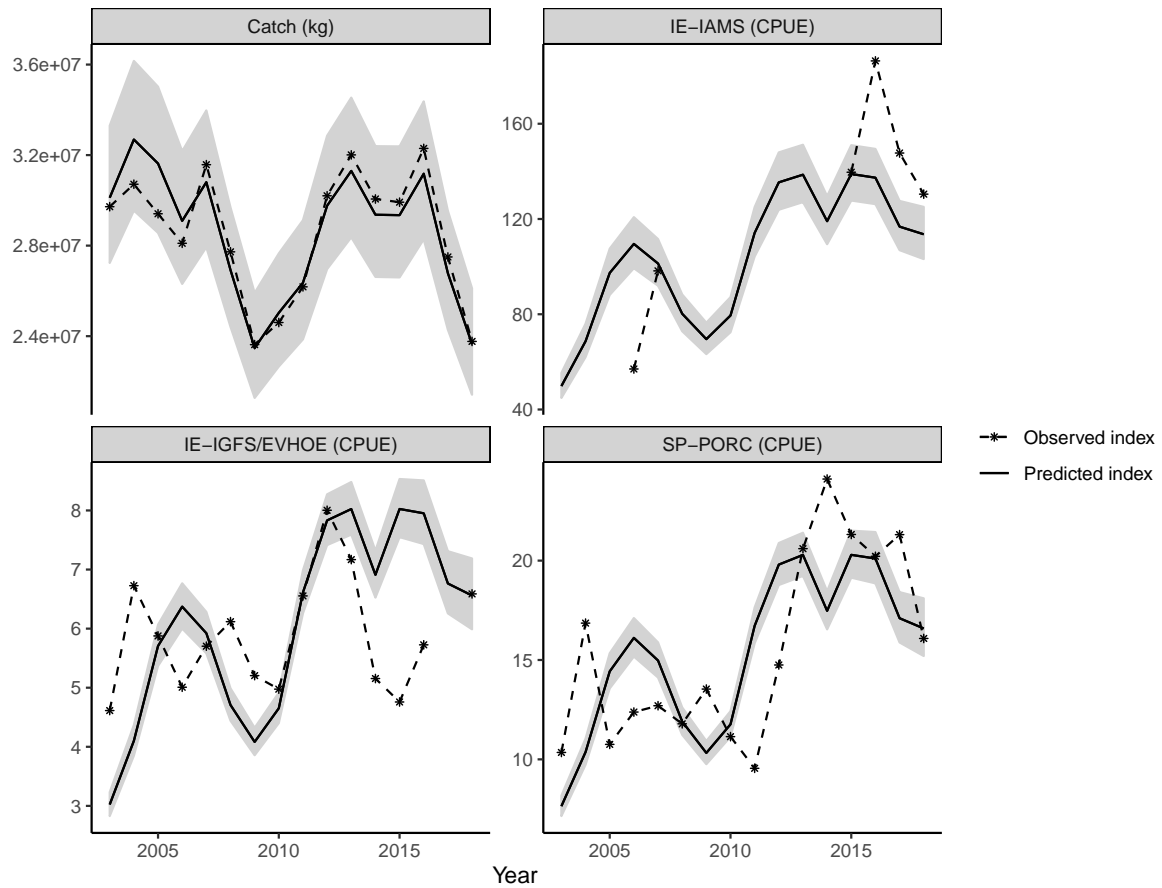


Figure 2: Model fits of observations in catch (kg) and survey indices (CPUE = kg per unit effort) compared to predicted values and their uncertainty for white-bellied anglerfish. Surveys are the combined index for the Irish groundfish and French EVHOE surveys (IE-IGFS/EVHOE), the Irish monkfish and megrim survey (IE-IAMS) and the Spanish Porcupine bank survey (SP-PORC). Shaded grey area is $2 \times$ standard error of the predicted values of the catch/indices.

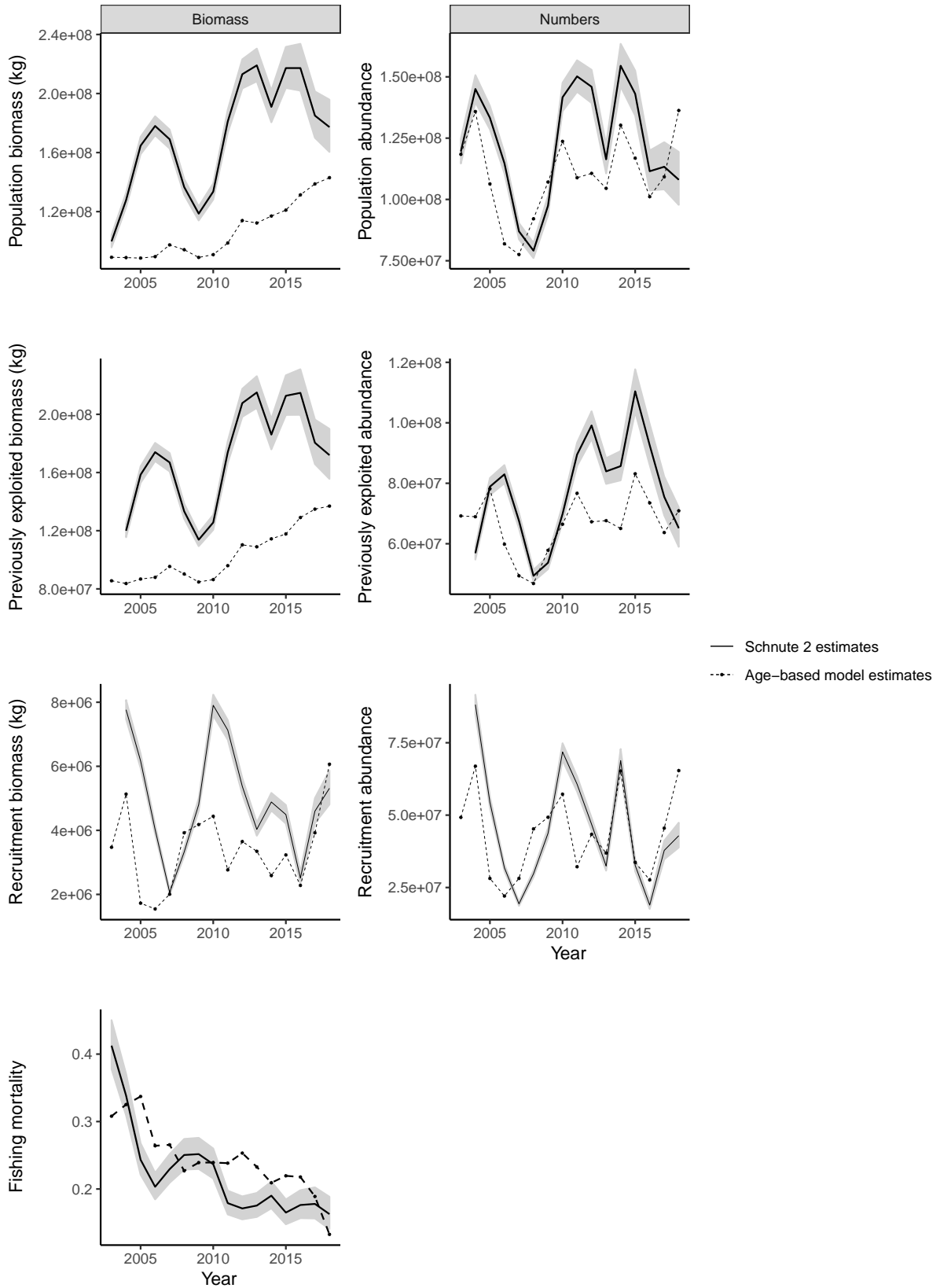


Figure 3: Comparison of the Schnute assessment and a4a assessment estimated values for white-bellied anglerfish. Total biomass and numbers, previously-exploited biomass and numbers, recruitment biomass and numbers, and fishing mortality are shown. Shaded grey area is 2*standard error of the estimated time series.

3.1.2 CSA model fit

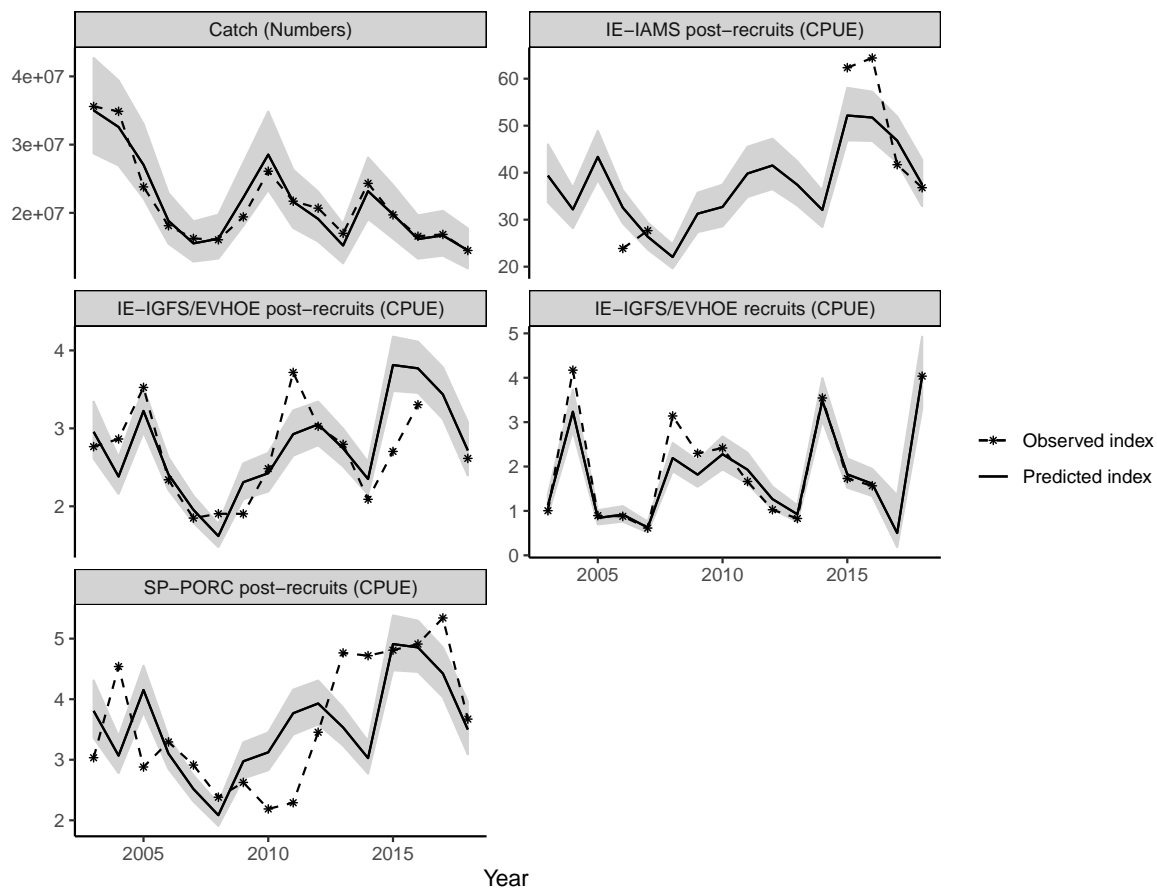


Figure 4: Model fits of observations in catch (numbers) and survey indices (CPUE = numbers per unit effort) compared to predicted values and their uncertainty for white-bellied anglerfish. Surveys are the combined index for the Irish groundfish and French EVHOE surveys (IE-IGFS/EVHOE), the Irish monkfish and megrim survey (IE-IAMS) and the Spanish Porcupine bank survey (SP-PORC). Shaded grey area is $2 \times$ standard error of the predicted values of the catch/indices.

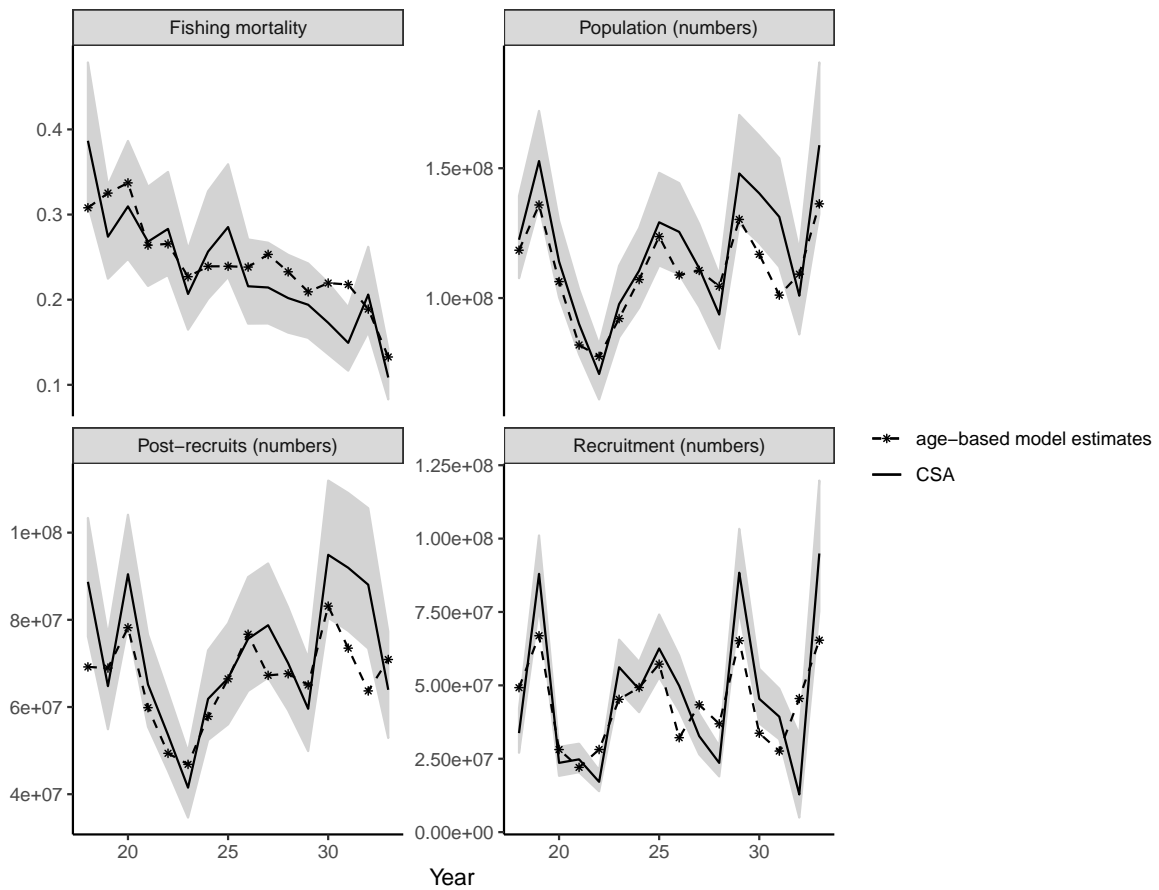


Figure 5: Comparison of the CSA assessment and a4a assessment estimated values for white-bellied anglerfish. Fishing mortality, total numbers, post-recruit numbers and recruitment numbers are shown. Shaded grey area is 2*standard error of the estimated time series.

3.2 Black-bellied anglerfish

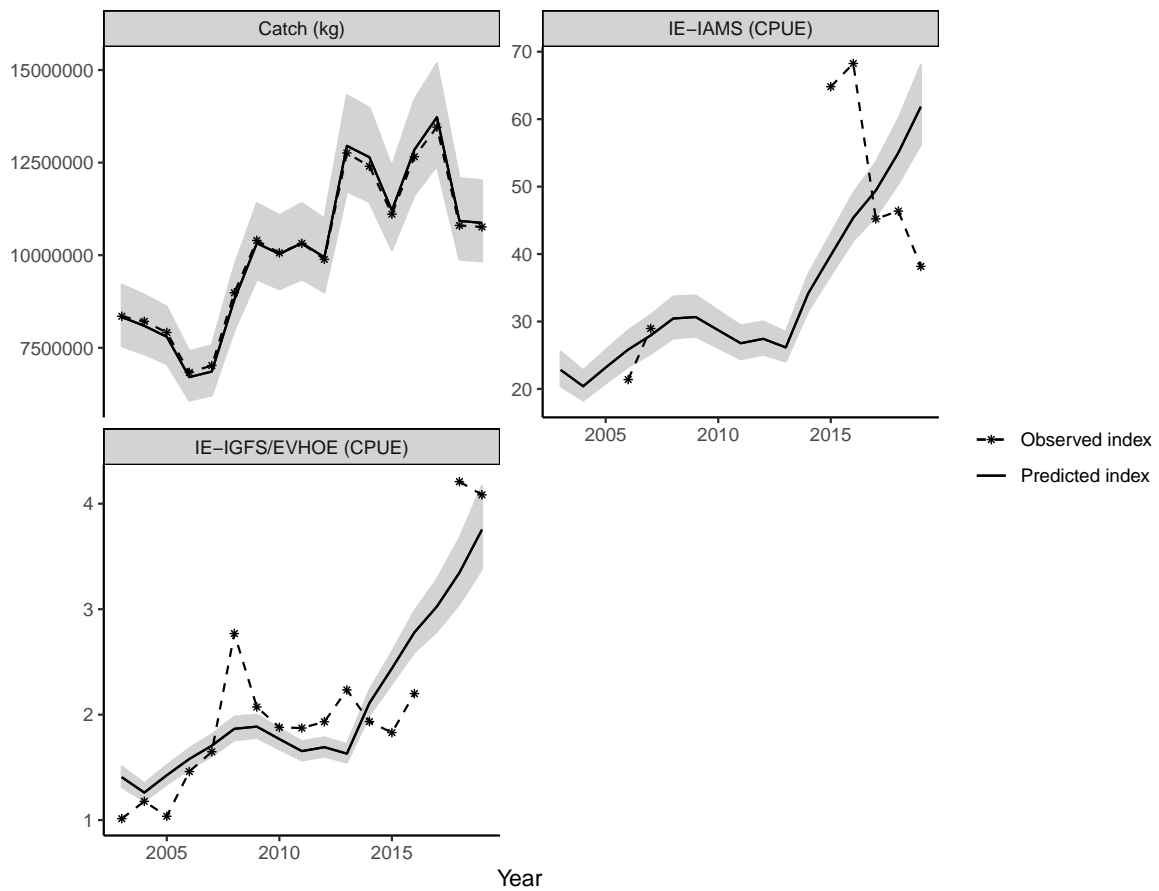


Figure 6: Model fits of observations in catch (kg) and survey indices (CPUE = kg per unit effort) compared to predicted values and their uncertainty for black-bellied anglerfish. Surveys are the combined index for the Irish groundfish and French EVHOE surveys (IE-IGFS/EVHOE) and the Irish monkfish and megrim survey (IE-IAMS) . Shaded grey area is 2*standard error of the predicted values of the catch/indices.

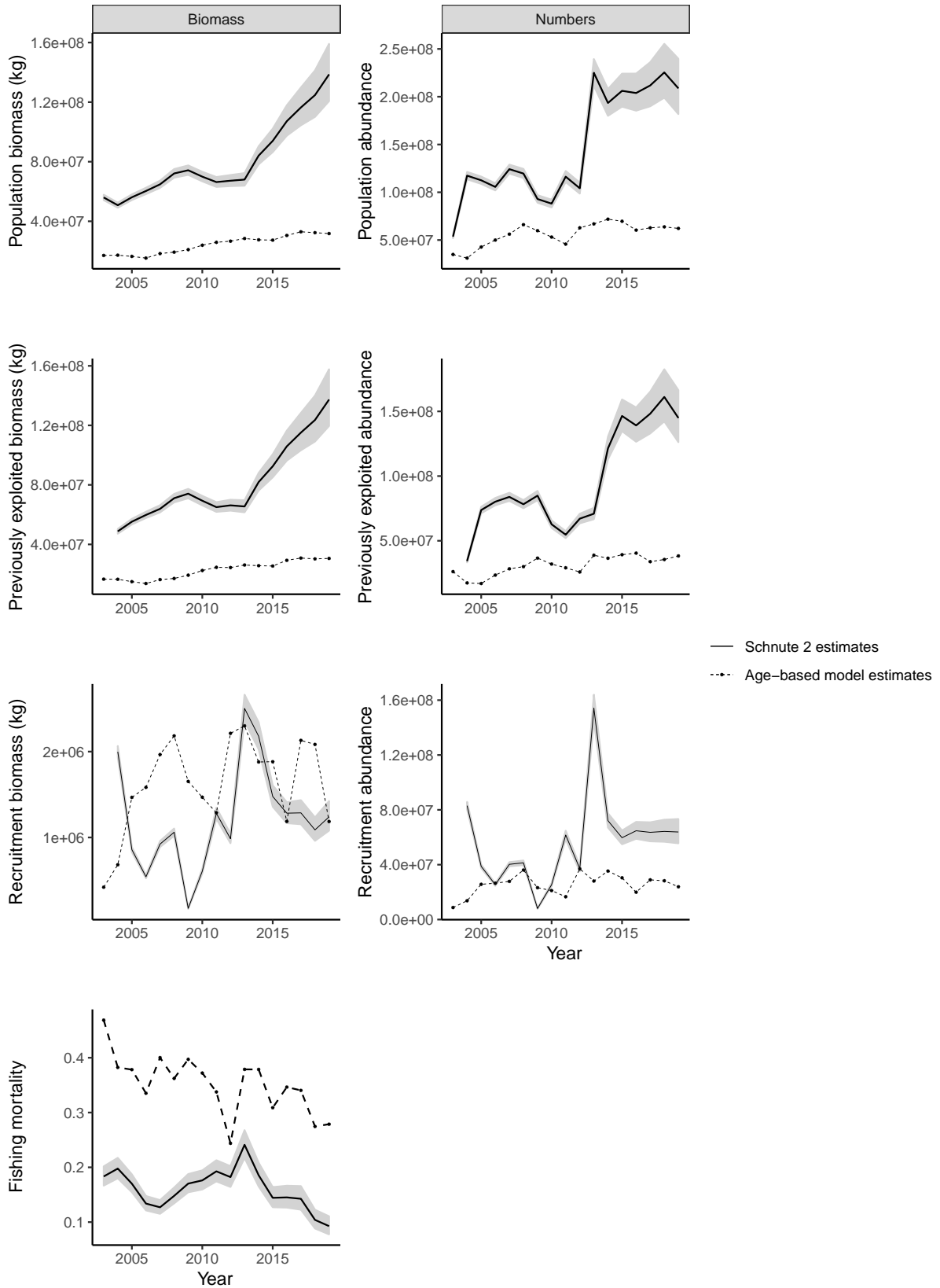


Figure 7: Comparison of the Schnute assessment and a4a assessment estimated values for black-bellied anglerfish. Total biomass and numbers, previously-exploited biomass and numbers, recruitment biomass and numbers, and fishing mortality are shown. Shaded grey area is 2*standard error of the estimated time series.

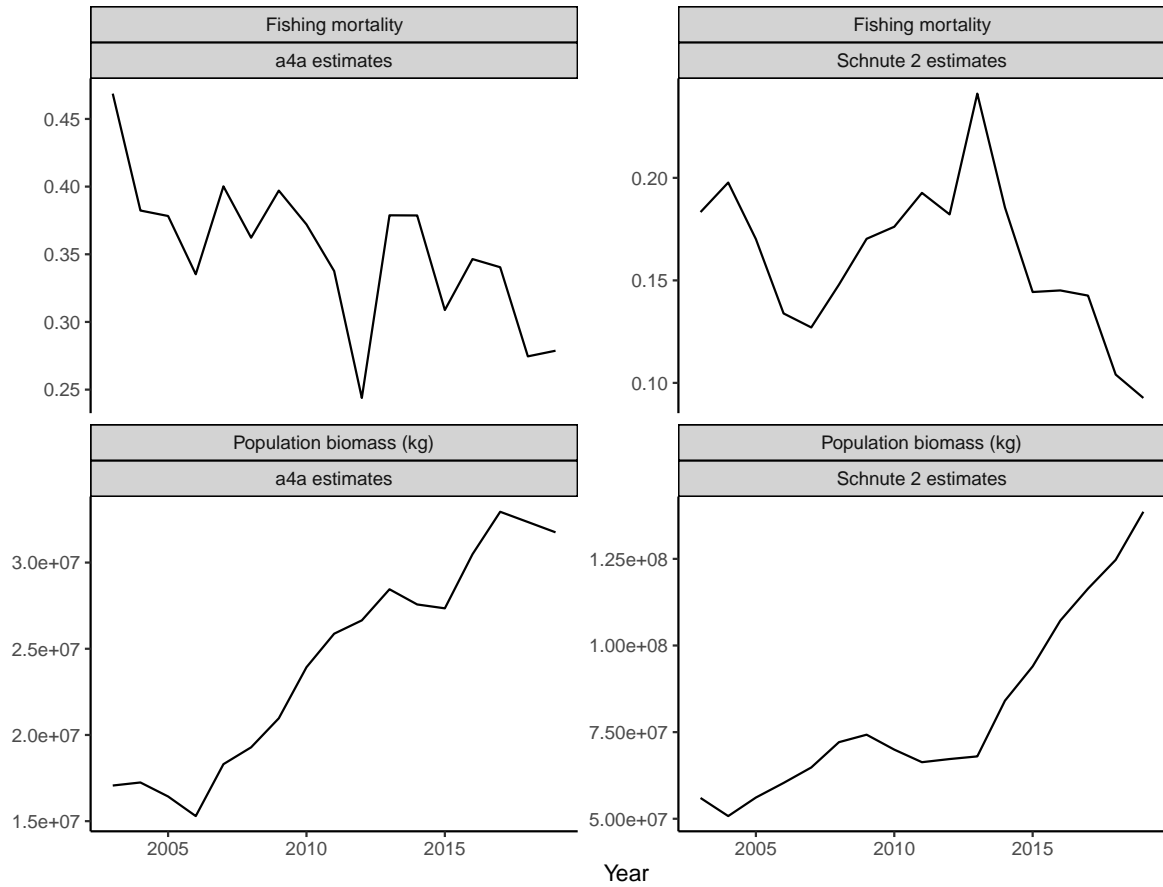


Figure 8: Comparison of the Schnute assessment and a4a assessment estimated values for black-bellied anglerfish. Total biomass and fishing mortality for each assessment are shown individually. Shaded grey area is 2*standard error of the estimated time series.

3.2.1 CSA model fit

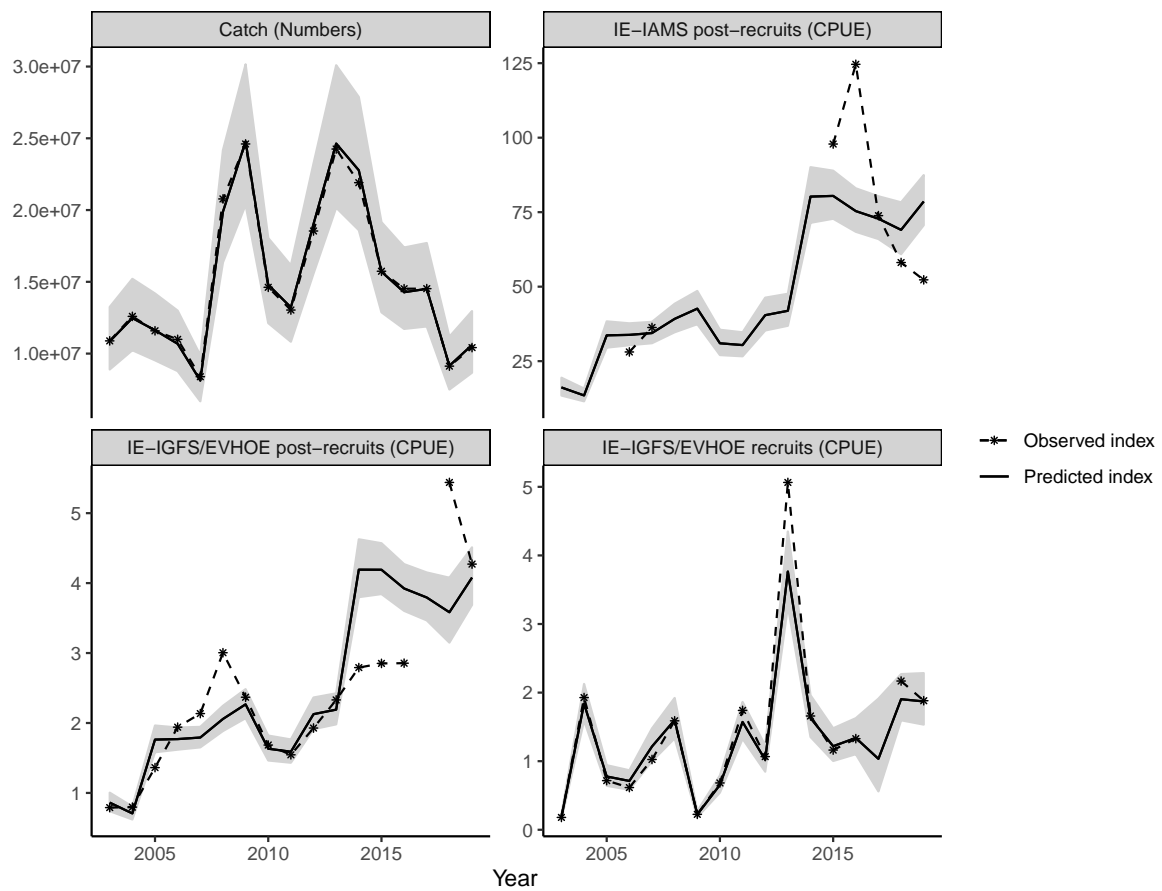


Figure 9: Model fits of observations in catch (numbers) and survey indices (CPUE = numbers per unit effort) compared to predicted values and their uncertainty for black-bellied anglerfish. Surveys are the combined index for the Irish groundfish and French EVHOE surveys (IE-IGFS/EVHOE) and the Irish monkfish and megrim survey (IE-IAMS). Shaded grey area is $2 \times$ standard error of the predicted values of the catch/indices.

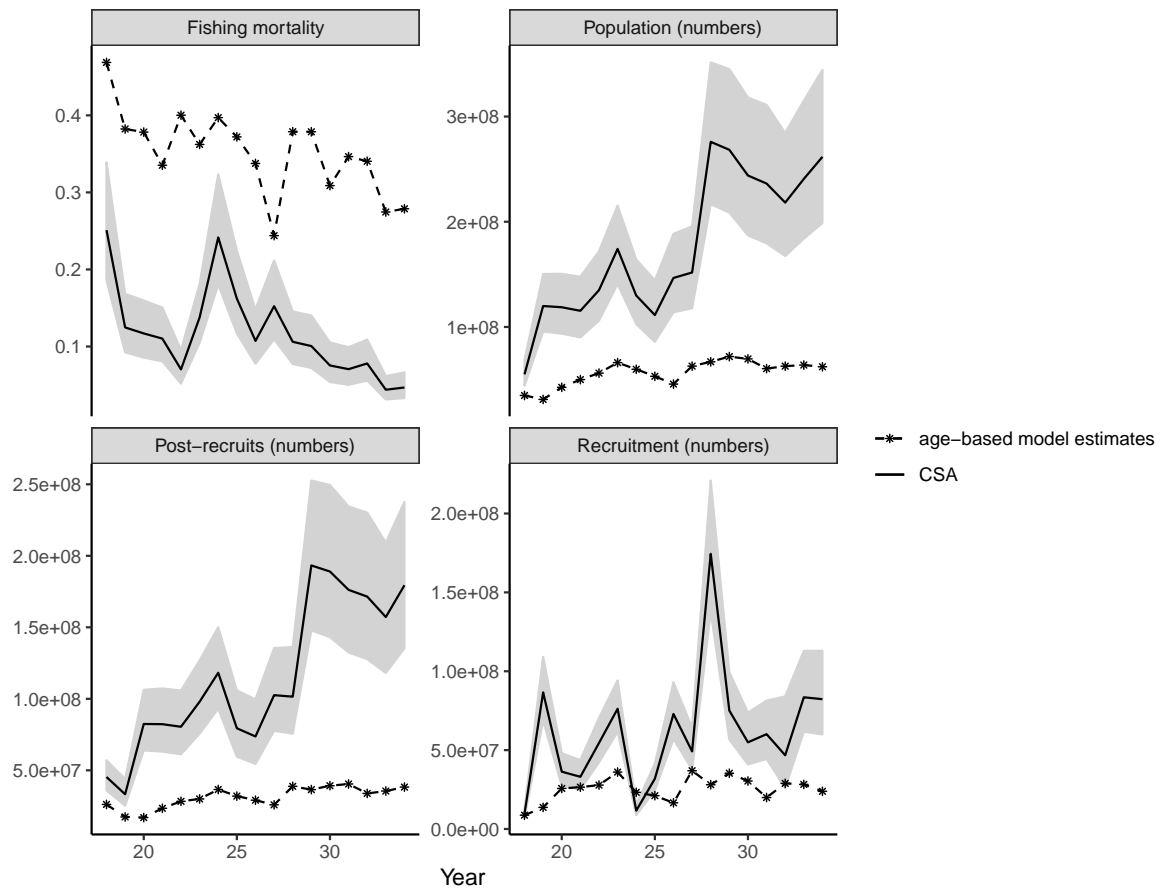


Figure 10: Comparison of the CSA assessment and a4a assessment estimated values for black-bellied anglerfish. Fishing mortality, total numbers, post-recruit numbers and recruitment numbers are shown. Shaded grey area is 2*standard error of the estimated time series.

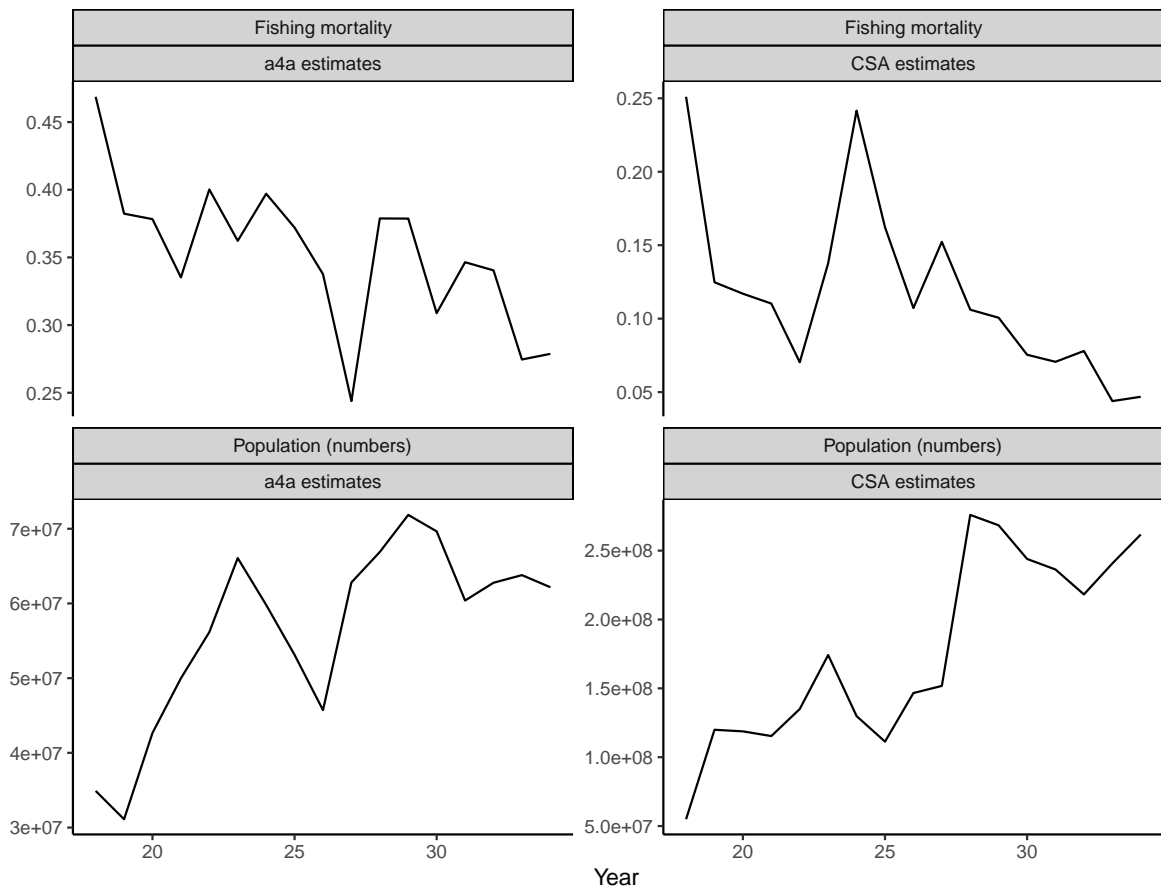


Figure 11: Comparison of the CSA assessment and a4a assessment estimated values for black-bellied anglerfish. Fishing mortality and total numbers for each assessment are shown. Shaded grey area is 2*standard error of the estimated time series.

4 Key points and discussion

4.1 Performance

- The Schnute model fits reasonably closely to both the white-bellied and black-bellied anglerfish catch/survey indices. However the CSA fit for white-bellied anglerfish is better in terms of closeness of fit.
- The CSA assessment also out performs the Schnute assessment for white-bellied anglerfish in terms of closeness to the age-based assessment estimates, however the Schnute assessment is not that far off, particularly when the biomass estimates are divided by their relevant mean weights to give numbers in the population. This discrepancy between biomass estimated in Schnute and biomass estimated in the age-based assessment may be due to differences in mean weights or differences in numbers of older fish.
- Both assessments gave similar results for the black-bellied anglerfish, estimating biomass/numbers at quite different absolute values to the a4a assessment. However, the overall trend in estimated time series was similar.

4.2 Survey class and selectivity

- It is worth considering that the Schnute model only takes entire biomass indices, whereas CSA can take total numbers, post-recruit and recruit indices. a4a is even more specific in its tailoring of indices for the assessment. Schnute still performs reasonably well even though significantly less information is given to the model and trends are very different between total biomass and recruit/post-recruit/age range specific indices. CSA performs very well and this is likely due to its flexibility in fitting to different stages.
- Selectivity is also a key consideration. Both assessments performed assume each stage is fully selected within the fishery (although differences can be entered as an input in CSA). Within the a4a assessment difference in selectivity can be taken into account within the model. We know that neither species is fully selected, particularly black-bellied anglerfish as the recruits at age 0 are somewhat smaller than white -bellied anglerfish recruits. This is likely to have an effect on the estimates from each of the assessments but particularly black-bellied anglerfish.
- CSA is a simpler model than a4a but still has the flexibility to emulate estimated values from a tailored age-based model, despite relaxed assumptions around selectivity in white-bellied anglerfish.
- Selectivity may be more of an issue for black-bellied anglerfish. Preliminary results (not shown in WD) indicate that if fish less than 15cm are removed from the data completely and $15cm \geq \text{recruits} \leq 30cm$, the Schnute model gives similar similar absolute estimates to the a4a assessment.

4.3 Overall

- Currently both these assessment do not include process error but in the future this could be incorporated.
- Forecasting from these assessments would be possible. Both assessments give yearly recruits to which a S-R model could be fitted or average taken. Yearly fishing mortality is estimated within both assessments.
- These simpler stage based assessments offer a credible alternative to more complex models. Taking into account the caveats discussed above, both stage-based methods perform surprisingly well considering the simpler model framework and are likely to be less influenced by ageing/growth uncertainty.

References

- Collie, J. S. and Sissenwine, M. P. 1983. Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences*, 40: 1871–1879.
- Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 37: 268–282.
- Fournier, D. A. and Doonan, I. J. 1987. A length-based stock assessment method utilizing a generalized delay-difference model. *Canadian Journal of Fisheries and Aquatic Sciences*, 44: 422–437.
- Hilborn, R. and Walters, C. J. 1992. *Quantitative fisheries stock assessment: choice, dynamics and uncertainty*. Springer Science & Business Media.
- Li, Y., Lee, L. M., and Rock, J. 2019. Modeling population dynamics and nonstationary processes of difficult-to-age fishery species with a hierarchical bayesian two-stage model. *Canadian Journal of Fisheries and Aquatic Sciences*.
- Schnute, J. 1985. A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences*, 42: 414–429.
- Schnute, J. 1987. A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences*, 44: 924–940.
- Smith, M. T. and Addison, J. T. 2003. Methods for stock assessment of crustacean fisheries. *Fisheries Research*, 65: 231–256.

Appendix F

sbar R package manual

Package ‘sbar’

January 30, 2022

Title Stage-based Assessments in R

Version 0.0.0.9000

Description

c(Implementations of the stage-based fisheries assessment models CSA and Schnute (1987). Models are those detailed and implemented in a manuscript currently under review.)

License GPL (>= 3)

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

URL <https://github.com/lbatts/sbar>

BugReports <https://github.com/lbatts/sbar/issues>

Depends R (>= 2.10)

Imports TMB (>= 1.7.18),

reshape2,
ggplot2,
abind,
TMBhelper

LinkingTo TMB

Remotes github::kaskr/TMB_contrib_R/TMBhelper

Suggests rmarkdown,
knitr

VignetteBuilder knitr

R topics documented:

ank78	2
csa	2
makesbarclass	5
plot.sbarclass	6
schnute_obserror	7
schnute_orig	10

Index**13**

ank78	<i>ank78</i>
-------	--------------

Description

Catch and survey data for black-bellied anglerfish stock in the Celtic Sea and northern Bay of Biscay

Usage

```
ank78
```

Format

a list of vectors and dataframes

Author(s)

Hans Gerritsen

Examples

```
data(ank78)
summary(ank78)
```

csa	<i>Prepare an object for a CSA assessment to be used in optimiser</i>
-----	---

Description

Create list with a TMB framework, including data, gradients and NLL function for a CSA assessment that can be optimised.

Usage

```
csa(
  catch_no,
  indices_no,
  indices_att,
  ts,
  selrec = 1,
  start_q = 1e-08,
  start_surveycv = 0.1,
  start_prec0,
  start_rec,
  start_nmort = 0.2,
```

```

start_f_calc = 0.3,
start_catchcv = 0.1,
fix_nmort = TRUE,
fix_prec0 = FALSE,
fix_surveycv = FALSE,
fix_catchcv = TRUE
)

```

Arguments

catch_no numeric vector of catch numbers over time period of assessment

indices_no matrix of survey indices (numbers) of dimensions: no. of indices x no.years

indices_att matrix of survey indices attributes of dimensions: no. of indices x 2. First column defines survey and second column defines survey type (1 = recruit index, 2 post-recruit index, 3 = whole assessed population index). For example the minimum needed to run CSA is one survey split into a recruit index and a post-recruit index, the attribute matrix should look like:

survey	type
1	1
1	2

ts numeric. Survey timing parameters

selrec matrix of selection proportions of the recruit indices (in comparison to the post-recruit index) if known. Dimensions: no. of recruit indices x no. years. Defaults to 1 for all years (i.e. no difference between recruit and post-recruit indices of a survey)

start_q Starting values for survey catchability parameters. Default is 1e-6

start_surveycv Starting values for survey cv parameters. Default is 0.1

start_prec0 Starting parameter value for post-recruit numbers at first time step. Default is $4 * \max(\text{catch.no})$.

start_rec Starting parameter values for estimated recruit numbers. Default is $2 * \max(\text{catch.no})$.

start_nmort Starting parameter value for natural mortality. Default is 0.2

start_f_calc Starting parameter values for estimated fishing mortality. Default is 0.3.

start_catchcv Starting parameter value for catch cv. Default is 0.1

fix_nmort logical. Should natural mortality be fixed in the model

fix_prec0 logical. Should post-recruit numbers in the first time step be fixed in the model.

fix_surveycv logical. Should survey CV be fixed in the model.

fix_catchcv logical. Should catch CV be fixed in the model.

Details

csa is simply a wrapper function that gives the output from `MakeADFun` from TMB, i.e. an objective function with derivatives, hessian etc. Optimisation and extraction of values from the assessment models has been kept separate to allow flexibility with optimisation methods as well as easy access to `MakeADFun` outputs such as the hessian, gradients etc. See http://kaskr.github.io/adcomp/_book/Introduction.html and TMB documentation for details. Users should refer to the sbar vignette by running `vignette("intro_to_sbar", "sbar")` for details on the csa function.

The table below gives the outputs and description of the values that can be extracted from this assessment model after optimisation and `summary(TMB::sdreport(x))` of the csa object.

Output	Description
logitqhat	logit transformed survey catchability
logphat1	log transformed post-rec numbers at first time step
logrhat	log transformed recruitment numbers
logf_calc	log transformed fishing mortality
log_surveycv	log transformed survey CV
log_catchcv	log transformed catch CV
lognmort	log transformed natural mortality
phat	post-recruit numbers
rhat	recruit numbers
bhat	total numbers
lnphat	log transformed post-rec numbers
lnbhat	log transformed total numbers
lnc	log transformed catch numbers
c_calc	catch numbers
logpred_survey	log transformed predicted survey indices
sdsurv	Survey index standard deviation
f_calc	fishing mortality
phat1	post-rec numbers at first time step
nmort	natural mortality
qhat	survey catchability

Value

List with components for optimiser in R. This output is that of the function `MakeADFun` from TMB

Access (transformed) starting values of parameters to be estimated with `x$par` (where `x` is the csa object) to see what parameters are to be estimated.

Examples

```
## Not run: obj <- csa(catch_n = catch.no, indices_no = obs,
indices_att = indices_att, ts = #'timing', #'start_nmort = nm)
## End(Not run)
```

makesbarclass	<i>Convert sdreport summary into an object of for plotting with class: sbarclass</i>
---------------	--

Description

Create an object with class: 'sbarclass', which can then be used to quickly plot fit and stock predictions from an sbar assessment

Usage

```
makesbarclass(x = matrix(), survey_names, cat, ind, years)
```

Arguments

x	the matrix output from <code>summary(sdreport(x))</code> , where x is an optimised sbar stock assessment model.
survey_names	character vector of survey names
cat	numeric vector of catch that was used in the sbar assessment. Used to plot the fit of the assessment
ind	matrix of survey indices that were used in the sbar assessment. Used to plot the fit of the assessment
years	numeric vector of years

Details

See vignette for a detailed example of this function in use.

Value

An object of class: sbarclass

See Also

[plot.sbarclass](#) to plot a sbar assessment after using this function.

Examples

```
## Not run: survnames<- c("IBTS recruits (CPUE)","IBTS post-recruits (CPUE)")
x <- makesbarclass(obs.srep,survnames,catch.no,obs,years)
## End(Not run)
```

plot.sbarclass	<i>Plot object of with class 'sbarclass'</i>
----------------	--

Description

Plot an object with class:'sbarclass', plot fit or stock predictions with ggplot functions

Usage

```
## S3 method for class 'sbarclass'  
plot(x, out = "fit", ...)
```

Arguments

x	object of class:sbarclass
out	a character that specifies whether the "fit" or "stock" should be plotted
...	No specific usage

Details

See vignette for a detailed example of this function in use.

Value

list

See Also

[makesbarclass](#) for to conversion into an object that can be used by this function.

Examples

```
## Not run:  
plot(x,out="fit")  
  plot(x,out="stock")  
## End(Not run)
```

schnute_obserror	<i>Prepare an object for Schnute Adapted Observation Error Model assessment for an optimiser</i>
------------------	--

Description

Create an object with TMB framework, including data, gradients and NLL function for a Schnute adapted observation error assessment that can be optimised.

Usage

```
schnute_obserror(
  version = 2,
  catch_b,
  indices_b,
  ts,
  mwts,
  tsp = 0,
  rho,
  W,
  ind_l_wt = 1,
  start_q = 1e-08,
  start_indexsigma = 0.1,
  start_B0,
  start_sigma = exp(-0.2),
  start_f_calc = 0.3,
  start_rec_a,
  start_rec_b,
  spawn_prop = 1,
  start_catchsigma = 0.1,
  fix_sigma = TRUE,
  fix_B0 = FALSE,
  fix_indexsigma = FALSE,
  fix_catchsigma = TRUE,
  adrep = FALSE
)
```

Arguments

version	numeric, either 1, 2 or 3. This controls what deterministic equations in the model are used to derive population biomass. 1 and 2 use the fraction of of total biomass in a given year due to newly recruited fish. This fraction is derived from mean weights and detailed in the schnute vignette. version = 3 is the more classical population dynamics.
1	whole biomass derived from recruit biomass
2	whole biomass derived from previously exploited biomass

3 whole biomass is a combination of recruit biomass and previously exploited biomass

catch_b	numeric vector of catch biomass over time period of assessment
indices_b	matrix of biomass surveys (CPUE) of dimensions: no. of surveys x no.years
ts	numeric. Survey timing parameters
mmts	matrix of mean weights from sampling with dimensions: 3 x no. years. recruit mean weights \bar{Y} (first row), previously exploited biomass mean weights \bar{Z} (second row) and entire assessed biomass mean weight \bar{X} (third row).
tsp	numeric. Timing of spawning. Default to 0 (start of year).
rho	numeric. Growth parameter, slope of linear growth model.
W	numeric. Growth parameter, intercept of linear growth model.
ind_l_wt	numeric. Survey weighting in the likelihood. Defaults to 1 fro each survey, ie.e. equal weighting
start_q	Starting values for survey catchability parameters. Default is 1e-6
start_indexsigma	Starting values for survey sd parameters. Default is 0.1
start_B0	Starting parameter value for biomass at first time step. Default is 5*max(catch_b)
start_sigma	Starting parameter value fraction of population that survives natural moratlity. Default is $e^{0.2}$
start_f_calc	Starting parameter values for estimated fishing mortality. Default is 0.3.
start_rec_a	Starting parameter value for the 'a' parameter of the Beverton-Holt stock-recruit function. The asymptotic biomass of recruits. Default is 1/5*max(catch_b).
start_rec_b	Starting parameter value for the 'b' parameter of the Beverton-Holt stock-recruit function. The spawning stock biomass needed to produce a/2 on average. Default is 4*max(catch_b).
spawn_prop	proportion of biomass that is mature. Defaults to 1 for each year.
start_catchsigma	Starting parameter value for catch sd. Default is 0.1
fix_sigma	logical. logical. Should survival be fixed in the model
fix_B0	logical. Should biomass in the first time step be fixed in the model
fix_indexsigma	logical. Should survey standard deviation be fixed in the model
fix_catchsigma	logical. Should catch standard deviation be fixed in the model
adrep	logical. Whether the user would like the ADreport variables (and their derivatives) reported for starting parameters.

Details

schnute_obserror is simply a wrapper function that gives the output from [MakeADFun](#) from TMB, i.e. an objective function with derivatives, hessian etc. Otimisation and extraction of values from the assessment models has been kept separate to allow flexibility with optimisation methods as well as easy access to [MakeADFun](#) outputs such as the hessian, gradients etc. See http://kaskr.github.io/adcomp/_book/Introduction.html and TMB documentation for details.

Users should refer to the sbar vignette by running `vignette("intro_to_sbar", "sbar")` for details on the `schnute_obserror` function.

The table below gives the outputs and description of the values that can be extracted from this assessment model after optimisation and `summary(TMB::sdreport(x))` of the `schnute_obserror` object.

Output	Description
logitq	logit transformed survey catchability
logB0	log transformed biomass in the first time step
logitsigma	logit transformed survival of natural mortality
logindex_sigma	log transformed survey standard deviation
logcatch_sigma	log transformed catch standard deviation
logf_calc	log transformed fishing mortality
lnb	log transformed total biomass
lnpr	log transformed previously-exploited biomass
lnr	log transformed recruit biomass
lnN	log transformed total numbers
lnPR	log transformed previously-exploited or post-recruit numbers
lnR	log transformed recruit numbers
lnC	log transformed predicted catch numbers
biomass	total biomass
B0	total biomass in the first time step
N	total numbers
ssb	spawning biomass
post_rec	previously-exploited or post-recruit biomass
PR	previously-exploited or post-recruit numbers
rec_bio	recruit biomass
rec_no	recruit numbers
C	predicted catch numbers
catch_pred	predicted catch biomass
logpred_survey	log transformed predicted survey indices
omega	fraction of total biomass in a given year due to newly recruited fish
mu	fraction of the catch removed before natural mortality
index_sigma	Survey indices standard deviation
catch_sigma	catch standard deviation
sigma	survival of natural mortality
logrec_param	log transformed recruit parameters from a Beverton-Holt SR function if estimated in versions 1 or 3
rec_param	the two recruit parameters fro a Beverton-Holt SR fonctionif estimated in versions 1 or 3
qhat	survey catchability

Value

List with components for optimiser in R. This output is that of the function [MakeADFun](#) from TMB
Access (transformed) starting values of parameters to be estimated with `x$par` (where `x` is the `schnute_obserror` object) to see what parameters are to be estimated.

Examples

```
## Not run: obj <- schnute_obserror(catch_b = catch_biomass, indices_b = obs,
```

```
ts = timing, mwts = mwts, rho = rho1, W = W1, start_sigma = sigma)
## End(Not run)
```

schnute_orig	<i>Prepare list for Schnute Original Process Error Model assessment in an optimiser</i>
--------------	---

Description

Create an object with TMB framework, including data, gradients and NLL function for a Schnute Original Process Error Model assessment that can be optimised.

Usage

```
schnute_orig(
  version = 2,
  catch_b,
  indices_b,
  ts,
  mwts,
  tsp = 0,
  mu = 0.5,
  rho,
  W,
  ind_l_wt = 1,
  start_q = 1e-08,
  start_indexsigma = 0.1,
  start_sigma = exp(-0.2),
  start_rec_a,
  start_rec_b,
  spawn_prop = 1,
  fix_sigma = TRUE,
  fix_indexsigma = FALSE,
  adrep = FALSE
)
```

Arguments

version numeric, either 1, 2 or 3. This controls what deterministic equations in the model are used to derive population biomass. 1 and 2 use the fraction of of total biomass in a given year due to newly recruited fish. This fraction is derived from mean weights and detailed in the **schnute** vignette. version = 3 is the more classical population dynamics.

- 1 whole biomass derived from recruit biomass
- 2 whole biomass derived from previously exploited biomass
- 3 whole biomass is a combination of recruit biomass and previously exploited biomass

catch_b	numeric vector of catch biomass over time period of assessment
indices_b	matrix of biomass surveys (CPUE) of dimensions: no. of surveys x no.years
ts	numeric. Survey timing parameters
mmts	matrix of mean weights from sampling with dimensions: 3 x no. years. recruit mean weights \bar{Y} (first row), previously exploited biomass mean weights \bar{Z} (second row) and entire assessed biomass mean weight \bar{X} (third row).
tsp	numeric. Timing of spawning. Default to 0 (start of year).
mu	numeric. proportion of catch taken before natural mortality.
rho	numeric. Growth parameter, slope of linear growth model.
W	numeric. Growth parameter, intercept of linear growth model.
ind_l_wt	numeric. Survey weighting in the likelihood. Defaults to 1 for each survey, i.e. equal weighting
start_q	Starting values for survey catchability parameters. Default is 1e-6
start_indexsigma	Starting values for survey sd parameters. Default is 0.1
start_sigma	Starting parameter value fraction of population that survives natural mortality. Default is $e^{0.2}$
start_rec_a	Starting parameter value for the 'a' parameter of the Beverton-Holt stock-recruit function. The asymptotic biomass of recruits. Default is $1/5 * \max(\text{catch}_b)$.
start_rec_b	Starting parameter value for the 'b' parameter of the Beverton-Holt stock-recruit function. The spawning stock biomass needed to produce a/2 on average. Default is $4 * \max(\text{catch}_b)$.
spawn_prop	proportion of biomass that is mature. Defaults to 1 for each year.
fix_sigma	logical. Should survival be fixed in the model
fix_indexsigma	logical. Should survey standard deviation be fixed in the model
adrep	logical. Whether the user would like the ADreport variables (and their derivatives) reported for starting parameters.

Details

schnute_orig is simply a wrapper function that gives the output from [MakeADFun](#) from TMB, i.e. an objective function with derivatives, hessian etc. Optimisation and extraction of values from the assessment models has been kept separate to allow flexibility with optimisation methods as well as easy access to [MakeADFun](#) outputs such as the hessian, gradients etc. See http://kaskr.github.io/adcomp/_book/Introduction.html and TMB documentation for details. Users should refer to the sbar vignette by running `vignette("intro_to_sbar", "sbar")` for details on the schnute_orig function.

The table below gives the outputs and description of the values that can be extracted from this assessment model after optimisation and `summary(TMB::sdreport(x))` of the schnute_orig object.

#'

Output	Description
logitq	logit transformed survey catchability

logitsigma	logit transformed survival of natural mortality
logindex_sigma	log transformed survey standard deviation
lnb	log transformed total biomass
lnpr	log transformed previously-exploited biomass
lnr	log transformed recruit biomass
lnN	log transformed total numbers
lnPR	log transformed previously-exploited or post-recruit numbers
lnR	log transformed recruit numbers
lnC	log transformed observed catch numbers
biomass	total biomass
N	total numbers
ssb	spawning biomass
post_rec	previously-exploited or post-recruit biomass
PR	previously-exploited or post-recruit numbers
rec_bio	recruit biomass
rec_no	recruit numbers
C	observed catch numbers
logpred_survey	log transformed predicted survey indices
omega	fraction of total biomass in a given year due to newly recruited fish
index_sigma	Survey indices standard deviation
sigma	survival of natural mortality
logrec_param	log transformed recruit parameters from a Beverton-Holt SR function if estimated in versions 1 or 3
rec_param	the two recruit parameters fro a Beverton-Holt SR functionif estimated in versions 1 or 3
qhat	survey catchability

Value

List with components for optimiser in R. This output is that of the function [MakeADFun](#) from TMB

Access (transformed) starting values of parameters to be estimated with `x$par` (where `x` is the `schnute_orig` object) to see what parameters are to be estimated.

Examples

```
## Not run: obj <- schnute_orig(catch_b = catch_biomass, indices_b = obs_fill,
ts = timing, mwts = mwts, rho = rho1, W = W1, start_sigma = sigma)
## End(Not run)
```

Index

`ank78`, [2](#)

`csa`, [2](#)

`MakeADFun`, [4](#), [8](#), [9](#), [11](#), [12](#)

`makesbarclass`, [5](#), [6](#)

`plot.sbarclass`, [5](#), [6](#)

`schnute_obserror`, [7](#)

`schnute_orig`, [10](#)

Appendix G

Supplementary material for chapter 5

Supplementary Appendices for *Comparative impacts of ageing bias on stage- and age-based fisheries stock assessment and management*

Luke Batts^{1,*}, C oil n Minto¹, Hans Gerritsen² and Deirdre Brophy¹

February 3, 2022

1 Appendix 1: Methods

1.1 Operating model

Stock-recruitment model for the operating model was a Beverton-Holt function:

$$R_{t+1} = \frac{aS_t^*}{b + S_t^*} \quad (1)$$

where S_t^* is spawning stock biomass and R_t is recruit numbers. a and b are the parameters of the stock-recruitment curve, equivalent to the maximum or asymptotic number of recruits (a) and the spawning stock biomass needed to produce recruitment equal to $a/2$ on average (Beverton and Holt, 1957). The S-R function was parameterised for virgin biomass and steepness, which were defined for each life-history strategy

$$a = \frac{4 \cdot vb \cdot h}{spr0(5 \cdot h - 1.0)} \quad (2)$$

$$b = \frac{vb(1.0 - h)}{5 \cdot h - 1.0} \quad (3)$$

where vb is virgin biomass, h is steepness and $spr0$ is the spawners-per-recruit when fishing mortality is zero.

1.2 Age-slicing bias

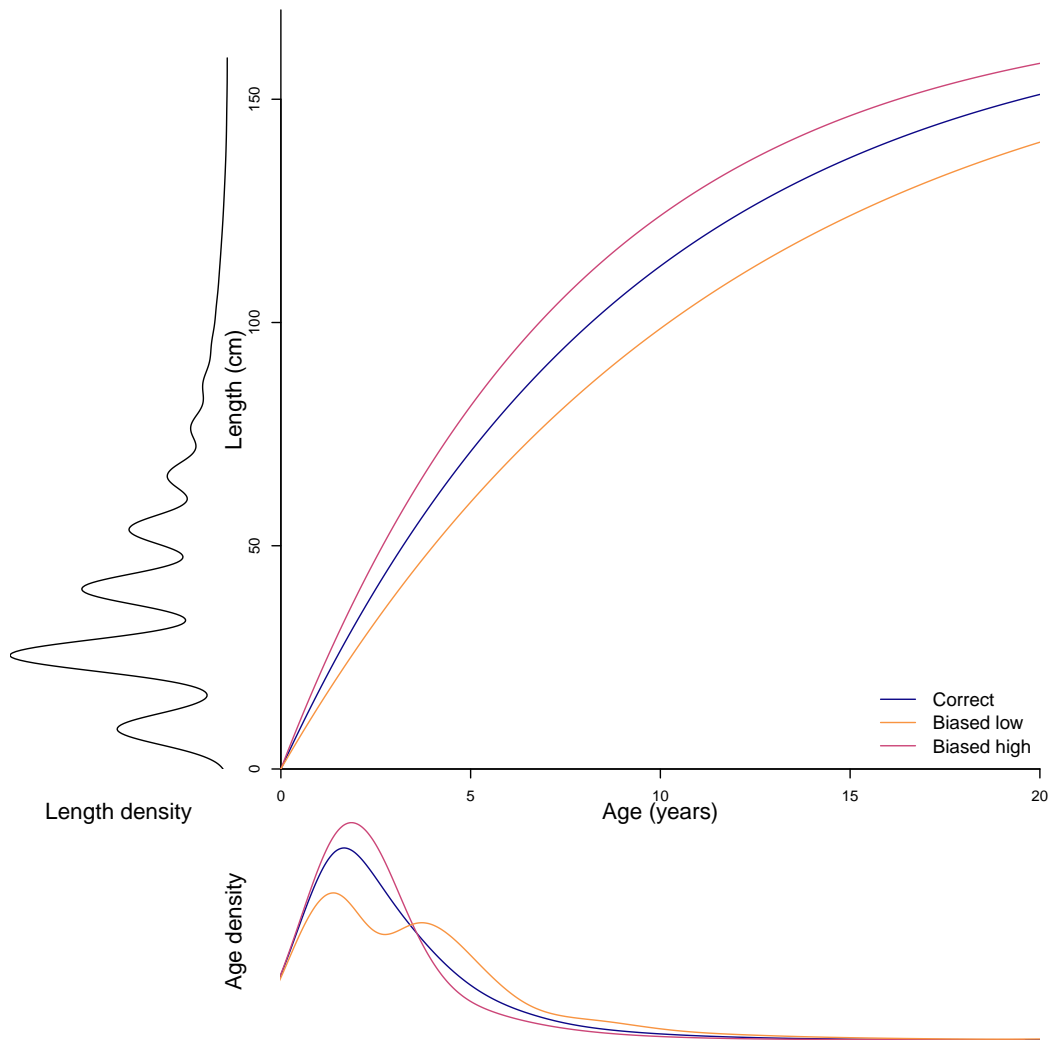


Figure 1: Effect of bias on von Bertalanffy growth curves, numbers-at-length density distribution (step 1 of age-slicing) and re-sliced numbers-at-age density distribution (step 2 of age-slicing) for a typical catch-at-age sample in a given year (biased high = $1.2K^c$ and biased low = $0.8K^c$). Age densities were calculated using a relatively large bandwidth to show larger shifts in density.

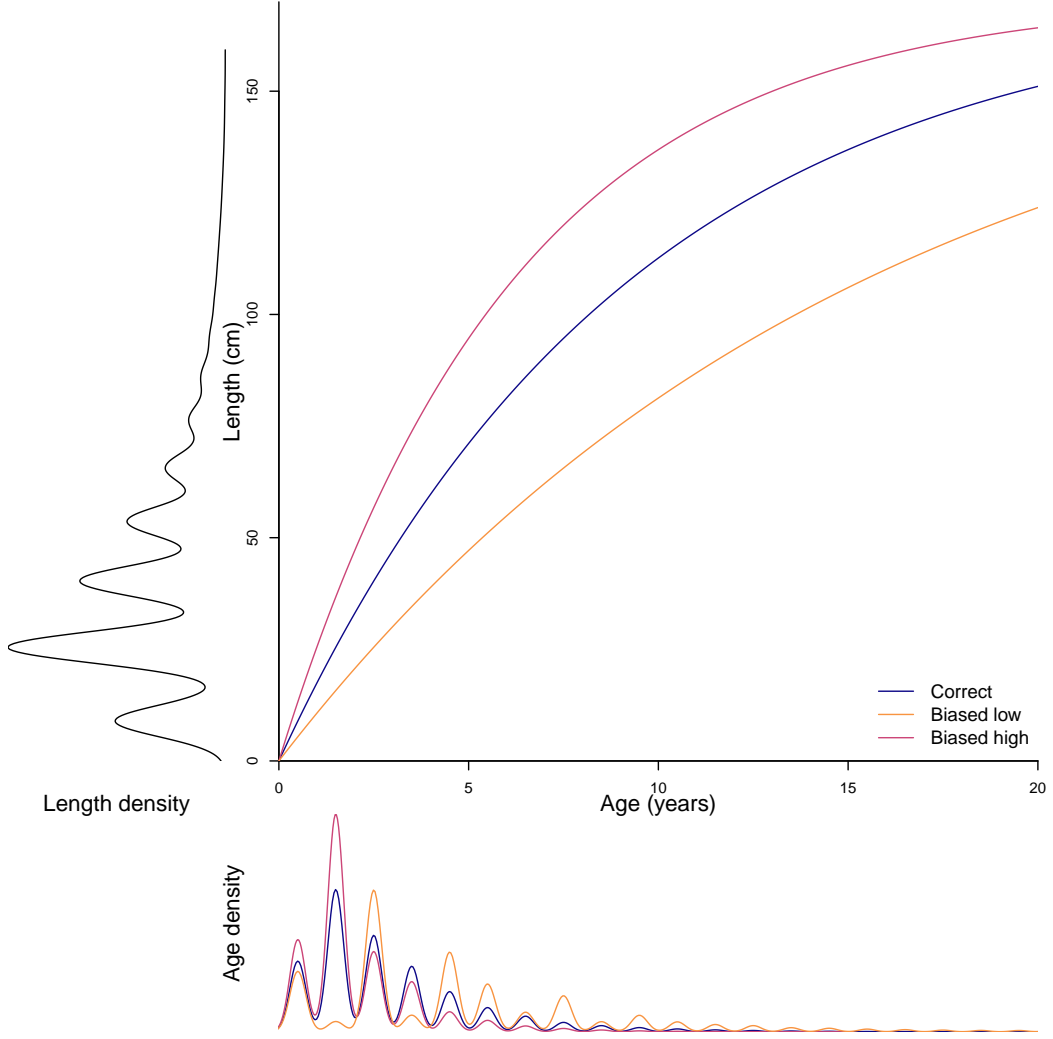


Figure 2: Effect of extreme bias on von Bertalanffy growth curves, numbers-at-length density distribution (step 1 of age-slicing) and re-sliced numbers-at-age density distribution (step 2 of age-slicing) for a typical catch-at-age sample in a given year (biased high = $1.5K^c$ and biased low = $0.6K^c$). Age densities calculated with a relatively small bandwidth.

1.2.1 F_{msy} and N_{msy}

To derive F_{msy} and N_{msy} for a CSA assessment, we first assumed a Beverton-Holt stock-recruit function (B-H) for recruitment numbers at equilibrium R^* ,

$$R^* = \frac{aN^*}{b + N^*} \quad (4)$$

where N^* is stock numbers at equilibrium and R^* is recruit numbers at equilibrium. a and b are the parameters of the stock-recruitment curve, equivalent to the maximum or asymptotic number of recruits (a) and the stock numbers needed to produce recruitment equal to $a/2$ on average (b) (Beverton and Holt, 1957).

Given CSA assumes the stock is fully selected and natural mortality M is constant stock numbers

at equilibrium was derived:

$$N^* = N^* e^{-(F+M)} + R^* \quad (5)$$

$$R^* = N^* - N^* e^{-(F+M)} \quad (6)$$

$$N^* = \frac{R^*}{1 - e^{-(F+M)}} \quad (7)$$

where F is a given fishing mortality for entire stock, M is natural mortality and N^* is stock numbers at equilibrium at a given value of F . Substituting equation 4 into equation 7,

$$N^* = \frac{\frac{aN^*}{b+N^*}}{1 - e^{-(F+M)}} \quad (8)$$

and solving equation 8 for stock numbers at equilibrium,

$$N^* = \frac{a - b(1 - e^{-(F+M)})}{1 - e^{-(F+M)}} \quad (9)$$

gives stock numbers at equilibrium N^* as a function of B-H recruit parameters, natural mortality M of the stock and a given fishing mortality F . Catch numbers at equilibrium were estimated using the Baranov catch equation,

$$C_n^* = \frac{F}{M + F} (1 - e^{-M-F}) N^* \quad (10)$$

where C_n^* is catch numbers at equilibrium. Initially we explored whether catch numbers could be used to estimate a maximum sustainable fishing mortality and maximum sustainable numbers. These preliminary investigations concluded that a completely numbers-based procedure was not possible as catch numbers at equilibrium was heavily weighted by recruitment, thus estimating F_{msy} much too high.

1.2.2 Mean weight at equilibrium

In order to estimate yield or catch biomass at a range of fishing mortality values mean catch weight:

$$C^* = C_n^* \cdot \bar{w} \quad (11)$$

where C^* is catch biomass at equilibrium and \bar{w} is mean catch weight. Mean weight changes with level of fishing mortality, so to account for this we introduce assumptions on age and growth. Again, assuming all fish are fully selected, equilibrium numbers-at-age can be written as:

$$N^* = R^* e^{-(F+M)(a-a_r)} \quad (12)$$

$$(13)$$

where a is a vector of the ages of fish in stock and a_r is the age at recruitment. Mean weight of the recruited stock at equilibrium can be given as:

$$\bar{w}^s = \frac{\sum_{a_r}^{a_{max}} N_a^* w_a^s}{\sum_{a_r}^{a_{max}} N_a^*} \quad (14)$$

$$= \frac{\sum_{a_r}^{a_{max}} (R^* e^{-(F+M)(a-a_r)}) w_a^s}{\sum_{a_r}^{a_{max}} R^* e^{-(F+M)(a-a_r)}} \quad (15)$$

$$= \frac{\sum_{a_r}^{a_{max}} (e^{-(F+M)(a-a_r)}) w_a^s}{\sum_{a_r}^{a_{max}} e^{-(F+M)(a-a_r)}} \quad (16)$$

Equation 16 shows that the proportion of each age that survives at equilibrium and the stock weight-at-age w_a^s determines the mean recruited stock weight \bar{w}^s . Using this methodology, substituting mean stock weight-at-age with catch mean weight-at-age, we can calculate mean catch weight \bar{w} and as a result catch biomass at equilibrium for a given value of fishing mortality.

2 Appendix 2: Results

2.1 a4a initial assessment

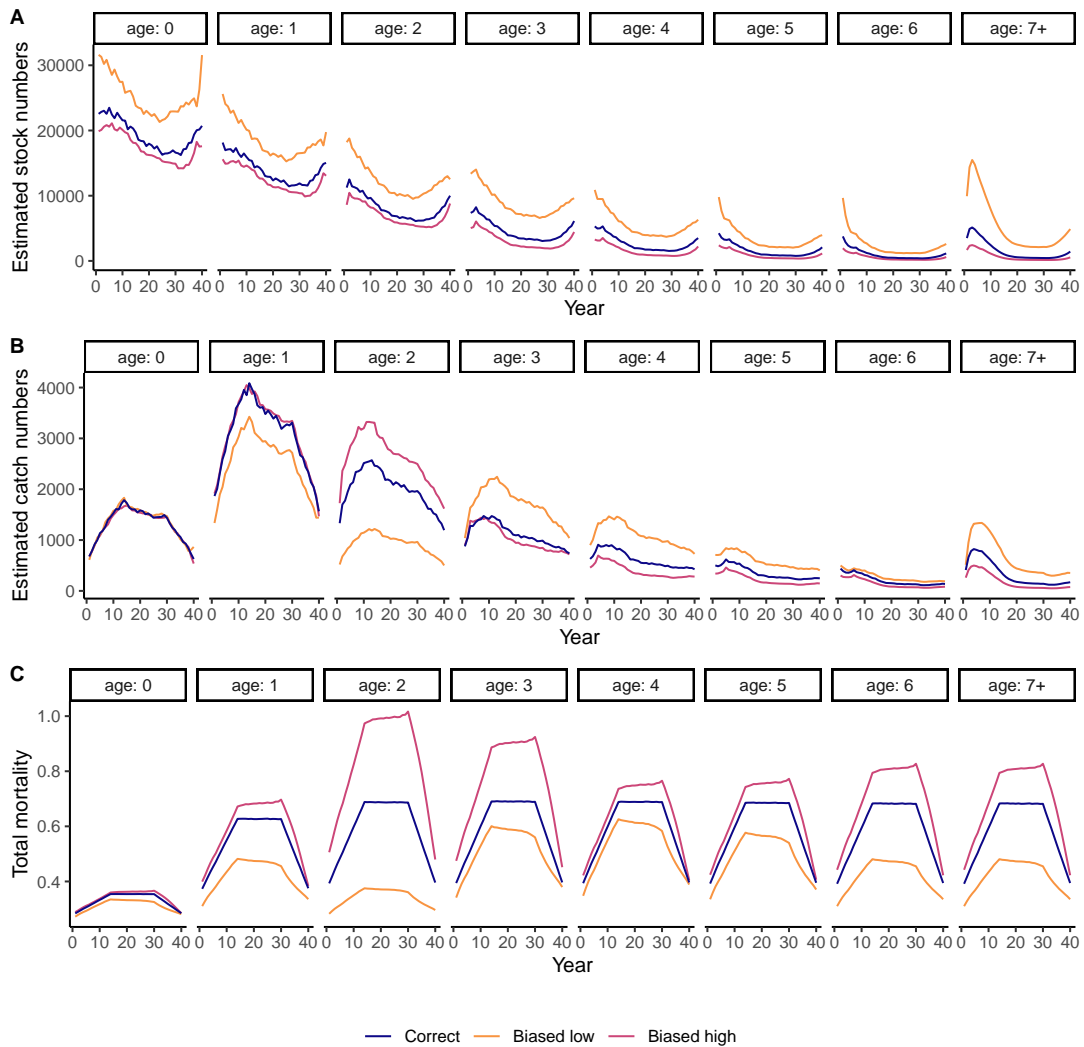


Figure 3: Summary of estimated stock numbers numbers at age, catch numbers at age and total mortality at age from a4a assessments on initial data years (2:41) of correct, biased high and biased low simulated stocks.)

2.2 a4a MSE

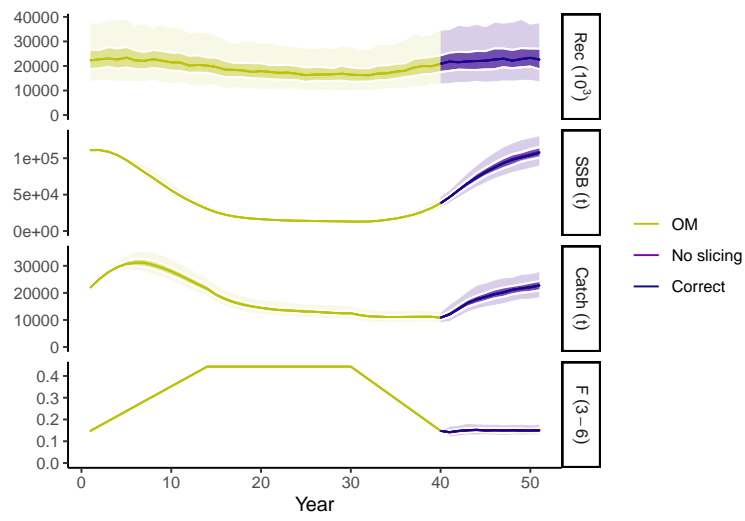


Figure 4: Results of MSEs using a4a as the estimator within the MP and EqSim estimated BRPs for the original stock ("No slicing") and the stock sliced but with no bias on the growth parameter K . Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

2.3 CSA MSE

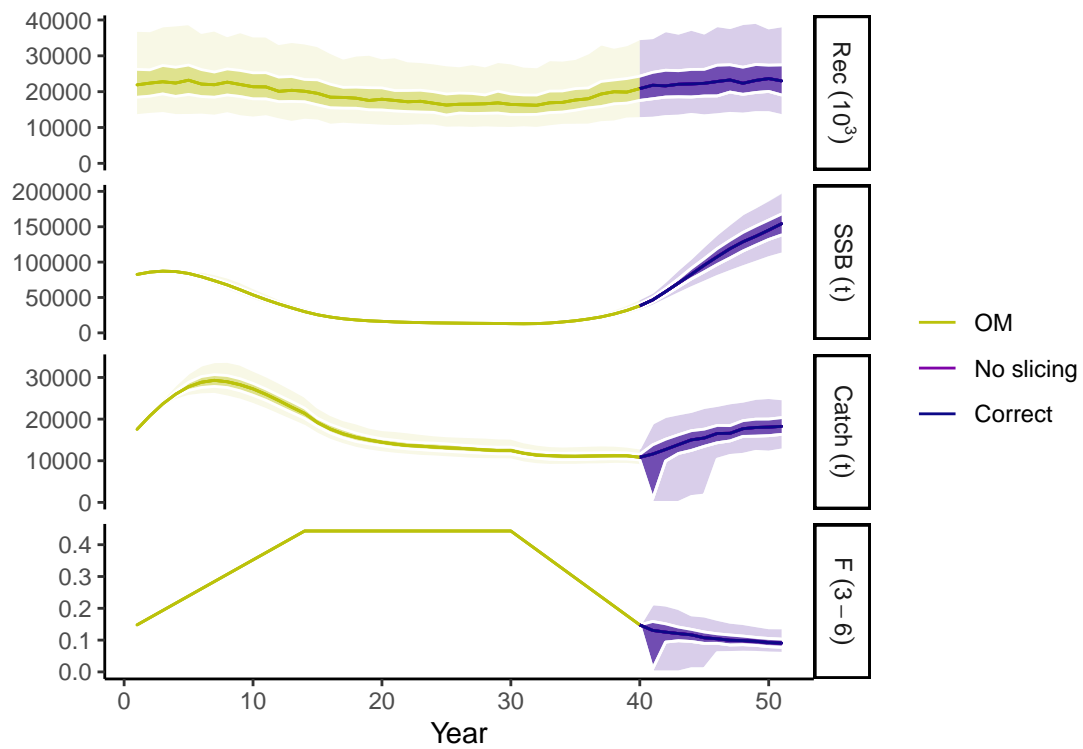


Figure 5: Results of MSEs using CSA as the estimator within the MP and the new methodology for estimating CSA BRPs for the original stock ("No slicing") and the stock sliced but with no bias on the growth parameter k . Solid lines are the median values across replicates, light shading represents 90% quantiles and dark shading 75% quantiles.

2.4 Performance statistics

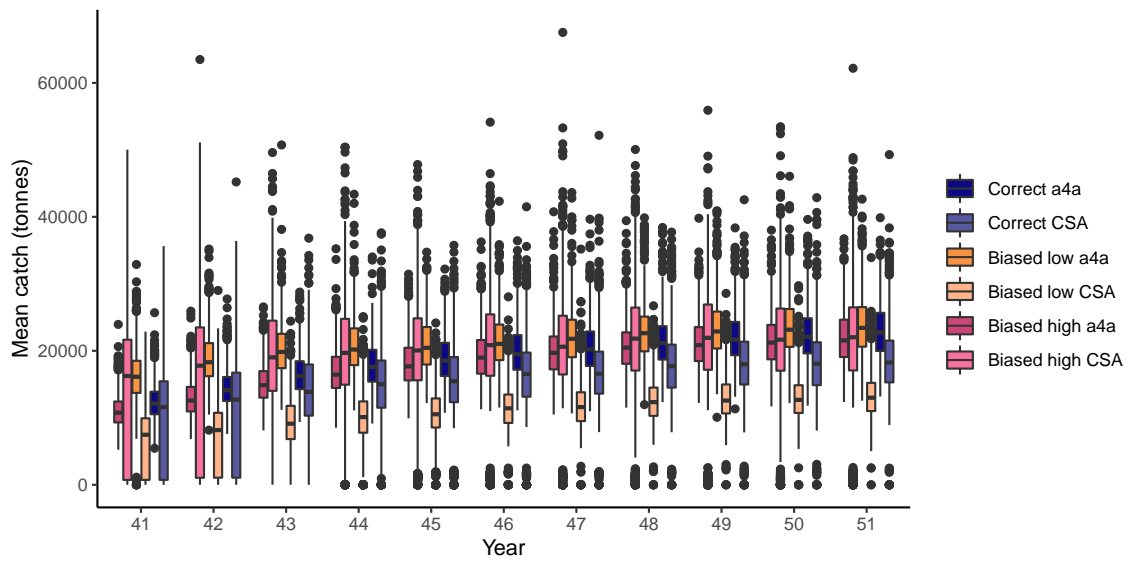


Figure 6: Mean catch biomass from MSE simulations across management procedures (a4a and CSA) and observation error models (*correct*, *biased high* and *biased low*).

References

Beverton, R.J. and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fisheries Investigation Series 2, volume 19, UK Ministry of Agriculture, Fisheries, and Food, London, UK .