Aligning stock assessment and harvest strategies in data-limited, multi-species fisheries: a hierarchical modeling approach to BC's flatfish fishery

by

Samuel Donald Nelson Johnson

M.Sc., Simon Fraser University, 2012
B. Math, University of Newcastle, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
School of Resource and Environmental Management
Faculty of Environment

SIMON FRASER UNIVERSITY
Fall 2021

Declaration of Committee

Name:	**Samuel Donald Nelson Johnson**
Degree:	**Doctor of Philosophy**
Thesis title:	**Aligning stock assessment and harvest strategies in data-limited, multi-species fisheries: a hierarchical modeling approach to BC's flatfish fishery**
Committee:	**Chair:** Frank Gobas Professor, Resource and Environmental Management
	Sean P. Cox Supervisor Professor, Resource and Environmental Management
	Duncan Knowler Committee Member Professor, Resource and Environmental Management
	Liangliang Wang Committee Member Associate Professor, Statistics and Actuarial Science
	Ashleen J. Benson Committee Member Adjunct Professor, Resource and Environmental Management
	Brett van Poorten Examiner Assistant Professor, Resource and Environmental Management
	Steven X. Cadrin External Examiner Professor, University of Massachussets, Marine Science and Technology

Abstract

Contemporary fisheries management involves setting quantitative objectives, and establishing decision rules that determine management actions in response to monitoring data. Within that system, feedback harvest strategies are developed via simulation, where various decision rule features are tested against realistic fishery conditions, and are ranked by their relative performance measured against the quantitative fishery objectives. While the adoption of formal harvest strategies has been increasing globally for some time, most contemporary harvest strategies, and their associated stock assessments, continue to be single-species oriented despite a high prevalence of technical interactions in fisheries. Over three research chapters, I use a simulation approach to investigate whether multi-species harvest strategies, based on hierarchical stock assessment models and multi-species reference points that incorporate technical interactions among species, are better suited to the management of data-limited, multi-species fisheries than traditional single-species approaches. First, I use simulation-evaluation to show that the estimation performance of hierarchical multi-stock surplus production models is more robust to declining data quality than a single-stock version of the same model, creating potential for improved information feedbacks in data-limited contexts. Next, I show that TACs set based on hierarchical model estimates of biomass and productivity are better able to maximise yield under all data quantity scenarios, thanks to negatively correlated biases in key management parameters and target harvest rates that acknowledge technical interactions. Finally, I use closed loop simulation to compare the economic and conservation risks of single-species maximum sustainable yield (MSY), multi-species MSY, or maximum economic yield (MEY) harvest strategies. Taken together, the results of my thesis support expanded usage of hierarchical models in fisheries stock assessment and management, which may lead to wider adoption of formal harvest strategies. Further, I show that yields, food security, and/or economic rent can be increased (not always at the same time) by including technical interactions in reference point calculations.

Keywords: fisheries science; multi-species stock assessment; data-limited; groundfish; harvest strategy

Dedication

Dedicated to Sweeney Constance, Ursa Isaac, and Silas Arthur

Acknowledgements

First, I have to thank my senior advisor, Sean Cox, for his patient guidance on my journey from abstract theoreticist to applied scientist. You included me in your research program and gave me the space to try a great many things, finding the right niche in the world of fisheries to develop my own expertise. You also appear to have infinite patience for my first drafts.

Thanks to my supervisory committee, Lianglang Wang, Ashleen Benson, and Duncan Knowler. You all asked insightful questions and made generous contributions to my work, even when some research paths turned into dead ends. I know I could have called more often.

I am grateful for the funding provided by multiple agencies for this research. For most of my degree, funding was provided by the Canadian Groundfish Research and Conservation Society, the Canadian Sablefish Association, and the Pacific Halibut Management Association through the Mitacs industry partnership program. Additional funding was also provided by NSERC discovery grants via Sean Cox.

I could not have stayed sane without the network of Quantitative Fisheries Research Group friends and colleagues. You all provided lots of interesting discussions, learning moments, and sometimes just gave me the time to complain about why my models weren't fitting. A special thanks goes to our lab-dog, Marco, for the pets.

Thanks also to the scientists at Fisheries and Oceans, Canada, past and present, for providing data, discussions, and hosting me on the last voyage of the Ricker. Special thanks to Rob Kronlund for great discussions and the inside scoop on working with the department, Malcolm Wyeth for organsing the Synoptic trawl survey in 2016, and Norm Olsen, Sean Anderson, Maria Surry at DFO for responding to my (sometimes comprehensive) data requests.

A special thanks goes to my friends and colleagues at Landmark Fisheries Research. In addition to being a great work environment, Landmark provided an opportunity for me to operationalise and apply my research in parallel to my education, which not every student has the chance to do. This has made me a better scientist and consultant, and I am happy to have joined as an official part of the team.

I can't even begin to thank my family enough for their support (and patience) while I completed this journey. Sweeney, you've been the foundation of our lives while I got to

follow my heart through grad school. For ten years! I promise you, this is my last degree. To my boys, Ursa and Silas, Daddy loves you, and I'm glad you both came along to help propel me towards the finish line. To my siblings and parents, you've all provided moral and material support along the way, and I'll always be grateful for that. Matthew, without your support after I first moved to Canada, I wouldn't have had as much of a chance to enjoy my new home, so thank you.

Finally, thanks to Kait, Gabby, Beau, Stephen, Todd, Midoli, Jess, Andy, Cory, Jens, Oliver, Jonathan, Lucas, and Taylor. You all provided valuable friendship over the years and a social outlet in the much needed breaks from research and writing.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Contemporary fisheries management involves setting quantitative objectives and establishing decision rules that determine management actions in response to monitoring data [30]. Within this system, feedback harvest strategies are developed via simulation, where various decision rule features are tested against realistic fishery conditions, and are ranked by their relative performance measured against the quantitative fishery objectives [38, 116]. Within feedback harvest strategies, harvest decisions generate fish stock responses, which feed back through observational data to affect future harvest decisions [65]. Harvest decisions are based on management procedures, which are formal rules for making management decisions under uncertainty, and represent the link between fishery dependent and independent data, stock assessment, and management actions (e.g., catch limits). Uncertain fishery observations, such as survey and catch data, flow through an estimation model (or empirical rule) to produce biomass and/or productivity estimates, which are in turn used to set total allowable catch (TAC) via a pre-defined harvest decision rule (Figure 1.1). TACs are then distributed to harvesters as quota, who apply fishing effort, produce catch, and deplete both quota and the fish stock. Feedback from management decisions often facilitates learning about the productivity of the stock, allowing managers to adapt the managment procedure and more closely achieve the objectives of the harvest strategy.

While the adoption of formal, simulation tested harvest strategies has been increasing globally [112, 90], most contemporary harvest strategies, and their associated stock assessments, continue to be single-species oriented despite a high prevalence of multi-species technical interactions in fisheries [131, 85]. In particular, non-selective harvesting in multi-species fisheries creates choke effects in output controlled fisheries [6]. Choke effects occur when less productive or easier to catch species have their catch limits filled faster than other *technically interacting* species, prematurely restricting access to fishing grounds. When harvest decisions ignore technical interactions, they are, by design, unlikely to produce desired fishery management outcomes [100].

The status quo for multi-species fisheries is to manage them like a collection of single-species fisheries (Figure 1.1(a)), which has two main disadvantages. First, when technical

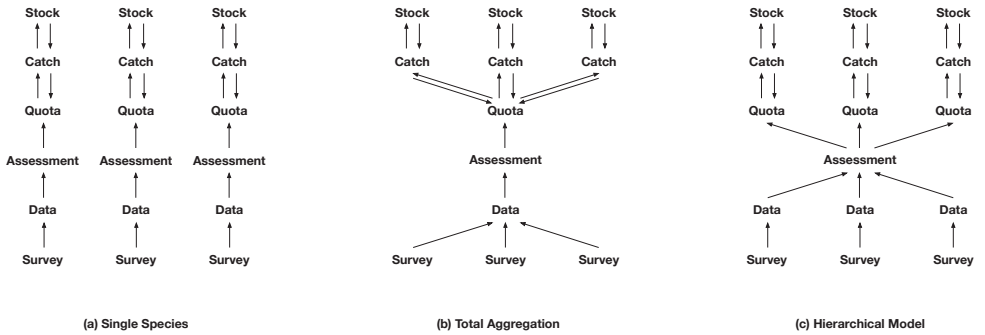

Figure 1.1: Three conceptual models of multi-species fishery management: (a) status quo single species approach, (b) total aggregation of species, and (c) hierarchical multi-species model. Each arrow indicates influence of one element on another, with arrows in two directions showing feedback effects. To reduce complexity, influences of catch on the data element, and stocks on the survey elements, are not shown.

interactions are ignored, catch and quota are often imbalanced because harvest rules are based on single-species population equilibria, or biological reference points, that are derived without technical interactions. When fishing effort is applied, technical interactions lead to bycatch (or byproduct when landed), and quota for bycatch/byproduct species is utilised faster than expected [6]. The resulting catch/quota imbalance then leads to aforementioned choke effects, when quota for easier to catch, or *choke*, species becomes more scarce, and therefore more expensive to acquire, limiting harvester access to fishing grounds and reducing profitability. Second, parallel single-species management systems are linear, or chain-like, with limited (or no) redundancy (Figure 1.1(a)), and so are vulnerable to interruptions at any point along the chain. For example, if a survey indexing stock biomass is discontinued for any reason, but fishing continues after the survey ends, then the information feedback required to link management actions (quota) to stock health via stock assessments is weakened or, when assessments become impossible, eliminated. Such *data limited* scenarios and their weakened feedbacks also amplify choke effects, when highly uncertain stock status estimates lead to more improperly scaled catch limits for all species [15].

Data-limitations in multi-species fisheries are sometimes reduced by aggregating (or pooling) data across species for assessment and management. Pooling effectively increases the data sample size, creating higher statistical power data. The pooled multi-species complex is then assessed as a single stock, producing total biomass and average productivity estimates, which can be used to set a quota for the complex as a whole (Figure 1.1(b)). Such methods are often used to manage tropical reef fisheries, where there are a high number of species with only catch and effort data available, and insufficient resources for single-species management [3, 31]. Technical interactions between species are no longer an issue when catch limits are not allocated to individual species, because catch is limited for the complex as a whole. However, a lack of individual species catch limits could lead to overfishing when

quotas are based on average productivity of the complex, meaning that half of the species are likely to become overfished. Furthermore, without any allocation of quota to individual species, there is no control over the composition of fish that harvesters catch to fulfil the complex quota. Therefore, conditions may arise (e.g., changes in landed values) that encourage targeting or avoidance of one or more species, which can also create overfishing.

There are a number of examples where hierarchical statistical models provide a compromise solution to the challenges of data-limited, multi-species fishery management outlined above [138, 69, 120]. Like single-species models, hierarchical models produce stock-specific management parameter estimates, allowing harvest decisions to be made on a stock-specific basis. Uncertainty in species-specific estimates is reduced under data-limited conditions by sharing information between species via shared, hierarchical prior distributions on model parameters that are assumed to be similar among stocks (e.g., productivity for similar species), rather than through data-pooling. Prior distributions draw parameter estimates towards a hypothetical prior mean via *parameter shrinkage*. Shrinkage produces similar benefits to data pooling, while also allowing parameter estimates to be based more on data than strong *a priori* assumptions, which are common in data-limited scenarios [138, 120, 69]. The statistical benefits of parameter shrinkage are well documented for linear models, such as Ricker stock-recruitment models [107, 118, 80], but there is little research into whether shrinkage effects are beneficial for fishery stock assessment models, where parameters are embedded in recursive, non-linear population dynamics, such as for iteroparous groundfish species. Further, it is unclear whether any resulting statistical benefit translates into improved management outcomes when harvest decisions are based on the outputs of hierarchical stock assessment models.

In this thesis, I investigate and compare the statistical and management performance of both single- and hierarchical multi-species stock assessments, and the risks of basing harvest decisions on those assessments, in a data-limited, multi-species fishery context. Over three chapters, I develop a multi-species, spatially explicit harvest strategy based on hierarchical multi-stock and multi-species stock assessment models (Figure 1.1(c)) and biological reference points that incorporate technical interactions. Using simulation, I show that multi-species harvest strategies produce superior outcomes to strategies based on single-species models and reference points, can eliminate (on average) costly choke effects, and either increase resource rents or yields, depending on the objectives, with acceptable levels of conservation risk.

Simulation models for each chapter are based on the spatial structure and life history of three right-eyed flounders (*Pleuronectidae Spp.*) in British Columbia, Canada - Dover sole (*Microstomus pacificus*), English sole (*Parophrys vetulus*), and (southern) Rock sole (*Lepidopsetta bilineata*) - which I refer to as the BC flatfish complex. Despite 60+ years of commercial exploitation, the BC flatfish complex has low statistical power data, characterised by short, noisy fishery independent series and longer, noisier fishery dependent

series. Given the low power data, stock assessments for BC flatfish complex species are infrequent, with no model based stock assessment for Dover sole ever having been completed [40], around 8-10 years between stock assessments for English sole [42, 135], and even longer between published stock assessments for Rock sole [61], which include both fully age-structured stock assessment models and simpler delay difference models.

For the simulations in this research, I split the BC flatfish complex data into three spatial stock areas. Areas are labeled, from north to south, Hecate Strait/Haida Gwaii (HSHG), Queen Charlotte Sound (QCS), and West Coast of Vancouver Island (WCVI). Each area aggregates multiple Pacific Fishery Management Authority major statistical areas, specifically 5CDE (HSHG), 5AB (QCS), and 3CD (WCVI) (Figure 1.2), and are roughly the same as the management unit boundaries used by Fisheries and Oceans, Canada, for setting catch limits for all three flatfish species [45]. There are two mismatches between the stock structure I assume and the real management boundaries for BC flatfish complex flatfish. First, for Rock sole, area 5E (west coast of Haida Gwaii) is considered an unmanaged separate stock with insufficient data for assessment and minimal removals (and therefore no catch limits) [61], and for English sole, the WCVI and QCS areas are managed as a single stock [45], presumably because of the low catch levels in those areas and the lack of biological data establishing any finer stock structure through phenotypic variation [12].

Despite the mismatch between the stock structure I assume and the management of the BC flatfish complex in practice, there is supporting evidence for my assumption. Tagging studies show very little inter-area movement of all three species [41, 44, 43, 73, 56, 137, 146, 46], with the majority of tagged fish (68% - 99%) found within 15 miles of the location of their release, despite several years at liberty. A notable exception is around 1.4% of Dover sole tagged in Hecate Strait (5D) recaptured off the north coast of Haida Gwaii (5E) during winter rockfish fishing, which is a known winter spawning ground for Dover sole [43]. Further, while it was very rare (i.e., $< 1\%$), English sole appear to be most prone to larger migrations, with fish tagged in northern Hecate Strait (5D) found as far south as Washington state (area 3B) and one fish tagged in area 5B was found as far south as northern California (area 1C) [46]. While the results of the tagging studies may be biased by the proximity of tagging locations to areas of high fishing intensity for BC flatfish complex species, there is also further supporting evidence for the assumed stock structure. Phenotypic differences exist at the assumed scale for Dover sole, for which length and age data exist in all three management areas, which when fit to produce distinct von Bertalanffy growth models with asymptotic lengths L_∞ for males ranging between 45 cm and 49 cm, while female fishes have a smaller range between 53 cm and 54 cm[1]. Similarly, von Bertalanffy asymptotic length L_∞ parameters for male Rock sole differ by around 4cm between the HSHG and QCS areas (no

[1]Unpublished data, GFBIO database hosted by Fisheries and Oceans, Canada, at the Pacific Biological Station in Nanaimo, BC

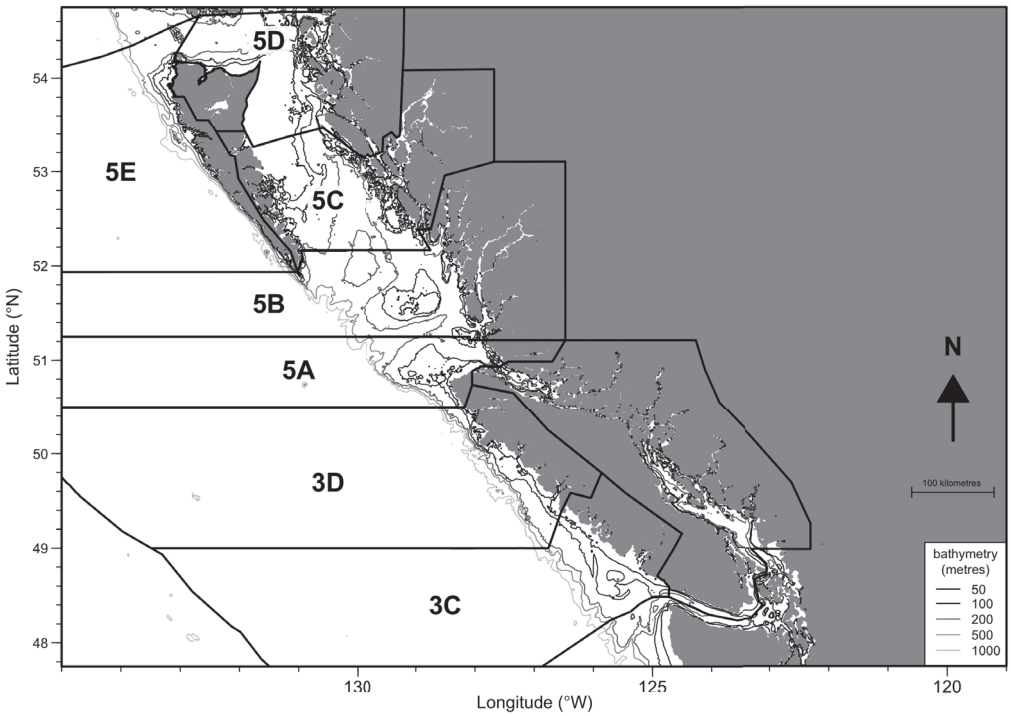

Figure 1.2: British Columbia PFMA boundaries and bathymetric contour lines

biological data exists for QCS English sole, or WCVI English or Rock soles). Additionally, there are deep gullies that run through areas 5A and 5B (QCS) from coastal inlets to the continental slope, which likely inhibit migration of flatfish between each of the areas (Figure 1.2). Finally, choosing to aggregate are 5E Rock sole with the 5CD Hecate Strait population is probably of no consequence, because the evidence points to Rock sole having finer resolution stock structure than both my assumption and the current management unit boundaries [73, 56].

In Chapter 2, I investigate strengthening the links between low-power data and stock status estimates via hierarchical multi-stock surplus production assessment models. Highly uncertain fishery data weakens feedback links between stock status and catch limits. Those weakened feedbacks lead to sub-optimal fishery management when status estimates are highly uncertain, or missing, given low statistical power data [16]. To improve estimation in the presence of low power data, hierarchical models are defined with shrinkage priors for sur-

vey catchability and/or stock productivity. Single-stock and hierarchical multi-stock models are fit to simulated data for a multi-stock Dover sole complex in a simulation/evaluation study, demonstrating links among data quality, parameter shrinkage, and model bias and precision for hierarchical multi-stock stock assessment models. The benefits of shrinkage are evaluated by comparing bias and precision in estimates from single-stock models to bias and precision in hierarchical multi-stock assessment models estimates, while varying data quality in high and low statistical power scenarios. Finally, hierarchical multi-stock and single-stock models are fit to real catch and index data for BC Dover sole complex, which has never had a model based stock assessment [70]. Results show that estimates from multi-stock models are more robust to simulated low statistical power scenarios. Furthermore, when fit to real Dover sole data, hierarchical multi-stock models are selected over single-stock models when ranked according to Akaike's Information Criterion.

In Chapter 3, I use closed-loop simulation to understand how parameter shrinkage affects the fishery management outcomes, by testing harvest strategies for BC flatfish complex stocks based on surplus production stock assessment models. Under hierarchical models, gains in precision and reductions in bias for stocks with low power data can come at the cost of increased bias and reduced precision for stocks with higher power data [119]. Therefore, it it is unclear if the improvements in information feedbacks for low power stocks offset the degrading of feedbacks for high power stocks, leading to a net benefit in fishery outcomes. On the other hand precise or unbiased models are not guaranteed to produce acceptable fishery management outcomes, due to irreducible error in all statistical models [134, 81]. To understand the management performance of hierarchical models, I use biomass estimates from five surplus production models, including hierarchical multi-stock and single-stock models, to set catch limits for the BC flatfish complex under high, medium, and low data quantity scenarios. Management performance is estimated via comparison to decisions by an omniscient manager [144, 86], showing that hierarchical multi-species stock assessments produce superior conservation and yield performance across all scenarios.

Finally, in Chapter 4, I compare conservation and economic risks of single-species harvest strategies (ignoring technical interactions) against multi-species harvest strategies (incorporating technical interactions). Choke effects caused by ignoring technical interactions can limit the resource rent (profit) available to harvesters by limiting revenue when fishing grounds are closed. A bio-economic model for the BC flatfish complex fishery is extended from the operating model in Chapter 3, with realistic fishing costs and revenues [5]. Harvest strategies based on multi-species maximum sustainable yield (MSY) and maximum economic yield (MEY) reference points, which incorporate technical interactions, are compared to single-species MSY strategies via economic and conservation performance. Multi-species MEY harvest strategies are shown to consistently out-perform single- and multi-species MSY based strategies in both conservation and economic outcomes.

1.1 Statement of interdisciplinarity

This research brings together the natural and social sciences, specifically fisheries science/management and economics. Chapter 4 quantifies economic and conservation risks of applying single- and multi-species harvest strategies for the BC flatfish complex multi-species fishery. To define harvest strategies maximising resource rent (MEY) and estimate economic risks, I defined a bio-economic model that integrates realistic economic models of demand and costs of fishing with a multi-species population dynamics and fishery model.

1.2 Contributions

A PhD thesis is not a solitary achievement. I am the first author of all chapters, having conducted all analyses and written all first drafts, and as such the thesis is written in the first person singular. However, Chapters 2 - 4 are edited from manuscripts that are either already published in the primary literature (Chapters 2 and 3), or are intended to be (Chapter 4), and all three chapters have benefited from discussions, editing, and comments from co-authors Sean Cox (Chapters 2 - 4) and Duncan Knowler (Chapter 4). While Chapters 1 and 5 are not intended to be peer-reviewed primary research and have no formal co-authors, they also benefited from discussions and review by colleagues and committee members, especially by my senior supervisor Sean Cox.

Chapter 2

Evaluating the role of data quality when sharing information in hierarchical multi-stock assessment models, with an application to Dover sole

This chapter is published in the Canadian Journal of Fisheries and Aquatic Sciences (2018), co-authored with Sean P. Cox under the same title [70].

2.1 Introduction

Fisheries stock assessment modeling uses catch and abundance monitoring data to estimate the status and productivity of exploited fish stocks [57]. Despite improvements in catch monitoring and increasing prevalence and quality of fishery-independent surveys of abundance, many fisheries remain difficult to assess because the data lack sufficient statistical power to estimate key quantities necessary for management [106]. Low power data may arise, for example, because time-series are short relative to the productivity cycles of exploited fish stocks, historical fishing patterns may be weak or uninformative, and monitoring data may simply be too noisy to extract biomass and productivity signals [79]. Where these situations occur, stocks are often deemed *data-limited* [77, 16].

An emerging approach to fisheries stock assessment is to use a hierarchical approach to assess data-limited stocks simultaneously with data-rich stocks. Data-limited stocks can "borrow information" from data-rich stocks, providing a compromise between data-intensive single-stock assessments and problematic data-pooling approaches [69, 68, 120]. The hierarchical multi-stock approach, which shares information between data-rich and data-poor stocks, treats multiple stocks of the same species as replicates that, to varying degrees, share environments, life history characteristics, ecological processes, and fishery interac-

tions [107, 118, 80]. Information present in the observations for data-rich replicates is shared with more data-poor replicates via hierarchical prior distributions on parameters of interest [120, 138]. Sharing information in this way could improve scientific defensibility of assessments for data-limited stocks, because stock status and productivity estimates are informed by data rather than strong *a priori* assumptions on population dynamics parameters.

Information-sharing properties of hierarchical models are realized as the shared hierarchical priors induce shrinkage of estimated parameters towards the overall prior mean [13, 51]. Although shrinkage can reduce bias in the presence of high uncertainty (e.g. very data-limited stocks), it may also increase bias for data-rich replicates by pulling estimated parameters closer to the group mean. Shrinkage properties are well understood for hierarchical linear models [67, 124], including those applied in fisheries. For example, when estimating productivity of Pacific salmon stocks, hierarchical Ricker stock-recruitment models are more successful at explaining variation in stock productivity when stocks are grouped at scales consistent with climatic variation [107, 94]. It is unclear, however, whether the benefits observed for linear models extend to iteroparous groundfish stocks, for which productivity parameters are deeply embedded within non-linear population dynamics and statistical models.

Parameter shrinkage has been observed in stock assessments for data-limited groundfish and shark species when grouped with data-moderate species [69, 68, 120], but it is unknown whether such shrinkage in reality increases or decreases bias in parameter estimates. Simulation tests of the hierarchical multi-stock approach to age-structured assessments revealed that bias reductions in one species often induce greater bias for others in the assessment group, indicating that shrinkage could imply unwanted trade-offs [119].

In this chapter, I use a simulation approach to investigate relationships between hierarchical model structure, bias, and precision for hierarchical multi-stock Schaefer stock assessment models. For the hierarchical multi-stock models, shared prior distributions are applied to survey catchability and optimal harvest rate (productivity) and then evaluated to find the combinations of shared priors that produce the most reliable estimates of key management parameters when fit to simulated data from high and low data quality multi-stock complexes. The best performing single and multi-stock models were then applied to real data for a Dover sole complex in British Columbia, Canada.

2.2 Methods

A multi-stock complex representing the Dover sole (*Microstomus Pacificus*) fishery in British Columbia, Canada is simulated. Simulated Dover sole observations are generated in data quality scenarios ranging from low to high statistical power. Under each scenario, bias and precision metrics are determined for key management parameters under both single-stock and hierarchical multi-stock Schaefer models. In the hierarchical multi-stock assessment

models, shared evolutionary history and a common scientific survey influence the definition of shared prior distributions. For example, stocks that share evolutionary history may have similar productivity at low stock sizes [69, 68], and a common trawl survey may induce correlations in catchability (trawl efficiency) observation errors.

2.2.1 Study system

British Columbia's Dover sole complex is divided into three distinct but connected stocks (Figure 2.1), distributed along the BC coast from the northern tip of Haida Gwaii, south through Hecate Strait into Queen Charlotte Sound, and on the west coast of Vancouver Island. Although the Dover sole fishery has operated since 1954, prior to 1970 it was very limited, increasing to present levels by the late 1980's (Figure 2.2).

Despite a long history of exploitation, Dover sole stocks have never been evaluated using model-based assessments. No observational data exists for the Queen Charlotte Sound (QCS) and west coast of Vancouver Island (WCVI) stocks prior to 2003, precluding a model based assessment before that time [39]. The Hecate Strait and Haida Gwaii (HSHG) stock was surveyed from 1984 - 2003 (Figure 2.2, Survey 1), but data was only used to perform catch curve analyses for total mortality rate estimates [39]. During 1984 - 2003, a fine-mesh trawl survey was used for the Vancouver Island stock and a portion of the Hecate Strait stock, but the survey was not designed for groundfish and produced stock indices that were highly variable. Since 2003, a newer bottom trawl survey has operated coast-wide, which samples all three stocks (Figure 2.2, Survey 2), but no assessment has been performed in that time.

Dover sole may be suitable for a hierarchical multi-stock assessment for 3 main reasons. First, the Hecate Strait stock has longer series of informative data than the other stocks, potentially providing information for the other two stocks. Second, modeling a single-species makes it likely that stock productivities and responses to the environment are similar. Lastly, all stocks are observed by Survey 2, making it likely that the observation model parameters for each stock are similar for that survey. By applying the hierarchical multi-stock approach, the similarities between stocks may be exploited to the benefit of the whole complex, extending model based stock assessments for Dover sole for the first time.

2.2.2 Simulation Framework

The simulation framework is composed of an operating model that simulates biological dynamics, catch, and observational data, and an assessment model that performs both single-stock and hierarchical multi-stock assessments of the simulated data. Both operating and assessment models use a process-error Schaefer formulation for biomass dynamics, where the biomass in each year deviates from the expected value using a log-normal process error term. Matching the operating model and assessment model enabled me to focus on the effects of hierarchical estimation and shrinkage without confounding among hierarchical priors and

the model structure. The operating model is specified in the R statistical software package, and the assessment models were written in the Template Model Builder (TMB) package [123, 76].

The simulation approach is described below in 3 main sections (i) the operating model, (ii) assessment models, and (iii) simulation experiments. The next section describes the operating model structure, including process errors, and how catch and survey observations were generated. Assessment models are then outlined, with details of the shared hierarchical prior distributions given in Appendix 2.7.A Finally, I present the experimental design and performance metrics for the simulations.

Operating model

Biomass dynamics are simulated for each stock s in the assessment complex on an annual time step t, using the process-error Schaefer model [111]

$$B_{s,t+1} = \left(B_{s,t} + r_s B_{s,t}\left(1 - B_{s,t}/B_{s,0}\right) - C_{s,t}\right) e^{\epsilon_{s,t}}, \tag{2.1}$$

where $B_{s,t}$ is the biomass of stock s at time t, r_s is the intrinsic rate of increase, $B_{s,0}$ is the unfished equilibrium biomass, and $\epsilon_{s,t}$ is the process error deviation for stock s at time t. Schaefer model process error deviations $\epsilon_{s,t}$ are decomposed via the sum of a shared (across stocks) mean year-effect $\bar{\epsilon}_t$, and a correlated (among stocks) stock-specific effect $\zeta_{s,t}$, which is the s component of the vector $\zeta_{\cdot,t}$, that is,

$$\begin{aligned} \epsilon_{s,t} &= \bar{\epsilon}_t + \zeta_{s,t}, \\ \bar{\epsilon}_t &\sim N(0, \kappa), \\ \zeta_{\cdot,t} &\sim N(\vec{0}, \Sigma). \end{aligned}$$

The covariance matrix Σ is specified as a diagonal decomposition $\Sigma = DMD$, where D is a diagonal matrix of stock-specific standard deviations σ_s, and M is the matrix of stock correlations. For simplicity, all stocks are simulated with identical pair-wise covariances, i.e., for a 3 stock complex

$$M = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix},$$

and all stocks experience the same magnitude of stock-specific process errors where $\sigma_s = \sigma$, implying

$$D = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma \end{pmatrix}.$$

The operating model values of κ and σ are chosen to give a total process error variance of $\sigma^2 + \kappa^2 = 0.01$, or roughly a 10% total relative standard error (Table 2.1).

Fishery history is simulated for 34 from 1984 ($t = 1$) to 2017 ($t = 34$). Each stock is initalized in 1984 at a pre-determined depletion level $d_{s,1}$ relative to unfished biomass, i.e., $B_{s,1} = d_{s,1} \cdot B_{s,0}$. Unless otherwise stated, $d_{s,1} = 1$, which is varied as an experimental factor (Table 2.2). Because the simulated populations were a single-species, multi-stock complex, the same base biological parameters $B_{s,0}$ (kilo-tonnes), and r_s are simulated for all stocks s (Table 2.1). While identical parameters may not faithfully represent the true Dover sole complex, they help to focus the results on the effects of shrinkage in parameter estimates, rather than differences in biological parameters. This choice also simplifies reporting and interpretation of the results, reducing the results to a smaller set of representative stocks through symmetry, rather than analysing every stock in the complex.

Fishery catch and fishery independent biomass indices are sampled from each stock each year. Catch is assumed to be perfectly implemented as $C_{s,t} = U_{s,t}B_{s,t}$, where $U_{s,t}$ is the harvest rate applied in a pulse fishing event following each year's production, and fully observed (i.e., no under-reporting). Harvest rates are simulated in three temporal phases and scaled to optimal fishing mortality as $U_{s,t} = U_t^{mult} \cdot U_{s,MSY}$, where U_t^{mult} is the piecewise linear function of t:

$$U_t^{mult} = \begin{cases} U_i + (t-1) \cdot \frac{U_d - 0.2}{t_d - 1} & 1 \leq t \leq t_d, \\ U_d + (t - t_d) \cdot \frac{U_m - U_d}{t_m - t_d} & t_d \leq t \leq t_m, \\ U_m & t_m \leq t \leq T; \end{cases} \tag{2.2}$$

where U_i, U_d and U_m are the initial, development, and managed phase harvest rates, respectively, t_d is the last time step of the development phase, and t_m is the beginning of the final managed phase (Figure 2.3). In the base operating model, I used values $U_i = 0.2$, $U_d = 4$ and $U_m = 1$ for harvest rate multipliers, with $t_d = 5$ (1988), and $t_m = 15$ (1998) for phase timing, to simulate a high initial development phase followed by a reduction in pressure, allowing the stock to recover. This formulation was designed to create more and less informative catch histories, depending on the parameter values [129].

Survey indices of biomass are simulated for each stock s and survey o via the observation model

$$I_{o,s,t} = q_{o,s} B_{s,t} e^{\delta_{o,s,t}},$$

where $q_{o,s}$ is stock-specific catchability coefficient for survey o. Observation errors are drawn from the distribution

$$\delta_{o,s,t} \sim N(0, \tau_o),$$

where τ_o is the survey observation error log scale standard deviation for survey o. Within each survey, stock-specific catchabilities $q_{o,s}$ are randomly drawn from a log-normal distribution with a mean survey catchability coefficient $\bar{q}_o$ and between-stock log-standard

deviation $\iota_{q,o}$ via

$$q_{o,s} \sim \log N(\bar{q}_o, \iota_{q,o}).$$

It is not always true that catchability will be similar between stocks. Indeed, it is only possible to model catchability as a hierarchical process between stocks here because biomass indices are swept area estimates. To see this, note that the general formula for catchability is $q = ca/A$, where c is gear efficency, a is the average area fished by the gear during the survey, and A is the total area of the surveyed stock's habitat [4]. Because the geographic boundaries of stocks may differ, it will usually be the case that $A \neq A'$ between 2 distinct stocks s and s', even if the average surveyed area a and gear efficiency c are the same. For a trawl survey, it is advantageous that the area swept by the fishing gear is often known exactly, with $a = t \cdot v \cdot w$, where t is the standard tow duration, v is the tow velocity and w is the door-width of the trawl net. Therefore, the total of randomly sampled survey catches $C_t = qE_tB_t$ from a total effort of $E_t = n_t$ tows can be transformed into biomass estimates when scaled by the reciprocal of the proportion of area swept, e.g. $B'_t = \frac{A}{n_t a} C_t = cB_t$. Then the effect of stock area is scaled out of the index, and catchability is reduced to gear efficency c, or the response of individual fish to the survey gear. A major assumption of this chapter is that individual response to the survey gear is similar within the same species. This argument is then extended to swept area biomass estimates calculated from a stratified survey, like the trawl survey used for Dover Sole.

Biomass indices are simulated from two surveys operating over different periods to emulate the current Dover sole complex history (Figure 2.2). The first ($o = 1$) represented Survey 1, which operated from 1984 to 2003 ($t = 1, \ldots, 20$), with observation model parameters $\tau_1 = 0.2$ for the observation errors, and a mean survey catchability of $\bar{q}_1 = 0.5$ with a standard deviation of $\iota_{q,1} = 0.1$. For survey 2 ($o = 2$), which operates from 2003 to 2017 ($t = 20, \ldots, 34$), an observation error standard deviation of $\tau_2 = 0.4$ is used, and a mean catchability of $\bar{q}_2 = 0.6$ with a standard deviation of $\iota_{q,2} = 0.1$.

Assessment model

Stock-specific biological and management parameters arre estimated using multi-stock and single-stock versions of a state-space Schaefer stock assessment model. To minimize the effect of assessment model mis-specification, the deterministic biomass dynamics in the assessment models and the operating models were defined to be the same, Equation (2.1). Details of the assessment model prior distributions are not presented in this section. Instead, the equations for each multi-level prior in the hierarchical multi-stock assessment model are given in Table 2.3, and the details of all prior distributions are given in Appendix 2.7.A.

Hierarchical multi-stock assessment models For the full hierarchical multi-stock model, shared prior distributions are defined for conditional maximum likelihood estimates

of stock-specific catchability $\hat{q}_{o,s}$ within each survey, and optimal harvest rate $U_{s,MSY}$, which is a surrogate for stock productivity (Table 2.3). In total, 4 configurations are defined, including a "null" multi-stock model. Each multi-stock model configuration is defined by whether each of the hierarchical priors is estimated along with the leading model parameters. When a hierarchical prior is "off", shared priors are bypassed and the model used the fixed hyperprior mean and standard deviation instead (Table 2.3, Single level priors). Full details of the single and multi-level priors are in supplemental material.

Single-stock assessment model The single-stock assessment model is defined as a special case of the multi-stock null model. Prior distributions on catchability and productivity use the single level priors (Table 2.3, q.4 and U.4).

Optimization Assessment models integrate the objective function over random effects using the Laplace approximation, obtaining a marginalized likelihood [76]. The marginalized likelihood was then maximized via the `nlminb()` function in `R` to produce parameter estimates and corresponding asymptotic standard errors [123]. An assessment model was considered to be converged when the optimisation algorithm report convergence, which was characterized by gradient components of the TMB model all having magnitude less than 0.001, and a positive definite Hessian matrix. Standard errors of derived parameters are estimated from the Hessian matrix using the δ-method. Random effects were process errors $\zeta_{s,t}$ for all model configurations, and stock-specific catchability parameters $\log q_{os}$ when the shared catchability prior was estimated.

2.2.3 Simulation experiments

An experimental design approach is used to investigate performance of the four hierarchical multi-stock assessment model configurations under different levels of statistical power in the simulated data. Multiple scenarios are defined to determine whether (and possibly to what extent) hierarchical multi-stock assessment methods can provide better estimates of key management parameters, compared to single-stock approaches, when fit to data with low statistical power.

Experimental factors are selected to increase and decrease the statistical power, or quality, of the simulated assessment data. The choice of factors determining high- and low-information scenarios was guided by previous studies of assessment models, as well as my own experience with production model behaviour [57, 79, 24]. Combinations of experimental factors were chosen according to a space-filling experimental design (Table B.1) [74]. Space filling designs improve the efficiency of large simulation experiments by reducing the number of individual runs, while still producing acceptable estimates of factor effects.

High and low statistical power scenarios were generated by varying 5 experimental factors: (1) historical fishing intensity; (2) the number S of stocks in the complex; (3) the

number L of low information stocks in the complex; (4) the initial year of stock assessment T_1 for the L low information stocks; and (5) the initial stock depletion levels $d_{s,1}$ for the L low information stocks (Table 2.2).

Two levels of historical fishing intensity were defined, varying U_i, U_d and U_m in Equation (2.2). Levels are chosen to produce one-way and two-way trip dynamics when the simulated biomass is initialised at unfished equilibrium in 1984. One way trips are produced by fishing at a constant rate of $U_{s,MSY}$ for the whole historical period (top row, Figure 2.3), while the two-way trips are produced by the base operating model settings (bottom row, Figure 2.3). The constant harvest rate scenarios have two significant disadvantages: first, it is impossible, in general, to estimate the optimal harvest rate without overfishing [60, Ch 1], which does not occur in these scenarios; second, when stocks are initialized at fished levels it was difficult to determine the stock size and initial biomass.

Complex sizes S test the intuitive notion that grouping more stocks together increases the benefit of shrinkage. The sensitivity of this notion to relative differences in the number of stocks is tested via the factor L, which determines how many of the S stocks are "low information". Low information stocks have short time series and fished initialisation at a pre-determined depletion d_s level, which together reduced or removed contrast in biomass time series and lower observational data quality. By beginning assessments of low information stocks when Survey 2 was initiated, and simulating Survey 2 as a shorter and noisier series of observations, I subject low information stocks to non-equilibrium starting conditions as well as poor quality survey data, a situation that is common for data-limited fisheries. When $L > 0$, initial biomass B_{s,T_1} was estimated for the low information stocks in addition to unfished biomass, optimal harvest rate and catchability.

The single-stock and each hierarchical multi-stock assessment model configurations are fit to simulated data under each combination of experimental factors. The distributions the single-level and multi-level hyperpriors (Table 2.3, q.2, q.4, U.2, and U.4) are given random mean values m_q and m_U in each simulation replicate, chosen from a log-normal distribution centred at the true mean value (across stocks, and possibly surveys) with a 25% coefficient of variation, testing the robustness of the assessment model to uncertainty in the prior distribution. The same initial seed value R is used across all experimental treatments so that variability in assessment error distributions is predominantly affected by the factor levels and model configurations, rather than random variation in the process and observation errors. Random variation is not completely avoidable, though, as some assessment models fail to converge under some combinations of treatment and random seed values. In these cases, the optimisation was restarted with jittered initial parameter values up to 20 times, after which the simulation moved on to a different random seed value. The total number of replicates for each experiment and prior configuration are shown in Table 2.6 (Appendix 2.7.B).

Performance metrics

Performance of both the single-stock and multi-stock assessment models was measured by their ability to estimate current biomass $\hat{B}_{s,2017}$, MSY level biomass $\hat{B}_{s,MSY}$, equilibrium optimal harvest rate $\hat{U}_{s,MSY}$, and relative terminal biomass $\hat{B}_{s,2017}/\hat{B}_{s,0}$. I also found catchability estimates $\hat{q}_{o,s}$ to be important in the analysis of these models, so I calculated performance metrics for catchability as well.

It is important to understand the effect of shrinkage on the bias and precision of estimates of the key parameters θ above, because such shrinkage may result in misleading harvest advice. For example, shrinkage may simultaneously increase both bias and precision for a given parameter (e.g. MSY), leading to confidence intervals that may not contain the true parameter value. Therefore, I used four performance metrics to represent these effects: (1) median relative errors (MREs); (2) ratios of median absolute relative errors (MAREs); (3) confidence interval coverage probability (IC); and (4) the predictive quantile. All metrics are defined in detail below. While MREs only indicate model bias, all other metrics are affected by both the bias and precision of the estimator, and can be better interpreted when the bias is known.

For MRE and MARE metrics, relative errors $RE(\hat{\theta}_{i,s})$ of the model estimate $\hat{\theta}_{i,s}$ were calculated for each replicate i and stock s, i.e.

$$RE(\hat{\theta}_{i,s}) = 100 \cdot \left(\frac{\theta_{i,s} - \hat{\theta}_{i,s}}{\theta_{i,s}} \right).$$

Estimator bias and precision were quantified by computing the median relative error $MRE(\theta_s) = \text{med}(RE(\hat{\theta}_{\cdot,s}))$ and median absolute relative error $MARE(\theta_s) = \text{med}(|RE(\hat{\theta}_{\cdot,s})|)$ of relative error distributions $RE(\hat{\theta}_{\cdot,s})$ over all replicates i. I chose to use MAREs because they are independent of scale and less sensitive to outliers than root mean square errors. Values closer to zero indicate better performance for both metrics, with lower MRE values indicating lower bias, and lower MARE values indicating lower bias, higher precision, or both.

In the simulation experiments, assesment models were compared via ratios of single-stock to multi-stock MARE statistics for each stock s and parameter θ, i.e.,

$$\Delta(\theta_s) = \frac{MARE_{ss}(\theta_s)}{MARE_{ms}(\theta_s)} - 1, \tag{2.3}$$

where ss and ms represent the MARE values for the single- and multi-stock hierarchical assessment model estimates, respectively. Using this definition, $\Delta(\theta_s) > 0$ occured when the multi-stock assessment model had a lower MARE value, indicating that multi-stock estimates had higher precision, lower bias, or both. Estimation performance for an assessment complex as a whole was indicated by an aggregate MARE ratio $\overline{\Delta}(\theta_s)$ for each stock's

parameter θ_s, i.e.,

$$\overline{\Delta}(\theta) = \frac{\sum_s MARE_{ss}(\theta_s)}{\sum_s MARE_{ms}(\theta_s)} - 1,$$

which allowed us to compare estimation performance of single and multi-stock assessment models over the whole assessment complex.

Interval coverage probability was calculated across reps i within each combination of experimental factors and model configuration. I calculated the realized interval coverage probability under an assumption of normality on the log scale, because all quantities of interest are constrained to be positive, and chose the nominal coverage probability as 50%, with a corresponding z-score of 0.67. These two choices defined our interval coverage probability metric as

$$IC_{50}(\log\theta_s) = \frac{1}{100}\sum_i I(\log\theta \in (\hat{\log}\,\theta_{i,s} - 0.67\hat{se}(\log\theta)_{i,s}, \hat{\log}\,\theta_{i,s} + 0.67\hat{se}(\log\theta)_{i,s})),$$

where I is the indicator function, $\hat{\log}\,\theta_i$ is the model estimate of $\log\theta$ in replicate i, and $\hat{se}(\log\theta)_i$ is the model standard error of $\log\theta$ in replicate i. For a 50% interval coverage, realized rates $IC_{50\%}(\log\theta_s)$ closer to the nominal rate 0.5 are better. The confidence interval is considered conservative when realized coverage rates are above the nominal rate, which could indicate either decreased bias of the parameter estimate or high uncertainty (larger standard errors). On the other hand, the confidence interval is considered permissive when realized rates are below the nominal rate, indicating that the uncertainty may be under-represented by the parameter estimate and its standard error.

Finally, for each parameter the distribution of predictive quantiles was calculated over replicates i, defined as

$$Q(\log\theta_{i,s}) = P(\hat{\log}\theta_{i,s} < \log\theta_{i,s}) = \int_{x=-\infty}^{x=\log\theta_{i,s}} f(x \mid \hat{\log}\theta_{i,s}, \hat{se}(\log\theta)_{i,s})dx,$$

where $f(x|m,s)$ is the normal probability density function with mean m and standard deviation s. The resulting distribution of quantiles is best interpreted graphically, and indicates how well the model is estimating parameter uncertainty. Well performing estimators will have a near-uniform distribution of Q values, because true values should be distributed randomly across the full domain of the parameter's sampling distribution. Estimators that under-represent uncertainty by produce standard errors that are too small and will, therefore, have excess density near $Q = 0$ and $Q = 1$ (i.e a $\bigcup$-shaped graphical distribution), indicating that true values have larger z-scores in the sampling distribution. Models that over-represent uncertainty have standard errors that are too large and will collect density near $Q = .5$ (i.e. a $\bigcap$-shaped graphical distribution), indiciating lower z-scores of true values in the sampling distribution.

An experimental design approach was used for simulation models to analyse the effects of experimental factors and assessment model configurations on the MARE and Δ performance metrics [74]. This method attmpts to simplify the complex response surfaces via a generalized linear meta-model of the response surface to simulation model inputs (i.e. factor levels and assessment model prior configurations) [89]. Meta-models are defined in Appendix 2.7.C.

2.2.4 Assessment for British Columbia Dover sole

All 8 multi-stock assessment model configurations and the single-stock assessment model were fit to the Dover sole data for the three stocks in Figure 2.2. All stocks were initialised in a fished state, beginning in 1984 for the HS stock, and 2003 for both QCS and WCVI stocks.

For $B_{s,MSY}$ and $B_{s,init}$, a prior with mean $m_{B,s} = 20$ and standard deviation $s_{B,s} = 20$ was applied, keeping the relative standard deviation at 100%. For the process error variances, two alternative hypotheses were defined on the strength of environmental effects on population dynamics, modeled as β parameters of the inverse-gamma prior distributions on process error variance terms when using $\alpha_\sigma = 3$. The first hypothesis was $\beta_\sigma = 0.16$, placing the prior mode at around 0.04, favouring process errors with a larger standard deviation around $\sigma = 0.2$. The second was to use $\beta_\sigma = 0.01$, reducing the prior mode to 0.0025, favouring process errors with a small standard deviation around $\sigma = 0.05$.

For each model fit, Akaike's information criterion was calcualted, corrected for the sample size (number of years of survey data) for each stock (AICc) [11]. Selected models were the group of multi-stock configurations that performed the best under both hypotheses according to their AICc values, and estimates of optimal harvest rate $U_{s,MSY}$, terminal biomass $B_{s,T}$, optimal biomass $B_{s,MSY}$, relative biomass $B_{s,T}/B_{s,0}$, and current fishing mortality relative to the optimal harvest rate $U_{s,T}/U_{s,MSY}$, as well as standard errors for all estimates. The sum of single-stock AICc values was used to compare to the complex aggregate AICc score for comparing single-stock and multi-stock model fits. While this may be a slight deviation in use of the AIC, I believe it is useful and satisfies the restrictions of the AICc, i.e., the collection of single-stock models is fit to the same data as the multi-stock models, and the process of adding AICc values is analogous to adding single-stock model log-likelihood values within a joint likelihood.

2.3 Results

Experimental results are presented for stock $s = 1$, a low information stock if $L > 0$ in the information scenarios, and identical to the remaining stocks otherwise. First, meta-model effects on MARE ratios $\Delta(\theta_s)$ and complex aggregate $\overline{\Delta}(\theta)$ are used to interpret model configuration effects, and the remaining metrics help to interpret factor effects.

2.3.1 Single-stock versus multi-stock assessments of the base operating model

As expected, shrinkage effects from hierarchical multi-stock assessment models often improved precision of key management parameter relative errors from multi-stock models compared to single-stock models, when fit to data from the base operating model (Figure 2.4). Although this pattern extended across most model configurations and variables, the effect was most noticeable for optimal harvest rate U_{MSY} and optimal biomass B_{MSY}, and weakest for absolute B_T and relative B_T/B_0 terminal biomass. Also, the effects of hierarchical priors were most noticeable for parameters that were subject to those priors, i.e. catchability had larger increases in precision under a model configurations that estimated a shared prior on catchability (Figure 2.4, q_1, q_2 under the q AM configuration).

Estimator bias was less sensitive to hierarchical multi-stock configurations, with sometimes very subtle effects. For example, optimal harvest rate U_{MSY}, optimal biomass B_{MSY}, and survey 1 catchability q_1 estimates were all relatively unbiased under the single-stock model, and all multi-stock model configurations had a negligible effect on the bias (Figure 2.4). In contrast, survey 2 catchability q_2, and absolute and relative terminal biomass B_T and B_T/B_0 were biased under the single-stock model, so were themselves very sensitive. As with precision, the bias of catchability q_2 was most reduced by the q and q/U_{MSY} configurations, and these improvements translated directly into reductions in absolute bias of the terminal biomass estimates B_T and B_T/B_0.

The other performance metrics indicated that the q and q/U_{MSY} configurations performed similarly under the base operating model. For the management parameters most useful in setting harvest advice, productivity U_{MSY} and current biomass $B_T, B_T/B_0$, the q/U_{MSY} configuration either improved all metrics, or kept metrics within a tolerable level of the ideal (Figure 2.5), e.g. interval coverage fell for U_{MSY}, but remained within 10% of the nominal level. Similarly, predictive quantile $Q(\theta)$ distributions were slightly more uniform under the q/U_{MSY} configuration than the single-stock model, indicating an improvement in estimator precision and bias, however the difference between q and q/U_{MSY} configurations was subtle.

Increased precision in catchability and biomass parameters under hierarchical multi-stock models was not always a benefit. Under a single simulation replicate, 95% confidence intervals of biomass estimates from joint models were generally more precise than single-stock estimates; however, increased precision occasionally created estimates that were over-precise, leaving true biomass values outside confidence intervals (Figure 2.6, Stock 2, q and Q/U_{MSY} models). Furthermore, hierarchical estimation appeared to falsely detect an increasing trend in biomass, where the single-stock model was more conservative (Figure 2.5, Stock 2), but corrected the same behaviour in the single-stock model for a different stock in the same complex (Figure 2.5, Stock 1).

2.3.2 Simulation Experiment Results

Model configuration effects

When comparing MARE values through the Δ metric, multi-stock model configurations that estimated the shared prior on survey catchability, denoted q and q/U_{MSY}, stood out as the most beneficial for parameters of the low data quality stocks (stock $s = 1$). Both of these configurations increased Δ values, or had effects that were within 1 standard error of zero (Table 2.4, Stock 1 Δ values), indicating that multi-stock model configurations produced MARE values at most equal to those produced by single-stock models.

As under the base operating model, according to the Δ metric the best performing hierarchical multi-stock model for providing harvest advice was q/U_{MSY}. Closer inspection of β_q and $\beta_{U_{MSY}}$ values indicated that estimation of the mean optimal harvest rate reduced the larger benefit to catchability in both surveys $q_{1,1}, q_{2,1}$ and optimal biomass $B_{1,MSY}$ (Table 2.4, β_q and $\beta_{q,U_{MSY}}$). On the other hand, while the U_{MSY} prior had no effect on terminal biomass ($\Delta(B_T)$), the effects on relative biomass $\Delta(B_T/B_0)$ were nearly tripled over the reference level β_0. The Δ values for optimal biomass B_{MSY} and catchability parameters were lower, but these parameters are not particularly critical for providing harvest advice.

The q and q/U_{MSY} configurations stood out at the complex level also, with higher meta-model coefficients than the U_{MSY} configuration (Table 2.4, Complex Aggregate $\overline{\Delta}$ Values). Under the aggregate MARE ratio $\overline{\Delta}$, it was more difficult to separate the two best models as the meta-model coefficients for both q and q/U_{MSY} were closer together, e.g. $\overline{\Delta}(B_T)$, and there was a reduction in $\overline{\Delta}(U_{MSY})$ under the q/U_{MSY} configuration. Unlike the stock-specific Δ values, the prior configuration had an effect on the $\overline{\Delta}(U_{MSY})$ response in the aggregate, where the q/U_{MSY} configuration produced the biggest reduction $\overline{\Delta}(U_{MSY})$. On the other hand, the largest increase over the null model reference level was also produced by the q/U_{MSY} configuration for the $\overline{\Delta}(B_T/B_0)$ response, indicating a tradeoff between estimates of stock status and productivity.

The U_{MSY} configuration tended to perform the worst according to the Δ metric. I expected to see a benefit to productivity parameter estimates but I were surprised to find there was no benefit to a low data quality stock. Moreover, meta-model coefficients for Δ and $\overline{\Delta}$ response variables were consistently smaller than the other configurations, and often negative or insignificant.

Factor effects

As expected, the effects of shrinkage were most beneficial under low-information scenarios, according to the Δ metrics. When the biomass was initialized in a fished state, Δ and $\overline{\Delta}$ values increased (Table 2.4, $\beta_{d_{s,1}} < 0$). Similarly, there were significant increases in Δ and $\overline{\Delta}$ values for all parameters when the assessments were initialized at the beginning of survey 2 (Table 2.4, $\beta_{T_1} > 0$). These improvements under low information conditions are largely

driven by a stabilising effect of shrinkage. That is, single-stock models produced relatively larger MARE values as data data quality was reduced. Under the same conditions, the hierarchical multi-stock models were restricted from increasing MARE values as fast by shrinkage (Table 2.4).

I found that the q and q/U_{MSY} configurations were sensitive to data quality and the choice of performance measure. For example, under a 1-way trip fishing history with 4 identical stocks (Figure 2.7), the q configuration eliminatedd bias in U_{MSY} and improved interval coverage from 62% to 56%, correcting an under-precise estimator. In contrast, the q/U_{MSY} configuration was over-precise, indicated by an interval coverage of 33% and the quantile distribution becoming slightly $\bigcup$-shaped, and also increased bias in U_{MSY} estimates (Figure 2.7, U_{MSY}).

On the other hand, the q/U_{MSY} configuration appeared to perform better under a 2-way trip fishing history, a short time series, and fished initialisation. The q/U_{MSY} configuration reduced bias for relative biomass B_t/B_0 and almost eliminated bias for U_{MSY} (Figure 2.8, U_{MSY}). Interval coverage also improved under the q/U_{MSY} configuration for terminal biomass estimates B_T and B_T/B_0, coming closer to the nominal rate of 50%. Although the U_{MSY} interval coverage fell to 36% under the q/U_{MSY} configuration, indicating an over-precise estimator, I viewed this as favourable compared to the q configuration, where U_{MSY} was under-precise by a similar amount, yet remained positively biased.

The effect of complex size S and the number of low information stocks L interacted in unexpected ways. According to the selected meta-model, the size of the complex S and the number of low information stocks L appeared to have little effect on response values. Indeed, all β_S and β_L effects on Δ and $\overline{\Delta}$ values were at most 0.09 in magnitude, if they were included at all. These weak effects indicated that the linear meta-model is probably too simple for these factors (Figure 2.9). Increasing the number of low-information stocks L was always an improvement for Δ values when moving from $L = 0$ to $L = 1$. This was was expected given that the Δ values were calculated for stock $s = 1$ (a data poor stock if $L > 0$), and I expected that multi-stock models and single-stock models would have similar estimates when fit to complexes of data-rich stocks. Beyond $L = 1$ any improvements in MARE values were dependent on the size of the complex. Generally, it appeared that keeping the number of low information stocks under half of the complex size, i.e. $L < S/2$, preserved the most benefit in terms of precision, though this pattern reversed for $L = 3$ and $S = 4$. Complex aggregate $\overline{\Delta}$ values were comparatively flatter in response to the levels of L. I didn't produce response surfaces for other factor combinations as these factors all had 2 levels each, meaning that a linear model should capture the average behaviour.

2.3.3 Assessments of British Columbia Dover sole

Multi-stock models defined by shared catchability q and shared catchability and optimal harvest rate configurations q/U_{MSY} performed best for the British Columbia Dover sole

complex based on AICc values. These same configurations also performed best in in the simulation experiments. The U_{MSY} configuration and the null model both had AICc scores more than 500 points higher than the best performing multi-stock configuration. The selected multi-stock models gave AICc scores between 100 and 200 units below the total single-stock model scores under both hypotheses (Table 2.5, AICc), indicating that the increase in estimated parameters was justified. All models had lower AICc values under the assumption of low process error variance.

Hierarchical multi-stock models reduced parameter uncertainties when compared to single-stock models. Multi-stock models with shared priors produced lower cofficients of variation, defined as $CV = \sqrt{e^{se^2} - 1}$, for estimates of optimal biomass and productivity parameters, reducing coefficients of variation below 100% in some cases (single-stock vs multi-stock models in Table 2.4). Similar reductions in uncertainty are visible in reconstructions of stock biomass time series (Figure 2.10).

Assessments of the Dover sole complex were qualitatively similar between model configurations and hypotheses. The major differences between assessment model configurations were the level of uncertainty in parameter estimates, and the scale of each individual stock's biomass, but the trends over time were the same (Figure 2.10). The Hecate Strait (HS) stock showed increasing biomass since 1984, with more or less process variation depending on the configuration and variance hypothesis (Figure S5). The Queen Charlotte Sound stock showed an initial depletion with increased landings between 2003 and 2006, followed by some growth that has continued until present day. Finally, the West Coast of Vancouver Island (WCVI) stock showed a flat biomass trend following initial depletion from 2003 to 2006. The flat trend in the WCVI stock may indicate that fishing was balancing annual production.

The multi-stock assessment model configuration q/U_{MSY} generally estimated all stocks as smaller and more productive than other assessments (Table 2.5). This was most noticable for the QCS stock biomass estimates by multi-stock models, where the single-stock model considered the optimal biomass to be close to 18 kt, with a terminal relative biomass between 7% and 13%, in contrast to the selected multi-stock configurations, where optimal biomass was between 3 kt and 6 kt, with a current relative biomass between 95% and 110%. Under the single-stock model configuration, the biomass scales corresponded to expected catchability values of $q_{2,HS} = 0.10$, $q_{2,QCS} = 0.74$ and $q_{2,WCVI} = 0.16$. I considered this distribution of catchability values between stocks of the same species unlikely, given that the biomass indices are relative biomass values and catchability corresponded to trawl efficiency. It was more likely that the single-stock assessment reduced the biomass parameter estimates for the QCS stock because of the fished initialisation in 2003. Starting in this state removed any depletion signal from the earlier catch history, and allowing the model to explain the stock indices catch with a smaller biomass.

No selected multi-stock model indicated that Dover sole stocks were overfished or experiencing overfishing, however, the uncertainty in relative terminal biomass and harvest rate was often very high. That is, current relative biomass estimates were always at least 60% of unfished, but their coefficients of variation were in some cases above 50% of the mean estimate (Table 2.5). Similarly, although relative harvest rate estimates were all at most 70% of the optimal harvest rate (Table 2.5), their coefficients of variation were at least 65%, and sometimes greater than 100%, of the mean estimate for each stock under some model configurations, most often under the high variance assumption.

The q/U_{MSY} hierarchical multi-stock model configuration had the best fit to the data, which is not surprising given that the Dover sole complex closely matches the scenario shown in Figure 2.8, with a fished initialisation and 2 stocks having short time-series of observations. Under those simulation experiments, the q/U_{MSY} configuration was considered over-precise, but essentially unbiased, for U_{MSY} estimates. In contrast, for assessments of Dover sole data with low process error variance, the precision seems be lower under the q/U_{MSY} configuration, indicated by larger coefficients of variation (Table 2.5).

2.4 Discussion

The simulation study results indicate that, as expected, shrinkage effects in hierarchical multi-stock assessment models are most beneficial when some data sets have low statistical power. Furthermore, both configurations that estimated a shared catchability prior performed best for estimating key management parameters. On the other hand, shrinkage does not always improve stock assessment performance relative to a single-stock approach. In particular, the benefits of joint estimation depend on several factors, including the information content of the data, the choices for hierarchical model priors, and the particular management parameters of interest.

Model configurations that shared prior distributions on survey catchability (i.e., the configurations q and q/U_{MSY}) stood out as the best options for improving parameter estimates for stocks with low data quality. This result may occur because catchability is a linear parameter within the assessment, while optimal harvest rate parameters are embedded within non-linear popoulation dynamics. Although this hypothesis does not explain how different configurations increase or reduce bias and precision, it may provide a template to guide expectations and generate hypotheses when testing other hierarchical model behaviour.

I found that simply adding a joint likelihood can have positive effects, which was surprising because there should be no mathematical difference between optimising a set of single-stock models independently vs binding them in a joint model by simply adding their negative log likelihoods together. This result may indicate a stabilising effect from the joint likelihood, where simply including data-rich species without shared priors improves the numerical performance of minimisation algorithm.

There was mixed evidence that increasing the size of the assessment complex produced better results under hierarchical multi-stock models. For instance, in the lower information scenarios, the effect of the complex size depended on the number of low-information stocks present in the system. The most benefit for the first stock $s = 1$ was realized when moving from no low information stocks ($L = 0$) to one low information stock ($L = 1$). This is counter-intuitive, as decreasing information should reduce precision, but represents the stability induced by the shrinkage from the multi-stock models. Looking at response surfaces averaged over all factor levels and configurations, I found that complexes of size $S = 7$ provided the most stable benefit (in terms of MARE values) for different numbers of low information stocks L; however, I weren't testing for an optimal size, which would require a new design with a finer resolution on L and S factors.

Some of our results may be caused by a discrepancy between the underlying assumption of normality for parameter distributions used in the Laplace approximation to the integrated likelihood and the true parameter distribution [76]. Despite the integrated likelihood, the approximation by a normal distribution means that there is potential for bias caused by disagreement between the modes of the assumed normal distribution and true parameter distribution [136].

Although I investigated a single-species, multi-stock complex, where stocks represented biologically identical management units within the Dover sole fishery, the hierarchical multi-stock approach could be extended to a multi-species approach by simulating stocks with different biological parameters $B_{s,0}$ and r_s. I suspect that differences in unfished biomass $B_{s,0}$ would not have a strong effect on overall performance. In a Schaefer model context, the unfished biomass parameter determines the absolute scale at which the dynamics operate, but has little effect on the dynamics themselves. Density dependence in annual production is driven by this parameter, but that effect is independent of absolute biomass and relies, instead, on the relative biomass B_t/B_0. In contrast, differences among intrinsic growth rates may improve estimates in assessment models that estimate shared productivity priors. More productive stocks would grow faster when fishing pressure is reduced, reducing uncertainty in productivity estimates for those stocks. Stocks with more precise estimates may then have a dominating effect on the hierarchical prior, improving hierarchical assessments but potentially biasing estimates of weaker stock productivities [124].

Multi-species extensions to the framework in this chapter may also provide deeper insights. For example, introducing age-structured population dynamics [48], or a delay-difference formulation [128], would differentiate multiple species further than a simple Schaefer model by allowing for different maturation delays, growth rates, and recruitment dynamics to affect stock production. If biological data were unavailable for informing life-history parameter estimates under more realistic population dynamics, meta-analyses of Beverton-Holt life history invariants within family groups could provide informative prior distributions [97]. Indeed, recent meta-analyses have shown that publically available data-bases of life his-

tory parameters can be useful for this type of application [139]. Similar meta-analyses of the same data-bases, comparing species that are evolutionarily related, improves the utility of life history invariants by estimating different ratios within taxa, improving their utility as informative priors and potentially providing inverse-gamma priors on hierarchical variance terms in the form of evolutionary covariance estimates [140].

Several simplifying assumptions were made about population dynamics for simplicity in design and interpretation. In addition to assuming that biological parameters are the same for stocks within the complex, I assumed fishing pressure was identical among stocks, and the magnitude of species-specific effects was identical. The choice of identical biology removed a "stock-effect" on management parameter estimates, as discussed above for productivity. With different biological parameters, the ability to identify hierarchical estimator effects may be reduced due to confounding with stock effects. Next, subjecting stocks to identical fishing pressure simplifed the generation of assessment data. Simplifying the simulations in this way may have increased the correlation between stocks, improving performance of the hierarchical multi-stock estimators relative to more realistic situations. For example, it would be more realistic to link fishing mortality to fishing effort through a stock-specific fishery catchability.

Other simplifying assumptions were made related to process errors in the assessment model. Identical standard deviations were assumed, which matched the simulated dynamics, and both correlation in $\zeta_{,t}$ process errors and the shared year effect $\bar{\epsilon}_t$ were not estimated, despite simulating these effects. Correlation was ignored to increase stability in simulation trials, as estimating the correlation matrices often produced nonsensical results. It may be possible to improve correlation estimates by applying an inverse Wishart prior for the full estimated covariance matrix, but this was not inside the scope of this chapter. Shared year effects were also not estimated, as it became clear that it was confounded with the individual process errors, and there was no benefit to partitioning the variance across an extra process error term. Adding another data stream, such as an environmental index [80], or forcing the year effects to resemble a periodic or trend-zero behaviour [143], may improve these estimates in other studies.

I did not conduct sensitivity analyses of the hyperpriors. Intuitively, I expect that more precise inverse-gamma hyperpriors on estimated variance parameters would increase the shrinkage effect, and thereby clustering stock-specific estimates closer to a biased mean value. Instead of focusing on the behaviour induced by hyperprior settings, I chose instead to focus on the behaviour induced by defining the shared priors, and left the hyperpriors on prior means sufficiently vague to emulate the true prior knowledge about the Dover sole complex, and on prior variances sufficiently informative to encourage a shrinkage effect.

Fitting the hierarchical multi-stock surplus production models assessment to Dover sole data showed that shrinkage effects carried over to a real system. Shrinkage effects reduced uncertainty when data had low statistical power, and provided more realistic estimates of

catchability parameters than single-stock models, especially for the Queen Charlotte Sound stock. While the resulting estimates were sometimes quite uncertain, and a full assessment would require more scrutiny or a different model structure than is provided here, my results indicate that all three Dover sole stocks are likely in a healthy state given recent rates of exploitation.

The results confirm that hierarchical multi-stock production models are a feasible data-limited approach to stock assessment in multi-stock fisheries. Under low statistical power conditions, hierarchical multi-stock assessment modeling is preferable to data-pooling approaches for at least two reasons. First, hierarchical multi-stock models are able to produce stock-specific estimates that allow management decisions to be made at a higher spatial resolution and based on data rather than strong *a priori* assumptions or management parameter values averaged over stocks. Despite the potential for bias under low-power conditions, stock-specific estimates of key management parameters can provide meaningful and important feedback in the fishery management system. Second, using a hierarchical multi-stock method ensures that an assessment framework is readily available for more and better data, making it much easier to update model estimates later when more data is available. Moreover, the type of additional data to be collected could be prioritized by examining the standard errors for observation model components of the hierarchical multi-stock assessment models, where higher uncertainty may indicate a better return on investments in improved monitoring.

The feasibility of hierarchical multi-stock surplus production models relies on catch and effort data being available, but I consider hierarchical multi-stock production models as an important bridge between catch-only methods and more data-intensive methods. For instance, some catch only methods require restrictive *a priori* assumptions, such as an estimate of relative biomass as a model input [77, 37]. More recently, a multi-species assessment method was derived that removes the need for relative biomass estimates, but requires restrictive assumptions about fishery-dependent catchability and that all species are initially in an unfished state [14]. Our approach avoids all of these assumptions. For instance, (i) joint model estimates of relative biomass were stable in practice, and in simulations despite absence of a current relative biomass estimate (or assumption); (ii) hierarchical multi-stock models have better precision when initialized in fished states; and (iii) fishery catchability assumptions are not required. Thus, while the data needs are higher for our approach, the potential applications are broader in scope.

On the other hand, hierarchical multi-stock models should be scrutinized closely via standard assessment performance measures (e.g., retrospective analysis) before application to real management systems. In particular, I found that shrinkage can have unexpected non-linear side-effects. Closed-loop simulations would be needed to determine the long-term implications of these types of errors on multi-stock harvest management systems [113].

2.5 Tables

Table 2.1: Operating model parameters and their values

Description	Symbol	Value
Unfished Biomass	$B_{s,o}$	40kt
Intrinsic Rate of Growth	r_s	0.16
Shared Process Error SD	κ	0.071
Stock-specific Process Error SD	σ_s	0.071
Simulation Historical Period	$(T_{init}, \ldots, T)$	$(1984, \ldots, 2016)$

Table 2.2: Experimental factors and their levels

Description	Levels	Notes
Fishing History	1-way, 2-way trips	Low/High contrast in biomass
Complex Size, S	4,7,10	
Low data quality stocks, L	0,1,2,3	
Initial Assessment Year 1984, 2003	Short or long series of observations ($t = 1$ or $t = 20$ of $T = 34$ years)	
Initial Relative Depletion	0.4, 0.7, 1.0	Fished or unfished initialisation

Table 2.3: Multi- and single level priors used in the assessment model.

No.	Distribution
Survey Catchability	
Multi-level prior	
q.1	$\hat{q}_{o,s} \sim \log N(\log \hat{\bar{q}}_o, \hat{\iota}_o)$
q.2	$\hat{\bar{q}}_o \sim N(m_q, s_q)$
q.3	$\hat{\iota}_o^2 \sim IG(\alpha_q, \beta_q)$
Single level prior	
q.4	$\hat{q}_{o,s} \sim N(m_q, s_q)$
Optimal Harvest Rate	
Multi-level prior	
U.1	$\hat{U}_{s,MSY} \sim \log N(\log \hat{\bar{U}}_{MSY}, \hat{\sigma}_U)$
U.2	$\hat{\bar{U}}_{MSY} \sim N(m_U, s_U)$
U.3	$\hat{\sigma}_U^2 \sim IG(\alpha_U, \beta_U)$
Single level prior	
U.4	$\hat{U}_{s,MSY} \sim N(m_U, s_U)$

Table 2.4: (Continued on following page)

Response	Ref Level	Prior Configuration			Experimental Factor				
					Init. Dep	Init. Assessment	Low Data Stocks	Complex Size	Fishing History
	β_0	β_q	$\beta_{U_{MSY}}$	$\beta_{q/U_{MSY}}$	$\beta_{d_{1,1}}$	β_{T_1}	β_L	β_S	β_U
Low Data Quality Stock ($s = 1$) Δ Values									
$\Delta(U_{1,MSY})$	0.60 (0.07)	0.25 (0.09)	0.04 (0.09)	0.09 (0.09)	-0.14 (0.04)	0.35 (0.04)	-	-	-
$\Delta(B_{1,T})$	-0.01 (0.04)	0.28 (0.05)	-0.02 (0.05)	0.28 (0.05)	-0.04 (0.02)	0.08 (0.02)	-	-	0.11 (0.02)
$\Delta(B_{1,MSY})$	0.16 (0.03)	0.10 (0.04)	-0.07 (0.04)	0.02 (0.04)	-0.06 (0.02)	0.11 (0.02)	0.06 (0.02)	-	-0.02 (0.01)
$\Delta(B_{1,T}/B_{1,0})$	0.32 (0.07)	0.29 (0.10)	0.13 (0.10)	0.63 (0.10)	-0.09 (0.04)	0.30 (0.04)	-	0.09 (0.04)	0.14 (0.03)
$\Delta(q_{1,1})$	0.06 (0.05)	0.46 (0.07)	-0.01 (0.07)	0.27 (0.07)	-	0.15 (0.03)	-	-	0.06 (0.03)
$\Delta(q_{2,1})$	-0.02 (0.02)	0.23 (0.03)	-0.05 (0.03)	0.10 (0.03)	-	-	0.03 (0.02)	-0.04 (0.01)	0.08 (0.01)
Complex Aggregate $\overline{\Delta}$ Values									
$\overline{\Delta}(U_{MSY})$	0.47 (0.04)	0.13 (0.04)	-0.07 (0.04)	-0.07 (0.04)	-0.10 (0.03)	0.21 (0.03)	0.04 (0.03)	-	-0.03 (0.02)
$\overline{\Delta}(B_T)$	0.04 (0.02)	0.22 (0.02)	-0.03 (0.02)	0.21 (0.02)	-0.05 (0.01)	0.11 (0.01)	-0.03 (0.01)	0.02 (0.01)	0.07 (0.01)
$\overline{\Delta}(B_{MSY})$	0.11 (0.02)	0.11 (0.02)	-0.05 (0.02)	0.01 (0.02)	-0.07 (0.01)	0.09 (0.02)	-	-	0.03 (0.01)
$\overline{\Delta}(B_T/B_0)$	0.31 (0.04)	0.25 (0.04)	0.03 (0.04)	0.39 (0.04)	-0.08 (0.02)	0.24 (0.03)	-	-	0.06 (0.01)
$\overline{\Delta}(q_1)$	0.08 (0.03)	0.31 (0.03)	-0.02 (0.03)	0.21 (0.03)	-0.05 (0.02)	0.15 (0.02)	-	-	0.08 (0.01)
$\overline{\Delta}(q_2)$	-0.06 (0.02)	0.26 (0.02)	-0.04 (0.02)	0.15 (0.02)	-0.03 (0.01)	-	-	-0.02 (0.01)	0.07 (0.01)
Single-Stock Assessment MARE values									
$U_{1,MSY}$	40.52 (1.22)	-	-	-	-6.64 (1.49)	4.44 (1.21)	3.90 (1.66)	-	-9.00 (1.08)
$B_{1,T}$	29.01 (0.56)	-	-	-	-0.96 (0.64)	2.62 (0.54)	-	1.01 (0.62)	2.65 (0.51)
$B_{1,MSY}$	26.61 (0.49)	-	-	-	-5.56 (0.60)	3.67 (0.49)	3.11 (0.67)	-0.77 (0.54)	-
$B_{1,T}/B_{1,0}$	56.13 (1.97)	-	-	-	-11.68 (2.28)	17.71 (1.93)	-	-	14.34 (1.81)
$q_{1,1}$	19.58 (0.44)	-	-	-	-	3.46 (0.44)	-	-	-
$q_{2,1}$	17.97 (0.41)	-	-	-	-	0.59 (0.41)	-	-0.94 (0.49)	-1.00 (0.40)
Multi-Stock Assessment MARE values									
U_{MSY}	24.96 (0.87)	-3.54 (1.20)	0.55 (1.20)	-0.13 (1.20)	-1.70 (0.59)	-0.88 (0.47)	1.13 (0.65)	-0.78 (0.52)	-5.14 (0.42)
B_T	29.10 (0.76)	-6.22 (1.06)	0.67 (1.06)	-5.80 (1.06)	-	0.64 (0.39)	-	0.76 (0.46)	-
B_{MSY}	22.85 (0.78)	-1.85 (1.06)	1.28 (1.06)	-0.32 (1.06)	-4.23 (0.52)	1.28 (0.42)	1.90 (0.58)	-	-
B_T/B_0	40.65 (1.87)	-8.77 (2.56)	-5.02 (2.56)	-13.88 (2.56)	-5.47 (1.26)	4.24 (1.02)	2.49 (1.40)	-1.67 (1.12)	6.22 (0.91)
q_1	18.51 (0.75)	-5.39 (1.04)	-0.06 (1.04)	-3.33 (1.04)	0.69 (0.46)	1.64 (0.39)	-	-	-1.37 (0.37)
q_1	18.51 (0.81)	-3.73 (1.15)	1.24 (1.15)	-1.54 (1.15)	-	-	-	-	-2.58 (0.41)

Table 2.4: (Previous page.) Meta-model coefficients for multi-stock assessment model prior configurations (columns 3-5) and experimental factors (cols 6-10). Response variables are $\Delta(\theta_s) = \frac{MARE_{MS}(\theta_s)}{MARE_{SS}(\theta_s)} - 1$ values for stock $s = 1$ (rows 1-6), complex aggregate $\overline{\Delta}(\theta) = \frac{\sum_s MARE_{MS}(\theta_s)}{\sum_s MARE_{SS}(\theta_s)} - 1$ values (rows 7-12), single stock assessment MARE values for stock 1 (rows 13-18), and multi-stock model MARE values for stock 1 (rows 19 - 24). The intercept (col 2) is the average value of the response across all factors, and represents the null model configuration in rows 1-12 and 19-24. Coefficients of multi-stock model prior configurations independently give the average contribution of that configuration to the response value, while coefficients for experimental factors are calculated based on rescaling factors to the interval $[-1, 1]$. This means the contribution of each factor to the response is equal to its coefficient at the maximum factor value, and the negative value of its coefficient at the minimum factor value. Response values are found by summing across the rows, *taking only one prior configuration coefficient*, and scaling factor coefficients as necessary.

Table 2.5: Selected management parameter mean estimates, their coefficients of variation in parentheses, and corrected Akaike's Information Criterion (AICc) values for selected stock assessments applied to the real dover sole data under the High and Low process error variance hypotheses. Model labels for multi-stock models indicate the shared priors used in the fitting process. Total AICc values for the Single-Stock model are given for direct comparison with the multi-stock models.

	High Process Error Variance					**Low Process Error Variance**			
Model Config	HS	QCS	WCVI	Total	Model Config	HS	QCS	WCVI	Total
U_{MSY}									
Single-Stock	0.147 (0.69)	0.066 (1.14)	0.122 (0.94)	-	Single-Stock	0.113 (0.77)	0.100 (0.71)	0.136 (0.84)	-
q	0.127 (0.64)	0.092 (0.88)	0.095 (0.90)	-	q	0.115 (0.63)	0.097 (0.83)	0.104 (0.83)	-
q/U_{MSY}	0.205 (0.64)	0.191 (0.73)	0.214 (0.78)	-	q/U_{MSY}	0.156 (0.76)	0.151 (0.74)	0.170 (0.86)	-
B_T									
Single-Stock	33.189 (1.04)	4.841 (1.27)	11.487 (0.89)	-	Single-Stock	29.641 (1.13)	2.868 (0.66)	10.956 (0.83)	-
q	27.112 (0.82)	13.843 (0.85)	13.616 (0.74)	-	q	25.498 (0.79)	9.873 (0.87)	11.618 (0.77)	-
q/U_{MSY}	21.067 (0.91)	11.246 (0.93)	11.685 (0.83)	-	q/U_{MSY}	18.124 (0.96)	7.553 (1.05)	9.457 (0.88)	-
B_T/B_0									
Single-Stock	0.968 (0.67)	0.131 (1.97)	0.718 (0.80)	-	Single-Stock	0.874 (0.48)	0.077 (1.53)	0.700 (0.59)	-
q	0.917 (0.62)	1.091 (0.73)	0.702 (0.92)	-	q	0.842 (0.47)	0.950 (0.47)	0.631 (0.70)	-
q/U_{MSY}	0.932 (0.57)	1.071 (0.56)	0.830 (0.52)	-	q/U_{MSY}	0.878 (0.41)	0.967 (0.41)	0.708 (0.54)	-
U_T/U_{MSY}									
Single-Stock	0.081 (1.03)	0.597 (1.12)	0.520 (1.11)	-	Single-Stock	0.117 (0.83)	0.669 (0.65)	0.488 (0.96)	-
q	0.114 (0.85)	0.151 (1.07)	0.560 (1.04)	-	q	0.134 (0.67)	0.199 (1.03)	0.602 (0.97)	-
q/U_{MSY}	0.091 (0.87)	0.089 (1.00)	0.291 (0.95)	-	q/U_{MSY}	0.139 (0.66)	0.168 (1.01)	0.453 (0.88)	-
B_{MSY}									
Single-Stock	17.143 (0.90)	18.415 (1.53)	8.004 (0.88)	-	Single-Stock	16.951 (1.11)	18.672 (1.48)	7.830 (0.72)	-
q	14.790 (0.73)	6.343 (0.91)	9.693 (0.95)	-	q	15.142 (0.80)	5.196 (0.81)	9.210 (0.72)	-
q/U_{MSY}	11.306 (0.83)	5.250 (0.86)	7.043 (0.75)	-	q/U_{MSY}	10.323 (0.94)	3.904 (0.97)	6.680 (0.78)	-
$AICc$									
Single-Stock	-102.06	-22.487	-23.655	-148.202	Single-Stock	-163.261	-54.035	-56.802	-274.098
q				-169.838	q				-343.312
q/U_{MSY}				-252.292	q/U_{MSY}				-417.676

2.6 Figures

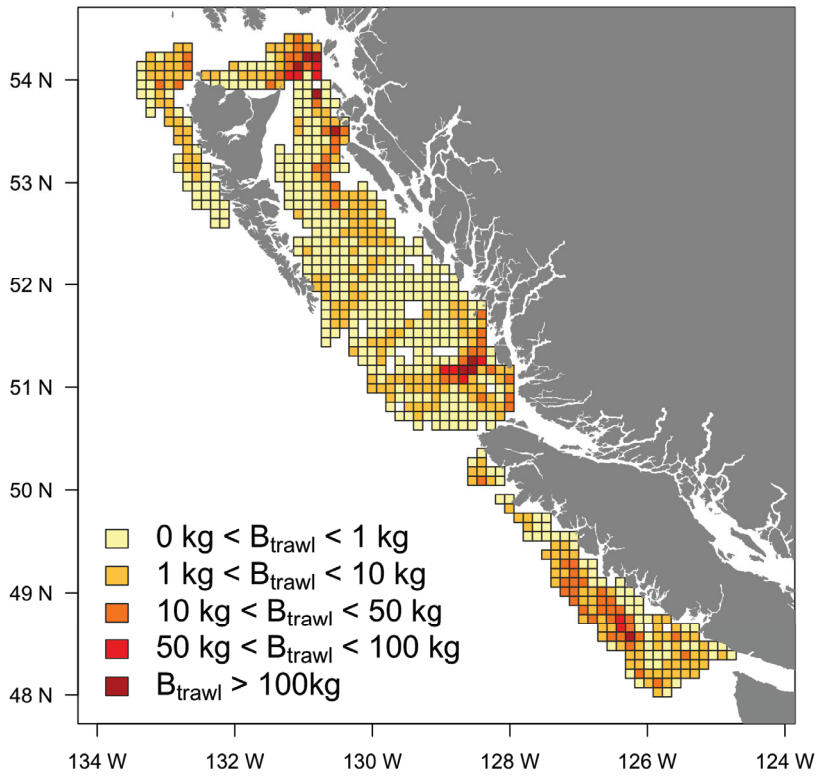

Figure 2.1: Mininum trawlable biomass B_{trawl} estimates for Dover Sole on the BC coast, aggregated to a 10km square grid. Estimates are produced by scaling average trawl survey (kg/m^2) density values in each grid cell by the cell's area in m^2. Locations that do not show a coloured grid cell do not have any survey blocks from which to calculate relative biomass. Survey density data is taken from the GFBio data base maintained at the Pacific Biological Station of Fisheries and Oceans, Canada.

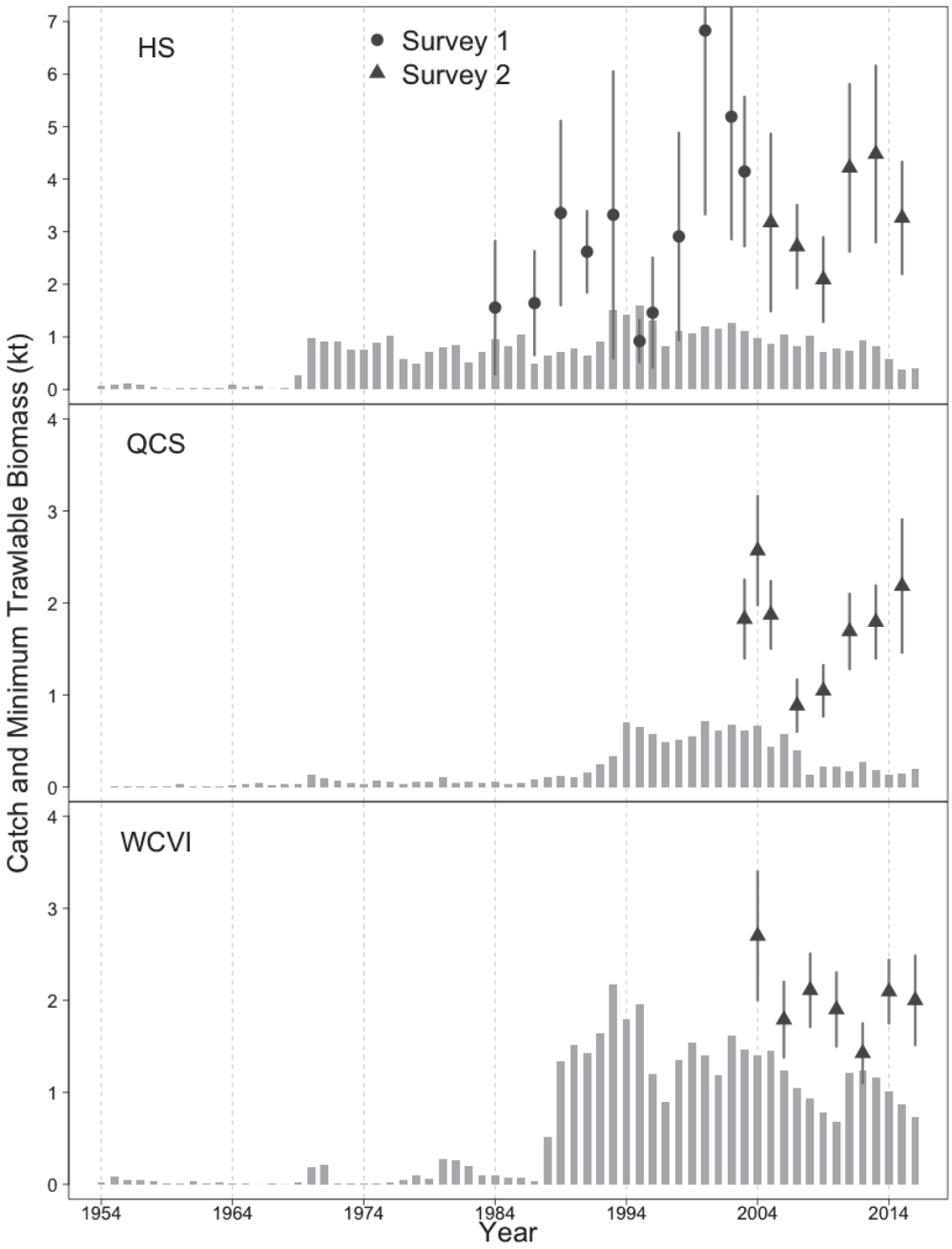

Figure 2.2: Time series of coastwide catch since 1954 (vertical bars) and relative biomass since 1984 (data points) for the three Dover Sole stocks: Haida Gwaii (HG), Queen Charlotte Sound (QCS) and West Coast of Vancouver Island (WCVI). The catch data are taken from the GFcatch, PacHarvTrawl and GFFOS data bases and trawl survey data were obtained from the GFBIO data base. All data bases are maintained at the Pacific Biological Station of Fisheries and Oceans, Canada.

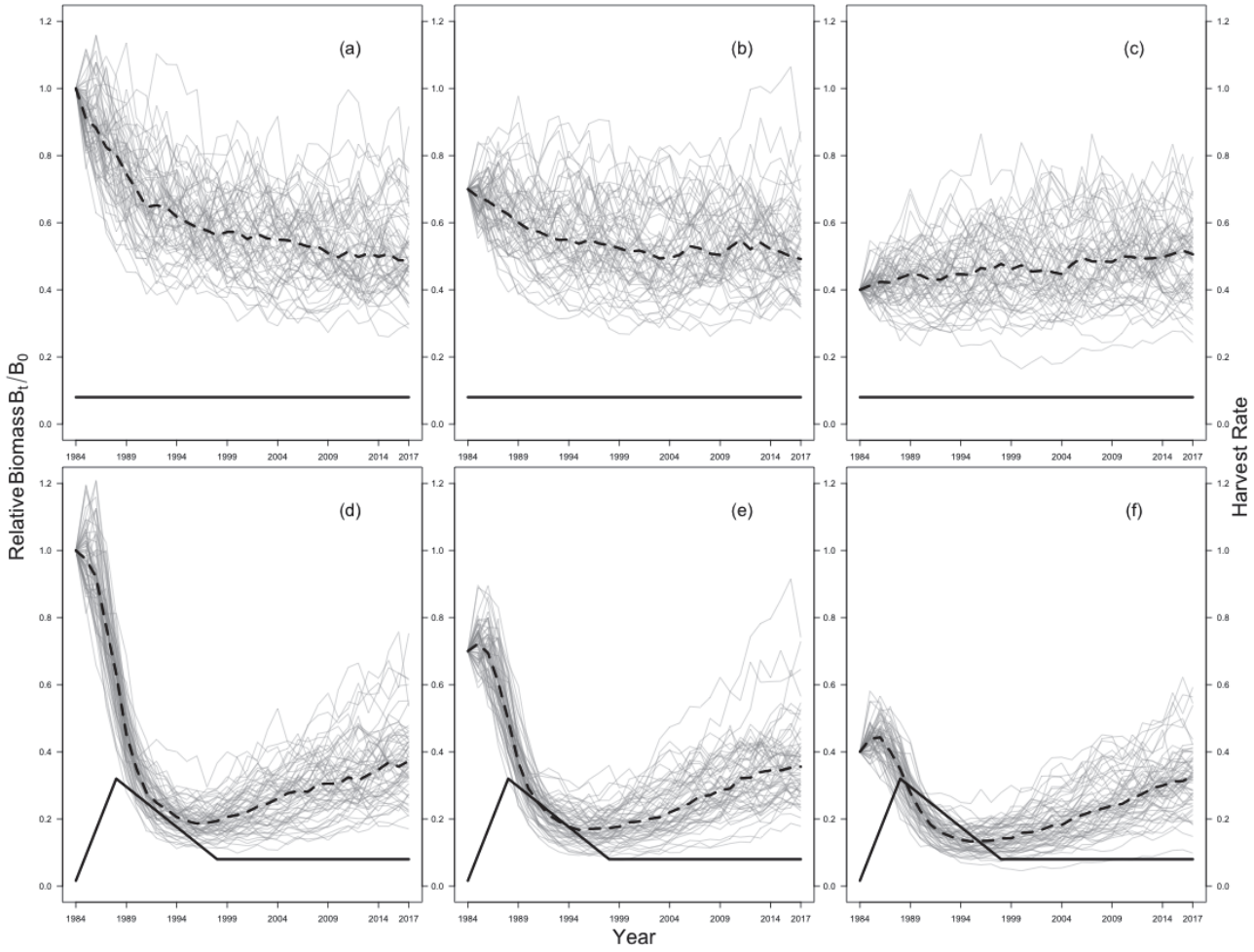

Figure 2.3: Biomass depletion trajectories of 60 random replicates under different historical fishing intensities and initial relative biomass. Plots (a) - (c) show the constant optimal harvest rate fishing history, which result in more one-way trips, and plots (d) - (f) show the two-way trip fishing history. Initial relative biomass of 100% (panels (a), (d)), 70% (panels (b), (e)), and 40% (panels (c), (f)) of B_0 are shown. The grey lines are traces from selected replicates, while the black dashed line is the median time series for those replicates, and the solid black line is the simulated harvest rate.

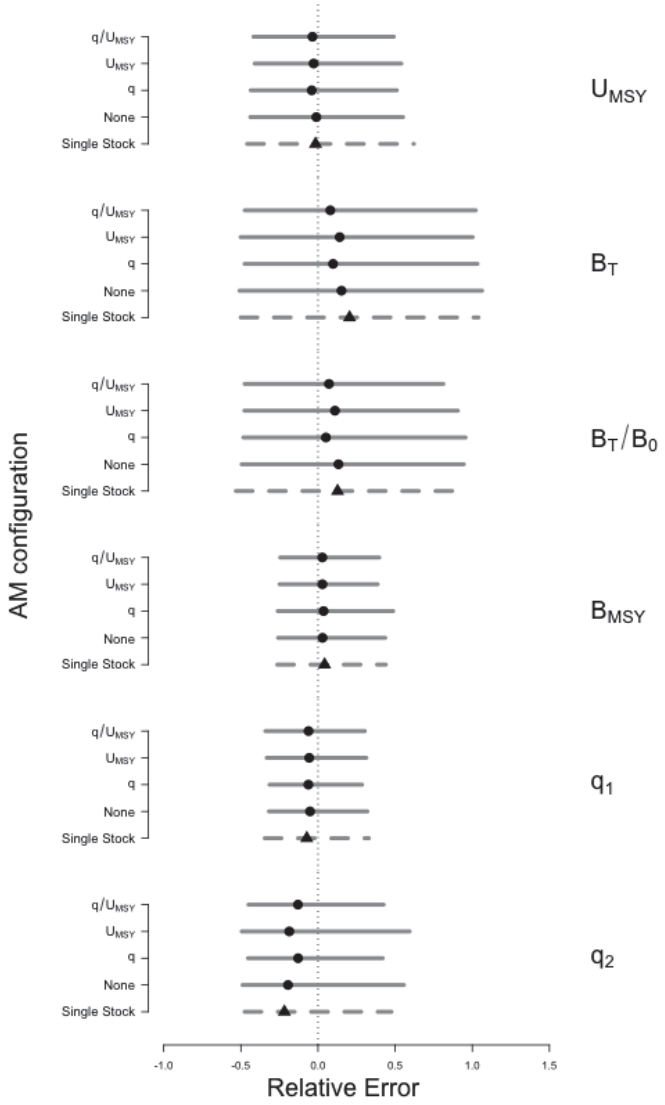

Figure 2.4: Relative error distributions for stock $s = 1$ leading and derived parameters estimated by the single stock (dashed lines and triangular points) and 4 multi-stock assessment models (solid lines and circular points) fit to data from the base operating model. Points indicate median relative errors and the grey lines the central 95% of the relative error distribution. From the top, parameters are optimal exploitation rate (U_{MSY}), terminal biomass (B_T), optimal equilibrium biomass (B_{MSY}), terminal biomass relative to unfished (B_T/B_0), and catchability from surveys 1 (q_1) and 2 (q_2). Assessment model (AM) configurations indicate the single stock model, or the parameters that had hierarchical prior distribution hyperparameters estimated in the multi-stock assessment model (e.g, q/U_{MSY} indicates that shared priors on both catchability and productivity were estimated).

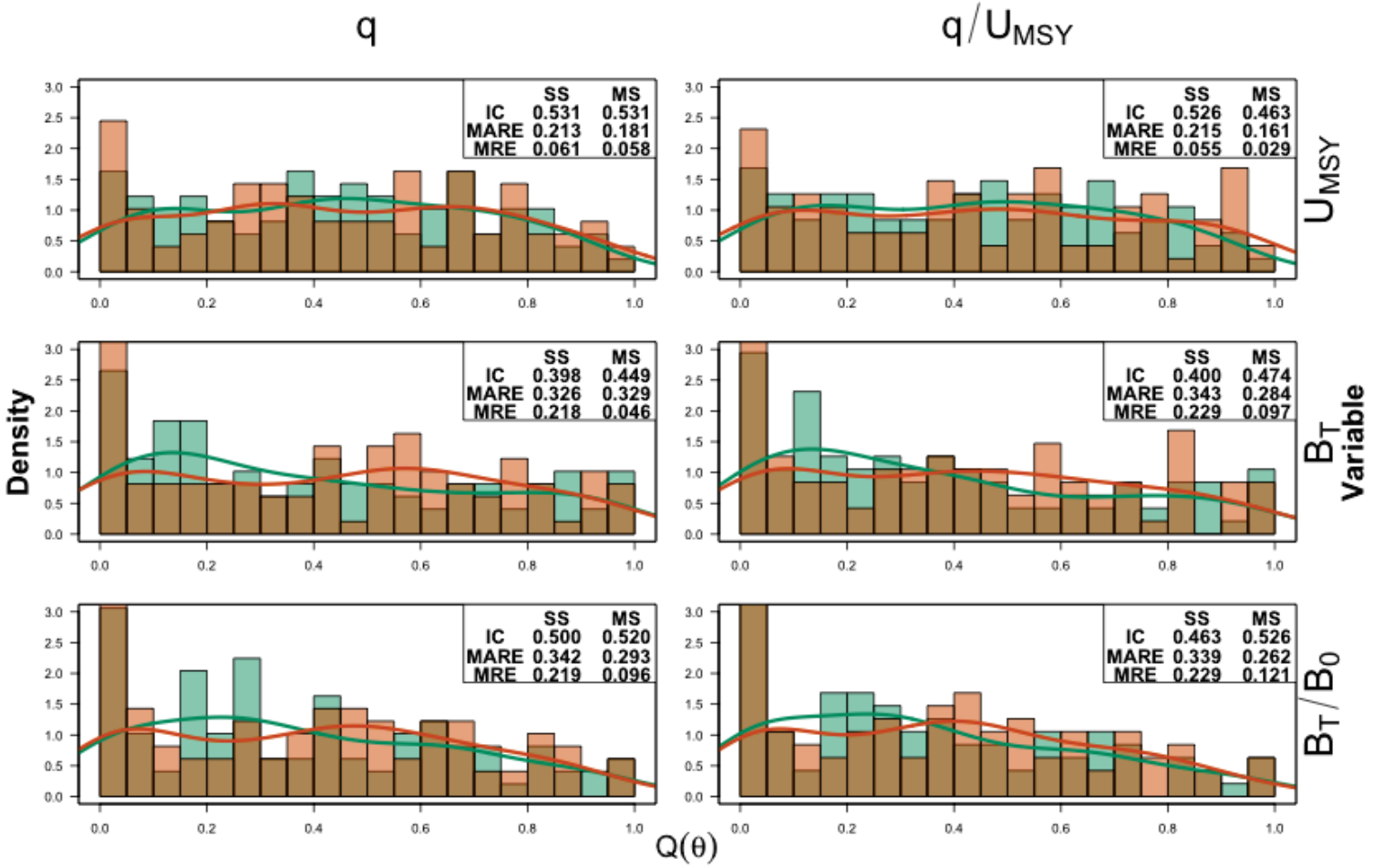

Figure 2.5: Density of predictive quantiles $Q(\theta)$ for estimates of key management parameters (rows) from single stock and q and q/U_{MSY} hierarchical multi-stock model configuration under the base operating model. Bars show probability density of Q distributions, with lines showing the kernel smoothed density for easier comparison between single stock (green) and multi-stock (orange) models. Top right hand corners of each panel show interval coverage (IC), median absolute relative error (MARE), and median relative error (MRE) for single stock (SS) and multi-stock models (MS).

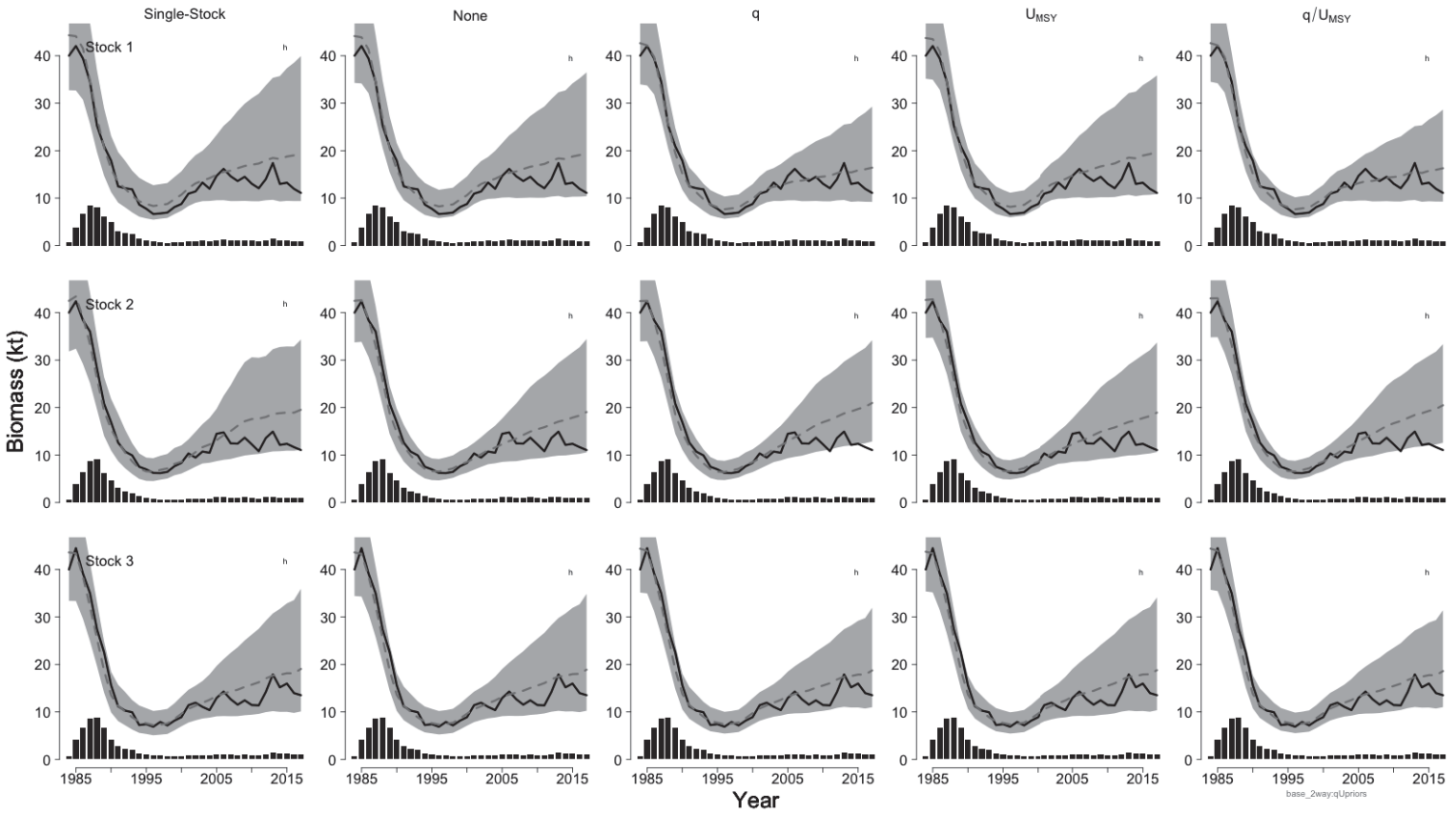

Figure 2.6: Time series of biomass and catch for a 3 stock complex, taken from a single simulation replicate using the base operating model. Thick unbroken lines indicate the simulated biomass values, while black vertical bars indicate the simulated catch. Assessment model estimated biomass is shown by dashed grey lines and 95% confidence intervals by shaded regions. Single-stock estimates are in the first column and the remaining columns show the four multi-stock model configurations, with titles corresponding to which shared priors are estimated. The 95% confidence intervals are calculated from the Hessian matrix for leading model parameters using the Δ-method by TMB's ADREPORT() function.

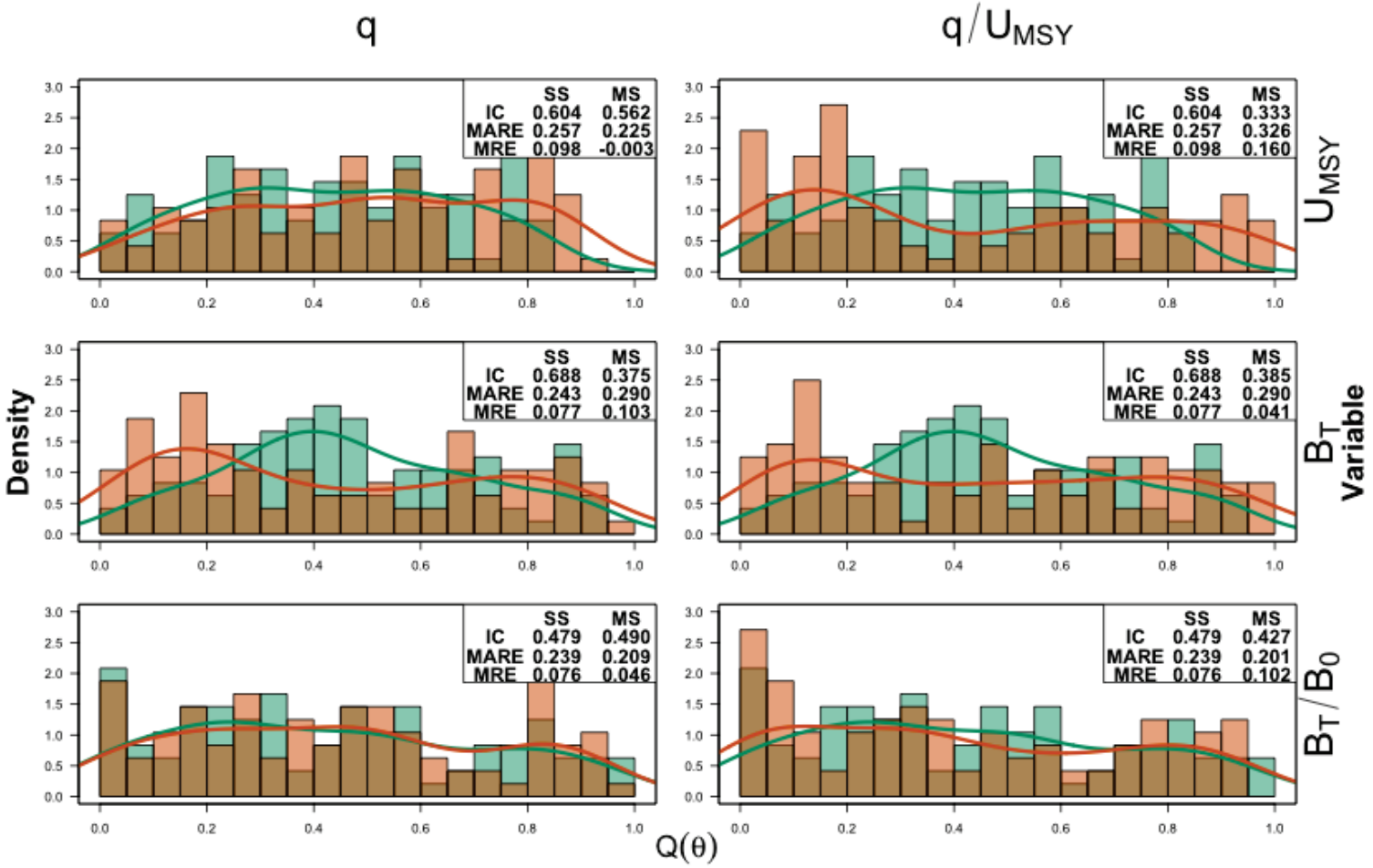

Figure 2.7: Density of predictive quantiles $Q(\theta)$ for estimates of key management parameters (rows) from single stock and q and q/U_{MSY} hierarchical multi-stock model configuration, fit to 4 identical stock under a 1-way trip fishing history over a long time-series of observations, initialised at unfished ($L = 0$). Bars show probability density of Q distributions, with lines showing the kernel smoothed density for easier comparison between single stock (green) and multi-stock (orange) models. Top right hand corners of each panel show interval coverage (IC), median absolute relative error (MARE), and median relative error (MRE) for single stock (SS) and multi-stock models (MS).

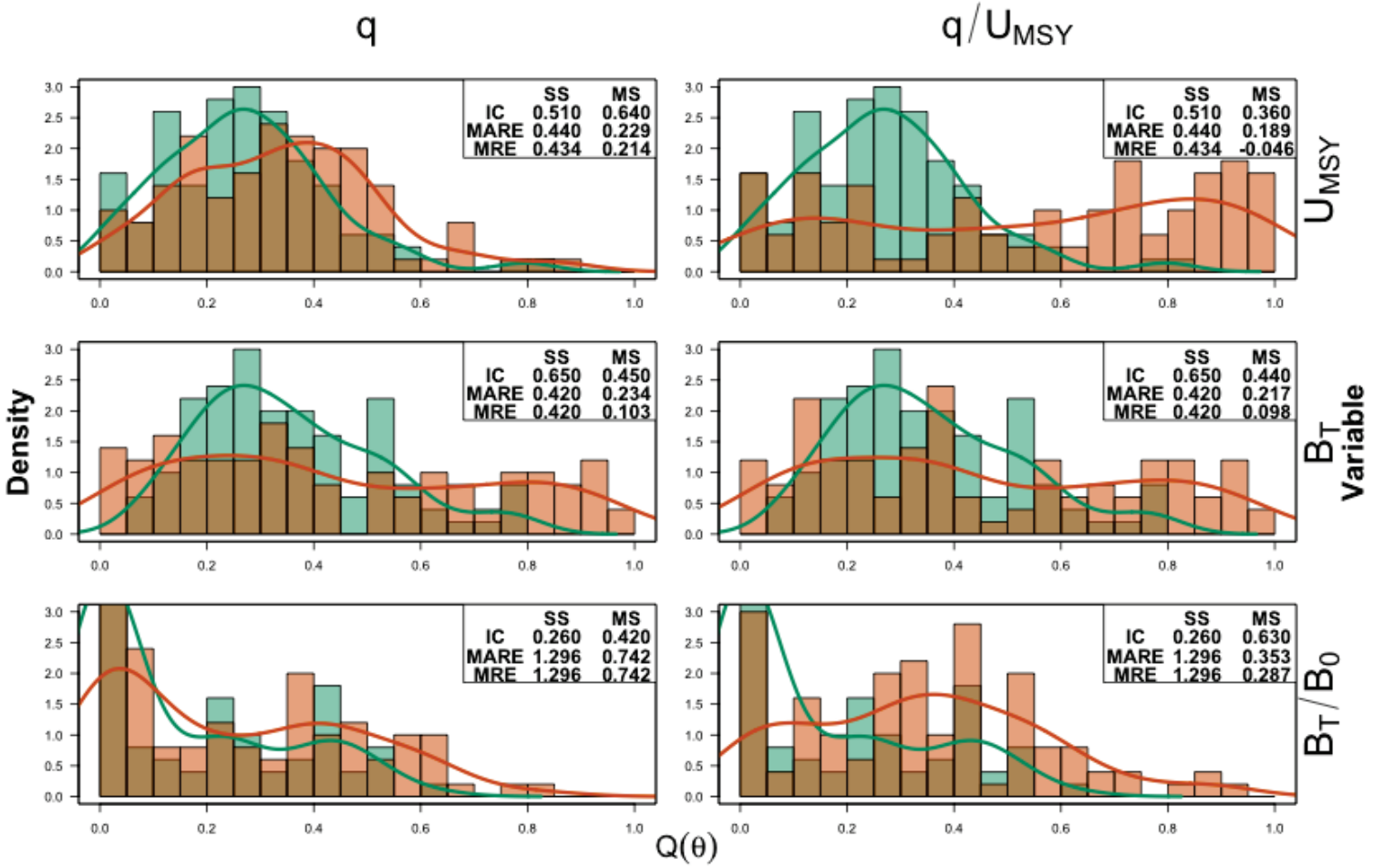

Figure 2.8: Density of predictive quantiles $Q(\theta)$ for estimates of key management parameters (rows) from single stock and q and q/U_{MSY} hierarchical multi-stock model configurations fit to a complex of four stocks with a 2-way trip fishing history with one low data quality stock ($L = 1$), which had a short time series of observations and was initialised at 40% of unfished. Bars show probability density of Q distributions, with lines showing the kernel smoothed density for easier comparison between single stock (green) and multi-stock (orange) models. Top right hand corners of each panel show interval coverage (IC), median absolute relative error (MARE), and median relative error (MRE) for single stock (SS) and multi-stock models (MS).

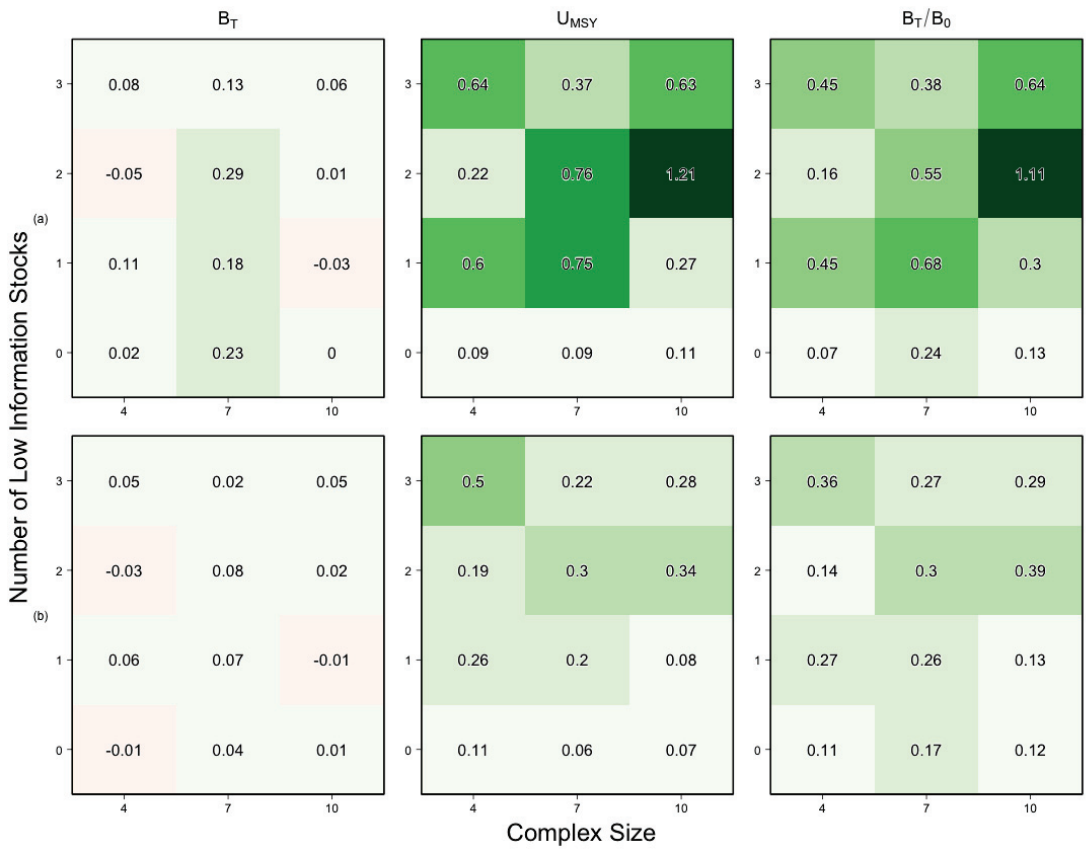

Figure 2.9: Response surface plots of (a) $\Delta(\theta_s) = \frac{MARE_{MS}(\theta)}{MARE_{SS}(\theta)} - 1$ and (b) $\overline{\Delta}(\theta) = \frac{\sum_s MARE_{MS}(\theta_s)}{\sum_s MARE_{SS}(\theta_s)} - 1$ values for $B_{1,T}$ (col. 1) and $U_{1,MSY}$ (col. 2) and $B_{1,T}/B_{1,0}$ (col. 3). Surfaces are plotted as responses to complex size S along the horizontal axis, and number of low information stocks L along the vertical axis. Colours represent the magnitude of the response value, with higher absolute values showing more saturation than absolute values closer to 0, and hue changing from red to green as responses pass from negative, indicating that the single stock performs better, to positive, indicating that the multi-stock model performs better. Response values in each cell are the mean response values for all experimental treatments where S and L took the corresponding values along the axes.

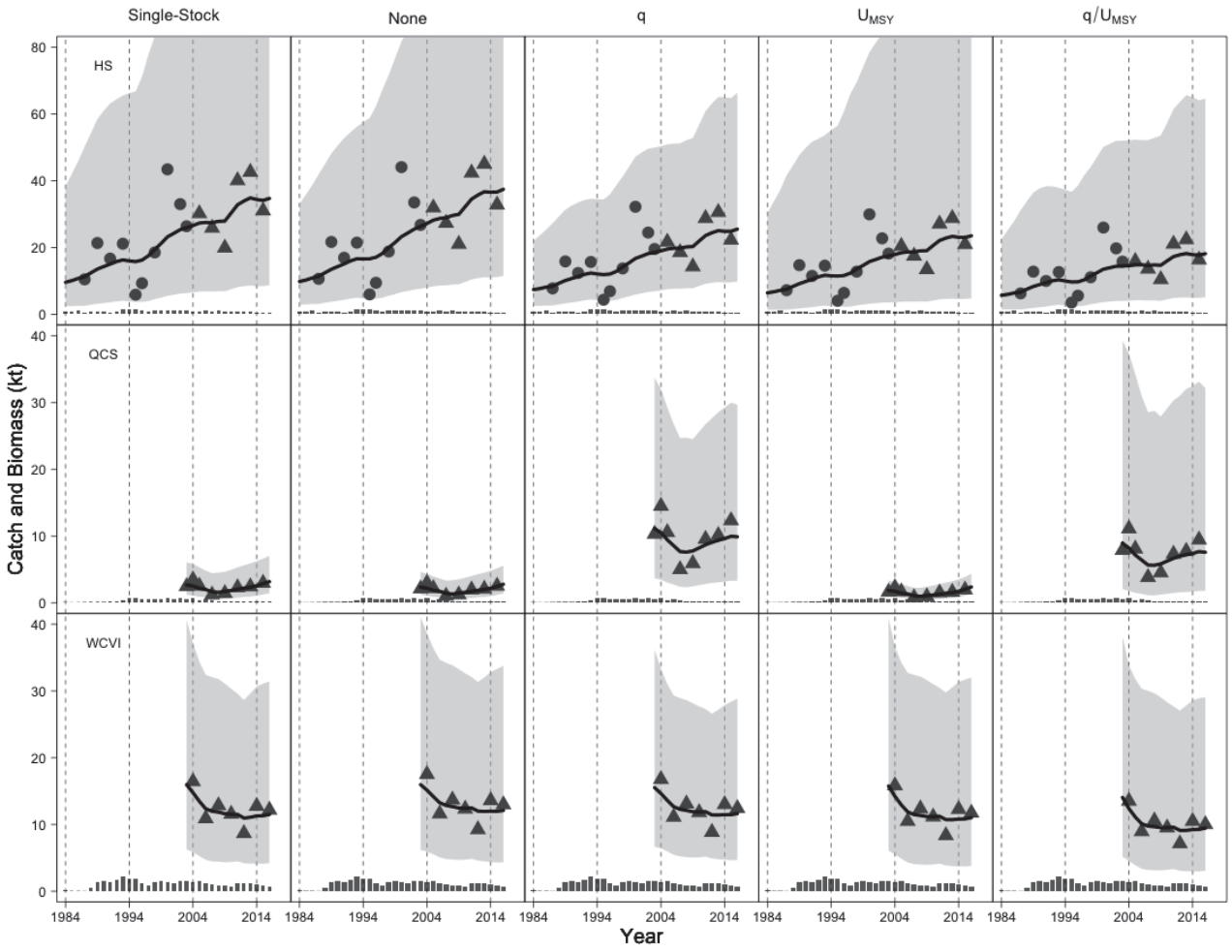

Figure 2.10: Estimated biomass time series for all three Dover Sole stocks. Estimates were produced by the single-stock and 4 top scoring multi-stock assessment model configurations under the low process error variance hypothesis. Grey regions indicate 95% confidence intervals around the maximum likelihood estimates, indicated by the black lines. Black vertical bars at the bottom of each plot show absolute landings and discards. Points indicate survey biomass data scaled by estimated catchability. Circular data points indicate Survey 1 (HS only), while triangular points indicate Survey 2.

2.7 Appendices

2.A Assessment Model Structure

In what follows, $\hat{x}$ denotes the estimate of a derived or leading model parameter x.

Biomass dynamics

The effect of assessment model mis-specification was minimized by matching the deterministic components of the biomass dynamics in the assessment models and the operating model (Eq. 1). In all assessment model configurations the leading biological parameters were $B_{s,MSY} = B_{s,0}/2$ and $U_{s,MSY} = r_s/2$. Biomass time series in the assessment models were initialized at time T_1, either at unfished levels $B_{s,T_1} = B_{s,0}$, or at a separately estimated non-equilibrium value $\hat{B}_{s,T_1}$ when $T_1 > 1984$ or the initial simulated biomass was below unfished levels. Biomass parameters were penalized by normal prior distributions centered at or near their corresponding true values, i.e.,

$$\hat{B}_{s,MSY} \sim N(B_{s,MSY}, B_{s,MSY}),$$
$$\hat{B}_{s,T_1} \sim N(B_{s,MSY}/2, B_{s,MSY}/2),$$

which allows estimates to vary within a realistic range by giving each prior a relative standard deviation of 100%.

When jointly modeled, stock-specific optimal harvest rates $U_{s,MSY}$ values are assumed to share a log-normal distribution with estimated hyperparameters (U.1, Table 3). The estimated prior mean $\overline{U}\hat{}_{MSY}$ followed a normal hyperprior (U.2, Table 3) where m_U was randomly drawn from a log normal distribution with a mean of 0.08 and a standard deviation corresponding to a 20% coefficient of variation. The hyperprior standard deviation $s_U = 0.08$ was chosen to give a roughly 100% CV in the hyperprior for the prior mean and allow the stock-specific values affect the estimate of $\overline{U}\hat{}_{MSY}$ more than the hyperprior. I chose a normal prior as this is the least informative while remaining continuous across the whole domain of the parameter space.

The estimated prior variance followed an inverse gamma distribution (U.3, Table 3) with $\alpha_U = 1$ and $\beta_U = 0.34$, to induce a log-normal coefficient of variation of 20% in the shared U_{MSY} prior. I chose this prior structure recognizing the shared biology of dover sole stocks implies productivities of similar magnitude (Myers et al. 1999). When I modeled the optimal harvest rates separately we assumed optimal harvest rate parameters followed the hyperprior directly (U.4, Table 3) with the same m_U and s_U values.

Observation models

Similar to the biomass dynamics, observational model structure matched between the operating and assessment models (Equation 3). A shared species-level prior distribution was

defined for stock-specific catchabilities $q_{o,s}$, with between stock variance $\iota^2_{q,o}$. Informative priors were also defined for survey observation error variances τ^2_o, with hyperparameters chosen so that the prior modes were equal to the simulated values for each survey.

When the hierarchical prior on catchability parameters was estimated, the model structure for the prior matched the simulated catchability model for each survey. Estimates of stock-specific catchability $\hat{q}_{o,s}$ were drawn from a shared log-normal distribution with estimated hyperparameters $\hat{\bar{q}}_o$ and $\hat{\iota}^2_{q,o}$ (q.1, Table 3). The estimated prior mean followed a normal hyperprior where m_q was drawn randomly as with the U_{MSY} prior, with an average of 0.55 (the average of the two surveys) and 20% CV, while $v_q = 0.55$ for a 100% hyperprior CV (q.2, Table 3), for similar reasoning as the U_{MSY} priors.

The prior variance followed an inverse gamma distribution (q.3, Table 3). Inverse gamma hyperparameters $\alpha_q = 1$ and $\beta_q = 0.34$ were chosen induce a $\hat{\iota}^2_{q,o}$ value that corresponds to a 20% CV, inducing a shrinkage effect. This prior structure reflects an assumption that dover Sole stocks have a similar availability to survey gear based on similar habitat preferences. Like the productivity prior, when catchability was estimated without a shared prior, I bypassed the mid-level prior, penalizing stock-specific catchability using the normal hyperprior with the same m_q and v_q values to signify no change in prior information (q.4, Table 3). I chose a normal hyperprior because it is less informative than a log-normal distribution, and the mean $m_q = .6$ and variance $v_q = 0.36$ are chosen to produce a relative standard deviation of 100%, allowing $\bar{q}_o$ and $q_{o,s}$ to vary in a realistic range, but informative enough to induce a shrinkage effect.

Observation errors for each survey were assumed to be drawn from a single log-normal distribution across stocks, with estimated log-standard deviation $\hat{\tau}_o$. To improve convergence in repeated simulation trials, I assumed the estimated log-variance $\hat{\tau}^2_o$ followed an inverse gamma prior distribution

$$\hat{\tau}^2_o \sim IG(\alpha_{\tau_o}, \beta_{\tau_o}). \tag{2.4}$$

Like the process error variance priors, the hyperparameter β_{τ_o} was chosen to place the mode of the inverse gamma distribution at the simulated values of τ^2_o when $\alpha_{\tau_o} = 0.1$.

2.B Experimental Design

Table 2.B.1: The space filling experimental design used for simulation experiments, with levels of fishing history U, initial year of assessment T_1, complex size S, initial depletion $d_{1,1}$, and number of stocks with low power data L (columns 1-5). The second set of columns shows the total number of simulation replicates required to get a set of 100 convergent simulations for statistically significant results under each assessment model configuration (columns 6-9).

Experimental Factor Levels					**Prior Configuration**			
U	T_1	S	$d_{1,1}$	L	None	q	U_{MSY}	q/U_{MSY}
(0.2,4,1)	1984	4	0.4	0	100	100	100	100
(1,1,1)	2003	7	0.4	0	102	102	102	102
(1,1,1)	1984	10	0.4	0	102	101	101	101
(0.2,4,1)	2003	10	0.4	0	100	100	100	100
(1,1,1)	1984	4	0.7	0	104	103	103	102
(0.2,4,1)	2003	4	0.7	0	100	100	100	100
(1,1,1)	2003	10	0.7	0	102	101	101	101
(1,1,1)	2003	4	1.0	0	104	103	103	103
(0.2,4,1)	1984	7	1.0	0	100	100	100	100
(1,1,1)	2003	4	0.4	1	104	103	102	102
(0.2,4,1)	1984	7	0.4	1	100	100	100	100
(1,1,1)	1984	7	0.7	1	104	102	101	102
(0.2,4,1)	2003	7	0.7	1	100	100	100	100
(0.2,4,1)	1984	10	0.7	1	100	100	100	100
(0.2,4,1)	1984	4	1.0	1	100	100	100	100
(1,1,1)	2003	7	1.0	1	104	102	102	102
(1,1,1)	1984	10	1.0	1	103	101	101	101
(0.2,4,1)	2003	10	1.0	1	100	100	100	100
(1,1,1)	1984	4	0.4	2	102	101	102	102
(0.2,4,1)	2003	4	0.4	2	101	100	100	100
(1,1,1)	2003	10	0.4	2	103	101	101	101
(1,1,1)	2003	4	0.7	2	102	102	101	102
(0.2,4,1)	1984	7	0.7	2	100	100	100	100
(1,1,1)	1984	7	1.0	2	102	102	102	101
(0.2,4,1)	2003	7	1.0	2	100	101	100	100
(0.2,4,1)	1984	10	1.0	2	100	100	100	100
(1,1,1)	1984	7	0.4	3	102	101	101	101
(0.2,4,1)	2003	7	0.4	3	100	101	100	100
(0.2,4,1)	1984	10	0.4	3	100	100	100	100
(0.2,4,1)	1984	4	0.7	3	100	100	100	100
(1,1,1)	2003	7	0.7	3	102	101	101	101
(1,1,1)	1984	10	0.7	3	103	101	101	101
(0.2,4,1)	2003	10	0.7	3	100	100	100	100
(1,1,1)	1984	4	1.0	3	104	103	102	102
(0.2,4,1)	2003	4	1.0	3	100	100	100	100
(1,1,1)	2003	10	1.0	3	102	101	101	101

2.C Meta-models for performance metrics

Performance metrics were modeled as responses to experimental factors and assessment model prior configurations using generalised linear "meta-models". Meta-modeling is a part of a formal approach to simulation experimentation, where outputs of a complex simulation model are viewed as responses to simpler functions of simulation model inputs [74]. The parameters of the meta-model are then used to facilitate interpretation of the results of complex simulation experiments. Generalised linear meta-models are used as they are robust to heterogeneous variance of response variable residuals [89], which were common outputs of the experimental treatment.

For performance metrics $y \in \{\Delta(\theta_s), \overline{\Delta}(\theta), MARE_{ss}(\theta_s), MARE_{ms}(\theta_s)\}$ the coefficients β of a generalised linear model

$$y = \beta_0 + \beta_{config} + \sum_i \beta_i x_i$$

were estimated for each experiment, where the β_0 is the intercept, β_{config} is the effect of the multi-stock assessment model prior configuration, and the coefficients β_i are the factor effects for factor levels x_i. Numerical explanatory variables, such as the year of initialisation and initial depletion, were scaled to $[-1, 1]$ to allow direct comparison of numeric effects with qualitative factor effects. To reduce the number of experimental treatments, I sampled factor levels using a space filling design (Kleijnen 2008). To reduce qualitative factors, I fit the historical fishing intensity and initial year of assessment as continuous variables even though they may not have continuous, or even approximately linear responses. Our reasoning for this is that both factors have 2 levels, and so a linear effect will capture the difference between the level effects sufficiently. For the historical fishing intensity, I regressed on the highest multiple of $U_{s,MSY}$ in the history, that is, $U_d = 1$ for one-way trips, and $U_d = 2$ for two way trips.

The intercept term β_0 represents the average response value at the reference levels of qualitative factors in the model. The only qualitative factor is the choice of multi-stock model configuration, so the intercept of the $\Delta(\theta_s)$ models is $\beta_0 = \beta_{None}$, representing the "null" multi-stock assessment configuration. When there are no qualitative factors in the meta-model, such as in the $MARE_{ss}(\theta_1)$ models, β_0 is simply the average response value over all factors.

Chapter 3

Hierarchical surplus production stock assessment models improve management performance in multi-species, spatially-replicated fisheries.

This chapter is published in Fisheries Research (2021), co-authored with Sean P. Cox under the same title [71].

3.1 Introduction

Managers of multi-species fisheries aim to balance harvest of multiple interacting target and non-target species that vary in abundance and productivity. Among-species variation in productivity implies variation in single-species optimal harvest rates, and, therefore, differential responses to exploitation. Single-species optimal harvest rates (e.g., the harvest rate associated with maximum sustainable yield) typically ignore both multi-species trophic interactions that influence species' demographic rates [53, 20], and technical interactions that make it virtually impossible to simultaneously achieve the optimal harvest rates for all species [108].

Technical interactions among species that co-occur in non-selective fishing gear are a defining characteristic of multi-species fisheries [108, 118] and, therefore, play a central role in multi-species fisheries management outcomes for individual species [100, 72]. Catch limits set for individual species without considering technical interactions subsequently lead to sub-optimal fishery outcomes [100, 115, 114]. For example, under-utilization of catch limits could occur when technically interacting quota species are caught at different rates (i.e., catchability) by a common gear, leading to a choke constraint in which one species quota is filled before the others [6]. Choke constraints are considered negative outcomes for multi-

species fishery performance, because they reduce harvester profitability as increasingly rare quota for choke species may limit access to fishing grounds, as well as driving quota costs above the landed value of the choke species [93].

Setting catch limits for individual species in any fishery usually requires an estimate of species abundance, which continues to be a central challenge of fisheries stock assessment [60, 121, 87], especially when species data are of low statistical power, such as short noisy time series of observations, or uninformative catch series [70]. Where such low power data exists, data pooling is sometimes used to extend stock assessments to complexes of similar, interacting stocks of fish [3]. Examples include pooling data for a single species across multiple spatial strata when finer scale data are unavailable or when fish are believed to move between areas at a sufficiently high rate [7, 117], and pooling data for multiple species of the same taxonomic group within an area when data are insufficient for individual species or during development of new fisheries [31]. Data-pooled estimates of productivity represent means across the species complex, implying that resulting catch limits will tend to overfish unproductive species and underfish productive ones [50].

In multi-species and/or multi-area contexts, hierarchical stock assessment models, which treat each area/species combination as a discrete yet exchangeable replicate, may represent a compromise between single-species and data-pooling approaches. For this chapter, I define a hierarchical stock assessment model as a model fit to multiple replicates (e.g. areas/species) simultaneously, using hierarchical hyper-priors on selected parameters to share information between replicates [138]. Hierarchical priors induce shrinkage effects in which parameter values are drawn towards an estimated overall mean value, thus improving model convergence for replicates with low statistical power data while still estimating replicate-specific parameters based on that data. Hierarchical methods based on data and hyper-priors stand in contrast to data-pooled methods that estimate a mean value only, or single-stock methods that usually rely on strong *a priori* assumptions about replicate specific parameters, forcing parameters to be identical among replicates, or using strongly informative priors, all of which will almost certainly increase assessment bias [69, 68, 120].

Although hierarchical stock assessments are expected to produce better estimates of species biomass and productivity than single-species methods in data-limited contexts, it remains unclear whether such improved statistical performance translates into better management outcomes [70]. Aside from some related simulations determining the benefits of manually sharing information gained when actively adaptively managing spatially replicated groundfish stocks [21], to my knowledge there are no evaluations of the management performance of hierarchical stock assessment models. Further, low assessment model bias and/or high precision, which are often unattainable outside of simulations, aren't necessary conditions for superior management performance, because biases can, in practice, sometimes compensate for each other (e.g., negative correlation in stock size and productivity), or be offset for by other parts of the management system, such as a reduction in harvest rate.

A modern fisheries management oriented paradigm is more concerned with the expected performance of a fisheries management system - made up of data, assessments, and harvest rules - despite the inherent, and at some point irreducible, uncertainties in the system [30, 134, 81].

In this chapter, I ask whether hierarchical stock assessment models improve management performance in a simulated multi-species, spatially replicated fishery. The simulated fishery is modeled on a spatially heterogeneous complex of Dover sole (*Microstomus pacificus*), English sole (*Parophrys vetulus*), and southern Rock sole (*Lepidopsetta bilineata*) off the coast of British Columbia, Canada, which is fished in three spatial management areas. Closed-loop feedback simulation is used to estimate fishery outcomes when catch limits are set based on biomass estimates from single-species, data-pooling, and hierarchical state-space surplus production models under high, moderate, and low data quantity scenarios. Assessment models are either fit to species-specific data as single-species or hierarchical multi-species models, or fit to data pooled spatially across management units, pooled across species within a spatial management unit, or totally aggregated across both species and spatial management units. Management performance under each assessment approach is measured by the risk of overfishing, and by cumulative absolute loss in catch, defined as the deviation from optimal catch trajectories. Optimal trajectories are generated by an omniscient manager, who sets annual effort to maximize total multi-species/multi-stock complex yield with perfect knowledge of all future recruitments [144, 86].

3.2 Methods

3.2.1 British Columbia's flatfish fishery

The multi-species complex of right-eyed flounders (*Plueronectidae*) in British Columbia (BC) is a technically interacting group of flatfishes managed over the BC coast (Figure 3.1). Although there are several right-eyed flounders in BC waters, I focus on the three species, indexed by s, Dover sole ($s = 1$), English sole ($s = 2$), and southern Rock sole ($s = 3$), which I hereafter refer to as Rock sole. Taken together, these species comprise a multi-stock complex (BC flatfish complex), which are part of the BC multi-species groundfish fishery.

The BC flatfish complex is managed in three spatially distinct stock areas, indexed by p (Figure 3.1) [45]. From north to south, the first stock area - Hecate Strait/Haida Gwaii (HSHG, $p = 1$) - extends from Dixon Entrance and north of Haida Gwaii, south through Hecate Strait. The second stock area - Queen Charlotte Sound (QCS, $p = 2$) - extends from the southern tip of Haida Gwaii to the northern tip of Vancouver Island. Finally, the third area - West Coast of Vancouver Island (WCVI, $p = 3$) - extends from the northern tip of Vancouver Island south to Juan de Fuca Strait. These areas are aggregates of PFMA major statistical areas, and are primarily delineated by management breaks, although there are

oceanographic features that may separate each of the stocks as defined here (e.g., gullies in the QCS area, and strong currents at the northern tip of Vancouver Island).

Each species/area combination has commercial trawl catch for the entire history (1956 - 2016), two commercial catch rate series, and at least one survey biomass index time series for each species (Figure 3.2), for a total of four distinct fleets indexed by f. The two commercial fishery catch-per-unit-effort (CPUE) series span 1976 - 2016, split between a Historical trawl fishery (1976 - 1995, $f = 1$) and a Modern trawl fishery (1996 - 2016, $f = 2$), with the split corresponding to pre- and post-implementation of an at-sea-observer programme, respectively. The fishery independent trawl survey biomass indices are the biennial Hecate Strait Assemblage survey in the HSHG area (1984 - 2002, $f = 3$), and the Multi-species Synoptic Trawl Survey, which operates every year but alternates between areas, effectively making it a biennial survey in all three areas (2003 - 2016, $f = 4$).

3.2.2 Closed-loop feedback simulation framework

Closed-loop simulation is often used to evaluate proposed feedback management systems. In fisheries, closed-loop simulation evaluates fishery management system components, such as stock assessment models or harvest decision rules, by simulating repeated applications of these components, while propagating realistic errors in monitoring data, stock assessment model outputs, and harvest advice [30, 27, 28]. At the end of the simulation, pre-defined performance metrics are used to determine the relative performance of the system components being tested.

The closed-loop simulation framework for the BC flatfish fishery includes a stochastic operating model, representing stock and fishery dynamics and generating observations with uncertainty, and an assessment model component that estimates stock biomass from simulated observations and fishery catches. The operating model simulates population dynamics of a spatially stratified, multi-species flatfish complex in response to exploitation by a multi-species trawl fishery in each of the three stock areas. Although total fishing effort is not restricted for the BC flatfish fishery, effort in each individual area is allocated such that no species-/area-specific catch exceeds the species-/area-specific total allowable catch (TAC), but fishing effort is allowed to increase until at least one species-/area-specific TAC is fully caught.

The simulation projected population dynamics for each species in each area (total nine stocks) forward in time for 32 years, or two Dover sole generations [130, longest generation time], with annual simulated assessments, harvest decisions, and catch removed from the population during 2017 - 2048. The following five steps summarise the closed-loop simulation procedure for each projection year t:

1. Update stochastic population dynamics and generate new realized catch $C_{s,p,t}(E_{p,t})$ in each area from effort (eq 3.1);

2. Generate new observation data $I_{s,p,f,t}$ (eqs 3.4-3.7)

3. Apply an assessment model (defined below) to estimate the spawning biomass for the upcoming year $\widehat{B}_{s,p,t+1}$ and an optimal harvest rate $\widehat{U}_{MSY,s,p}$ (eq 3.3);

4. Scale the estimated optimal harvest rate to an estimated multi-species optimal harvest rate, and use it with the biomass forecast to generate a total allowable catch $TAC_{s,p,t+1}$, allocated among species/areas if using a pooled method (eqs 3.8 - 3.13);

5. Allocate effort $E_{p,t+1}$ to fully realise at least one TAC from step 4 in each stock area (eq 3.2).

Operating Model

The operating model (OM) is a multi-species, multi-stock age- and sex-structured population dynamics model (Appendix 3.7.A). Population life-history parameters for the operating model are estimated by fitting a hierarchical age- and sex-structured model to data from the real BC flatfish complex.

Fishing mortality for individual stocks is scaled to commercial trawl effort via species-specific catchability parameters, i.e.,

$$F_{s,p,f,t} = q^F_{s,p,f} \cdot E_{p,f,t}, \tag{3.1}$$

where $F_{s,p,f,t}$ is the fishing mortality rate applied to species s in stock-area p by fleet f in year t, and $q^F_{s,p}$ is the commercial catchability coefficient scaling trawl effort $E_{p,t}$ in area p to fishing mortality (Table 3.1). The relationship in equation 3.1 implies a multi-species maximum yield $MSY_{MS,p}$ within each area, from which the model equilibria are derived, including effort $E_{MSY,MS,p}$, species yield $MSY_{MS,s,p}$, the biomass $B_{MSY,SS,s,p}$ at which multi-species maximum yield is achieved, and the optimal harvest rate $U_{MSY,MS,s,p} = MSY_{MS,s,p}/B_{MSY,MS,s,p}$ that produces multi-species maximum yield, all of which are given in Table 3.1 [108]. The single-species optima, which are also in Table 3.1, are derived in the usual way by maximising single-species yield $MSY_{SS,s,p}$, produced at the optimal biomass $B_{MSY,SS,s,p}$ by applying the optimal harvest rate $U_{MSY,SS,s,p} = MSY_{SS,s,p}/B_{MSY,SS,s,p}$.

For simulated fishing in the projection time period, the fully realised TAC depends on choke effects determined by relative catchabilities, and absolute biomass levels of each species within an area. To simulate the choke effects, the fishing effort allocated to each area is the maximum amount required to fully utilize the TAC of at least one species, but never exceed the TAC of any one of the three species, i.e.,

$$E_{p,t+1} = \max\{E \mid C_{s,p,t+1}(E) \leq TAC_{s,p,t+1} \forall s\}, \tag{3.2}$$

where $C_{s,p,t+1}(E)$ is the catch of species s when effort E is applied in area p in year $t+1$ (the method for determining TACs is explained below). I use maximum effort instead of an explicit effort dynamics model because the former captures choke effects present in the real fishery system, while reflecting the output controlled BC groundfish fishery, for which there is no explicit limit on the total amount of fishing effort expended by license holders.

Surplus production stock assessment models

At each time step t, simulated annual assessments are used to estimate the expected future spawning biomass estimate $\widehat{B}_{t+1}$ via a state-space Schaefer production model [126, 111], modified to better approximate the biomass-yield relationship underlying the age-/sex-structured operating model [105, 148]. I extend the hierarchical state-space model from Chapter 2 to fit to data from the multi-species, spatially stratified BC flatfish complex, as well as fit to single-stock data from individual or data-pooled stocks, via the biomass equation

$$B_{s,p,t+1} = \left[B_{s,p,t} + U_{MSY,s,p} \cdot \frac{m_{s,p}}{m_{s,p}-1} \cdot B_{s,p,t} \cdot \left(1 - \frac{1}{m_{s,p}} \cdot \left(\frac{B_{s,p,t}}{B_{MSY,s,p}} \right)^{m_{s,p}-1} \right) - C_{s,p,t} \right] e^{\zeta_{s,p,t}}, \tag{3.3}$$

where the management parameters $MSY_{s,p}$ (optimal single-species yield) and $U_{MSY,s,p}$ (optimal single-species harvest rate) are the leading model parameters, $B_{MSY,s,p} = MSY_{s,p}/U_{MSY,s,p}$ is the biomass at which $MSY_{s,p}$ is acheived under the optimal harvest rate, $m_{s,p}$ is the Pella-Tomlinson parameter controlling skew in the biomass/yield relationship (derived from operating model yield curves), and $\zeta_{s,p,t}$ are annual process error deviations.

In total, I define the following five potential assessment model configurations (Figure 3.3):

1. Total Pooling (1 management unit);
2. Species Pooling (3 management units, independent parameters);
3. Spatial Pooling (3 species, independent parameters);
4. Single-species (9 management units, independent parameters);
5. Hierarchical Multi-species (9 management units, hierarchical priors);

where the number of management units is shown in parentheses (i.e., delete subscripts in eq. 7 for species or stock-area as appropriate, e.g., Spatial Pooling models have no p subscript).

Prior distributions on optimal yield $MSY_{s,p}$, optimal harvest rate $U_{MSY,s,p}$, catchability $q_{s,p,f}$, and process error deviations $\zeta_{s,p,t}$ are defined for each assessment model, with unbiased prior means based on the true operating model values (Table 3.C.2, Appendix 3.7.C). While unbiased prior distributions are not possible to define in practice, my research question is

focused on the management performance when using estimates of biomass and productivity from each of the five AMs to set TACs. While priors certainly affect those estimates, there is considerable variation in how an analyst might choose prior mean values for biological parameters, and those choices become more uncertain as data are reduced. For example, an analyst might choose a prior mean of average catch over an agreed historical period as the MSY prior mean, and choose either a conservatively low value based on a review of stock assessments for similar species, or if it is available, the ratio of catch to an index of biomass in the same historical period to inform a prior mean U_{MSY} value. In cases where time-series of data are short, there may be no appropriate period for formulating such priors, while noisy data (e.g., commercial CPUE) lack sufficient statistical power to draw parameter estimates away from biased mean values, especially when priors may be informative in the presence of data limitations. I based our unbiased priors on the OM to keep priors as similar as possible among AMs, thereby avoiding variation in performance caused by different prior choices.

Data generation for assessment models

Time series of catch, commercial CPUE, and relative biomass indices are simulated in the historical and projection periods for fitting assessment models. While it is more usual to use the actual historical data in most closed loop simulations, given that pooled data must be simulated for all time periods, simulating all data series in the history and projections removes any effect of simulated and real data on performance of pooled and non-pooled AMs. The Historical fishery CPUE and Hecate Strait Assemblage survey are simulated only for the periods shown in Figure 3.2, but the Modern fishery CPUE and Synoptic trawl survey biomass index are also simulated in the projection period. All biomass indices are simulated with log-normal observation errors, with precision based on the conditioning assessment (Table 3.A.1).

Biomass indices for individual stocks are all simulated as relative biomass with observation error

$$\bar{I}_{s,p,f,t} = q_{s,p,f} \cdot B_{s,p,f,t}, \tag{3.4}$$

where $\bar{I}_{s,p,f,t}$ is the index without observation error, $q_{s,p,f}$ is the catchability coefficient for species s and stock-area p, and $B_{s,p,f,t}$ is the biomass of species s in stock-area p vulnerable to fleet f in year t. For the Assemblage and Synoptic fishery independent surveys ($f = 3, 4$), the catchability coefficients $q_{s,p,f}$ are trawl efficiency estimates from the conditioning assessment. For commercial CPUE, catchability coefficients are estimated in the conditioning assessment by assuming that catch rates are an unbiased relative index. While the assumption that catch rates are an unbiased relative index likely increases the information contained in the simulated CPUE indices over real CPUE data, the larger CVs estimated by the conditioning assessment retain realistic variation. Moreover, simulating

commercial data as relative biomass reduces differences in the variability and bias of the data simulated in the OM history and data simulated in the projection.

The method for generating pooled catch and biomass index data is analogous to the unpooled data, combined via summation. Catch data is summed without any scaling, and biomass indices without observation error are summed after scaling by catchability (trawl efficiency in the surveys), i.e.,

$$\bar{I}^{pooled}_{s,f,t} = \sum_{p} \mathbf{1}(I_{s,p,f,t} > 0) \cdot q_{s,p,f} \cdot B_{s,p,f,t} \tag{3.5}$$

$$\bar{I}^{pooled}_{p,f,t} = \sum_{s} \mathbf{1}(I_{s,p,f,t} > 0) \cdot q_{s,p,f} \cdot B_{s,p,f,t} \tag{3.6}$$

$$\bar{I}^{pooled}_{f,t} = \sum_{s,p} \mathbf{1}(I_{s,p,f,t} > 0) \cdot q_{s,p,f} \cdot B_{s,p,f,t} \tag{3.7}$$

where $f = 1, .., 4$, $\bar{I}^{pooled}_{s,f,t}$ is the spatially pooled index for species s, $\bar{I}^{pooled}_{p,f,t}$ is a species pooled index for area p, $\bar{I}^{pooled}_{f,t}$ is the totally aggregated index (all without error), and $\mathbf{1}(I_{s,p,f,t} > 0)$ is the indicator function that takes value 1 when survey f in area p took samples in year t, and 0 otherwise.

Target harvest rates and total allowable catch

Simulated harvest decision rules apply a constant target harvest rate to generate TACs from one-year ahead biomass forecasts obtained from each assessment model, i.e.,

$$TAC'_{s,p,t+1} = \widehat{U}_{s,p,t} \cdot \widehat{B}_{s,p,t+1}, \tag{3.8}$$

where $\widehat{U}_{s,p,t}$ is the estimated target harvest rate for species s in stock-area p, and $\widehat{B}_{s,p,t+1}$ is the year $t+1$ biomass forecast from the assessment model applied in year t. The estimated target harvest $\widehat{U}_{s,p,t}$ is defined to incorporate realistic assessment errors in productivity estimates while simultaneously targeting maximum multi-species yield $MSY_{MS,s,p}$. To avoid incorporating technical interactions directly into the AM dynamics, the target harvest rate scales the AM estimate of $\widehat{U}_{MSY,s,p}$ from equation 3.3 by the ratio of multi-species and single-species optimal harvest rates from the operating model, i.e.,

$$\widehat{U}_{s,p,t} = \widehat{U}_{MSY,s,p} \cdot \frac{U_{MSY,MS,s,p}}{U_{MSY,SS,s,p}} \tag{3.9}$$

where $U_{MSY,MS,s,p}, U_{MSY,SS,s,p}$ are the optimal harvest rates maximising multi-species and single-species yield, respectively, taken from the operating model (Table 3.1), and $\widehat{U}_{MSY,s,p}$ is the assesment model estimate of the single-species optimal harvest rate applied at time t.

Inter-annual increases in TAC are limited to 20% for all individual stocks, i.e.,

$$TAC_{s,p,t+1} = \min\{TAC'_{s,p,t+1}, 1.2 \cdot TAC_{s,p,t}\}, \tag{3.10}$$

where $TAC'_{s,p,t+1}$ is the proposed TAC determined above, and $TAC_{s,p,t}$ is the previous year's TAC. The constraint on inter-annual TAC changes reflects the constraint on inter-annual changes in fishing effort in the objective function used for the omniscient manager solutions described below, which are meant to simulate gradual investment in additional fishing effort.

Pooled TACs are set analogously to the stock-specific case above, with pooled target harvest rates applied to biomass projections from pooled assessments. For a spatially pooled assessment of species s, we define the operating model spatially pooled optimal harvest rates as

$$U_{MSY,MS,s} = \frac{\sum_p MSY_{MS,s,p}}{\sum_p B_{MSY,MS,s,p}}, \tag{3.11}$$

$$U_{MSY,SS,s} = \frac{\sum_p MSY_{SS,s,p}}{\sum_p B_{MSY,SS,s,p}}, \tag{3.12}$$

where the notation is as defined above, with species pooled and totally pooled rates defined analogously. For setting TACs under pooled assessments, the harvest rate scalar from single- to multi-species U_{MSY} in equation 3.9 uses the ratio of pooled $U_{MSY,MS}$ and $U_{MSY,SS}$ values, as defined in equations 3.11 and 3.12. Assessment model estimates of pooled optimal harvest rates were then scaled by the ratio of the pooled operating model optimal harvest rates.

Pooled TACs are split within an area or across spatial strata proportional to Synoptic trawl survey indices for the individual stocks. For example, if the TAC for area p is set by a species pooled assessment, then the proposed TAC for species s is defined as

$$TAC'_{s,p,t+1} = \frac{I_{s,p,t^*}}{\sum_{s'} I_{s',p,t^*}} TAC_{p,t+1}, \tag{3.13}$$

where $\bar{I}_{s,p,t^*}$ is the most recent individual biomass index from the Synoptic survey for species s in area p. The most recent index is used because the synoptic survey alternates between areas each year, so not all individual indices are present in a given year.

3.2.3 Simulation experiments and performance

A total of 15 simulation experiments are conducted, comprising five assessment models and three data quality scenarios. Simulations are integrated over stochastic errors by running a total of 200 random replicates of each combination, where each simulation is initialized with the same set of random seeds to eliminate random effects among combinations of assessments and data scenarios. Assessment convergence is defined as a positive definite Hessian

matrix and a maximum gradient component less than 10^{-3} in absolute value. Replicates are considered significant when assessments converge in 95% of time steps, chosen to reflect the fact that fitting models becomes more difficult as data quality is deliberately reduced, and a simulated assessment can not always be tuned like a real assessment performed by a real-life analyst. Results are then calculated based on the first 100 random seed values that produced significant replicates for all species and stocks considered. The operating model projections are run for two Dover sole generations [130, 32 years], as that species has the longest generation time.

Operating model population dynamics are identical among replicates for each stock during the historical period (1956 - 2016), except for the last few years near the end, where the operating model simulates recruitment process errors that were not estimated in the conditioning assessment. Simulated log-normal observation and process errors in the projection are randomly drawn with the same standard deviations as the errors used in the historical period, and bias corrected so that asymptotic medians match the expected values, i.e., for the two fishery independent surveys ($f = 3, 4$), the species/area specific biomass indices are simulated as

$$I_{s,p,f,t} = \bar{I}_{s,p,f,t} \cdot \exp(\tau_{s,p,f} \cdot \delta_{s,p,f,t} - 0.5\tau^2_{s,p,f}) \tag{3.14}$$

where $\bar{I}_{s,p,f,t}$ is the index without error defined above, $\tau_{s,p,f}$ is the log-normal observation error standard deviation, $\delta_{s,p,f,t}$ is the annual standard normal observation error residual, and subscripts s, p, f, t are for species, stock, fleet and year, respectively. Recruitments are simulated as in Section 3.A. Error is added to survey biomass indices for pooled data independently of the error added to individual indices, i.e.,

$$I_{s,f,t} = \bar{I}_{s,f,t} \cdot \exp(\tau_{s,f} \cdot \delta_{s,f,t} - 0.5\tau^2_{s,f}) \tag{3.15}$$

$$I_{p,f,t} = \bar{I}_{p,f,t} \cdot \exp(\tau_{p,f} \cdot \delta_{p,f,t} - 0.5\tau^2_{p,f}) \tag{3.16}$$

$$I_{f,t} = \bar{I}_{f,t} \cdot \exp(\tau_{f} \cdot \delta_{f,t} - 0.5\tau^2_{f}) \tag{3.17}$$

where $\tau_{s,f}, \tau_{p,f}, \tau_f$ are averaged over the components of the pooled index.

Operating model data quantity scenarios

The three data quantity scenarios transition from high to low by successively removing the earlier CPUE index series from the full set, i.e.,

1. **High**-data: Historical CPUE (1956 - 1996), Modern CPUE (1996 onwards), Assemblage survey (1984 - 2002, biennial, HSHG only), Synoptic survey (2003 onwards, biennial);

2. **Mod**erate-Data: Modern CPUE, Assemblage survey, Synoptic survey;

3. **Low**-Data: Assemblage survey, Synoptic survey.

To improve convergence, the Hierarchical Multi-stock and Single-species assessment models are initialised later under the **Mod** and **Low** data scenarios, with the starting year of the assessments set to the first year with index data, which was 1984 in HSHG for both scenario, and 1997 or 2003 for other areas under the **Mod** and **Low** scenarios, respectively.

Performance evaluation

Omniscient manager simulations Assessment model performance is measured against a simulated omniscient fishery manager aware of all the future consequences of harvest decisions and, therefore, able to adapt the management to meet specific quantitative objectives under any process error conditions [144]. Omniscient manager solutions are used rather than equilibrium based metrics [113] because most stocks are in a healthy state in 2016 (i.e. above single-species B_{MSY}, Table 3.2) and, therefore, the time-path of fishery development was important [144].

The omniscient manager is an optimisation of future fishing effort by area (Appendix 3.B), with the objective function

$$\mathcal{O} = \left[\sum_{s,p} -\log(\bar{C}_{s,p,\cdot})\right] + \mathcal{P}_{diff}\left(\sum_{p} E_{p,\cdot}\right) + \mathcal{P}_{init}\left(\sum_{p} E_{p,2017}\right) + \mathcal{P}_{overfished}\left(B_{s,p,\cdot}\right), \quad (3.18)$$

where $-\log \bar{C}_{s,p,\cdot}$ is the negative log of total future catch for species s in area p over the projection period (equivalent to maximising catch). Penalty functions $\mathcal{P}$ (eq. 3.20) are applied for annual changes in total effort across all three areas being above 20% ($\mathcal{P}_{diff}$) to match the TAC smoother in stochastic experiments, differences greater than 10% between the last year of historical effort and the first year 2017 of simulated effort ($\mathcal{P}_{init}$), and a penality when biomass drops below a lower threshold of 40% of single species $B_{MSY,SS}$ in more than 5% of time-steps ($\mathcal{P}_{overfished}$). The threshold of $0.4B_{MSY,SS}$ is commonly suggested as a limit reference point for Canadian fisheries [34], below which the stock could experience irreparable harm and recruitment could become impaired.

An omniscient manager solution is obtained for each stochastic trajectory in the stochcastic management simulations. Each replicate is run for 80 years to produce several years free of end effects, such as transient dynamics at the beginning of the projection, or reduced penalties for overfishing at the end of the projection.

Cumulative catch loss For each stochastic trajectory, the cumulative absolute loss in catch is calculated as [144]:

$$L_{i,s,p} = \sum_{t=T_1}^{T_2} |C_{i,s,p,t,sim} - C_{i,s,p,t,omni}|, \tag{3.19}$$

where the $C_{i,s,p,t,\cdot}$ values are commercial trawl catch for replicate i, species s, and stock p from stochastic simulations (sim) or the omniscient manager simulation ($omni$) simulation, with $C_{i,s,p,t,sim} - C_{i,s,p,t,omni}$ defined simply as catch loss, which is positive when catch is higher than the omniscient manager's, and negative when lower. When repeated over all significant random seed values, the loss functions generate a distribution of cumulative absolute catch loss, which are used to determine each assessment model's relative performance under the three data scenarios. Cumulative absolute catch loss is calculated for the ten year period $T_1 = 2028$ to $T_2 = 2037$, chosen in the middle of the projection period because dynamics in the earlier period were dominated by the smoothers on effort and catch for the omniscient manager and TACs, respectively.

A paired analysis is used to rank AM configurations across species, stocks and replicates. Within each replicate i, each AM's cumulative absolute catch loss determines the relative rank of each AM under a species/stock/OM scenario combination, where lower loss ranks higher. Any replicates with less than 95% convergence rates for any AM on any management units are excluded from the aggregate rankings, and any species/stock combinations that fail to reach 100 replicates meeting the convergence criteria for all AMs are also excluded, to reduce variability in ranking distributions caused by varying sample sizes and random seeds. Rankings are then pooled across remaining species and stocks within an OM/AM combination, from which the modal rank and average rank are calculated.

Biomass and overfishing risk Biomass risk is measured by the probability of stock biomass being below both 80% and 40% of single-species operating model $B_{MSY,SS,s,p}$. The threshold of 80% of $B_{MSY,SS}$ is generally considered to be the level where a fish stock transitions from an optimally fished state to an overfished state [58], while 40% of B_{MSY} is commonly suggested as a limit reference point for Canadian fisheries, below which a stock may be considered critically overfished as recruitment may become impaired and rebuilding may be required [34]. For both biomass levels, the probability is calculated as

$$P(B_{s,p,t} < \lambda B_{MSY,SS,s,p}) = \frac{1}{3200} \sum_{i=1}^{100} \sum_{t=2016}^{2047} \mathbb{1}(B_{i,s,p,t} < \lambda B_{MSY,SS,s,p})$$

where i, s, p, t are replicates meeting the 95% convergence criteria, species, stocks and time steps, respectively, $\lambda = 0.4, 0.8$, and $\mathbb{1}$ is the indicator function that takes value 1 when

its argument is true, and zero otherwise. Overfishing risk is similarly calculated as the probability of fishing mortality exceeding single-species operating model $F_{MSY,SS,s,p}$.

3.2.4 Sensitivity analyses

Parameter prior distributions are a key feature of most contemporary stock assessment models, even in data-rich contexts. Moreover, prior distributions are a defining feature of hierachical multi-species stock assessment models. Therefore, I focus most sensitivity analyses on fixed prior standard deviations for AM leading parameters MSY and U_{MSY}, and the hierarchical shrinkage prior SDs τ_q and $\sigma_{U_{MSY}}$ (Table 3.4). Sensitivity of performance to future increases in fishery independent survey precision is also tested, where the Synoptic survey observation error variance is reduced with a linear ramp-down over the first 5 years of the projection period to simulate a gradual increase in survey effort.

3.3 Results

3.3.1 Omniscient Manager Performance

As expected, the omniscient manager was able to achieve the theoretical multi-species optimal yield in the presence of technical interactions during the middle of the projection period (Figure 3.4, blue closed circle). Median biomass, catch, and fishing mortality reach the equilibrium after a transition period of about 20 years. During the transitionary period, effort is slowly ramped up in each area from the end of the historical period, stabilising around area-specific E_{MSY} after about 12 years (Figure 3.5, blue closed circles). As is common when maximising catch over a finite time horizon, a reduction in biomass and an increase in effort was observed towards the end of the 80 year projection period, where penalties from lower biomasses and overfishing the stock relative to the multi-species optimal harvest rate are lower than the catch increases; however, the omniscient manager avoids a complete crash of the complex thanks to the penalty on avoiding the critically overfished state at 40% of single-species $B_{MSY,SS}$.

Each area had similar relationships between single-species and multi-species optimal biomass levels, with one species overfished, one underfished, and one close to optimally fished, relative to single-species reference points. In HSHG and QCS, the overfished stocks were both Rock sole, fished down to 74% (HSHG) and 81% of $B_{MSY,SS}$ to increase fishery access to Dover and English soles in those areas (Table 3.1). In WCVI, English sole was slightly overfished relative to single-species at 91% of $B_{MSY,SS}$, but this would be of little concern in a real fishery if the absolute size were not so small. Despite the tendency of the omniscient manager toward overfishing at least one species in each area, very few optimal solutions risked severe overfishing below 40% of B_{MSY} (Table 3.3), indicating that lost yield from more intense overfishing relative to single-species optimal levels is not compensated by increased harvest from other species, which are sometimes larger populations with higher

TACs. The probability of being critically overfished in the period 2028 - 2037 was 1% for English and Rock sole stocks, and 0% for all Dover sole stocks.

Although DER stocks begin the simulations in an overall healthy state, the omniscient manager reduced fishing effort to 0 in all areas early in the projection period in some replicates (Figure 3.5, 2016-2020). In these cases, anticipatory feedback control by the omniscient manager reduced fishing effort to avoid low spawning stock biomasses, thus ensuring higher production in later time steps where recruitments were lower than average for sustained periods.

3.3.2 Assessment model performance

Rankings by catch loss

Due to low convergence rates for HSHG and QCS Rock sole and WCVI Dover sole in the High scenario, and WCVI Rock sole under the Mod scenario (Figure D.1, Appendix 3.7.D), results from those management units are excluded from catch loss rankings.

Hierarchical Multi-species assessment models ranked highest under all data quantity scenarios (Table 3.4). Under the Low data quantity scenario, the Hierarchical model had the lowest cumulative absolute catch loss in over 40% of replicate/stock/species combinations (Figure 3.6), ranking highest by average and modal rank in the aggregate. When split across individual species/areas, the modal rank for the hierarchical AM remained at 1, but there was more variability in the ranking distributions in the HSHG area (e.g. Dover sole), leading to a drop in average rank (Figure D.1, Appendix 3.7.D). As data quantity increased for the Mod scenario, the modal rank and average rank of the hierarchical model remained the highest out of all AMs in the aggregate, ranking 1st in just under 40% of cases (Figure 3.6), while the variability in rankings in the HSHG area increased (Figure D.1, Appendix 3.7.D), pushing the average rank a littler lower than under the Low scenario. Results under the High scenario were similar to the Mod scenario, where the Hierarchical model was ranked first in about 40% of cases, and a slightly higher average rank than the Mod scenario, driven largely by a reduction in catch loss for WCVI English sole (Figure D.1).

When ordered by average rank, Species Pooling AMs came second after the hierarchical AMs under all data quantity scenarios (Table 3.4). Total and Spatial Pooling methods performed similarly under the High scenario, with only a small difference between average ranks, but quite different rank distributions, where Spatial pooling catch loss rankings were almost uniformly distributed, but Total pooling ranks had a clear mode in 4th place (Figure 3.6, High scenario), which was reflected in individual stock rank distributions (Figure D.1, Appendix 3.7.D). The Total Pooling AM's 4th place rank was observed in around 40% of cases in the aggregate (Figure 3.6), and was also the modal rank for most individual stock distributions (Figure D.1). As data were removed for the Mod data quantity scenario, the Spatial Pooling AM dropped from 3rd to 4th place, driven largely by increased catch loss in the QCS area. The Single-species method had the worst rank under all data quantity

scenarios, with a consistent modal rank of 5th across all scenarios (Figure 3.6), and most consistently inferior performance observed under the High data quantity scenario, where the average rank was 4.63 (Table 3.4).

Choke effects both increased and reduced catch loss, depending on the assessment errors. On the one hand, there were several cases where AM underestimates of biomass and/or productivity produced TACs that produced a realised harvest rate much lower than the target associated with maximum multi-species yield. This behaviour was most prominent under Single-species AMs for the Low and Mod scenario, where the fished initialisation led the AMs to estimate a larger and less productive stock in equilibrium, estimating a steep early decline in biomass followed by apparent equilibration to a low biomass state where catch balances production (Figure 3.7), producing very low TACs despite the large positive error in unfished biomass. Low TACs of more catchable species constrained the TACs of the remaining species, increasing catch loss across the whole complex (Figure 3.8). On the other hand, when TACs pushed harvest rates very close to the target level, choke effects could reduce catch loss by constraining TACs that were too high compared to the target, protecting against overfishing relative to the multi-species harvest rate. These "protective" choke effects were observed under the pooling AMs and the hierarchical AM, and occured even when there were large assessment errors. For example, under the hierarchical AM in the Mod scenario, AMs perceived a larger and less productive stock than the OM, similar to the Single-species AM, but shrinkage priors on productivity parameters were able to reduce the magnitude of the error in U_{MSY} so that, despite having some of the largest errors in biomass forecasts and the B_{MSY}, the AM did not perceive a stock that quickly declined from its high initial biomass (Table 3.7, Appendix 3.7.E). Biomass estimates were then positively biased with respect to the operating model biomass, which were compensated by the negatively biased productivity estimates, producing more appropriately scaled TACs for all stocks, with choke effects occasionally protecting against overfishing. More optimal TACs then reduced catch loss overall, with median catch loss close to zero for the 2028 - 2037 period (Figure 3.10).

Biomass and overfishing risk

There was little chance of any stock being in an overfished state under any data scenario and AM combination (Table 3.4). No AM pushed stocks into a critically overfished state where $B_{s,p,t} < 0.4B_{MSY,SS,s,p}$ under the High and Mod data scenarios, while under the Low scenario a small probability of being critically overfished was observed under the Species Pooling (max 1%) and Spatial Pooling (max 6%) AMs. As expected, there were higher probabilities of being overfished ($B_{s,p,t} < 0.8B_{MSY,SS,s,p}$) under all scenarios, with the highest observed at 22% under the Spatial Pooling AM and Low scenario. The Single-species AM rarely pushed any stock into an over-fished state under any data scenario, as

biomass was usually underestimated for at least one species in each area, producing a choke constraint and higher catch loss as discussed above.

As expected, there was a higher risk of overfishing relative to single-species optimal fishing mortality rates $F_{MSY,SS,s,p}$ when attempting to maximise multi-species catch. Under the High and Mod scenario, Spatial Pooling and Single-species AMs were the only AMs with positive probabilities across the entire BC flatfish complex, with Spatial Pooling AMs overfishing more often. Unsurprisingly, overfishing was observed under all AMs for the Low scenario, reflecting the greater difficulty in estimating species productivity with low power data. Similar to the other scenarios and biomass risk above, the Spatial Pooling AM exceeded $F_{MSY,SS}$ more often (max 50%), while the Single-species AM was the most conservative given that it often underestimated biomass and caused choke constraints (max 1%). Hierarchical AMs were moderate compared to the other AMs, falling in the middle of the pack and exceeding $F_{MSY,SS}$ in 14% of years on average (max 29%).

Catch-Biomass trade-offs

Distributions of catch and biomass relative to MSY_{SS} and $B_{MSY,SS}$, respectively, were produced for the time period $2028 \leq t \leq 2037$ for each assessment model and Scenario combination. The medians of those relative catch and biomass distributions were visually compared to each other and to the central 95% of the omniscient manager's trajectories over the same time period (Figure 3.11) in order to understand the biomass and catch trade-offs between different model choices.

The Hierarchical Multi-species assessment model median catch between 2028 and 2037 came closest (i.e., smaller Euclidean distance) to the omniscient manager median catch under the Low data quantity scenario outside the HSHG area, and for HSHG Dover sole (Figure 3.11, compare points to the horizontal segment in black crosshair). For HSHG English and Rock soles, while the Hierarchical model tended to take a little more catch than the omniscient manager under the Low scenario, a large biomass surplus relative to the omniscient manager was present, indicating that realised harvest rates from TACs set by the hierarchical AM were smaller than those set by the omniscient manager.

Although the range of biomass-catch trade-offs were quite broad for each stock, the majority of Scenario/AM combinations lie inside the central 95% distributions of the omniscient manager (Figure 3.11, black crosshairs). Notable exceptions to this were the Single-species models for all species/areas under the High scenario and in HSHG under the Low scenario for Dover and English, the Hierarchical models in QCS and WCVI under the High scenario, and the Total Pooling method for WCVI Rock sole under the High data quantity scenario. As described above, the Single-species methods tended to under-estimate biomass and productivity under the Low scenario, so the tendency of the Single-species median biomass-catch to be towards the lower right corner of the range is expected. For the Single-species and Hierarchical methods under the High scenario, the reason is a combination of

large negative assessment errors in some or all years for choke species, and the 20% limit on increases in catch, producing a low-catch/high-biomass dynamic. Finally, the Total Pooling outlier in WCVI Rock Sole is caused by a persistent small negative assessment error, and an underallocation of the pooled TAC to WCVI Rock thanks to the differing Synoptic survey catchabilities. The Total Pooling method was also outside the central 95% of the omnisceint manager in HSHG under the High scenario, but only marginally so.

Catch-biomass trade-offs were approximately collinear under both High and Low data quantity scenarios for all stocks. All spawning stock biomasses were well above both the single-species and multi-species optimal levels at the beginning of the projection period (e.g., Table 3.1, Stock Status), meaning that the catch limits set by all methods were depleting a standing stock and benefited from its surplus production. Under these conditions, an increase in catch almost linearly caused a decrease in biomass as the compensatory effect of density dependence was minimal.

During the 2027 - 2036 period, median catch and biomass under the omniscient manager were higher than the optimal levels for both the multi-species and single-species maximum yield (green diamond crosshair and blue circle, respectively, Figure 3.11). As discussed above, this is because the transitionary period from fishery development to equilibrium dynamics is takes about 20 years under the omniscient manager. While catch is higher than MSY during the development period, this is not necessarily overfishing as the biomass is also higher than B_{MSY}; however, there was overfishing relative to single-species F_{MSY} occured for all stocks under the omniscient manager, and, as expected, with higher probability for all stocks where $B_{MSY,MS} \leq B_{MSY,SS}$ (Table 3.3).

3.3.3 Sensitivity of results to prior standard deviations

Average model sensitivities were summarised by fitting linear regressions to the distributions of median cumulative absolute catch loss (Figure 3.12). To remove the effect of absolute catch scales on the regressions, median loss distributions were standardised across assessment models, stratified by species, stock-area, and data scenario. Regressions with positive slopes had increasing catch loss with increasing uncertainty, and negative or zero slopes indicated a decrease or no change in catch loss with increasing prior uncertainty. Linear model slope parameter p-values for each AM were also calculated to determine if an effect was significant or not, where I define significance as $p < 0.1$.

While all AMs except the Total Pooling AM were slightly sensitive to MSY prior CVs under the high scenario, there was no change in the relative rankings of the AMs (Figure 3.12, first column). Under the Low scenario, only the Single-species and Species Pooling AMs were sensitive to MSY CVs, with catch loss dropping about a standard deviation as CVs decreased from 1.0 to 0.1 for both AMs, with Species Pooling AMs moving from fourth to second lowest mean catch loss, but the relative ranking of all other AMs remained the same.

Hierarchical and Single-species AMs were sensitive to the U_{MSY} prior standard deviation under both scenarios, and the Species Pooling AM was also sensitive under the Low scenario (Figure 3.12, second column, $p > 0.05$). However, under the High scenario, no sensitivity led to a change in ranking for any AMs over the range of U_{MSY} prior standard deviations tested, as most AMs showed a catch loss change of less than 0.5 standard deviations, except for the Single-species AM, which was again an outlier. Under the Low scenario, the Total Pooling and Single-species AMs remained at the bottom of the rankings, despite a decrease off around 1.5 standard deviations for the Single-species AM. At the same time, Spatial Pooling, Species Pooling, and Hierarchical AMs catch losses under the Low scenario appeared to converge as prior standard deviations increased from 0.1 to 1.0.

Interestingly, the Hierarchical Multi-species assessment models reacted differently to changing catchability and U_{MSY} hierarchical shrinkage priors under the High and Low scenarios (Figure 3.12, third column). As $\sigma_{U_{MSY}}$ and τ_q increased from 0.1 to 0.5, average catch loss increased by about 1 standard deviation under the High scenario ($p < 0.01$), mostly driven by assessment errors in the QCS and WCVI areas, leading to a widening range of catch trajectories as hierarchical prior standard deviations increased. Under the Low scenario, median absolute catch-loss dropped by about 1 standard devation, indicating that TACs were closer to the omniscient manager as hyper-prior SDs increased. The improved performance with increasing hyper-prior uncertainty was caused by a combination of subtle effects, with some assessments becoming more biased, and others less biased, which occasionally switched the choke species in an area. The combination of raised and lowered catch limits and technical interactions produced beneficial choke effects, reducing catch loss overall.

Finally, there was low sensitivity to the reduction in Synoptic survey standard errors (Figure 3.12, fourth column). Under the High scenario, the all but the Spatial Pooling AM had significant improvements of around 0.3-0.5 standard deviations with decreasing standard error ($p < 0.05$), while the Spatial pooling had a non-significant change ($p = 0.18$). Under the Low scenario, only the Species Pooling AM catch loss improved significantly, dropping around 0.25 standard deviations with decreasing Synoptic survey standard error ($p < 0.01$), and switching places with the Hierarchical AM for the lowest average catch loss. For both scenarios, the decreases in catch loss appear to be caused by a slight improvement in pooled TAC allocation with increased survey precision, but choke constraints were still present.

3.4 Discussion

In this chapter, I demonstrated that hierarchical stock assessment models may improve management performance in a spatially-replicated multi-species flatfish fishery. When available data quantity was moderate or low (indicated here by time-series length), biomass and har-

vest rate estimates from hierarchical stock assessment models resulted in catches that were closer to an omniscient manager's optimal reference series compared to catch limits derived from single-stock and data-pooling assessment methods. Under high data quantity scenarios, data-pooling methods outperformed hierarchical models, but the latter still outperformed single-stock assessment methods. This suggests that hierarchical assessment methods could be a better approach to making catch limit decisions than conventional single-species methods under typical fisheries data quality conditions, such as short or noisy time series of fishery-independent observations or uninformative catch series.

Ranking performance according to catch loss highlighted a fundamental trade-off between multi-species and single-species fishery objectives. Management performance was measured via comparison to an omniscient manager simulation with the objective of maximising complex yield, implying that the simulated BC flatfish complex management objectives were targeting the multi-species optima derived from technical interactions. While targeting multi-species optima does increase complex yield overall, there are cases where the fishing mortality rates exceed single-species F_{MSY}, which is defined as overfishing in a single-species context. An obvious question is, would the rankings continue to hold if catch limits targeted individual stocks' single-species U_{MSY}? Further, would the rankings change if ranking of AM performance was based on the risk of overfishing, or being in an overfished state, relative to single-species reference points? Targeting single-species harvest rates would likely lead to higher instances of choke effects since technical interactions are being ignored. Further, targeting single-species yields would probably not affect the rankings based on AM performance, since the estimated optimal harvest for hierarchical models had compensatory biases in biomass and harvest rates which would hold for the single-species harvest rates as well, while single-species and pooled models lacked those compensatory biases, and pooled models had issues achieving target TACs given the allocation mode. If ranking by the probability of overfishing or being in an overfished state, our results show that under the High and Mod scenarios the Hierarchical model would still tie for first, but would fall to third under the Low scenario, with the single-species model taking first place. However, even at the highest probability of exceeding single-species F_{MSY} under the hierarchical AMs, there was no chance of pushing stocks into a critically overfished state, and at most 6% chance of pushing any BC flatfish complex stocks into an overfished state of less than 80% of $B_{MSY,SS}$, which would be acceptable management performance in most fisheries. Furthermore, choosing single-species AMs would forgo a large amount of catch, which may align with the objectives of some stakeholders, but probably not harvesters themselves. This highlights that measuring multi-species fisheries performance according to single-species objectives (i.e., reference points) may be overly conservative, and that while some overfishing may be required to achieve multi-species yield objectives, that overfishing can be sustainable. While it is certainly possible to come up with a set of commercial catchabilities that would produce critically overfished stocks while maximising complex yield, that was not

the case for the DER complex simulated here, and it is unclear if such a mix would be a realistic scenario.

The above results arise from models that are necessarily a simplification of the real stock-management system. The harvest rules applied to DER complex species were relatively simple and may require more detail or complexity for practical applications. All harvest rules were constant target harvest rates, which do not include precautionary "ramping-down" of catch towards a limit biomass level [34, 28]. Including a ramped harvest rule would reduce the probability of stocks being critically overfished in some cases, probably at some further cost of choke effects, but there was very little chance of critically overfishing any stock under our simulations anyway. Second, catch limits for the simulated BC flatfish complex were set based on estimated target harvest rates that were scaled by *a priori* known scalars derived from multi-species yield curves, which may positively bias results towards lower catch loss in general. Incorporating multi-species yield curve calculations to the assessment model ouput into the harvest decision would be simple to do, but would require either a model of increased complexity to link fishing effort to single-species yield, or an extra assumption linking effort to surplus production model yield calculations, which would likely increase assessment model errors. Finally, the TAC allocation model for data-pooled methods was only one example from a large set of potential options. Understanding the relative risks of data-pooling would require testing alternative allocation methods, which was beyond the scope of this paper.

Replacing commercial fishery catch rates with relative biomass in the simulations increased the statistical power of commercial index data and skewed results under the High scenario. Under previous versions of the simulations where commercial indices were simulated as catch rates, but without observation error, the Hierarchical method ranked lower than the three Pooling AMs under the High scenario, but still higher than the Single-species model. The change in performance is largely caused by learning or targeting behaviour of harvesters in Historical (1956 - 1996) trawl fishery that caused indices to increase for some BC flatfish complex stocks while estimated biomass was decreasing in the model, indicating time-varying catchability for that fleet. Time-varying catchability was less of an issue for Pooled AMs because data pooling is intended to increase sample size and reduce variability. Replacing catch rates with relative biomass indices simulated with a constant catchability reduced the advantage of data pooling under the High scenario, allowing Hierarchical models to achieve the lowest catch loss under all three scenarios.

The only multi-species interactions considered were technical interactions, which although an important part of exploited system dynamics, are not the entire story. Although there is limited evidence for ecological interactions among BC flatfish complex species [108, 142], what does exist may influence the multi-species yield relationship with fishing effort or, as with technical interactions, inhibit the ability of the management system to meet target catch levels. For example, individual survival or growth may change in response

to varied fishing pressure due to unmodeled linkages [20]. Yet, including such ecological interactions would imply a highly data rich scenario, which is counter to our focus on surplus production models applied to multi-species fisheries. Furthermore, accounting for potential ecological interactions would require multiple OMs to test performance against a range of plausible hypotheses, since ecological uncertainties are much broader in complexity and scope than technical interactions alone. Nevertheless, future work combining technical interactions with minimum realistic models for ecological interactions could help determine the extent to which assessment approaches affect these more complex multi-species fisheries outcomes [110]. For example, while diet overlap between the three species is small off the coast of Oregon, the major Rock sole prey was recently settled pleuronectiform fishes, which may include Dover and English sole young and therefore shift the complex equilibrium as fishing pressure is applied, reducing predation mortality for Dover and English sole young, and reducing prey availability for Rock sole [142, 20].

The BC flatfish complex trawl fishery effort model was also a simplification of reality, where effort was limited only by the TACs in each area. Limiting by TACs was intended to reflect the management of the real BC groundfish fishery in which harvester decisions drive TAC utilisation among target species [115, via increasing catchability], and non-target or choke species [10, via decreasing catchability]. Changing catchability for targeting or avoidance could be simulated as a random walk in the projections, with correlation and variance based on the historical period, or perhaps simulated via some economic sub-model that accounted for ex-vessel prices and variable fishing costs. These economic factors could affect targeting and avoidance behaviour among species [115, 114], as well as effort allocation among stock-areas [59, 145]; however, it is not clear that the median results above would be significantly different given the potential magnitude of assessment model errors in data-limited scenarios. Impacts of a detailed effort dynamics sub-model would probably be more important in more extreme data-limited scenarios that relied solely on fishery CPUE as an index of abundance, which I did not test here. In fact, it would be interesting to determine whether the hierarchical information-sharing approach would exacerbate assessment model errors in the (common) context where fishery CPUE is the main abundance index.

Our assessment models were all different versions of a state-space surplus production model, and rankings of AMs may vary when other model configurations with more biological realism are included. For example, the aggregate productivity parameter U_{MSY} could be separated into growth, natural mortality, and stock-recruit steepness by using a delay-difference or age-structured model formulation [32, 128, 47], which may partially offset the advantages of data-pooling and hierarchical assessment methods exhibited above. However, in the contexts where biological data are missing or of low sample size, there would be even greater reliance on strong *a-priori* assumptions for additional parameters in models of higher complexity, which I predict would lead to similar results.

Despite the limitations above, this chapter's results indicate that even in fisheries with long time series of catch and effort data, hierarchical multi-species assessment models may be preferable over typical single-species methods. The poor performance of the single-species models in all scenarios highlights the difference between data-rich (i.e., a higher quantity of data) and information-rich (i.e., data with higher statistical power) fisheries. The high data quantity scenario differed from the moderate and low scenarios by the inclusion of a historical series of fishery dependent CPUE, which was quite noisy and subject to the effects of changing harvester behaviour like targeting (variable catchability), and therefore, additional historical CPUE data had little effect on cumulative catch loss under the single-species models. In contrast, the data-pooling procedures all ranked higher than single-species and multi-species models under the data-rich scenario, as they were able to leverage additional statistical power from the historical CPUE by effectively increasing the sample size through data aggregation. The superior performance of the hierarchical model over the single-species model under the high data scenario indicate that shared priors partially compensate for low statistical power when setting TACs, but not as much as data-pooling.

Superior management performance of the hierarchical models was primarily caused by compensatory (negatively correlated) biases in biomass and productivity. Biases in biomass estimates were comparable between Single-species and Hierarchical models, but the hierarchical shrinkage prior structure defined for productivity parameters led to target harvest rate estimates that, while biased, combined with biased biomass estimates to produce TACs that were closer to the omniscient manager's. Therefore, while improved performance relative to single-species models under lower data quantity conditions is consistent with our previous study, where statistical performance of hierarchical multi-stock assessments improved with decreasing data quantity and quality [70], the improved management performance of Hierarchical methods was due to a fundamentally different mechanism. This difference may be explained by a different assessment model parameterisation, increased complex size, and a different and simplified experimental design.

The benefits of parameter shrinkage induced by the hierarchical stock assessments relies on the similar life histories of BC flatfish complex species. Similar life histories allow joint distributions of productivity and catchability parameters to be more precise, drawing estimates close to the prior mean when data have low statistical power. Using hierarchical models to improve individual stock (replicate) parameter estimates assumes that those parameters are exchangeable, that is, any permutation of the replicates would leave the joint probability distribution of those parameters unchanged [138, 51]. If, for example, a species with a very different life history was included, such as a slow growing and late-maturing rockfish from the Sebastes genus, then that species would have much lower productivity (U_{MSY}) than any of the BC flatfish complex flatfish, and would likely have a much different survey trawl efficiency as well, depending on its preferred habitat. Such disparate life histories and catchabilities would require a decrease in the precision of the joint distri-

bution of the hierarchically modeled parameters to preserve exchangeability, and therefore may reduce or completely eliminate the benefits of parameter shrinkage for any complex containing species with disparate life histories.

Hierarchical and Single-species models were sensitive to changes in prior precision for assessment model productivity parameters under the Low data quantity scenario. The Hierarchical model went from lowest median catch loss (i.e., ranked first) to highest (ranked last) as prior precision on the complex mean productivity was reduced, indicating that the compensatory biases that gave hierarchical models the advantage under low data conditions were dependent on this prior. Under the same scenario, the Single-species method improved slightly as prior precision on the stock-specific productivity parameter was reduced, allowing more compensatory bias into the harvest rate estimates. Similarly, under the same low-data scenario, Hierarchical models achieved significantly lower catch loss as hierarchical shrinkage prior precision was reduced. Lower catch loss under reduced hierarchical precision could be attributed to higher variability in catchability and productivity estimates, allowing the stock-specific estimates to achieve more optimal TACs despite the constrained complex or species mean values.

The data-pooled methods performed better under the High scenario and were generally insensitive to priors, indicating that the data were more influential than priors on the TACs and allocation. This may be because data-pooled observation errors are biased low, being simulated independently of the observation errors in the component indices, and using the average standard deviation of the components. If aggregate indices pooled errors from each component index, then the resulting observation error variance would be additive in the components, especially if those errors were positively correlated, which may be the case under a common survey or fishery.

One might expect that data-pooled methods would outperform other methods as future precision in the Synoptic survey was increased, amplifying the pooling advantage, but this was not the case. The lack of dominance by the pooled methods was caused by the allocation model and, when pooled over space, the pooling method. The allocation model allocated TAC in proportion to Synoptic survey index, which was biased away from absolute stock size by the Synoptic trawl efficiency parameter, leading to under-allocation for some species/areas and choke constraints. When data were pooled over areas, the alternating biennial observations caused a sawtooth pattern as low and high biomass areas dominated the pooled index, which inflated the observation errors independent of the survey precision.

I showed that choke effects are not a uniformly negative outcome for multi-species fisheries, and may indicate a mismatch between the target harvest rate and optimal complex yield. The usual assumption is that choke species restrict access to fishing grounds, decreasing profitability through lost yield of target species, and higher quota prices for choke species [93]; however, we found that choke species sometimes prevented overfishing when TACs for the non-choke species were set too high, allowing harvest strategies to meet multi-species

objectives despite large assessment errors for individual species in the complex. In reality, choke effects would likely be lessened by changing species catchability via harvester targeting and avoidance, creating a more complex relationship between effort and complex yield; but, the existence of a choke species would still indicate a mismatch between an individual species' TAC and the optimal exploitation rate for meeting the management objectives for the multi-species complex.

3.4.1 Conclusion

Hierarchical multi-species surplus production assessment models can outperform single-species production models in meeting multi-species harvest objectives across high, moderate, and low data quantity scenarios, while avoiding states of conservation concern with high probability. While hierarchical model estimates of biomass used for setting TACs were often more biased than those produced by other methods, negatively correlated bias in biomass and productivity was better matched under hierarchical models than the other methods as data quantity was reduced, translating into better management performance across the multi-species flatfish fishery. I recommend that assessment and management of multi-species fisheries include hierarchical models that acknowledge technical interactions when designing harvest strategies and management procedures for data-limited, multi-species fisheries. Otherwise, management procedures based on single-species approaches that rely heavily on prior knowledge (inducing bias) and ignore technical interactions (making objectives impossible to achieve) may give a misleading picture of the expected management performance in multi-species, spatially heterogeneous fisheries.

3.5 Tables

Table 3.1: Unfished biomas B_0, single-species MSY based reference points $B_{MSY,SS}$, MSY_{SS}, and $U_{MSY,SS}$, stock status as absolute biomass in 2016 B_{2016}, depletion relative to single-species optimal biomass $B_{2016}/B_{MSY,SS}$, commercial trawl catchability scalar q^F, and multi-species reference points including technical interactions $B_{MSY,MS}$, MSY_{MS} and $U_{MSY,MS}$ for all nine BC flatfish complex stocks in 2016. Biomass quantities are given in kilotonnes, and depletion levels and harvest rates are unitless.

		SS Reference Points			Stock Status			MS Reference Points		
Stock	B_0	$B_{MSY,SS}$	MSY_{SS}	$U_{MSY,SS}$	B_{2016}	$B_{2016}/B_{MSY,SS}$	q^F	$B_{MSY,MS}$	MSY_{MS}	$U_{MSY,MS}$
Dover sole										
HSHG	16.51	4.34	1.22	0.28	8.36	1.92	0.022	5.77	1.18	0.20
QCS	5.45	1.46	0.42	0.28	3.36	2.30	0.015	1.91	0.40	0.21
WCVI	13.59	3.58	1.14	0.32	8.45	2.36	0.025	3.60	1.14	0.32
English sole										
HSHG	8.60	2.21	0.87	0.39	4.85	2.20	0.026	2.31	0.87	0.38
QCS	0.57	0.15	0.06	0.39	0.41	2.78	0.016	0.16	0.06	0.36
WCVI	0.86	0.22	0.09	0.39	0.50	2.23	0.020	0.20	0.09	0.43
Rock sole										
HSHG	12.34	3.78	1.07	0.28	7.68	2.03	0.025	2.81	1.03	0.37
QCS	4.33	1.29	0.40	0.31	2.13	1.65	0.014	1.04	0.39	0.38
WCVI	1.12	0.34	0.10	0.29	0.55	1.62	0.012	0.35	0.10	0.28

Table 3.2: Summary of sensitivity analyses, showing the total number of experiments, the factor being varied, the levels of that factor, and the data scenarios and AMs included in the analysis.

N	Factor	Levels	Scenarios	AMs
30	MSY prior CV	$0.1, 0.5, 1.0$	High, Low	All
30	U_{MSY} prior SD	$0.1, 0.5, 1.0$	High, Low	All
6	Hierarchical prior SDs τ_q, $\sigma_{U_{MSY}}$	$0.1, 0.2, 0.5$	High, Low	Hierarchical only
30	Synoptic survey SD τ	0.1,0.5,1.0	High, Low	All

Table 3.3: Probability of being overfished and experiencing overfishing with respect to single-species reference points, and catching less than the historical miniumum during the time period 2028 - 2037 for all nine BC flatfish complex stocks when managed by the omniscient manager.

	Prob. of being overfished		Prob. of overfishing	Prob. of low catch
Stock	$P(B_t < .4B_{MSY,SS})$	$P(B_t < .8B_{MSY,SS})$	$P(F_t > F_{MSY,SS})$	$P(C_t < \min C_{1951:2016})$
Dover sole				
HSHG	0.00	0.01	0.07	0.01
QCS	0.00	0.01	0.02	0.01
WCVI	0.00	0.07	0.54	0.01
English sole				
HSHG	0.01	0.10	0.39	0.02
QCS	0.01	0.05	0.20	0.02
WCVI	0.01	0.23	0.78	0.02
Rock sole				
HSHG	0.01	0.59	0.91	0.01
QCS	0.01	0.41	0.84	0.02
WCVI	0.01	0.08	0.42	0.02

Table 3.4: Summary of AM catch loss rankings and biomass risk under each scenario. The catch loss rankings show the modal and average ranks, calculated across replicates, species, and stocks. Biomass risk columns show the average probability, with range in parentheses, of stocks being in a critically overfished (i.e. below $0.4B_{MSY,SS}$) or overfished (i.e., below $0.8B_{MSY,SS}$) state, calculated across species and stocks. Within a scenario, AMs are ordered by average rank.

	Catch Loss Rankings		Biomass		Fishing Mortality
AM	Modal Rank	Average Rank	$P(B_t < 0.4B_{MSY,SS})$	$P(B_t < 0.8B_{MSY,SS})$	$P(F_t > F_{MSY,SS})$
High					
Hierarchical Multi-species	1	2.08	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Species Pooling	2	2.35	0.00 (0.00, 0.00)	0.00 (0.00, 0.01)	0.03 (0.00, 0.08)
Spatial Pooling	4	2.88	0.01 (0.00, 0.02)	0.03 (0.00, 0.09)	0.13 (0.01, 0.29)
Total Pooling	4	3.07	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.02)
Single-species	5	4.63	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Mod					
Hierarchical Multi-species	1	2.30	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Species Pooling	1	2.40	0.00 (0.00, 0.00)	0.01 (0.00, 0.02)	0.08 (0.02, 0.15)
Total Pooling	3	2.88	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.01 (0.00, 0.04)
Spatial Pooling	4	2.99	0.01 (0.00, 0.04)	0.04 (0.00, 0.12)	0.16 (0.02, 0.38)
Single-species	5	4.42	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)
Low					
Hierarchical Multi-species	1	1.94	0.00 (0.00, 0.00)	0.02 (0.00, 0.06)	0.12 (0.00, 0.25)
Species Pooling	3	2.70	0.01 (0.00, 0.01)	0.03 (0.00, 0.06)	0.14 (0.05, 0.25)
Total Pooling	2	3.12	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.01 (0.00, 0.05)
Spatial Pooling	4	3.45	0.02 (0.00, 0.04)	0.06 (0.00, 0.15)	0.21 (0.04, 0.44)
Single-species	5	3.78	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.01)

3.6 Figures

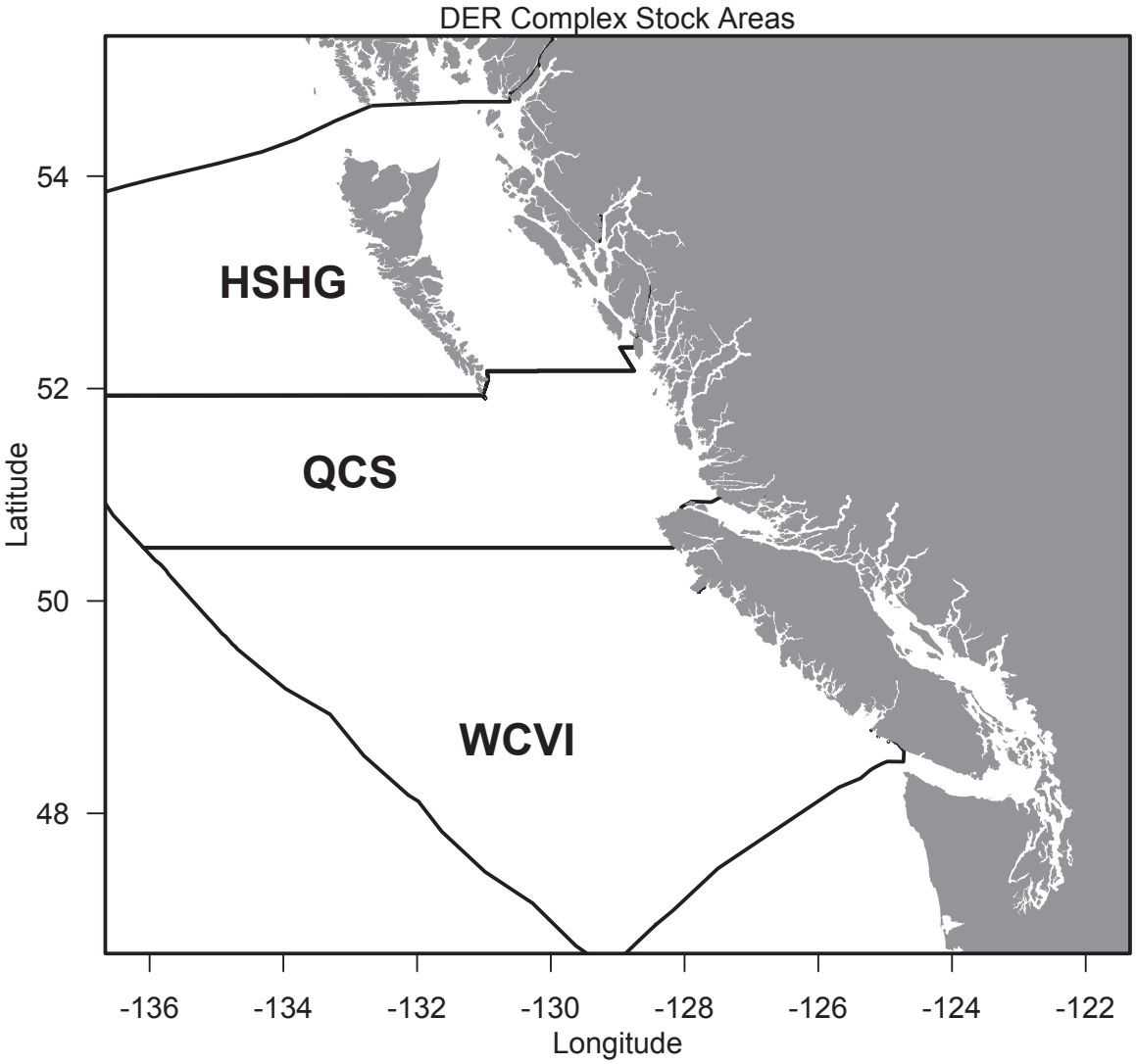

Figure 3.1: Boundaries for each of the BC flatfish complex stock areas on the BC coast, showing, from north to south, Hecate Strait/Haida Gwaii (HSHG), Queen Charlotte Sound (QCS), and West Coast of Vancouver Island (WCVI).

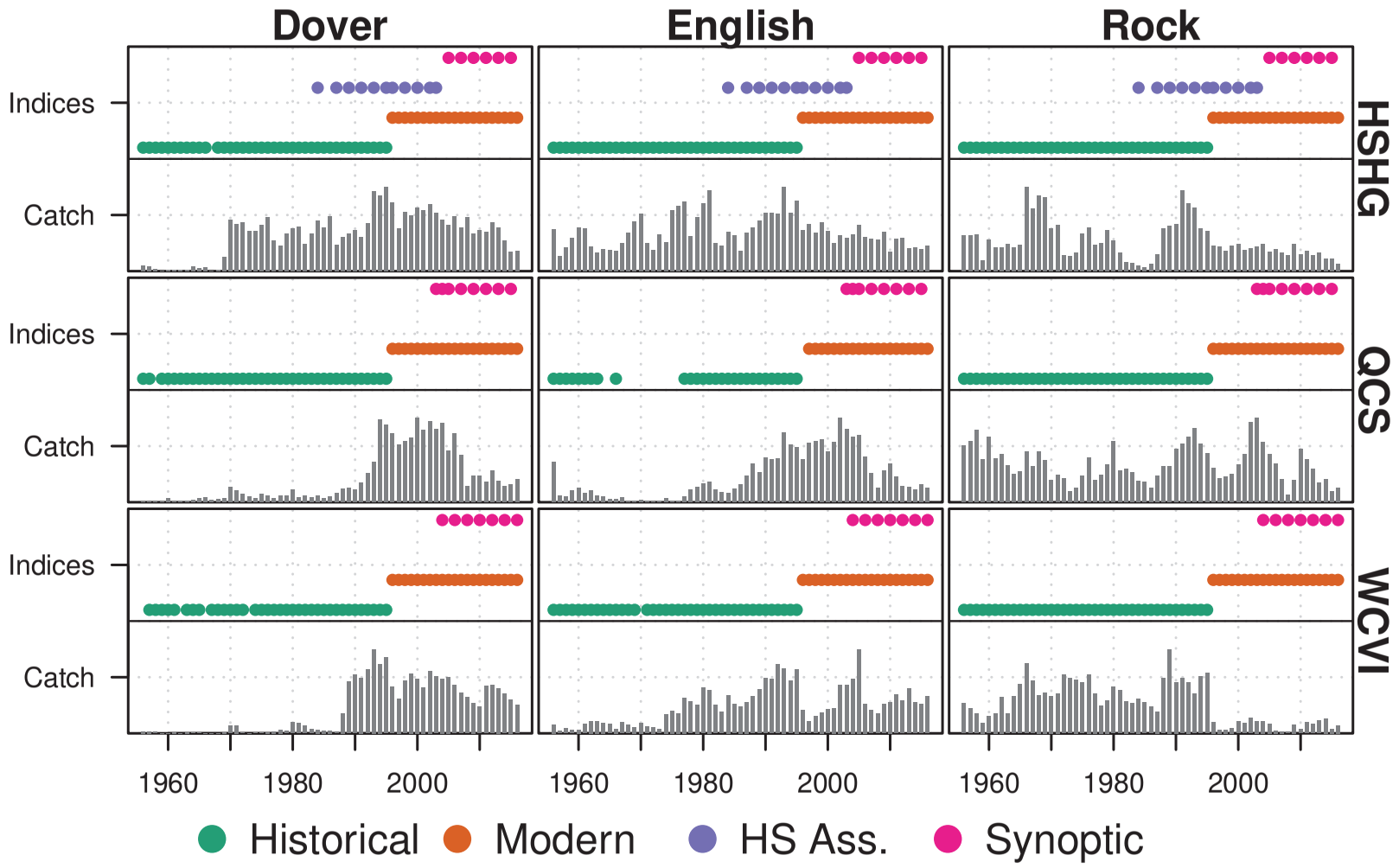

Figure 3.2: A summary of BC flatfish complex data available for fitting the assessment model. Panels show species from left to right, and rows show stock areas from north to south. Each panel is split into two cells, with the top showing a coloured filled circle when a commercial CPUE or biomass index observation exists, with the colours corresponding to particular fleets as labeled in the legend. The bottom cell shows commercial catch relative to the maximum catch over the model history in grey bars.

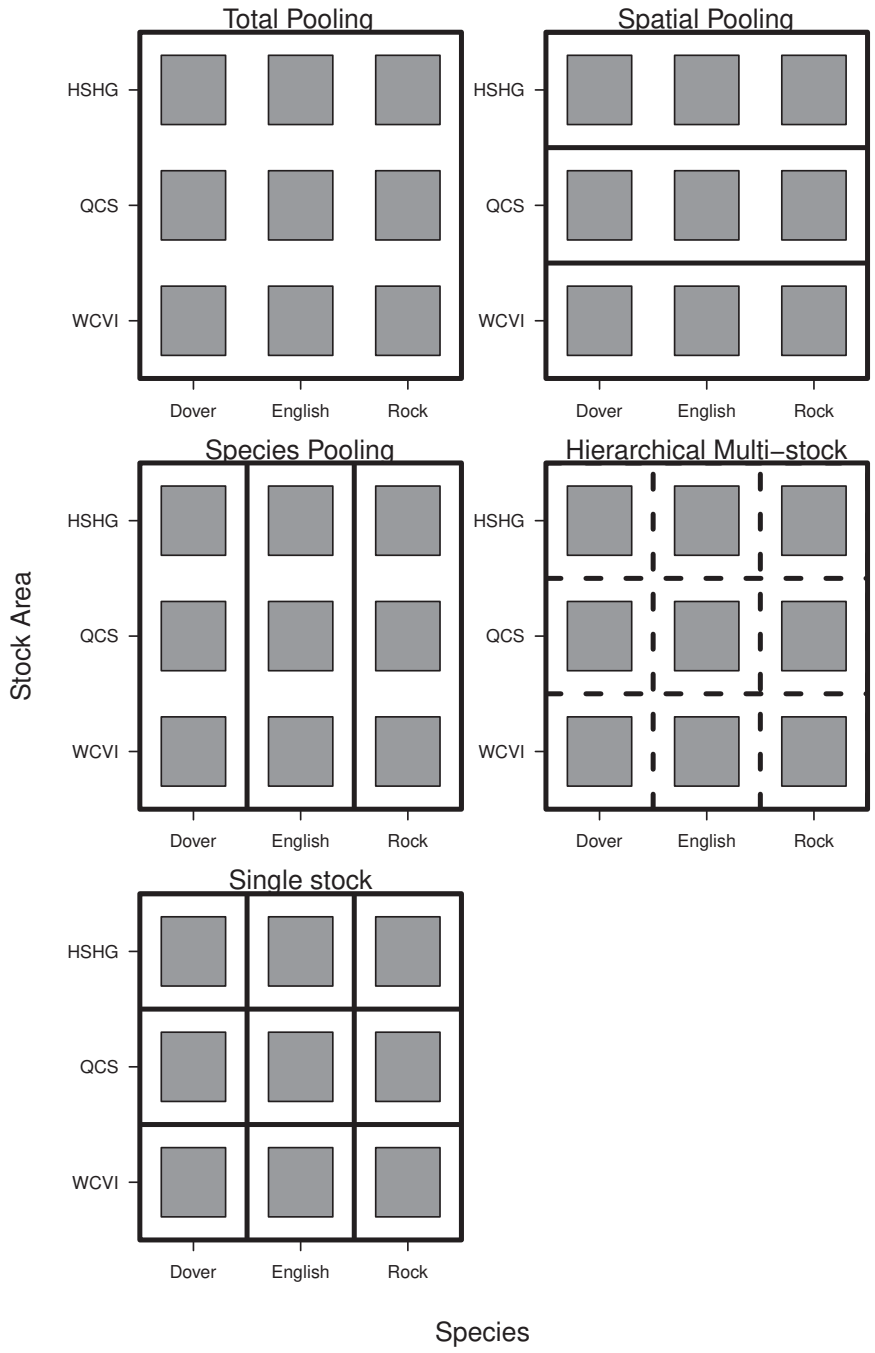

Figure 3.3: Conceptual models of the five assessment model configurations. In each panel, the nine grey boxes represent each BC flatfish complex population, as indicated by the axis labels. Data are pooled for any population not separated by a black line, e.g., all nine are pooled in the total pooling case. In the hierarchical model, the broken lines indicate that data are separated, but information is shared between populations via the statistical model's hierarchical prior distributions.

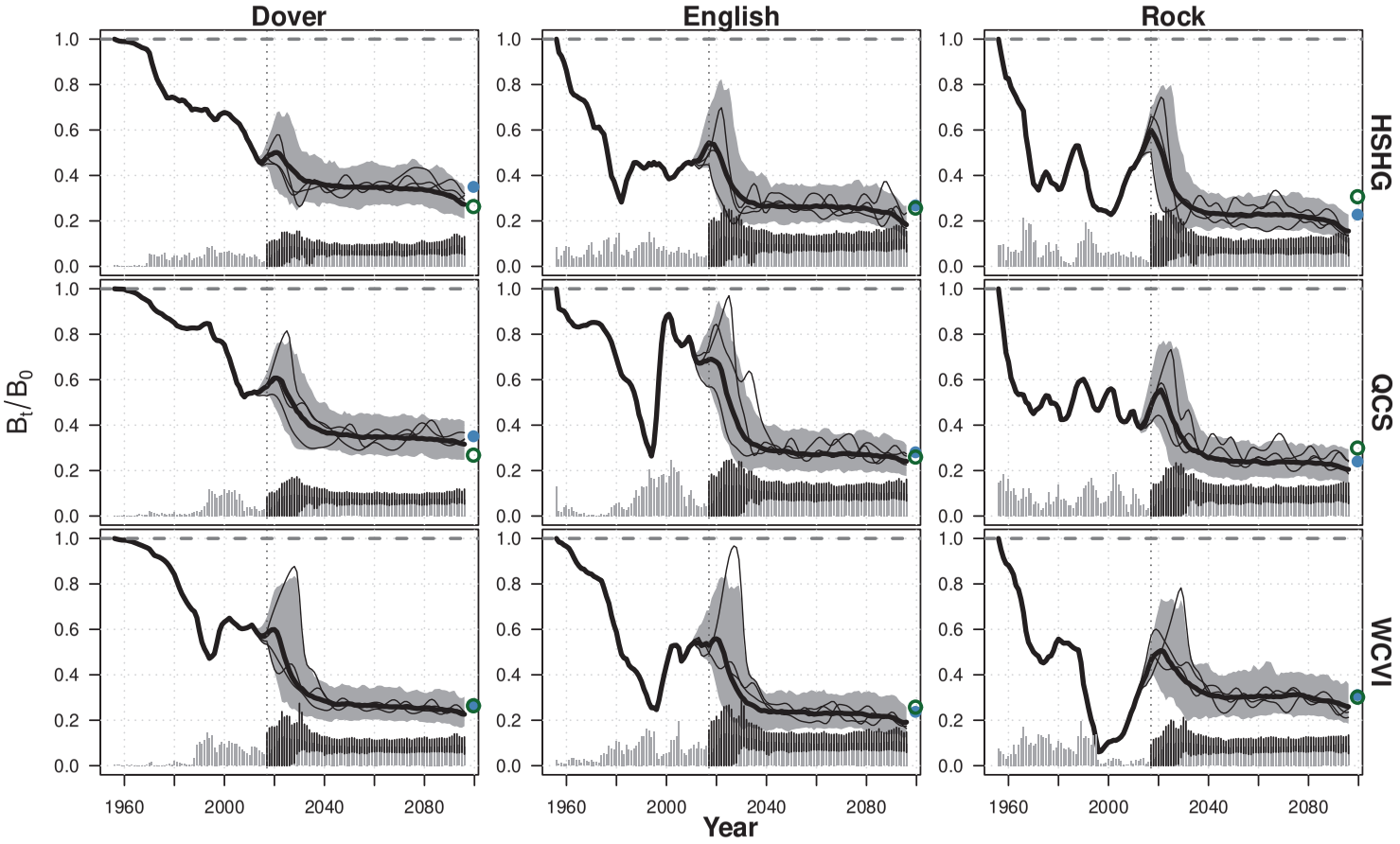

Figure 3.4: Spawning biomass depletion and relative catch simulation envelopes for all nine BC flatfish complex management units from the omniscient manager simulations. Median biomass is shown by the thick black line, with the grey region showing the central 95% of the distribution of spawning biomass, and thin black lines showing three randomly selected simulation replicates. Catch is shown as grey bars in the historical period, which represent median catch in the projection, with thin vertical line segments showing the central 95% of the catch distribution. The depletion level associated with the traditional single species optimal biomass $B_{MSY,SS}$ is shown as an open green circle on the right-hand axis, while the depletion level associated with the multi-species maximum yield is shown as a blue closed circle on the right-hand axis.

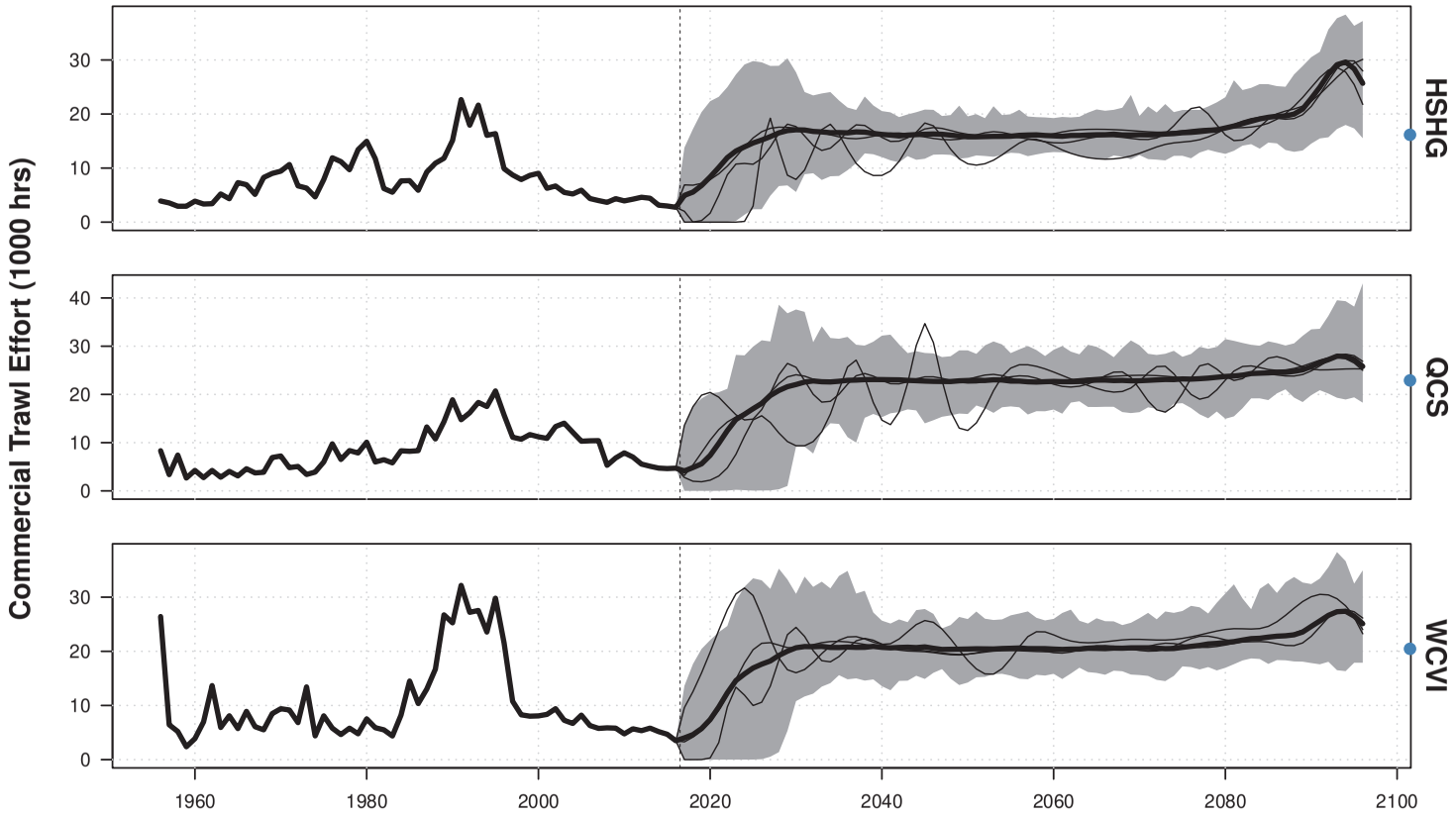

Figure 3.5: Commercial fishing effort simulation envelopes for each stock area. Historical and median simulated effort in the projection period are shown by a thicker black line, while the central 95% of the distribution of simulated effort in the projection period is shown as grey shaded region, with single simulation replicates shown as thinner black lines. The effort level that achieves multi-species maximum yield is shown as a closed blue circle on the right-hand axis.

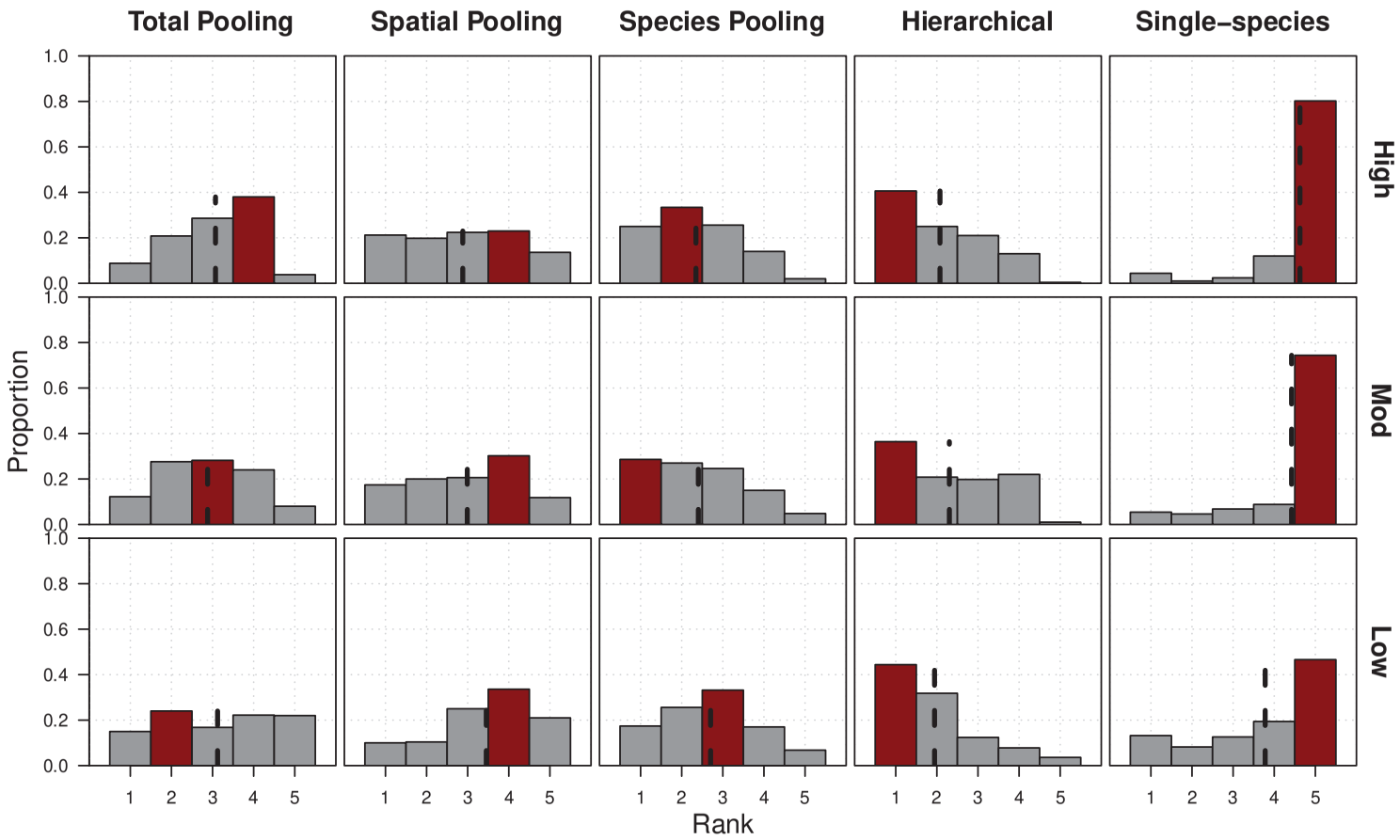

Figure 3.6: Distributions of AM rankings by cumulative absolute loss in catch (kt) for the projection years 2028 to 2037 under each assessment model (columns) and OM data scenario (rows). Ranks are calculated for each species/stock combination within a replicate, then distributions of ranks are across species, stocks, and replicates that met the 95% convergent AMs criterion. The modal rank is shown as a dark red bar, and the average rank as a vertical dashed line.

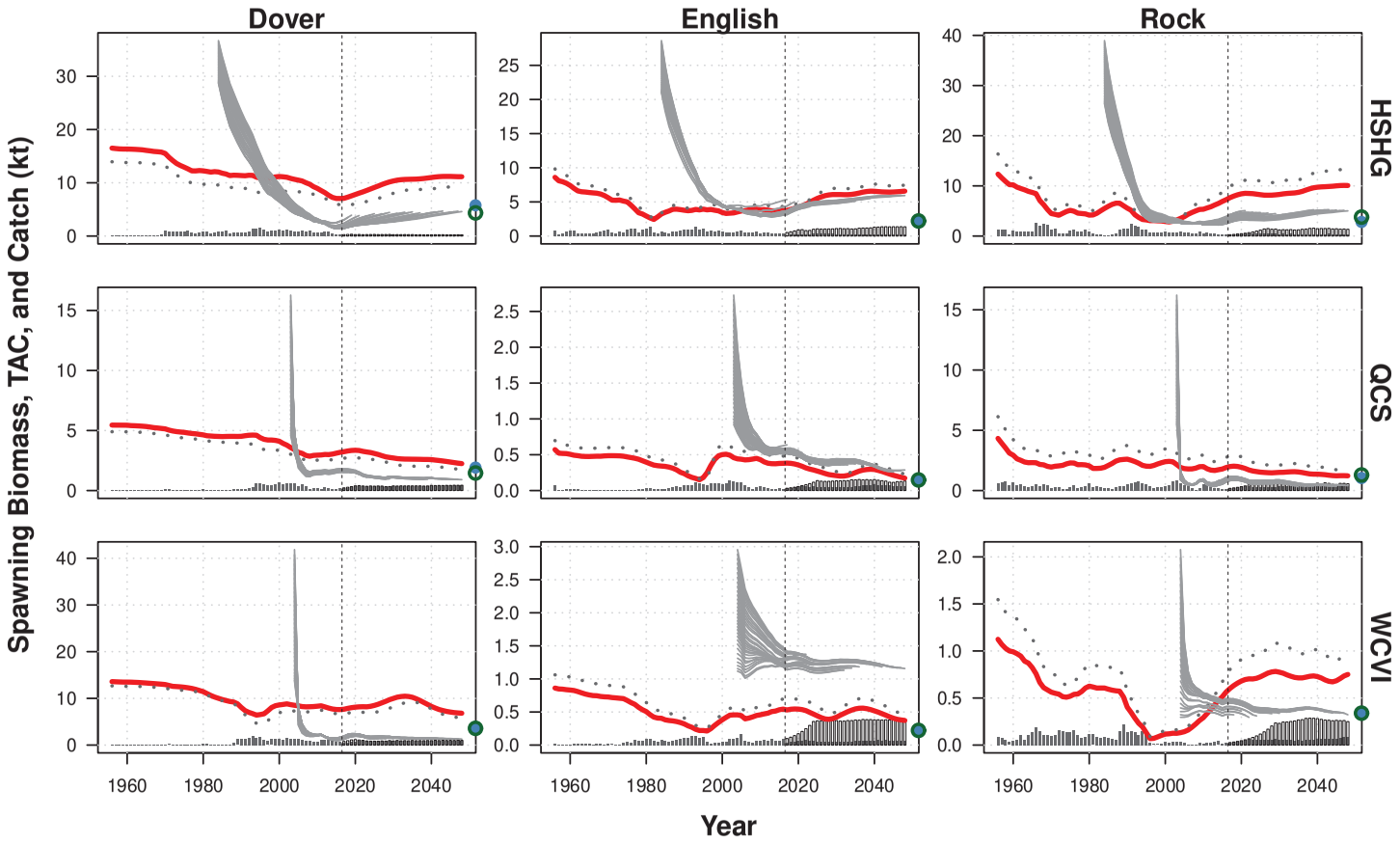

Figure 3.7: Operating model spawning stock biomass (red line), commercial trawl vulnerable biomass (grey dotted line), retrospective assesment model estimates of spawning stock biomass (thin grey/purple lines), and catch and TACs (grey bars) from the first simulation replicate in the Low data-quality scenario and under the Single-stock assessment model. Catch bars show realised catch in grey for the whole simulation period, and unfilled bars in the projection period show the difference between MP set TACs and realised catch. Coloured circles on the right hand vertical axis show the biomass level associated with the multi-species (closed blue circle) and single-species (open green circle) maximum sustainable yield. Purple lines indicate an assessment with a non-positive-definite Hessian matrix.

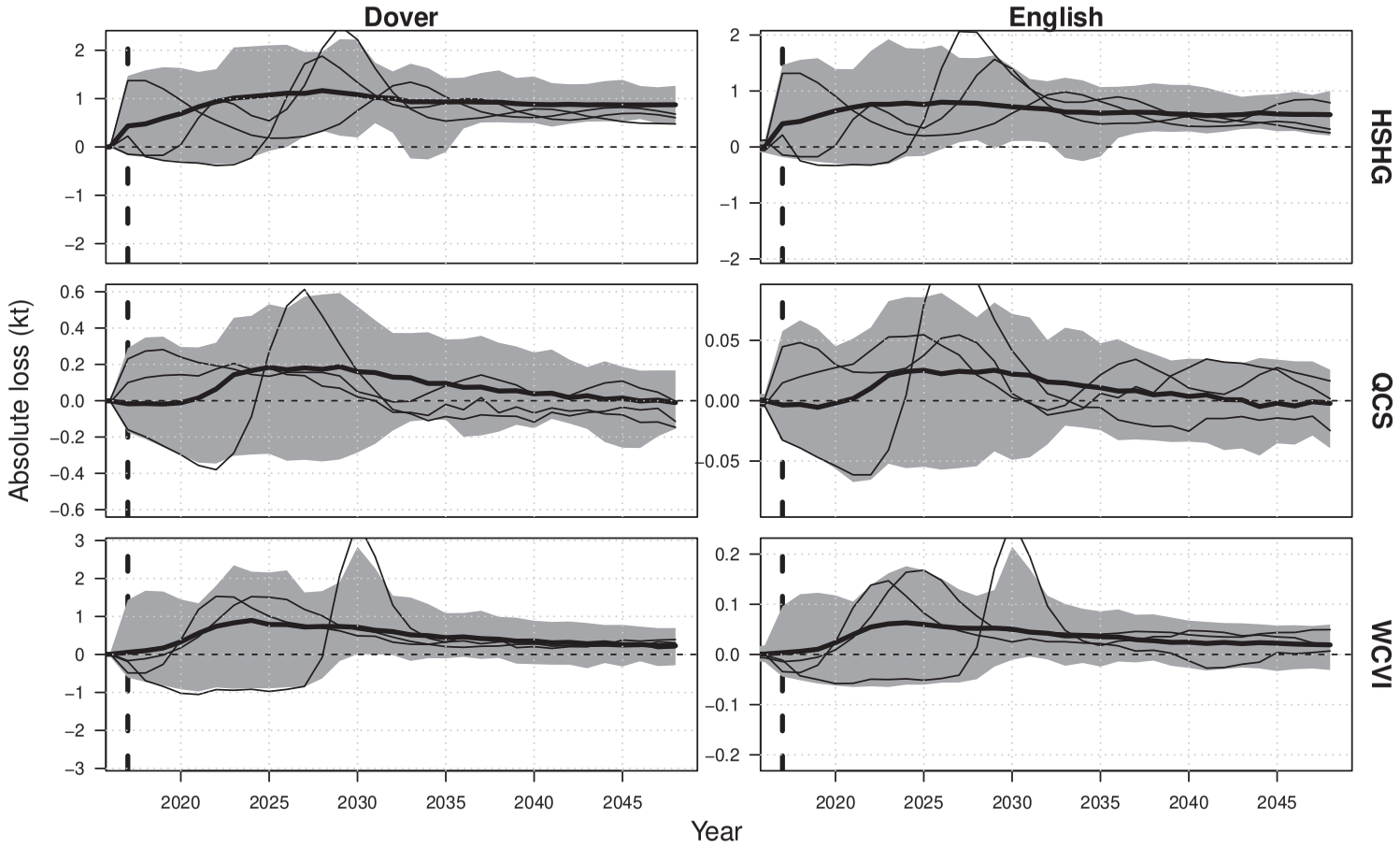

Figure 3.8: Catch loss simulation envelopes for the six BC flatfish complex management units included in the catch loss ranking, when assessed using the Single-species assessment model under the Low data-quality scenario. Median catch loss is shown by the thick black line, the central 95% of the catch loss distribution is shown as the grey shaded region, and the thin black lines show three randomly selected simulation replicates. The thick vertical dashed line shows the beginning of the projection period, and the horizontal thin dashed line shows zero catch loss.

Figure 3.9: Operating model spawning stock biomass (red line), commercial trawl vulnerable biomass (grey dotted line), retrospective assesment model estimates of spawning stock biomass (thin grey/purple lines), and catch and TACs (grey bars) from the first simulation replicate in the Low data-quality scenario and under the Hierarchical Multi-species assessment model. Catch bars show realised catch in grey for the whole simulation period, and unfilled bars in the projection period show the difference between MP set TACs and realised catch. Coloured circles on the right hand vertical axis show the biomass level associated with the multi-species (closed blue circle) and single-species (open green circle) maximum sustainable yield.

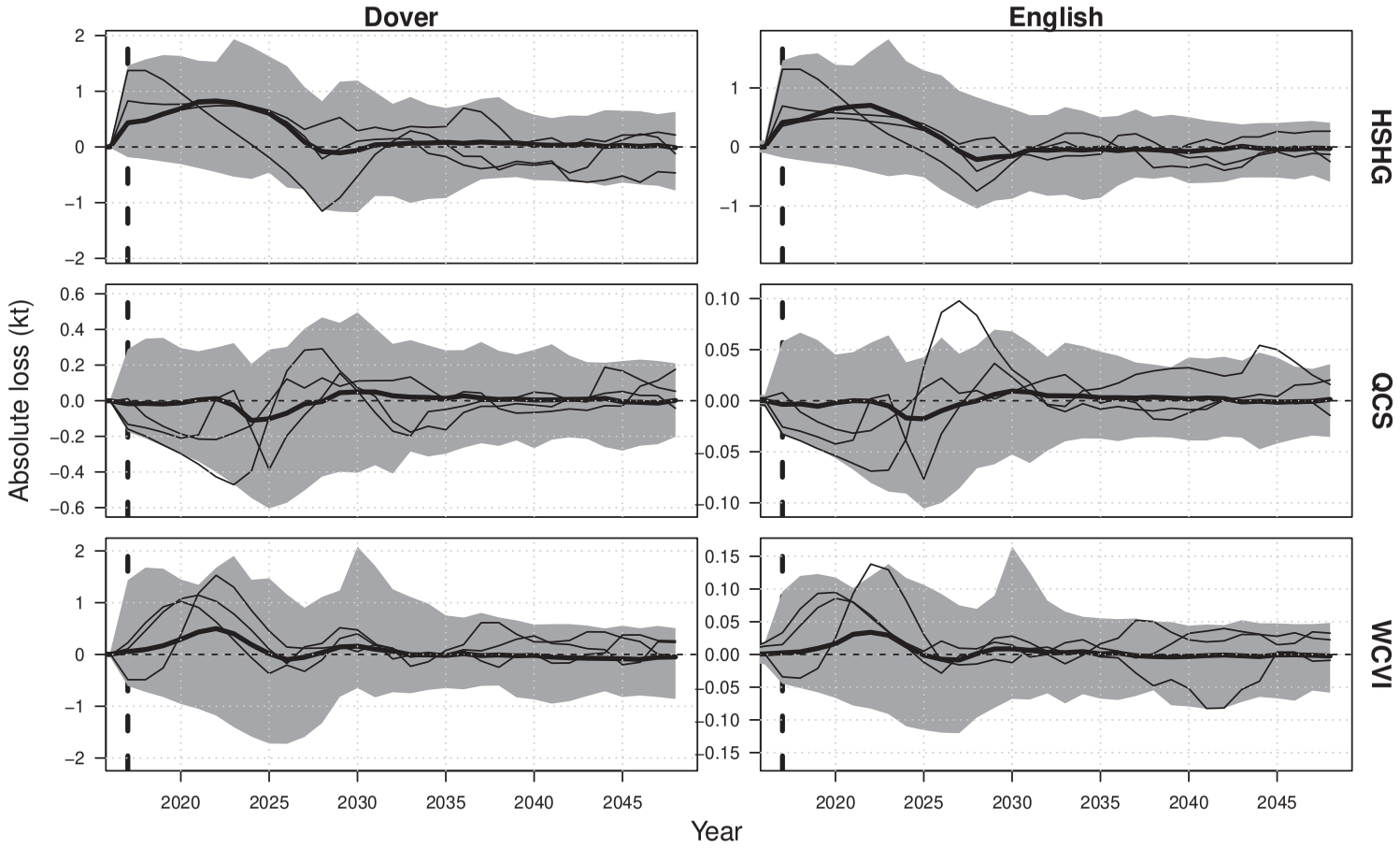

Figure 3.10: Catch loss simulation envelopes for the six BC flatfish complex management units included in the catch loss ranking, when assessed using the Hierarchical Multi-species assessment model under the Low data-quality scenario. Median catch loss is shown by the thick black line, the central 95% of the catch loss distribution is shown as the grey shaded region, and the thin black lines show three randomly selected simulation replicates. The thick vertical dashed line shows the beginning of the projection period, and the horizontal thin dashed line shows zero catch loss.

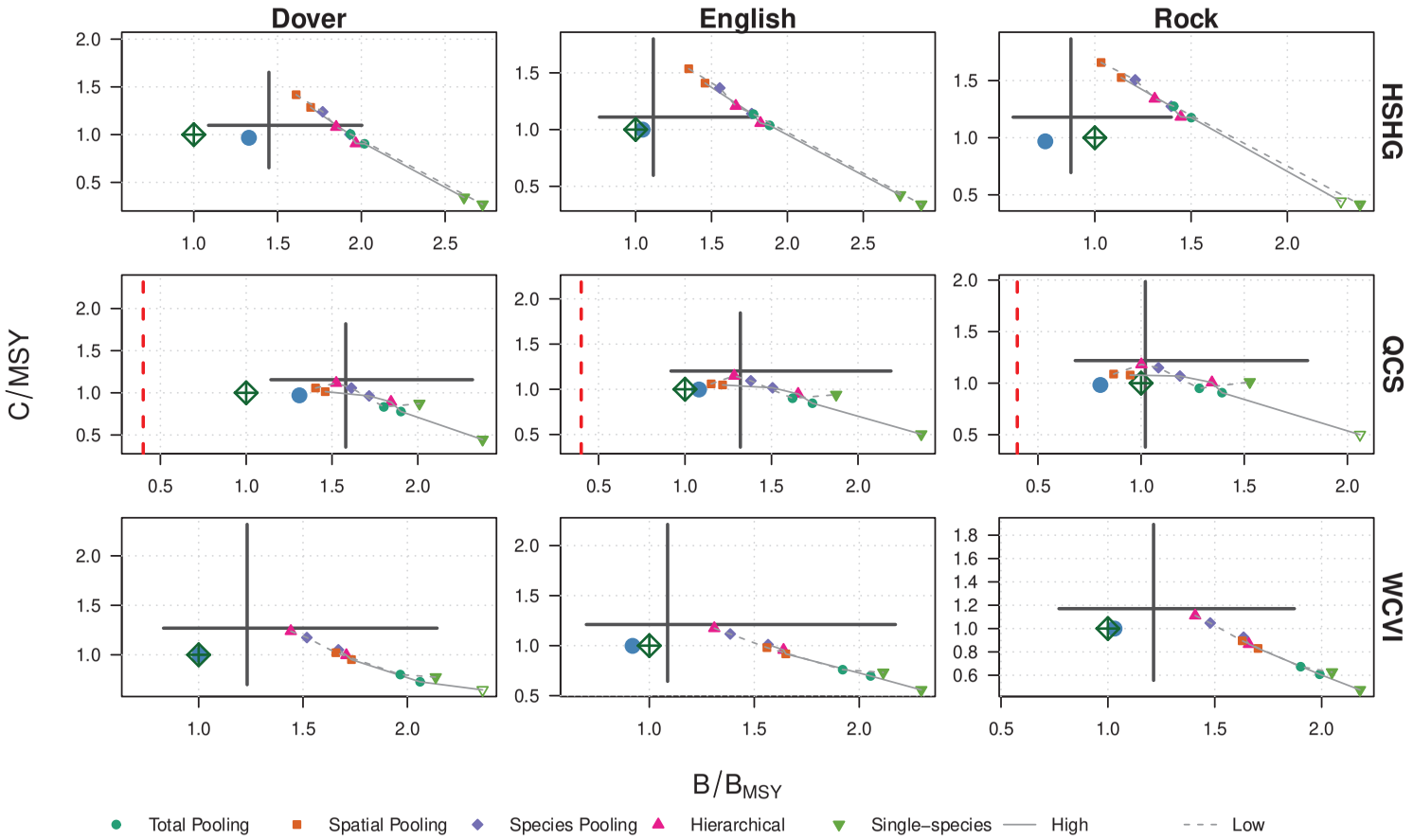

Figure 3.11: Tradeoff between catch and biomass during the 2028 - 2037 period implied by switching between different assessment models under High and Low data-quality scenarios. Panels are gridded by species (columns) and stocks (rows), with biomass relative to $B_{MSY,MS,s,p}$ on the horizontal axis, and catch relative to $MSY_{MS,s,p}$ on the vertical axis. Distributions of biomass and catch under the omniscient manager are shown by the black crosshair, with points indicating optimal biomass and yield for single species maximum yield (green crossed diamond) and multi-species maximum yield (closed blue circles). The biomass level at which a stock is critically overfished is shown as a vertical red dashed line. Coloured point symbols show median biomass and catch for over all replicates for different assessment models, with assessment models under the same data-quality scenario joined by a solid line (High) or dashed line (Low).

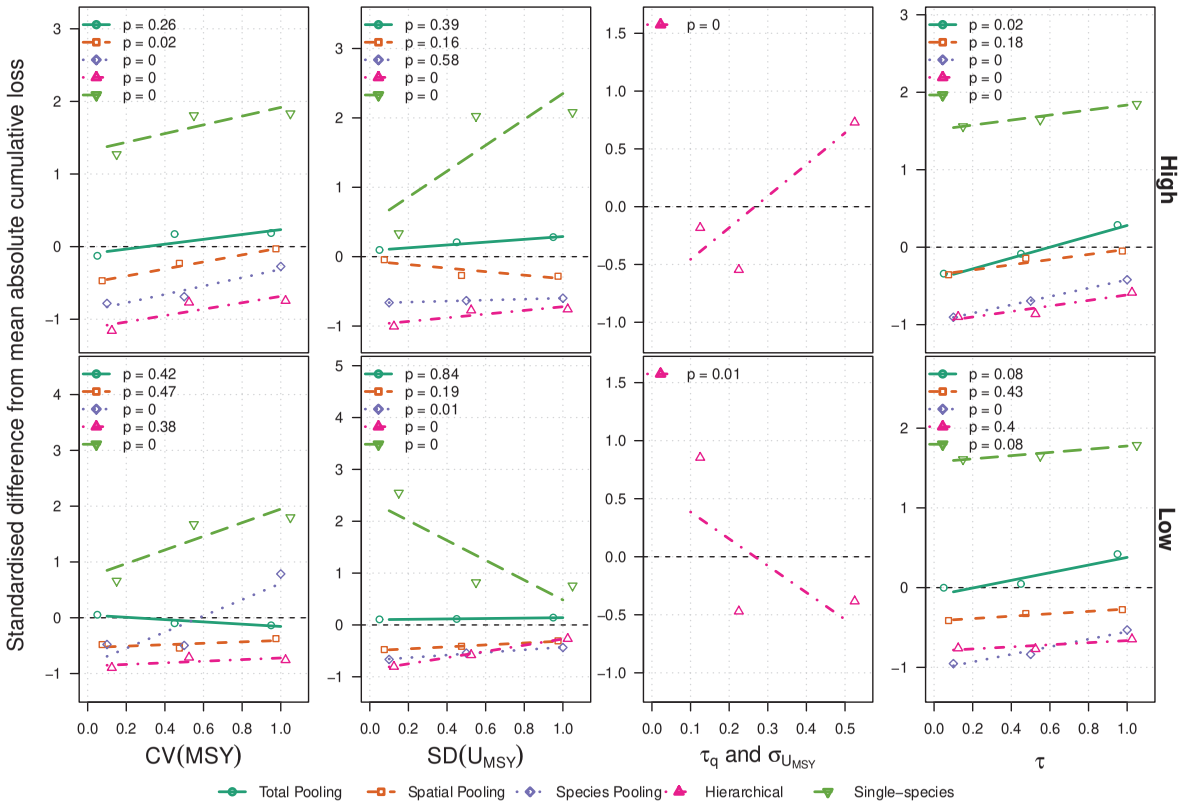

Figure 3.12: Regressions showing the average sensitivity of cumulative catch loss to the prior standard deviations under each assessment model (colours, line types) for the data-rich (left column) and data-poor (right column) scenarios. The horizontal axis on each plot shows the prior standard deviation (CV for B_{MSY}), while the vertical axis shows the standardised difference between median cumulative loss for an assessment model and the mean of median cumulative loss values over AMs, stratified by species and area. See the online version of the journal for a full colour version of the plot.

3.7 Appendices

3.A The operating model

The operating model was a standard age- and sex-structured operating model, with additional structure for multi-species and multi-stock population dynamics. BC flatfish complex species and stocks were simulated assuming no ecological interactions or movement between areas. The lack of movement may be unrealistic, especially for Dover Sole given their extent, but this is how the the BC flatfish complex stocks are currently managed in practice. The lack of ecological interactions is more realistic for Dover and English soles, as although both species are benthophagus, there is evidence that they belong to different feeding guilds [108].

BC flatfish complex abundance $N_{a,x,s,p,t}$ for age a, sex x, species s and stock p at the start of year t was given by

$$N_{a,x,s,p,t} = \begin{cases} 0.5R_{s,p,t} & a = 1, \\ N_{a-1,x,s,p,t-1} \cdot e^{-Z_{a-1,x,s,p,t-1}} & 1 < a < A, \\ N_{a-1,x,s,p,t-1} \cdot e^{-Z_{a-1,x,s,p,t-1}} + N_{a,x,s,p,t-1} \cdot e^{-Z_{a,x,s,p,t-1}} & a = A^{(s)}, \end{cases}$$

where R_t is age-1 recruitment in year t, $Z_{a,x,s,p,t}$ is the instantaneous total mortality rate, and $A^{(s)}$ is the plus group age for species s.

Numbers-at-age were scaled to biomass-at-age by sex/species/area- specific weight-at-age. Weight-at-age was an allometric function of length-at-age

$$w_{a,x,s,p} = \alpha_{x,s,p} \cdot L_{a,x,s,p}^{\beta_{x,s,p}}$$

where $\alpha_{x,s,p}$ scaled between cm and kg, $\beta_{x,s,p}$ determined the rate of allometric growth, and $L_{a,x,s,p}$ was the length in cm of a fish of age a, sex x, species s and stock p. Length-at-age was given by the following Schnute formulation of the von-Bertalanffy growth curve [127, 49]

$$L_a = \overline{L}_{A_1} - (\overline{L}_{A_2} - \overline{L}_{A_1}) \cdot \left(\frac{e^{-kA_1} - e^{-ka})}{e^{-kA_1} - e^{-kA_2}} \right)$$

where A_1 and A_2 are well spaced reference ages, $\overline{L}_{A_1}$ and $\overline{L}_{A_2}$ are the mean lengths in cm of fish at ages A_1 and A_2, and k is the growth coefficient. Note that in the growth model I dropped the sex, species and stock subscripts for concision.

The maturity-at-age ogive was modelled as a logistic function

$$m_{a,s,p} = \left(1 + e^{-\frac{\ln 19 (a - a^{mat}_{50,s,p})}{a^{mat}_{95,s,p} - a^{mat}_{50,x,s,p}}} \right)^{-1},$$

where $m_{a,s,p}$ was the proportion of age-a female fish of species s in stock p that were mature, and $a^{mat}_{50,s,p}$ and $a^{mat}_{95,s,p}$ are the ages at which 50% and 95% of fish of age-a, species s and stock p were mature.

Female spawning stock biomass was calculated as

$$B_{s,p,t} = \sum_a N_{a,x',s,p,t} m_{a,s,p} w_{a,x',s,p},$$

where x' denotes female fish only. Spawning stock biomass was used to calculate expected Beverton-Holt recruitment, which then had recruitment process errors applied

$$R_{s,p,t+1} = \frac{R_{s,p,0} \cdot 4h_{s,p} \cdot B_{s,p,t}}{B_{s,p,0} \cdot (1 - h_{s,p}) + (5h_{s,p} - 1) \cdot B_{s,p,t}} \cdot e^{\epsilon_{s,p,t+1} - 0.5\sigma^2_{R,s,p}},$$

where $R_{s,p,0}$ is unfished equilibrium recruitment, $B_{s,p,t}$ is the spawning stock biomass at time t, $B_{s,p,0}$ is unfished spawning stock biomasss, $h_{s,p}$ is stock-recruit steepness (average proportion of $R_{s,p,0}$ produced when $B_{s,p,t} = .2B_{s,p,0}$), and $\epsilon_{s,p,t}$ is the recruitment process error with standard deviation $\sigma_{R,s,p}$.

The operating model was initialised in 1956 at unfished equilibrium for all species s and areas p, with numbers-at-age in 1956 given by

$$N_{a,x,s,p,1956} = \begin{cases} 0.5R_{s,p,0} & a = 1, \\ N_{a-1,x,s,p,1956} \cdot e^{-M_{x,s,p}} & 1 < a < A, \\ N_{a-1,x,s,p,1956} \cdot \frac{e^{-M_{x,s,p}}}{1-e^{-M_{x,s,p}}} & a = A^{(s)}, \end{cases}$$

Fishery removals were assumed to be continuous throughout the year, with fishing mortality-at-age

$$F_{a,x,s,p,f,t} = S_{a,x,s,p,f} \cdot F_{s,p,f,t},$$

where $F_{s,p,f,t}$ is the fully selected fishing mortality rate for fleet f at time t, and $S_{a,x,s,p,f}$ is the selectivity-at-age a for sex x in species s and area p by fleet f. Selectivity-at -age was modeled as a logistic function of length-at-age

$$S_a = \left(1 + \exp\left(\frac{-\ln 19(L_a - l^{sel}_{50})}{l^{sel}_{95} - l^{sel}_{50}}\right)\right)^{-1},$$

where L_a is length-at-age, defined above, and l^{sel}_{50} and l^{sel}_{95} are the length-at-50% and length-at-95% selectivity, respectively; stock, species and fleet subscripts are left off for concision. Catch-at-age (in biomass units) was then found via the Baranov catch equation

$$C_{a,x,s,p,f,t} = (1 - e^{-Z_{a,x,s,p,f,t}}) \cdot N_{a,x,s,p,t} w_{a,x,s,p} \frac{F_{a,x,s,p,f,t}}{Z_{a,x,s,p,f,t}},$$

where total mortality-at-age is defined as

$$Z_{a,x,s,p,f,t} = M_{x,s,p} + S_{a,x,s,p,f} \cdot F_{a,x,s,p,f,t}.$$

Observation error standard deviations

Operating model observation error standard deviations were derived from estimates from fitting a hierarchical age-structured model to DER complex data.

Table 3.A.1: Log-normal observation error standard deviations for all BC flatfish complex biomass indices

	Observation Error SD			
Stock	Historical	Modern	HS Ass.	Syn
Dover sole				
HSHG	0.549	0.548	0.527	0.349
QCS	0.530	0.534		0.355
WCVI	0.569	0.530		0.315
English sole				
HSHG	0.521	0.551	0.491	0.360
QCS	0.507	0.586		0.403
WCVI	0.520	0.626		0.336
Rock sole				
HSHG	0.509	0.571	0.524	0.324
QCS	0.512	0.620		0.360
WCVI	0.520	0.845		0.405

3.B Omniscient Manager Optimisation

I defined penalty functions so that inside their respective desired regions the penalty was zero, and otherwise the penalty grew as a cubic function of distance from the desired region. For example, a penalty designed to keep a measurement x above a the desired region boundary ϵ is of the form

$$\mathcal{P}(x,\epsilon) = \begin{cases} 0 & x \geq \epsilon, \\ |x-\epsilon|^3 & x < epsilon. \end{cases} \tag{3.20}$$

This form has a several advantages over simple linear penalties, or a logarithmic barrier penalty [133]. First, the cubic softens the boundary threshold ϵ, effectively allowing a crossover if doing so favours another portion of the objective function. Second, unlike lower degree polynomials, cubic functions remain closer to the x-axis when $|x-\epsilon| < 1$. Third, zero penalty within the desirable region stops the objective function from favouring regions far from the boundaries of penalty functions. In contrast, a logarithmic function would favour overly conservative effort series to keep biomass far from a lower depletion boundary. Finally, the cubic penalty function and its first two derivatives are continuous at every point x, allowing for fast derivative-based optimisation methods.

I used a cubic spline of effort in each area to reduce the number of free parameters in the optimisation. For each area, 9 knot points were distributed across the full 40 year projection, making them spaced by 5 years. I padded the omniscient manager simulations by an extra eight years over the stochastic simulations to avoid any possible end effects of the spline entering the performance metric calculations. Effort splines were constrained to be between 0 and 120 times the operating model $E_{MSY,p}$, by replacing any value outside that range with the closest value inside the range (i.e. negative values by *zero*, large values by $120E_{MSY_p}$).

3.C Assessment model priors

Each AM included prior distributions on leading biological parameters MSY and U_{MSY}, process error deviations, and - in the case of the hierarchical AM - multi-level shrinkage priors on Synoptic survey catchability $q_{s,p,f=4}$ and U_{MSY}. All prior means were unbiased, being derived from the operating model values, which were aggregated depending on the level of pooling the AM used. A summary of the priors is given in Table S.1.

All AM configurations had the same process error prior standard deviation of $\sigma_{s,p} = 0.05$ (Table S.1, 1.1). While this process error standard deviation was very precise compared to the standard deviation of recruitment process errors ($\sigma_R = .4$), we found it was necessary to stop the AMs producing wildly variably biomass estimates that were overfit to the observation errors in the index data. This is because longer-lived groundfish species with slow demographic rates tend to have much lower variance in spawning biomass than recruitment, as each new fish has to pass through several years of survival, growth, and maturity to join the spawning biomass, all of which dampen recruitment variability.

Single-stock model

The single-stock model used operating model single-species reference points for prior mean MSY and U_{MSY} values, i.e.

$$\begin{aligned} m_{MSY_{s,p}} &= MSY_{SS,s,p}, \\ m_{\log U_{MSY,s,p}} &= \log(MSY_{SS,s,p}/B_{MSY,SS,s,p}) \end{aligned}$$

where $B_{MSY,SS,s,p}$ is the biomass at which single-species yield is maximised. The MSY parameter had a normal prior distribution with a 20% coefficient of variation (CV) (Table S.1, 1.2), and the U_{MSY} parameter had a log-normal prior with a log-standard deviation of $S_{U_{MSY}} = 0.3$ (Table S.1, 1.3).

Catchability/survey trawl efficiency prior means differed by fleet. Operating model catchability for the historical commercial CPUE series was time-varying, as there were significant targeting effects to account for when fitting the OM. To avoid the need for a time-varying catchability in the AM, I estimated the conditional maximum likelihood estimate of catchability for the historical fleet from the operating model, i.e.

$$\log \hat{q_{s,p,f=1}} = \frac{1}{|\mathcal{T}_{s,p,f=1}|} \sum_{t \in \mathcal{T}_{s,p,f=1}} (\log I_{s,p,f=1,t} - \log B_{s,p,f=1,t}),$$

where $\mathcal{T}_{s,p,f=1}$ is set of time steps where the Historical trawl fishery CPUE observations exist for species s in stock area p, $B_{s,p,f=1,t}$ is the biomass selected by the Historical trawl fishery, and $I_{s,p,f=1,t}$ is the Historical commercial CPUE index value. Then the prior mean

value was set to

$$m_{\log q_{s,p,f=1}} = \log \hat{q}_{s,p,f=1}.$$

For the remaining indices, the unbiased prior mean value was

$$m_{\log q_{s,p,f}} = \log q^*_{s,p,f},$$

where $f > 1$ is any of the remaining fleets, and $q^*_{s,p,f}$ is the operating model estimate of catchability for the modern fishery CPUE, and survey trawl efficiency for the Assemblage and Synoptic surveys. Prior standard deviations indicated vague prior knowledge for the commercial indices ($s_{q_{s,p,f}} = 0.5,\ f = 1, 2$), and more informative prior knowledge for the fishery-independent survey biomass indices ($s_{q_{s,p,f}} = 0.2,\ f = 3, 4$) (Table S.1, 1.4 - 1.7).

Hierarchical Multi-stock model

The hierarchical model had the same prior distribution as the Single-stock model for the species/area specific MSY parameters, with an unbiased prior mean based on the operating model value, and a 20% coefficient of variation (Table S.1, 1.8).

The U_{MSY} multi-level shrinkage prior had two levels, where individual stock log productivities $\log U_{MSY,s,p}$ (Table S.1, 1.9) were shrunk towards estimated species log-mean values $\overline{\log U_{MSY,s}}$ (Table S.1, 1.10), which in turn were shrunk towards a single BC flatfish complex log-mean value $\overline{\log U_{MSY}}$ (Table S.1, 1.11). Both levels of shrinkage priors assumed the same prior standard deviation of $\sigma_{U_{MSY}} = 0.1$. Finally, the BC flatfish complex mean $\log U_{MSY}$ value had a fixed hyper-prior mean set to the average optimal harvest rate from the operating model

$$m_{\log U_{MSY}} = \log\left(\frac{\sum_{s,p} MSY_{SS,s,p}}{\sum_{s,p} B_{MSY,SS,s,p}}\right),$$

with a hyper-prior standard deviation of $s_{U_{MSY}} = 0.3$, as in the Single-stock model.

The Synoptic survey had a single-level log-normal shrinkage prior for catchability, where individual stock log-catchability $\log q_{s,p,f=4}$ was shrunk towards a species log-mean value $\overline{\log q_{s,f=4}}$ with a fixed prior standard deviation of $\tau_q = 0.4$ (Table S.1, 1.12). The species *log*-mean value had an unbiased *log*-normal hyperprior mean equal to the OM estimate $\log q^*_{s,f=4}$, which was estimated as part of the OM's shrinkage prior on Synoptic survey catchability, and a hyper-prior standard deviation of $s_{q,f=4} = 0.2$, again matching the Single-stock model (Table S.1, 1.13).

Data-pooled Models

Data-pooled models had the same prior structure to the Single-stock model, but the prior means were aggregated according to the pooling method. Aggregated MSY prior means

were found by simple summation across the pooled species, areas, or both, i.e.,

$$m_{MSY_s} = \sum_p MSY_{s,p},$$
$$m_{MSY_p} = \sum_s MSY_{s,p},$$
$$m_{MSY} = \sum_{s,p} MSY_{s,p},$$

for spatial pooled (Table S.1, 1.14), species pooled (Table S.1, 1.17), and totally pooled configurations ((Table S.1, 1.20), with prior CVs of 20%. Similarly, log-normal prior mean $\log U_{MSY}$ values were given by

$$m_{\log U_{MSY,s}} = \log(\sum_p MSY_{s,p} / \sum_p B_{MSY,s,p}),$$
$$m_{\log U_{MSY,p}} = \log(\sum_s MSY_{s,p} / \sum_s B_{MSY,s,p}),$$
$$m_{\log U_{MSY}} = \log(\sum_{s,p} MSY_{s,p} / \sum_{s,p} B_{MSY,s,p}),$$

for spatial-pooled (Table S.1, 1.15), species-pooled (Table S.1, 1.18), and totally pooled configurations (Table S.1, 1.21), respectively. Prior standard deviations were the same as the Single-stock model, with $s_{MSY} = 0.2$ and $s_{U_{MSY}} = 0.3$.

Prior mean values for both commercial CPUE series and the Synoptic survey biomass indices were derived by taking an equal-weighted mean across the pooled strata, i.e.,

$$m_{q_{s,f}} = \frac{1}{3} \sum_p m_{q_{s,p,f}},$$
$$m_{q_{p,f}} = \frac{1}{3} \sum_s m_{q_{s,p,f}},$$
$$m_{q_f} = \frac{1}{9} \sum_{s,p} m_{q_{s,p,f}},$$

for spatial pooled (Table S.1, 1.16), species pooled (Table S.1, 1.19), and totally pooled configurations (Table S.1, 1.22), respectively, where $f \neq 3$, and $m_{q_{s,p,f}}$ is the same mean value used in the Single-stock and Hierarchical multi-stock cases. Because pooled indices were unequally contributed to by each stock, the prior standard deviations for catchability were increased over the Single-stock model to allow for more variability. For all pooled methods, $s_{q_{\cdot,\cdot,f=1}} = 10.0$ for the historical fishery CPUE, $s_{q_{\cdot,\cdot,f}} = 1$ for the modern fishery CPUE ($f = 2$) and the Assemblage survey biomass index ($f = 3$), and $s_{q_{\cdot,\cdot,f=4}} = 0.5$ for the Synoptic survey index (example catchability standard deviations above are shown for the Total Pooling method, with no species or stock subscript).

Table 3.C.2: Prior distributions used for each assessment method.

Eq. No.	Prior
All AMs	
(1.1)	$\zeta_{s,p,t} \sim N(0, 0.05)$
Single-species AM	
(1.2)	$MSY_{s,p} \sim N(m_{MSY_{s,p}}, 0.2 \cdot m_{MSY_{s,p}})$
(1.3)	$\log U_{MSY,s,p} \sim N(m_{\log U_{MSY}}, 0.3)$
(1.4)	$\log q_{s,p,1} \sim N(m_{\log q_{s,p,1}}, 1)$
(1.5)	$\log q_{s,p,2} \sim N(m_{\log q_{s,p,2}}, 1)$
(1.6)	$\log q_{s,p,3} \sim N(m_{\log q_{s,p,3}}, .5)$
(1.7)	$\log q_{s,p,4} \sim N(m_{\log q_{s,p,4}}, .25)$
Hierarchical multi-species AM	
(1.8)	$MSY_{s,p} \sim N(m_{MSY_{s,p}}, 0.2 \cdot m_{MSY_{s.p}})$
(1.9)	$\log U_{MSY,s,p} \sim N(\log \overline{U}_{MSY,s}, 0.1)$
(1.10)	$\log \overline{U}_{MSY,s} \sim N(\log \overline{U}_{MSY}, 0.1)$
(1.11)	$\log \overline{U}_{MSY} \sim N(m_{\log U_{MSY}}, 0.3)$
(1.12)	$\log q_{s,p,4} \sim N(\log \overline{q}_{s,4}, \tau_q)$
(1.13)	$\log \overline{q}_{s,4} \sim N(m_{\log q^*_{s,f}}, s_f)$
Spatial Pooled AM	
(1.14)	$MSY_s \sim N(m_{MSY_s}, 0.2 \cdot m_{MSY_s})$
(1.15)	$\log U_{MSY,s} \sim N(m_{\log U_{MSY,s}}, 0.3)$
(1.16)	$\log q_{s,f} \sim N(m_{\log q_{s,f}}, s_{q_{s,f}})$
Species Pooled AM	
(1.17)	$MSY_p \sim N(m_{MSY_p}, 0.2 \cdot m_{MSY_p})$
(1.18)	$\log U_{MSY,p} \sim N(m_{\log U_{MSY,p}}, 0.3)$
(1.19)	$\log q_{p,f} \sim N(m_{\log q_{p,f}}, s_{q_{p,f}})$
Totally Pooled AM	
(1.20)	$MSY \sim N(m_{MSY}, 0.2 \cdot m_{MSY})$
(1.21)	$\log U_{MSY} \sim N(m_{\log U_{MSY}}, 0.3)$
(1.22)	$\log q_f \sim N(m_{\log q_f}, s_{q_f})$

3.D Catch loss rankings for individual species/stocks

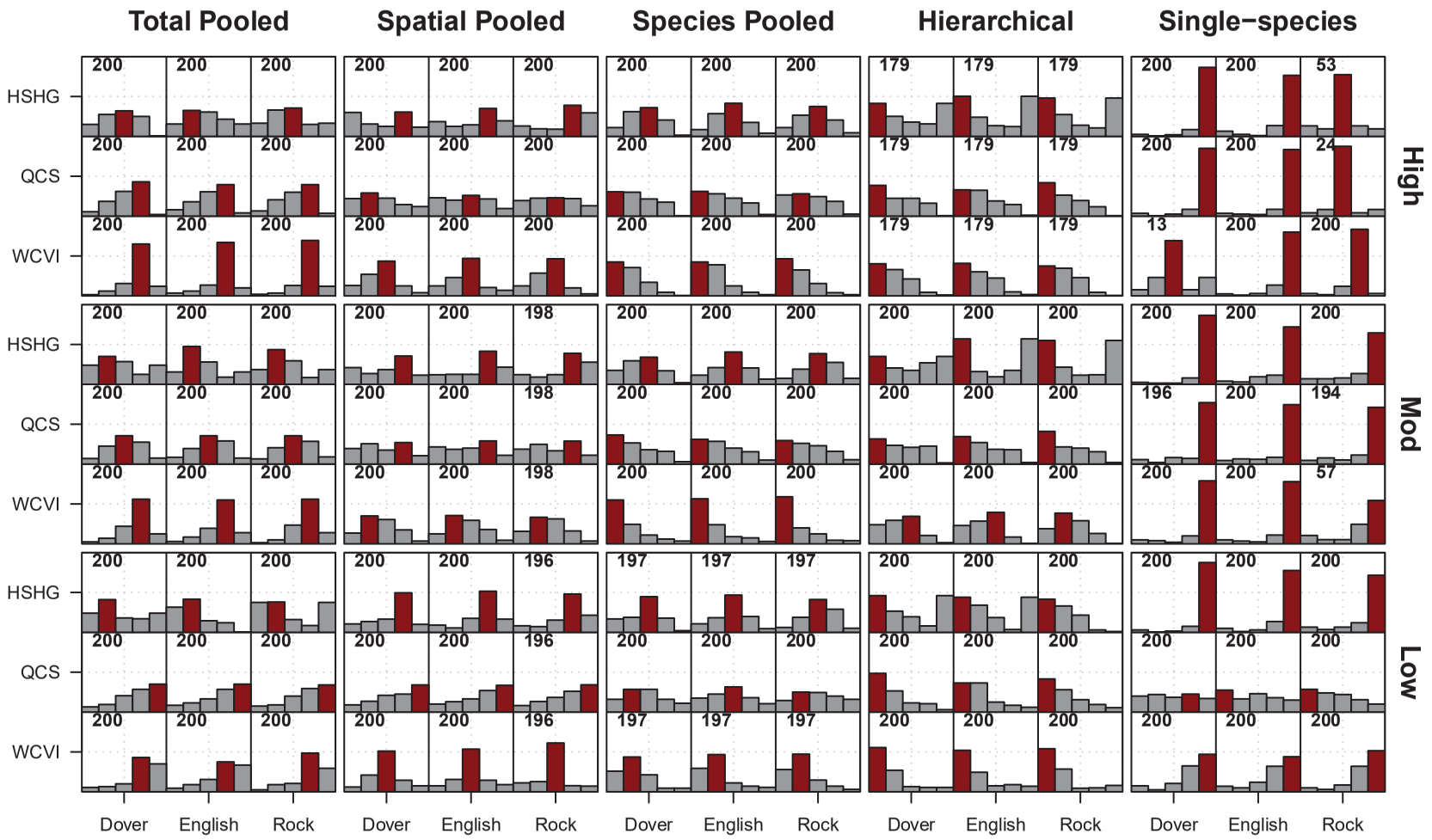

Figure 3.D.1: Distributions of AM rankings by cumulative absolute loss in catch for the projection years 2028 to 2037 under each assessment model (columns) and OM data scenario (rows), broken down by species/stock. Ranks are calculated for each species/stock combination within replicates that met the 95% convergent AMs criterion. The modal rank is shown as a dark red bar. The number of significant replicates exceeding the 95% AM convergence criterion is shown in the top-left corner of each cell.

3.E Assessment model MAREs

I calculated median absolute relative errors (MAREs) in biomass forecasts $\hat{B}_{s,p,t+1}$ and single-species reference points B_{MSY}, MSY, and U_{MSY}. Medians of absolute relative errors were taken across replicates and projection time steps for each AM/Scenario combination. For data-pooled AMs, operating model true values were found by summing over B_{MSY} and MSY for the appropriate species/areas, and taking the ratio of pooled MSY and B_{MSY} for pooled U_{MSY} estimates. In order to present only one MARE value per Scenario/AM combination, the mean MAREs were calculated over species and areas as appropriate.

Table 3.E.3: Median Absolute Relative Errors (MAREs) for each AM's estimates of forecast biomass B_{t+1}, and single-species production model equilibria optimal biomass B_{MSY}, optimal harvest rate U_{MSY}, and optimal yield MSY, under the different data scenarios. For AMs that estimate parameters for more than one management unit, the MAREs are averaged over species/areas. AMs are ordered within a scenario by the accuracy of MSY estimates.

	Model Accuracy			
AM	$\overline{MARE(\hat{B}_{t+1})}$	$\overline{MARE(B_{MSY})}$	$\overline{MARE(U_{MSY})}$	$\overline{MARE(MSY)}$
High				
Total Pooling	0.09	0.04	0.06	0.05
Spatial Pooling	0.85	0.54	0.28	0.08
Hierarchical Multi-species	0.37	0.36	0.30	0.09
Species Pooling	0.57	0.41	0.22	0.09
Single-species	0.44	0.38	0.38	0.27
Mod				
Total Pooling	0.11	0.04	0.04	0.03
Spatial Pooling	0.96	0.54	0.25	0.11
Hierarchical Multi-species	1.34	0.52	0.39	0.12
Species Pooling	0.76	0.48	0.22	0.15
Single-species	1.41	0.45	0.50	0.24
Low				
Total Pooling	0.16	0.06	0.04	0.04
Hierarchical Multi-species	1.62	0.72	0.40	0.05
Spatial Pooling	1.07	0.58	0.26	0.12
Species Pooling	0.96	0.61	0.28	0.16
Single-species	1.54	0.48	0.83	0.26

Chapter 4

A bio-economic modelling approach to estimating risks from multi-species and single-species harvest strategies for a spatially heterogeneous, technically interacting, flatfish fishery.

4.1 Introduction

Most fisheries management policy frameworks continue to be single-species oriented, despite the overwhelming prevalence of *technical interactions* in which multiple fish species are caught in non-selective fishing gear [131, 85, 108, 95, 6]. In multi-species fisheries, technical interactions affect fishery reference points and management targets derived from both maximum sustainable yield (MSY) and maximum economic yield (MEY) paradigms [88, 95, 108, 20, 63, 103, 55]. Thus, policies ignoring technical interactions are, by design, likely to produce sub-optimal outcomes in multi-species fisheries [100].

Reference points based on MSY aim to maximize equilibrium biological yield from a fish stock [78], while MEY reference points seek to maximize resource rent from a fishery [36, 63]. The main difference between MSY and MEY is that the latter aims for higher catch per unit of effort (catch rates) via higher target biomass and lower harvest rates on the most valuable species/stocks [63]. From a single-species perspective, greater resource rents obtained under an MEY paradigm increase economic welfare and improve conservation outcomes through higher standing stock biomass. However, despite these advantages, fishery policies targeting MEY are uncommon, perhaps because most fisheries involve multiple species where MEY would imply difficult trade-offs among multiple species and multiple stocks [36]. Specifically, trade-offs that likely require that some (e.g., low value)

species/stocks remain chronically overfished for the benefit of the overall multi-species fishery [36]. The challenge for multi-species fisheries managers, then, is finding approaches to assessing and managing these trade-offs without causing unacceptable conservation risks to any individual species and/or stocks.

In general, MSY and MEY reference points estimate fishery equilibrium states assuming a single, homogeneous unit-stock [12]. In reality, most fish populations are, to varying degrees, a collection of semi-isolated sub-stocks with unique productivities, biomass densities, and carrying-capacities. Ignoring such *spatial heterogeneity* in productivity and biomass density when estimating reference points will produce spatially averaged equilibria. Harvest strategies based on such spatially averaged reference points tend to over-harvest roughly half of all sub-stocks, and under-harvest the rest [8, 54].

In spatially heterogeneous fisheries, fishing activity is presumably attracted to higher density patches to take advantage of higher catch rates and therefore, reduced search costs [52]. If catch limits are based on spatially averaged reference points but fishing is concentrated on one particalar sub-stock, then that sub-stock will tend to be over-harvested relative to its stock-specific reference points. This mismatch between the scale of population dynamics (i.e., spatial heterogeneity) and the scale of control given by the harvest strategy is further complicated in multi-species fisheries, where local-scale productivities and densities differ among species occupying the same area.

Even where biological reference points such as MSY or MEY are matched to the spatial scale of population dynamics, technical interactions further complicate the implementation of those reference points within harvest strategies. Within multi-species fisheries, each species has somewhat unique availability to fishing gear (catchability), meaning that fishing pressure felt by each species differs for a given amount of fishing activity. Furthermore, unless the fishing activity (effort) is strictly controlled in space or time, the applied fishing pressure is rarely scaled to individual species productivities. Often, multi-species spatial heterogeneity attracts (or repels) fishing activity in ways that are unrelated to each individual stock's reference points. For example, harvesters may be drawn to a fishing ground to target one species, or to avoid another species, while at the same time attempting to maximise catch rates and lower fishing costs [52].

The trade-offs associated with ignoring technical interactions are unclear for spatially heterogeneous, multi-species fisheries. The attractiveness of each area to harvesters is intrinsically linked to the mix of species in each area, via the expected resource rent from that mix. Resource rent from a fishery captures the difference between the costs and revenues of fishing [63]. Some fishing costs are variable, such as fuel and crew share, while other costs, such as on capital and/or debt service, are fixed. Revenues depend primarily on the relationship between the size of the landed catch (quantity supplied) and the unit price, as captured by the demand curve. Quantity demanded in the market is determined by consumer preferences and societal factors such as household income, the price of substitutes

(e.g., similar species or other proteins), and the quantity supplied by competitors, which are captured in demand curves through price elasticities [66, 17, 5]. Thus, demand dynamics often operate at broader spatial scales than the spatial heterogeneity of a fish stock, so that catches from each stock will affect the price differently, implying different optimal harvest rates for each stock. Therefore, while conservative harvest strategies that ignore technical interactions may produce acceptable conservation performance, they are likely to create sub-optimal distributions of fishing effort with lower resource rents.

In this chapter, I investigate conservation and economic trade-offs arising when multi-species vs single-species harvest strategies are applied to a spatially heterogeneous, multi-species fishery. Harvest strategies are defined based on both single- and multi-species MSY and MEY reference points applied to a complex of three right-eyed flounders (*Pleuronectidae Spp.*) comrising Dover sole (*Microstomus pacificus*), English sole (*Parophrys vetulus*), and (Southern) Rock sole (*Lepidopsetta bilineata*). This three species complex is simulated in three spatially heterogeneous areas, where they are fished by a common trawl fishery off the coast of British Columbia (BC), Canada. Simulated fishing effort in each area has realistic constraints caused by technical interactions among the three species [6, 104]. Those same technical interactions are incorporated into multi-species MSY and MEY reference points, which, along with single species MSY reference points, are estimated analytically via equilibrium analysis [9, 95, 108, 63]. Multi-species MSY and MEY reference points are compared to long-run average simulated efforts derived from an 80-year projection with recruitment process variability, where either the catch or net-present-value (NPV) of the fishery was maximised. I show that, after accounting for own-price elasticity of demand and spatial heterogeneity in population dynamics - two realistic features of actual fisheries - MEY harvest strategies consistently produce better conservation and economic outcomes than either single-species or multi-species MSY policies. While MEY harvest stratgies have their shortcomings [36], it is surprising that MEY-based strategies are not more common in contemporary fisheries management, especially given their prominence in ecosystem based fisheries policy [91].

4.2 Methods

The following sections introduce the multi-species fishery system, then describe a bio-economic simulation model of the fishery, and describe multi-species MSY and MEY reference points derived via yield-per-recruit analysis and stochastic optimisation. Finally, we present a closed loop simulation approach, which I use to compare management performance among the harvest strategies in the presence of uncertainty.

4.2.1 Study system

BC's multi-species complex of right-eyed flounders is a technically interacting, spatially heterogeneous group of flatfishes managed across the whole BC coast. Although there are several right-eyed flounders in BC waters, I focused here on Dover sole, English sole, and Rock sole because these species together comprise a multi-stock complex (BC flatfish complex) managed in three spatially heterogeneous stock areas (Figure 4.1), Hecate Strait/Haida Gwaii (HSHG, management areas 5CDE), Queen Charlotte Sound (QCS, areas 5AB), and West Coast of Vancouver Island (WCVI, 3CD), as described in Johnson and Cox [71].

The BC flatfish complex is primarily fished by trawl vessels managed under an individual transferable quota system, with negligible levels of catch by other gear types[1]. Catch is supplied to local BC markets (48% of all seafood), as well as exported interprovincially (9% of all seafood) and internationally (43% of all seafood)[2]. International exports of Soles are primarily to the US (10%) and China (87%)[3].

A demand analysis shows that the BC flatfish complex species may be inferior goods (Appendix 4.7.A). An inferior good is one where demand decreases with increasing income, as people with more disposable income buy higher quality substitutes. This makes sense for BC flatfish complex species, as flatfish are inexpensive and most substitutes are priced higher. Potential substitutes were not formally included in demand analyses, but include Sablefish (*Anoplopma fimbria*), Salmon (*Onchorynchus Spp.*), and Halibut (*Hippoglossus stenolepis*), which are all higher priced and considered more popular protein sources.

4.2.2 Bio-economic simulation model

The bio-economic simulation model for the BC flatfish complex comprises a biological operating model representing the population dynamics of the BC flatfish complex, and an economic model representing the fishing costs and revenues resulting from trawl fishing effort dynamically allocated in time and space.

Biological operating model

The BC flatfish complex is modeled as nine biologically independent populations, three species in each of three areas, with no movement among areas and no ecological interactions

[1]Unpublished catch data, hosted at the Pacific Biological Station of Fisheries and Oceans, Canada, in Nanaimo, BC

[2]BC Fisheries and Aquaculture Sector, 2016 Edition; `https://www2.gov.bc.ca/assets/gov/farming-natural-resources-and-industry/agriculture-and-seafood/statistics/industry-and-sector-profiles/sector-reports/british_columbias_fisheries_and_aquaculture_sector_2016_edition.pdf`

[3]2018 British Columbia Agrifood and Seafood International Export Highlights; `https://www2.gov.bc.ca/assets/gov/farming-natural-resources-and-industry/agriculture-and-seafood/statistics/market-analysis-and-trade-statistics/2018_bc_agrifood_and_seafood_export_highlights.pdf`

within an area. Biological independence is somewhat realistic, given the presence of deep gullies in the QCS area that pose barriers to flatfish migration between areas, along with evidence of limited ecological interactions [108, 142, 71, 20].

The biological operating model is an age- and sex-structured, multi-species, multi-stock (area) population dynamics model [71]. For the remainder of the chapter, species are indexed by s, and areas by p. Population dynamics of each stock is driven by a compensatory Beverton-Holt stock-recruitment model, where the stock is defined as mature female spawning biomass, with unfished biomass B_0 producing unfished recruitment R_0, and stock-recruit steepness h, which is the fraction of R_0 expected when biomass is 20% of unfished (Table 4.1) [82]. Each species in the BC flatfish complex has a unique maturity-at-age schedule. Natural mortality is stock- and sex-specific, but constant over age and time (M_m, M_f, Table 4.1). Fishing mortality-at-age is the product of the proportion of individual fish of each age susceptible to the fishing gear (selectivity), trawl catchability $q^F_{s,p}$ for each population, and the trawl effort E_p in each area. The population dynamics and fishery parameters are estimated by fitting the model to fishery dependent and independent catch, catch-per-unit-of-effort (CPUE or catch rates), biomass indices, catch-at-age, and catch-at-length data sampled from the real BC flatfish complex over the 1956-2016 period. For model equations, see section 3.A.

Economic sub-model

The economic sub-model determines the resource rent from the BC flatfish complex fishery at each annual time step. Resource rent is defined as revenues to license holders minus variable costs resulting from the expenditure of fishing effort to each area, i.e.,

$$\pi(\vec{E}|\{\vec{B}_{s,p}\}_{s=1,..,S,\ p=1,...,P}) = \sum_{p=1}^{3}\left((1-\psi)\sum_{s} v_s\left(\sum_{p'=1}^{3} C_{s,p'}(E_{p'}, B_{s,p})\right)\cdot C_{s,p}(E_p, B_{s,p})\right) - c_p \cdot E_p, \tag{4.1}$$

where π is the resource rent for the whole BC flatfish complex fishery, $\vec{E} = \langle E_p | p = 1, 2, 3\rangle$ is a vector of trawl effort allocated to each area p, each $\vec{B}_{s,p}$ is a vector of biomass-at-age for species s and area p, ψ is the assumed crew share of revenue, $v_s(\sum_p C_{s,p})$ is the inverse demand curve relating the unit price to the quantity landed $C_{s,p}$, and c_p is the fuel cost-per-unit-effort of fishing in area p (or p' for the inner summation). The catch $C_{s,p}(E_p, \vec{B}_{s,p})$ is calculated from the Baranov catch equation given effort E_p and biomass-at-age $\vec{B}_{s,p}$ [71], i.e.,

$$C_{s,p}(E_p, \vec{B}_{s,p}) = \sum_{a=1}^{A_s} \frac{Sel_{a,x,s,p} q^F_{s,p} E_p}{Z_{a,x,s,p}} B_{a,s,p} e^{-Z_{a,x,s,p}}, \tag{4.2}$$

where A_s is the plus-group age for species s, $Sel_{a,x,s,p}$ is selectivity-at-age, and $Z_{a,x,s,p}$ is total mortality-at-age, with age indexed by a and sex indexed by x.

The inverse demand curves (Figure 4.2) are fit to BC flatfish complex catches, ex-vessel prices, and BC household income data (Appendix 4.7.A). While the BC flatfish complex species are sold outside of BC, household income in all markets is highly correlated. Therefore, while the exact income elasticity estimates will differ for each market, the sign of the elasticity would still be negative, implying that the BC flatfish complex species are inferior goods. Area-specific fuel costs per trawl hour were derived from the average cost of fuel to catch one kilotonne of groundfish catch in 2009 [99], adjusted to 2016 prices (Table 4.2). The fuel cost per kilotonne was then converted to cost per unit of fishing effort in each area by scaling by the ratio of catch to fishing effort from each area in 2009, i.e.,

$$c_p = .193 \cdot \frac{\sum_s C_{s,p,2009}}{E_{p,2009}}.$$

Fishing costs included only variable costs such as fuel and crew share, and excluded all fixed costs (e.g., repairs and debt service). Fixed costs were ignored, since they are difficult to assign to individual species or multi-species complexes. As a result, the above definition of resource rent provides an upper bound for economic yield.

4.2.3 Multi-species MSY and MEY reference points

Optimal yield and reference points via equilibrium analysis

Equilibrium yield reference points establish an equilibrium relationship between fishing effort, biomass-per-recruit, and average annual yield (i.e., a yield curve). Yield curves are computed via standard equations for both single-species and multi-species fisheries with technical interactions [132, 95, 108]. Any derivation starts from the assumption that, at equilibrium, recruitment balances mortality and leads to replacement of the spawning stock [122]. At unfished, the replacement assumption gives

$$B_{0,s,p} = \sum_{a=1}^{A_s-1} R_{0,s,p} e^{-(a-1)M_{s,p}} \cdot w_{a,s,p} \cdot m_{a,s} + R_{0,s,p} \frac{e^{-(A_s-1)M_{s,p}}}{1-e^{-M_{s,p}}} \cdot w_{A_s,s,p} \cdot m_{A_s,s}$$

where $w_{a,s,p}$ is weight-at-age for species s and stock p, and $m_{a,s}$ is the proportion mature-at-age a, and other parameters are given in Table 4.1. By factoring out $R_{0,s,p}$ the spawning-stock biomass-per-recruit is derived as

$$\phi_{s,p,0} = \frac{B_{0,s,p}}{R_{0,s,p}} = \sum_{a=1}^{A_s-1} e^{-(a-1)M_{s,p}} \cdot w_{a,s,p} \cdot m_{a,s} + \frac{e^{-(A_s-1)M_{s,p}}}{1-e^{-M_{s,p}}} \cdot w_{A_s,s,p} \cdot m_{A_s,s},$$

This expression can be extended to a general $\phi_{s,p,E}$, where E is any constant fishing effort, to produce a fished equilibrium spawning-stock biomass-per-recruit. First, it is algebraically

simpler to define survivorship-at-age for a given fishing effort E as

$$S_{a,s,p,E} = \begin{cases} e^{-\sum_{a'=1}^{a} Z_{s,p}} & a < A_s \\ \frac{e^{-\sum_{a'=1}^{a} Z_{s,p}}}{1-e^{-\sum_{a'=1}^{a} Z_{s,p}}} & a = A_s, \end{cases}$$

where $Z_{s,p} = M_{s,p} + s_a q_{s,p} \cdot E_p$, s_a is selectivity at age, and $q_{s,p}$ is the commercial catchability scalar transforming fishing effort to fishing mortality. Then, fished equilibrium spawning stock biomass-per-recruit is

$$\phi_{s,p,E} = \frac{B_{s,p,E}}{R_{s,p,E}} = \sum_{a=1}^{A_s} S_{a,s,p} w_{a,s,p} \cdot m_{a,s}$$

Similarly, equilibrium yield-per-recruit at a fishing effort E can be calculated via the Baranov equation as

$$YPR_{s,p}(E) = \sum_{a=1}^{A_s} S_{a,s,p,E} \cdot w_{a,s,p} \cdot (1 - e^{-Z_{a,s,p}}) \cdot \frac{s_a q_{s,p} E}{Z_{a,s,p}}.$$

Converting $\phi_{s,p,E}$ and $YPR_{s,p}(E)$ into equilibrium spawning biomass and yield, and thereby establishing MSY and B_{MSY}, requires the equilibrium recruitment $R_{s,p,E}$, which until now has been unnecessary. For a Beverton-Holt stock recruitment function parameterised by steepness and B_0, equilibrium recruitment is

$$\overline{R}_{s,p,E} = \frac{R_{s,p,0} \cdot 4h_{s,p} \cdot B_{s,p,E}}{B_{s,p,0} \cdot (1 - h_{s,p}) + (5h_{s,p} - 1) \cdot B_{s,p,E}},$$

within which spawning biomass can be substituted as $B_{s,p,E} = R_{s,p,E} \cdot \phi_{s,p,E}$ giving

$$R_{s,p,E} = \frac{R_{s,p,0} 4h_{s,p} \phi_{s,p,E} - B_{s,p,0}(1 - h_{s,p})}{5h_{s,p} - 1}.$$

For a given fishing effort E, equilibrium yield is then found by taking the product

$$Y_{s,p}(E) = R_{s,p,E} \cdot YPR_{s,p}(E)$$

and equilibrium spawning biomass as

$$B_{s,p,E} = R_{s,p,E} \cdot phi_{s,p,E}.$$

Single-species MSY_{SS} reference points (ignoring technical interactions) are found by maximising individual species s yield within each area p over fishing effort to derive the optimal species-specific effort $E_{MSY,SS,s,p}$ that should be applied to each area, ignoring any by-product/bycatch of other species (Table 4.1), and calculating the equilibrium yield and

biomass at that effort [122]. Similarly, multi-species MSY_{MS} reference points with technical interactions are derived by maximising the sum $Y_p(E) = \sum_s Y_{s,p}(E)$ of individual species yields to get maximum complex yield in area p [95, 96], producing $E_{MSY,MS,p}$ (Figure 4.3, Table 4.3). The resource rent produced at MSY_{MS} is found by substituting equilbrium effort and biomass into (4.1) using the 2016 version of the inverse demand curves.

The method for deriving economic yield reference points for a multi-species fishery with technical interactions depends on the relationship between unit price and landings. In general, MEY is derived by substituting the equilibrium yield relationship for a given effort into equation 4.1, which, similar to deriving rent for MSY_{MS} reference points, is then evaluated to give the equilibrium resource rent produced by a given fishing effort allocated among areas [55, 63]. If demand is perfectly elastic (i.e., a constant price with respect to landings), as is commonly assumed for internationally traded fish, then the surface may be optimised individually for each area. If, instead, a finite own-price elasticity is assumed (Appendix 4.7.A), then the unit price for each species changes with the quantity landed, and spatial heterogeneity in species compositions implies spatial heterogeneity in rents. The resulting relationship between rents and effort defines a surface over three dimensional effort space (one for each area), and optimal effort and MEY reference points are found via Newton-Raphson optimisation of the resource rent over the 3-dimensional effort vector $\vec{E}$.

Area-specific equilibrium rent curves are produced via a grid search over fishing effort, assuming the 2016 version of the inverse demand curves v_s. Effort grids included 100 levels of coastwide effort $E = \sum_p E_p$, equally spaced from 1,000 - 100,000 hours of trawling. At each grid point, the allocation of total effort to each area is optimised assuming that some fishing would occur in all areas, and that effort will be allocated to maximise rent. Under those assumptions, the optimal allocation of effort to each area is determined via Newton-Raphson optimisation of the allocation, numerically producing area-specific equilibrium curves.

Optimal yield and reference points via stochastic simulation

While the equilibrium methods above provide exact values for MSY reference points, they will approximate MEY reference points when the relationship between catch and price varies. When demand is dependent on a time-varying factor, such as household income as it is here (Appendix 4.7.A), E_{MEY} varies from year-to-year, as the unit-price varies, changing the total rent [22]. To estimate the series of NPV maximising efforts for each area, MEY reference points are estimated, and MSY reference points validated, by stochastic simulation. The simulated BC flatfish complex is projected from 2016 conditions and fishing effort is optimised under two alternative objective functions. The first maximises total multi-

species catch over an 80-year time horizon, i.e. estimate MSY_{MS}

$$C^{obj} = \sum_{t=2017}^{2096} \sum_{s} \sum_{p} C_{s,p,t},$$

to estimate MSY_{MS}. The second objective function estimates MEY by maximising the net present value (NPV) of the BC flatfish complex resource rent over the same time horizon, i.e.,

$$\pi^{obj} = \sum_{t=2017}^{2096} (1+d)^{-(t-2016)} \sum_{p} \left[\left(\sum_{s} v_{s,t} C_{s,p,t} \right) - c_p E_{p,t} \right],$$

where d is the annual discount rate (Table 4.2).

Simulations are repeated over 100 trials, where each trial involves a unique realization of a stochastic future recruitment via the stochastic age-1 recruitment series

$$R_{s,p,t} = \overline{R}_{s,p,t} \cdot e^{\sigma_{s,p} \cdot \epsilon_{s,p,t} - 0.5\sigma_{s,p}^2},$$

which deviates, by log-normal process deviations, the expected Beverton-Holt recruitment at time t

$$\overline{R}_{s,p,t} = \frac{R_{s,p,0} \cdot 4h_{s,p} \cdot B_{s,p,t}}{B_{s,p,0} \cdot (1-h_{s,p}) + (5h_{s,p}-1) \cdot B_{s,p,t}}, \tag{4.3}$$

where life-history parameters taken from Table 4.1, $B_{s,p,t}$ the mature female spawning biomass of species s in area p at time t, $\epsilon_{s,p,t} \sim N(0,1)$ is the standardised recruitment process deviation, and $\sigma_{s,p}$ is the recruitment process deviation standard error. Note that recruitment process deviations are simulated for all stocks from earlier than 2017, because historical recruitment could not be estimated reliably past 2014.

Equilibrium derivations of MSY and MEY reference points above are compared to distributions of effort, yield, and rent produced by the optimisations. Stochastic optima were defined as median values over the 2040 - 2060 period under the yield (C^{obj}) and NPV π^{obj} maximising objective functions, and denoted with an asterisk to differentiate from the equilibrium derivation, e.g. optimal efforts were $E^*_{MSY,MS,p}$ and $E^*_{MEY,p}$. Optima were calculated over 2040 - 2060 to allow enough time for the effort to stabilise after any initial transient period, while also avoiding end effects (e.g., high fishing mortality) caused by a finite time horizon.

While total effort is unconstrained in the long-run, short-run (inter-annual) changes in coastwide effort were constrained to +/-20% per year to reflect the potential for adding/subtracting new licenses in the fishery, as well as the lag time required for a capital investment (e.g., in new and larger vessels). While the effort is ultimately unbounded, the simulation treats the fishery as a single firm so it is possible that fishery rents will not be dissipated, unlike an

open access fishery. The optimisation takes the same approach as the omniscient manager simulation in Johnson and Cox [71].

4.2.4 Harvest strategy performance

While optimal fishing efforts have theoretical value, it is difficult to control total fishing mortality via fishing effort alone, since harvesters will continually adapt and improve harvest efficiency [18]. Instead, most contemporary fisheries employ output control quota systems involving a total allowable catch (TAC), set by taking the product of a target harvest rate and a biomass estimate. This approach does not require controlling the total amount of effort (e.g. number of fishing vessels, sets, or trawl hours) via regulation. For example, an MSY based harvest strategy will use MSY reference points for optimal effort E_{MSY} or fishing mortality F_{MSY} to estimate an optimal harvest rate U_{MSY}. Then, catch limits are derived from applying that harvest rate to estimates of biomass. Given the uncertainty in estimates of biomass, the risks and tradeoffs in fishery outcomes from repeated application of a harvest strategy are commonly tested in closed loop feedback simulation, or management strategy evaluation [113].

The tradeoffs among harvest strategies based on MSY_{SS}, MSY_{MS} and MEY reference points applied to the BC flatfish complex are identified via closed loop feedback simulations over an 80-year time horizon. At each time step, TACs are set using a constant target harvest rate based on the reference points being tested, i.e.

$$TAC_{s,p,t} = \hat{B}_{s,p,t} \cdot U_{s,p},$$

where $\hat{B}_{s,p,t}$ is a biomass forecast for species s in area p at time t, and $U_{s,p}$ is the harvest rate derived from either MSY_{SS}, MSY_{MS}, or MEY reference points, or from maximising long-run yield (MSY^*_{MS}) or NPV (MEY^*) when stochasticity is explicitly taken into account. Uncertainty in biomass forecasts $\hat{B}_{s,p,t}$ is simulated via auto-correlated log-normal assessment errors

$$\hat{B}_{s,p,t} = B_{s,p,t} \cdot e^{\delta_{s,p,t}}$$

to approximate the observation uncertainty inherent in fish stock biomass estimates. Parameters for the $\delta_{s,p,t}$ distribution are derived from the distribution of estimation errors for a hierarchical multi-species surplus production assessment model applied to the simulated BC flatfish complex (Appendix 4.7.B) [71]. Once TACs are set, annual fishing effort is allowed to increase until at least one species' TAC in each area is fully utilised, after which the area is closed to fishing [71]. Interannual differences in catch are limited to +/-20% for each stock-area combination, reflecting the limit on interannual differences in fishing effort used for the stochastic optimisation.

Fishery performance is measured by comparing the conservation and economic outcomes among the alternative harvest strategies (i.e., differing via their target harvest rates). Con-

servation performance is based on a comparison of spawning biomass distributions (over all simulation replicates and projection years) to fractions of single-species $B_{MSY,SS}$ over the period 2040 - 2060. Economic performance is measured as the NPV of the BC flatfish complex fishery over the whole 80-year time horizon under each harvest strategy. Risk in economic outcomes was measured via percentile $p^* \in \{0.05, 0.5, 0.95\}$, or *Value-at-risk-*$p^*$ ($VAR - p^*$), which is a non-parametric definition of risk commonly used in economic analyses [29, 92]

4.2.5 Sensitivity Analyses

Net present values and economic risks depend on the resource rents, which in turn depend on prices for each species (via inverse demand curves) and discount rates applied in NPV calculations. The sensitivity of stochastic optima under the rent objective (π^{obj}) is tested against demand and discount rate uncertainties in seven additional scenarios (Table 4.3). Three scenarios test alternative demand assumptions, specifically a perfectly elastic demand relationship, simulated by refitting demand curves with an infinite price elasticity of demand (i.e., zero price flexibility, Appendix A), and a constant demand over time by setting the growth in household income $\gamma = 0$. The remaining four scenarios test sensitivites to assumed discount rates δ and growth rates γ of household income (Table 4.3). Alternative discount rates to the 5% value used in the baseline simulation were alsot tested, and the alternative rates of growth in household income considered were +/-1 standard deviation of the recent mean rate of 3.9%[4] (Table 4.3).

4.3 Results

4.3.1 Equilibrium multi-species MSY

Multi-species yield-curves and the harvest rates that maximise multi-species complex yield reflect the spatial heterogeneity in species compositions and productivities (Figure 4.3). In HSHG and QCS areas, the optimal effort $E_{MSY,MS,p}$ produces lower optimal harvest rates for Dover sole and English sole when compared to $U_{MSY,SS}$, while Rock sole $U_{MSY,MS}$ harvest rates were 20% - 30% higher than Rock sole $U_{MSY,SS}$ ($U_{MSY,MS}/U_{MSY,SS}$, Table 4.4). In the WCVI area, English sole $U_{MSY,MS}$ harvest rate is 8% higher than English sole $U_{MSY,SS}$, while Dover sole $U_{MSY,MS}$ and $U_{MSY,SS}$ are practically identical. Despite large differences in optimal harvest rates for some stocks, optimal yields MSY_{SS} and MSY_{MS} are similar, as lower harvest rates are applied to larger equilibrium biomasses, and decreases from MSY_{SS} to MSY_{MS} were at most 5% for any one stock. The total of MSY_{MS} for the

[4] https://www2.gov.bc.ca/assets/gov/british-columbians-our-governments/government-finances/financial-economic-review/financial-economic-review-2020.pdf

BC flatfish complex fishery was 5.26 kt, which was only 110t, or 3%, lower than the sum of single-species MSY_{SS} yields.

Annual equilbrium rent from MSY_{MS} also reflected spatial heterogeneity in the BC flatfish complex. Total coast-wide annual rent is around CAD \$278,000 yr^{-1} (2016 dollars) across the three stock areas. The area with the highest rent is HSHG, with an average of almost CAD \$160,000 yr^{-1}, attained at the lowest effort of the three areas (Table 4.4), which is expected given that both catchability and total BC flatfish complex biomass is highest in HSHG. Rent for QCS (\$110,000 yr^{-1}) is dominated by a roughly equal mix of highly catchable Dover sole with lower own-price elasticity (i.e., more stable prices) and the more valuable Rock sole, while the lowest rent (\$80,000 yr^{-1}) occured in WCVI area, where catch is dominated by Dover sole, which has an upper bound on unit price from the exogenous US catch.

4.3.2 Equilibrium multi-species MEY

Like multi-species MSY, equilibrium resource rent curves also reflected the spatial heterogeneity of the BC flatfish complex (Figure 4.4), with total coastwide equilibrium rent at MEY of approximately CAD \$1.36 million yr^{-1} (2016 dollars) (Table 4.4). Rent maximising fishing effort E_{MEY} is estimated at slightly more than half of the effort level achieving MSY_{MS} for all areas. The difference in total equilibrium catch between MSY_{MS} and MEY is around 1.1 kt, with individual stock catches lower under MEY than MSY_{MS} by an average of 27%.

Similar to the above, resource rent is highest in the HSHG area at the lowest effort, reflecting the higher biomass of all species in that area (Table 4.4). While optimal effort E_{MEY} in QCS is about 20% higher than in WCVI, the resource rent is about 50% higher in the WCVI area. As with the MSY_{MS} case, higher WCVI rent is caused by higher ratios of Dover sole biomass relative to English and Rock sole biomasses, higher catchability, and less elastic demand for Dover sole, which largely compensate for higher fuel costs of fishing in WCVI and lower price of Dover sole. In short, WCVI operates as an efficient, single-species fishery for Dover sole.

The rent maximising allocation of fishing effort among areas is independent of the total effort available, and instead depends mostly on catchability scalars within an area and to a lesser extent on the unit price of each species. For any fixed level of coast-wide effort E, the rent maximising allocation to each area is always $\langle E_1, E_2, E_3 \rangle = \langle 0.25E, 0.42E, 0.33E \rangle$ for HSHG, QCS, and WCVI, respectively (Table 4.5), which appears related to catchability (Appendix 4.7.C).

4.3.3 Validating reference points via spatial effort optimisation

Stochastic optimisations spatially allocate fishing effort to each area in each year, either maximising catch to validate MSY_{MS} reference points, or maximising NPV to estimate MEY reference points that account for future changes in demand (Appendix 4.7.A).

As expected, stochasticity had no significant effect on MSY^*_{MS} reference points. Insensitivity of MSY_{MS} reference points to recruitment stochasticity is because the major time-varying factor in projected BC flatfish complex dynamics is the unit price, which are not included in the MSY^*_{MS} objective function. Stochastic approximations $E^*_{MSY,MS}$ are within 1 decimal place $E_{MSY,MS}$ in QCS and WCVI, and slightly higher for HSHG, which has a longer transitionary period before median dynamics settle close to the equilibria in each area (Figure 4.5). Similarly, harvest rate distributions under the stochastic optimisation centred almost exactly on $U_{MSY,MS}$ for all stocks (Figure 4.6), differing by approximately 0.5% of $U_{MSY,MS}$ (Table 4.5, $U^*_{MSY,MS}/U_{MSY,MS}$).

Despite the insensitivity of the biological reference points to changes in demand, resource rents from catch maximisation in the projections declined to less than zero (Table 4.6). Rents declined due to BC flatfish complex species being inferior goods. Annual rents dropped by CAD $1.3 million in the HSHG area, where the negative income elasticity was highest (Appendix 4.7.A), by $220,000 in QCS, and $360,000 in WCVI. In all three areas, catch maximisation in the presence of dropping demand made rents from fishing the BC flatfish complex negative (Table 4.6).

Negative income elasticity had a strong effect the effort series maximising NPV in stochastic simulations. As time goes on, household income increases at a rate of 3.9% per year, so negative income elasticity reduces prices of all BC flatfish complex species. Given the revenue reduction, the stochastic optimisation of the NPV for the BC flatfish complex fishery reduces efforts in all three areas so that rents from fishing remain positive through own-price elasticity. The largest effort reductions are in HSHG and WCVI, where Dover sole stocks are the largest, and the smallest change is in the QCS area, where Dover sole dominated the biomass but Rock sole was also a significant portion of the catch. Dependence on Dover sole is based on the assumed US west coast catches of around 8 kt (Figure 4.2 top panel, vertical dashed line), which creates an upper bound on Dover sole near the minimum price for any year, irrespective of the BC catch. In contrast, there is no exogenous US west coast catch of Rock sole, allowing prices to rise more than Dover sole as efforts decrease in other areas (Figure 4.2).

Harvest rates maximising NPV (i.e., U^*_{MEY}) are 20% - 40% lower than the corresponding U_{MEY} values under the 2016 version of the inverse demand curves (Table 4.6, $U^*_{MEY}/U_{MSY,SS}$). Reduced harvest rates in the stochastic optimisation reflect the negative income elasticity, creating conditions where effort (fishing costs) must be reduced so that rents remain positive. Unit prices decline as the demand curves shift downward with increas-

ing income (Appendix 4.7.A), which is usually explained as switching to more expensive proteins as they become affordable [33]. The decline in fishing efforts occurs after an initial development period, where effort increases to take advantage of higher initial prices before household incomes drive demand down (Figure 4.5). Despite consistently lower median harvest rates, the simulated distributions U^*_{MEY} contain the corresponding equilibrium U_{MEY} analogues for both QCS and WCVI, where Dover sole is at least 50% of the catch (Figure 4.6). The reduction in market price also depresses annual rents expected during the 2040 - 2060 period, dropping these by 42% (WCVI) to 67% (HSHG) (Table 4.5).

4.3.4 Conservation and economic performance via closed-loop feedback simulations

Conservation performance

None of the target harvest rates tested pose any conservation risk when applied in closed loop simulations (Figure 4.7). For all harvest rates, there was less than a 5% chance of pushing biomass below 40% of $B_{MSY,SS}$ (grey distributions, Figure 4.7), which is a commonly accepted limit reference point for biomass in Canadian fisheries [34]. Only a small fraction of simulations show biomass below 80% of $B_{MSY,SS}$, which is acceptable conservation performance in any Canadian fishery [34]. Further, there was no significant difference between 2040 - 2060 biomass distributions for optimal harvest rates derived from equilibrium and stochastic analyses, i.e., $U_{MSY,MS}$ and $U^*_{MSY,MS}$, which came out slightly above the equilibrium $B_{MSY,MS}$ and $B^*_{MSY,MS}$ values, respectively given in Tables 4.3 and 4.4.

Target harvest rates for single-species ($U_{MSY,SS}$), which ignore technical interactions, tend to fully exploit the species with the highest catchability, i.e. Rock sole in HSHG and QCS, and English sole in WCVI, making them the choke species under the single-species harvest strategy (Figure 4.8). Thus, single-species harvest rates produced lower conservation risks than the multi-species harvest rates, because access to other species was restricted once the high-catchability species TAC was attained. Therefore, TACs of other species were under-utilised and realised harvest rates were lower than their corresponding $U_{MSY,SS}$. This led to 2040 - 2060 biomass being higher than the implicit $B_{MSY,SS}$ target for the non-choke species.

The U_{MEY} and U^*_{MEY} harvest rates that maximise rent, and associated spawning biomasses B_{MEY} and B^*_{MEY}, are inherently more conservative compared to the $U_{MSY,SS}$ and $U_{MSY,MS}$. Therefore, conservation risk is negligible when maximising resource rent. Furthermore, strategies based on stochastic optimisation (i.e., U^*_{MEY}) are more conservative than strategies based on maximising 2016 resource rent via U_{MEY}, given the negative income elasticity (Figure 4.7).

Economic performance

Relative health of the BC flatfish complex stocks at the beginning of the projection period provide conditions under which catch maximisation gave an economic windfall in the short term. While median BC flatfish complex fishery NPV under MSY-based harvest strategies was lower than under MEY-based strategies (Table 4.7), there was a short transient period during which annual rents climbed higher under $U_{MSY,MS}$ and $U_{MSY,SS}$ compared to either U_{MEY} or U^*_{MEY} harvest rates (Figure 4.9). After the initial peak, rents under the multi-species and single-species MSY harvest policies declined quickly, becoming negative with 50% probability around 2034 in the HSHG and WCVI areas, and around 2040 in the QCS area. Rents under U^*_{MEY} and U_{MEY} followed a similar pattern, albeit with a lower peak sooner in the projection, and a slower decline to become negative somewhere between 2070 and 2090 (Figure 4.9). Declines in rent match the reduction in fish prices as household income increases (Figure 4.2), with faster declines in WCVI and HSHG, as before, due to the proportionally smaller Rock sole biomass in those areas.

As expected, targeting U_{MEY} and U^*_{MEY} produced the highest NPV of the fishery (Table 4.7). The MSY_{SS}-based harvest strategy had the second highest NPV, and the MSY_{MS}-based strategies the lowest. While higher NPV under single-species MSY strategies is somewhat counter to conventional wisdom, as choke effects are commonly cited as leading to economic hardship, it is a case-specific phenomenon caused by the current healthy state of the BC flatfish complex, and the protective nature of the choke effects as outlined above leading to lower realised harvest rates allowing the building of non-choke species biomass above $B_{MSY,SS}$. Such conditions lead to higher CPUE, which lowers fishing costs and increases rent. However, like the multi-species harvest rates, rents under $U_{MSY,SS}$ quickly fall after the initial windfall (Figure 4.9), with rents ultimately declining below zero under both catch maximisation strategies. Rents under the MSY_{MS} strategy declined more slowly in HSHG, given the choke effects and higher CPUE (Figure 4.9).

The Value-at-risk-5% (i.e., $VAR-5\%$) gives the 5^{th} percentile of the stochastic NPV distribution over all 100 replicates; in other words, the VAR-5% gives an indication of the worst expected economic performance of a harvest strategy, occuring with a probability of 5%. For $U_{MSY,MS}$ and $U^*_{MSY,MS}$, financial risk was quite high, with a VAR-5% around \$570,000 (an average rent of \$28,000 yr^{-1}, when adjusted for the discount rate of 5%) under $U_{MSY,MS}$, and \$140,000 (\$7,000 yr^{-1}) when targeting $U^*_{MSY,MS}$ (Table 4.7). It is likely that neither $U_{MSY,SS}$ nor $U_{MSY,MS}$ harvest strategies represent a viable fishery in the long term, given that the rent calculations did not include several sources of fishing costs, and rents decline after an initial short transient period (Figure 4.9). The $U_{MSY,SS}$ harvest rates had lower risk, with a $VAR-5\%$ of around \$3 million (\$150,000 yr^{-1}; Table 4.7).

Variability, and therefore risk, in NPV was reduced overall under MEY based strategies. Both U_{MEY} and U^*_{MEY} have $VAR-5\%$ at around \$11 million (\$550,000 yr^{-1}; Table 4.7),

which is higher than 95% of simulated NPVs under any MSY-based strategy. Further, the central 90% of NPVs for the BC flatfish complex under both MEY-based harvest strategies ranged from around $11 million at the low end to around $13 million at the high end. In contrast, the 5th and 95th percentiles under MSY harvest rates differed by over $5 million, indicating much higher uncertainty in future economic yield when attempting to maximise catch. Variability in economic yield when maximising catch or rents is similar to that observed in the stochastic optimisation, where efforts optimised to maximise rent had a much tighter distribution than efforts optimised to maximise BC flatfish complex total yield over the 80 year projection (Figure 4.5).

4.3.5 Sensitivity to economic assumptions

As expected, reducing the dependence of unit price (demand) on catch or household income leads to higher NPVs for the BC flatfish complex (Table 8). Significantly higher NPVs under the two perfect own-price elasticity assumptions are based on no drop in price the at higher levels of catch (quantity supplied), which in turn encourage higher fishing effort (compare E_{MEY} columns, Table 4.8 and 4.5). For the perfectly elastic and growing income scenario, NPV is even higher because the Rock sole has positive income elasticity, leading to more fishing over time in QCS (Appendix 4.7.A). The NPV for the finite own-price elasticity and growing income scenario was the lowest among the three demand curve sensitivities (Table 4.8). Fixed incomes created a stable price/catch relationship that does not reduce prices over time, allowing fishing effort levels to remain higher over the projections.

Time series of fishing effort for discount and income growth rate sensitivity scenarios sorted into two groups. The high discount rate ($d = 0.1$) and low income growth ($\gamma = 0.02$) scenarios had very similar optimal effort dynamics, and the low discount rate ($d = 0.025$) and high income growth ($\gamma = 0.058$) scenarios were also very similar (Figure 4.11). Similarities in the high-discount/low-growth scenarios continued across most of the projection period after an inital difference in all areas, where the higher discount rate initially attracted more fishing effort to take advantage of high initial CPUE (Figure 4.11). For the low discount/high income growth scenarios, median optimal effort differed by area. Median efforts converged early in QCS and WCVI, and stayed close for the entire projection. In contrast, in the HSHG area, the two scenarios diverged in the middle of the projection period in the HSHG area, where the lower discount rate scenario applied higher fishing effort, while the high income growth scenario led to a quicker decline in effort. Dropping effort as income grows is expected, as a the unit price of all fish is dropping faster through negative income elasticity, and the species composition in HSHG is roughly equal among the BC flatfish complex. The Dover sole effect meant rent was less sensitive to increased effort, so combined with a lower discount rate (i.e., more value for future rents), efforts could remain higher.

4.4 Discussion

In this chapter, I investigate the trade-offs between conservation and economic outcomes under single- and multi-species harvest strategies for spatially heterogeneous, multi-species fisheries. BC's three-species flatfish fishery is simulated over three areas, and five harvest strategies are defined based equilibrium maximum sustainable yield (MSY_{SS}, MSY_{MSY}) and maximum economic yield (MEY) reference points, as well as stochastically optimised multi-species reference points MSY^*_{MS} and MEY^*. For each harvest strategy, economic and conservation performance is estimated via closed loop simulation. For MSY_{MS} reference points, deterministic and stochastic optima are not substantially different, given that stochastic optima were medians, and recruitment process errors are bias corrected. In contrast, equilibrium MEY harvest rates are significantly higher than their stochastic analogues, as future increases in household incomes drove down the unit price, indicating that BC flatfish complex flatfish are considered an inferior good, and families will switch from flatfish to more expensive proteins as incomes increase [33].

For equilibrium multi-species reference points, catchability appears to be the major determinant of effort allocation in space. Indeed, the allocation is similar among MSY_{MS}, MSY^*_{MS}, and MEY reference points. Although species composition and own-price elasticity affect the total effort expended, their effect on the allocation of that effort among areas was minor. Further, the allocation of equilibrium E_{MEY} was insensitive to the value of own-price elasticity, with both flat (i.e., perfectly elastic) and downward sloping demand curves producing similar allocations. Allocations were insensitive to the shape of demand curves despite differences in yields and rents.

In contrast, market demand had a strong effect on the spatial allocation of effort under the stochastic optimisations. As demand for inferior goods declines with increasing household incomes, total effort was reduced, and the highest proportion shifted to the QCS area to target Rock sole, which had no upper bound on its unit price thanks to a lack of exogenous catch in the USA. Such effects of temporal variability in market demand highlight a common argument against using MEY reference points: that the assumptions are only valid for a short time and, therefore, MEY targets require constant updating to stay relevant [36]. On the other hand, frequent model updates to check misspecification is a basic requirement of adaptive management [143], and updating economic assumptions/models with more recent data at the same time is of low marginal cost.

While choke effects are commonly assumed to be inherently negative outcomes, this was not obviously the case for the BC flatfish complex fishery. Choke effects limit harvesters' ability to fully catch a species' TAC, and the assumption is that any foregone yield represents an economic loss to the fishery [6], especially when a choke species are of lower value. However, in the BC flatfish complex, all stocks are similarly priced, so choke effects more often limit excessive TACs caused by positively biased biomass estimates and the effect on

rent is more subtle [71]. Such "positive" choke effects are evident under the single-species MSY_{SS}-based harvest strategy, which produces lower probabilities of overfishing than the multi-species MSY_{MS}-based strategy. Under the single-species strategy, choke species were fished (on average) to the target biomass of $B_{MSY,SS}$ implied by the harvest rate, while remaining species started, and remained, well above $B_{MSY,SS}$, as TACs were underutilized and realised harvest rates were lower. On the other hand, choke effects represent quota that was allocated by managers according to the harvest strategy, but never utilised. This mismatch still represents lost revenue in the form of landings, or unrecoverable fixed costs in the form of lease price. In some BC fisheries, lease prices exceed 50% of the landed value of the fish [109], which may be a bitter pill to swallow for harvesters that risk leasing quota that they are choked out of utilising. Such secondary costs of choke effects are not included in this work, potentially making single-species management seem more optimistic than in reality.

Under the single-species MSY_{SS}-based strategy, there is a perception of lower economic risk. This perception is based on higher median NPV than under the multi-species maximum yield strategy. Higher NPV stems from an increase in rents early in the projection period, as a result of lower fishing costs associated with higher CPUE from under-utilised TACs caused by choke effects, and higher unit prices from lower landings, combining to transfer surplus from consumers to producers. However, while the NPV is higher, the rents still eventually decline to negative under the MSY_{SS} strategy after the initial period with positive rents, but the effect is less noticeable given the early positive rents and assumed discount rate. Rents decline below zero as unit prices are reduced with increasing household incomes, and MSY strategies do not account for changes in price. Combined with the secondary costs of choke effects outlined above, the economic performance of the MSY_{SS} based strategy is much less desirable than the NPV indicates.

The economic model used for the BC flatfish complex is fairly simple, and produces upper bounds on resource rents that may positively bias the net-present-values in the simulations. For example, fuel and crew share were the only variable costs included, but quota lease prices, which are a significant expense for many harvesters in this fishery, were ignored. While crew share can be considered a constant rate as it was treated here, the price of fuel is time-varying, meaning that the marginal cost of fishing probably changes year by year. However, time-varying fuel costs would require a model to project the price into the future, adding variability to the results[5], and would require several assumptions, given volatility and growth observed in fuel prices that has low correlation with GDP or CPI. Further, some choices would simply add noise, leaving the main findings based on average/median behaviour practically unchanged. Adding quota lease prices would change the allocation of

[5]Statistics Canada. Table 18-10-0256-01 Consumer Price Index (CPI) statistics, measures of core inflation and other related statistics - Bank of Canada definitions; `https://doi.org/10.25318/1810025601-eng`

effort over the course of a year, as when price/demand drops, the lease price also drops, except for choke species. Choke species quota actually becomes more valuable as TAC is decreased, as it is required for access to higher value species. Negative correlation between the landed price and quota lease price would lead to interesting secondary market dynamics and may lead to arbitraging behaviour [101]; however, including such dynamics would require more detailed effort dynamics, likely at the individual vessel level.

The modeling framework developed in this chapter can be extended to help answer broader ecosystem based fishery management questions. For example, the "30x30 initiative" aims to use marine spatial planning processes to protect 30% of global oceans in marine protected areas by 2030 [19]. However, there is considerable uncertainty about the effect that spatial planning has on the distribution of effort [62, 125]. One approach would be to link effort allocation to the economic model to more closely simulate how harvesters choose where to expend fishing effort [59, 83, 141]. A simple, emergent version of reallocating fishing effort to target different species compositons when prices changed was observed in the stochastic optimisations of rent in the BC flatfish complex fishery, where effort moved to QCS to target higher proportions of Rock sole as prices dropped. Another extension could evaluate predictors of choke effects, and how they relate to catch/quota balancing (TAC utilisation) observed in the past [84, 10]. TAC utilisations could be correlated with biomass indices and catch series to evaluate observable conditions that may lead to choke effects. Those predictors may then have practical utility as management 'meta-rules' that allow TACs to be scaled to permit sustainable overfishing of choke species to mitigate undesirable choke effects on high value species, within reason. Finally, an extension to multi-sector fisheries would be very valuable. Multiple gear types would create complex economic and catch yield surfaces, where the costs of fishing, catchability, and selectivity are different for different sectors [75, 55]. Economic demand may also be sector dependent, as different sectors sometimes sell in different markets based on the condition or size of landings (e.g., BC Sablefish long-line trap and hook fleets sell to Japan, while the trawl fleet sells their landings in continental North America).

4.4.1 Conclusion

I demonstrate significant benefits of including technical interactions in multi-species harvest strategies. Harvest strategies based on multi-species MSY eliminate choke effects, producing higher catches while avoiding states of conservation concern when harvest strategies are repeatedly applied to the BC flatfish complex fishery. Low conservation concern and higher catches increases food security, which is a growing need for the increasing global population. While rents under multi-species MSY-based strategies appear decline below zero over time, this is largely a function of BC flatfish complex species being inferior goods, which lowers the annual rents from all strategies. The same price reductions caused a reallocation of fishing effort under MEY based harvest strategies, highlighting that the optimal fishing effort

for a spatially heterogeneous fishery is not constant, depending on time as well as space. Finally, while choke effects in single-species MSY-based strategies protected the BC flatfish complex from overfishing and maintained higher NPV, the behaviour is strongly dependent on the relative health, catchabilities, and similar unit price of the BC flatfish complex stocks, meaning that the result is not general, and should not be taken as supporting evidence for business as usual. Despite the protective role of choke effects for the simulated BC flatfish complex fishery, real choke effects will still cause frustrations for harvesters when TACs are underutilized, as some harvesters will be left holding expensive quota that they are unable to utilise or sell.

4.5 Tables

Table 4.1: Estimates of biological population parameters, single-species reference points, and stock status in 2016 for all nine BC flatfish complex stocks.

Stock	**Species**	B_0	R_0	M_m	M_f	h	B_{MSY}	U_{MSY}	MSY	B_{2016}	B_{2016}/B_0	B_{2016}/B_{MSY}	q^F
HSHG	Dover	16.51	17.71	0.15	0.14	0.67	4.34	0.28	1.22	8.36	0.51	1.93	0.022
	English	8.60	15.27	0.19	0.17	0.68	2.21	0.39	0.87	4.85	0.56	2.19	0.026
	Rock	12.34	15.54	0.22	0.22	0.66	3.78	0.28	1.07	7.68	0.62	2.03	0.025
QCS	Dover	5.45	6.43	0.16	0.15	0.67	1.46	0.29	0.42	3.36	0.62	2.30	0.015
	English	0.57	0.98	0.19	0.17	0.68	0.15	0.40	0.06	0.41	0.72	2.73	0.016
	Rock	4.33	5.02	0.22	0.21	0.66	1.29	0.31	0.40	2.13	0.49	1.65	0.014
WCVI	Dover	13.59	18.23	0.17	0.16	0.67	3.58	0.32	1.14	8.45	0.62	2.36	0.025
	English	0.86	1.46	0.18	0.17	0.68	0.22	0.41	0.09	0.50	0.58	2.27	0.020
	Rock	1.12	1.17	0.20	0.20	0.66	0.34	0.29	0.10	0.55	0.49	1.62	0.012

Table 4.2: Table of economic sub-model variables, their assumed values, and descriptions.

Variable	Value	Description
f_{2009}	\$193,000	Fuel cost per kt of catch in 2009, in 2016 CAD
ψ	0.3	Crew share of revenue
d	0.05	Discount rate
$v_s(C_s)$		Unit price per kg of species s (CAD \$) at quantity demanded C_s (kt)
c_p		Fuel cost per unit of fishing effort (CAD \$ million)
$\pi(\vec{E})$		Equilibrium resource rent produced by annual fishing effort $\vec{E}$
γ	0.039	Annual growth rate of household income

Table 4.3: Descriptions of the sensitivity analyses of results to economic model assumptions about income, discount rate, and economic demand.

Sensitivity	Variable	Description
PED only	$\lambda_s = \infty$	Price independent of catch, but decreases with income
Income only	$\gamma = 0$	Price decreases with catch, but independent of income
PED and Income	$\lambda_s = \infty$ and $\gamma = 0$	Price independent of catch and income
High discount	$\delta = 0.1$	Discount rate of 10%
Low discount	$\delta = 0.025$	Discount rate of 2.5%
Low income growth	$\gamma = 0.02$	Income growth at 2%
High income growth	$\gamma = 0.058$	Income growth at 5.8%

Table 4.4: Maximum likelihood estimates of maximum sustainable multispecies yield (MSY_{MS}) and maximum economic yield (MEY) reference points for the British Columbia BC flatfish complex of flatfish. Effort (1000s of trawl hours) and rent (millons CAD$) are calculated by area, while catch (MSY, C_{MEY}, kt), biomass (B_{MSY}, B_{MEY}, kt), and harvest rate (U_{MSY}, U_{MEY}, unitless) are calculated by species within an area.

		BC flatfish complex maximum yield						**Maximum Economic Yield**					
Stock	**Species**	$E_{MSY,MS}$	$\pi\left(\vec{E}_{MSY,MS}\right)$	MSY_{MS}	$B_{MSY,MS}$	$U_{MSY,MS}$	$\frac{U_{MSY,MS}}{U_{MSY,SS}}$	E_{MEY}	$\pi\left(\vec{E}_{MEY}\right)$	MEY	B_{MEY}	U_{MEY}	$\frac{U_{MEY}}{U_{MSY,SS}}$
HSHG		16.16	0.16					5.35	0.82				
	Dover			1.18	5.77	0.205	0.73			0.82	9.96	0.082	0.29
	English			0.87	2.31	0.375	0.95			0.65	4.66	0.139	0.35
	Rock			1.03	2.81	0.368	1.30			0.91	6.37	0.144	0.51
QCS		22.90	0.11					8.96	0.22				
	Dover			0.40	1.91	0.211	0.74			0.30	3.13	0.095	0.33
	English			0.06	0.16	0.364	0.92			0.05	0.30	0.152	0.39
	Rock			0.39	1.04	0.380	1.22			0.35	2.18	0.159	0.51
WCVI		20.46	0.08					7.17	0.32				
	Dover			1.14	3.60	0.316	1.00			0.92	6.86	0.134	0.42
	English			0.09	0.20	0.427	1.08			0.07	0.43	0.161	0.41
	Rock			0.10	0.35	0.277	0.97			0.07	0.68	0.105	0.37

Table 4.5: Proportions of inverse average catchability and the estimated optimal allocation of effort among areas.

Area	$1/\bar{q}^{(F)}$	$E_{MSY,MS}$	E_{MEY}	$E^*_{MSY,MS}$	E^*_{MEY}
HSHG	0.256	0.272	0.249	0.275	0.189
QCS	0.416	0.385	0.417	0.383	0.475
WCVI	0.328	0.344	0.334	0.343	0.336

Table 4.6: Median values of dynamic approximations to maximum sustainable multispecies yield (MSY_{MS}) and maximum economic yield (MEY) reference points for the British Columbia BC flatfish complex of flatfish derived via stochastic optimisation. Effort (1000s of trawl hours) and annual resource rent ($M CAD) are calculated by area, while catch (MSY, C_{MEY}, kt), biomass (B_{MSY}, B_{MEY}, kt), and harvest rate (U_{MSY}, U_{MEY}, %) are calculated by species within an area.

		BC flatfish complex maximum yield						**Maximum Economic Yield**					
Stock	**Species**	E^*_{MSY}	$\pi\left(\vec{E}^*_{MSY}\right)$	MSY^*_{MS}	$B^*_{MSY,MS}$	$U^*_{MSY,MS}$	$\frac{U^*_{MSY,MS}}{U_{MSY,SS}}$	E^*_{MEY}	$\pi\left(\vec{E}^*_{MSY}\right)$	MEY^*	B^*_{MEY}	U^*_{MEY}	$\frac{U^*_{MEY}}{U_{MSY,SS}}$
HSHG		16.48	-0.86					3.00	0.18				
	Dover			1.23	5.81	0.210	0.75			0.55	11.36	0.048	0.17
	English			0.85	2.21	0.384	0.98			0.43	5.47	0.079	0.20
	Rock			1.09	2.90	0.375	1.34			0.66	7.91	0.083	0.30
QCS		22.95	-0.11					7.55	0.06				
	Dover			0.42	1.99	0.214	0.74			0.26	3.26	0.082	0.28
	English			0.06	0.16	0.367	0.92			0.04	0.32	0.128	0.32
	Rock			0.41	1.08	0.383	1.24			0.31	2.28	0.134	0.43
WCVI		20.54	-0.28					5.34	0.05				
	Dover			1.17	3.73	0.320	1.00			0.75	7.35	0.102	0.32
	English			0.09	0.21	0.431	1.05			0.06	0.49	0.121	0.30
	Rock			0.10	0.36	0.279	0.96			0.06	0.73	0.079	0.27

Table 4.7: Median and central 90% of the distribution of net present value of the BC flatfish complex fishery (in millions $CAD) in the projection period, median 2040 - 2060 effort $\vec{E}_p$ (1000s hours) and median annual rent $\pi(\vec{E}_p)$ (millions $CAD), taken over all simulation replicates for the closed loop simulations with simulated assessment errors. Net present value is discounted by an annual rate of 5%.

Target HR	Area	NPV	$\vec{E}_p$	$\pi(\vec{E}_p)$
$U_{MSY,SS}$		5.04 (2.92, 7.93)		
	HSHG	2.34 (0.65, 3.99)	11.33	-0.35
	QCS	1.56 (0.94, 2.08)	17.62	-0.04
	WCVI	1.27 (0.16, 2.43)	17.30	-0.19
$U_{MSY,MS}$		3.15 (0.57, 6.28)		
	HSHG	0.49 (-1.13, 2.4)	13.69	-0.56
	QCS	1.35 (0.83, 1.95)	19.19	-0.06
	WCVI	1.18 (0.05, 2.36)	17.76	-0.21
$U^*_{MSY,MS}$		2.75 (0.14, 5.94)		
	HSHG	0.18 (-1.52, 2.11)	14.06	-0.60
	QCS	1.32 (0.78, 1.94)	19.44	-0.07
	WCVI	1.12 (-0.01, 2.33)	17.95	-0.21
U_{MEY}		12.04 (11.13, 13.12)		
	HSHG	6.55 (5.91, 7.55)	4.68	0.15
	QCS	2.34 (2.02, 2.65)	7.64	0.06
	WCVI	3.14 (2.58, 3.63)	6.35	0.05
U^*_{MEY}		11.71 (10.84, 12.79)		
	HSHG	6.23 (5.71, 6.99)	2.62	0.20
	QCS	2.41 (2.12, 2.75)	6.44	0.07
	WCVI	2.99 (2.5, 3.51)	4.73	0.07

Table 4.8: BC wide and area-specific NPV (CAD $ million), area specific median commercial trawl effort E^*_{MEY} (1000s hours), and area-specific undiscounted annual resource rents MEY^* (CAD $ million) under the constant price sensitivity analyses. Median effort and rents are taken over all simulation replicates and the period 2040 - 2060, and NPV is discounted by an annual rate of 5%.

Scenario	Area	NPV	E^*_{MEY}	MEY^*
Infinite PED, Income growth		48.63		
	HSHG	30.99	8.99	1.78
	QCS	9.76	14.33	0.56
	WCVI	7.72	8.94	0.31
Infinite PED, Fixed Income		35.99		
	HSHG	21.81	8.00	0.99
	QCS	6.00	12.27	0.28
	WCVI	8.01	8.97	0.33
Finite PED, Fixed Income		24.62		
	HSHG	14.38	5.91	0.73
	QCS	3.97	9.20	0.20
	WCVI	6.25	7.47	0.27

Table 4.9: BC wide and area-specific NPV (CAD $ million), area specific median commercial trawl effort E^*_{MEY} (1000s hours), and area-specific undiscounted annual resource rents MEY^* (CAD $ million) under the discount rate and income growth rate sensitivity analyses. Median effort and rents are taken over all simulation replicates and the period 2040 - 2060, and NPV is discounted by an annual rate of 5%.

Scenario	Area	NPV	E^*_{MEY}	MEY^*
Low Discount $d = 0.025$		13.67		
	HSHG	7.82	2.45	0.17
	QCS	2.31	5.93	0.07
	WCVI	3.58	3.98	0.06
High Discount $d = 0.1$		13.01		
	HSHG	7.46	4.49	0.15
	QCS	2.17	8.45	0.05
	WCVI	3.42	6.35	0.04
Low Income Growth $\gamma = 0.02$		17.99		
	HSHG	10.31	4.37	0.38
	QCS	2.96	8.38	0.11
	WCVI	4.74	6.37	0.13
High Income Growth $\gamma = 0.058$		10.52		
	HSHG	6.02	2.10	0.07
	QCS	1.78	6.30	0.03
	WCVI	2.73	4.18	0.01

4.6 Figures

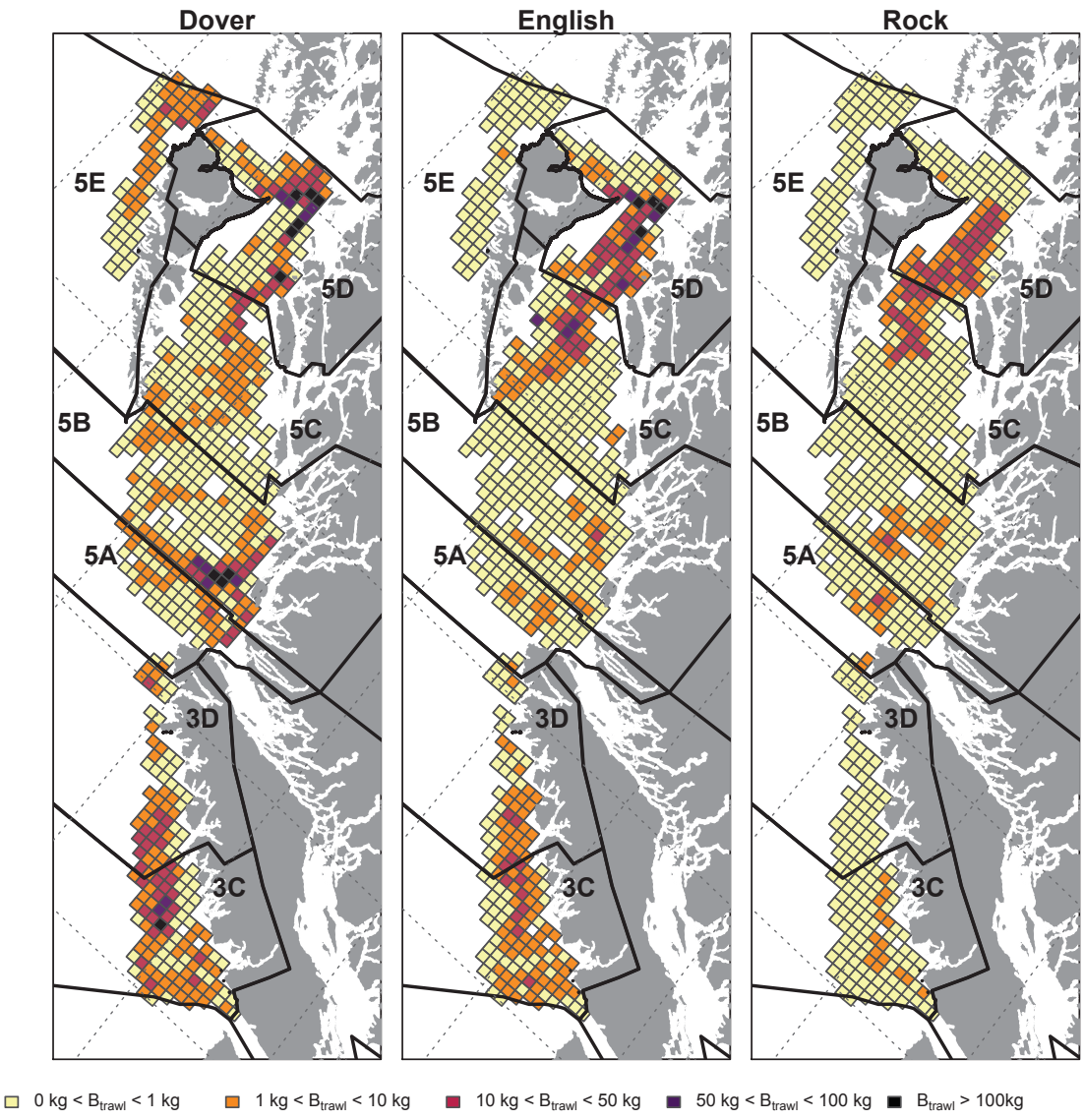

Figure 4.1: Mininum trawlable survey biomass B_{trawl} estimates for DER complex species on the BC coast, aggregated to a 10km square grid. Estimates are produced by scaling average trawl survey (kg/m^2) density values in each grid cell by the cell's area in m^2. Locations that do not show a coloured grid cell do not have any survey blocks from which to calculate relative biomass. Survey density for each grid cell is calculated from data for the Hecate Strait Assemblage Survey and the BC Groundfish Trawl Synoptic Survey, stored in the GFBio data base maintained at the Pacific Biological Station of Fisheries and Oceans, Canada. Thick black lines delineate the major statistical areas 3CD and 5ABCDE used for groundfish management BC, while the dashed grey lines makr out latitude and longitude, indicating the rotation of the coordinates to save space. The full colour figure is available in the online version of the article.

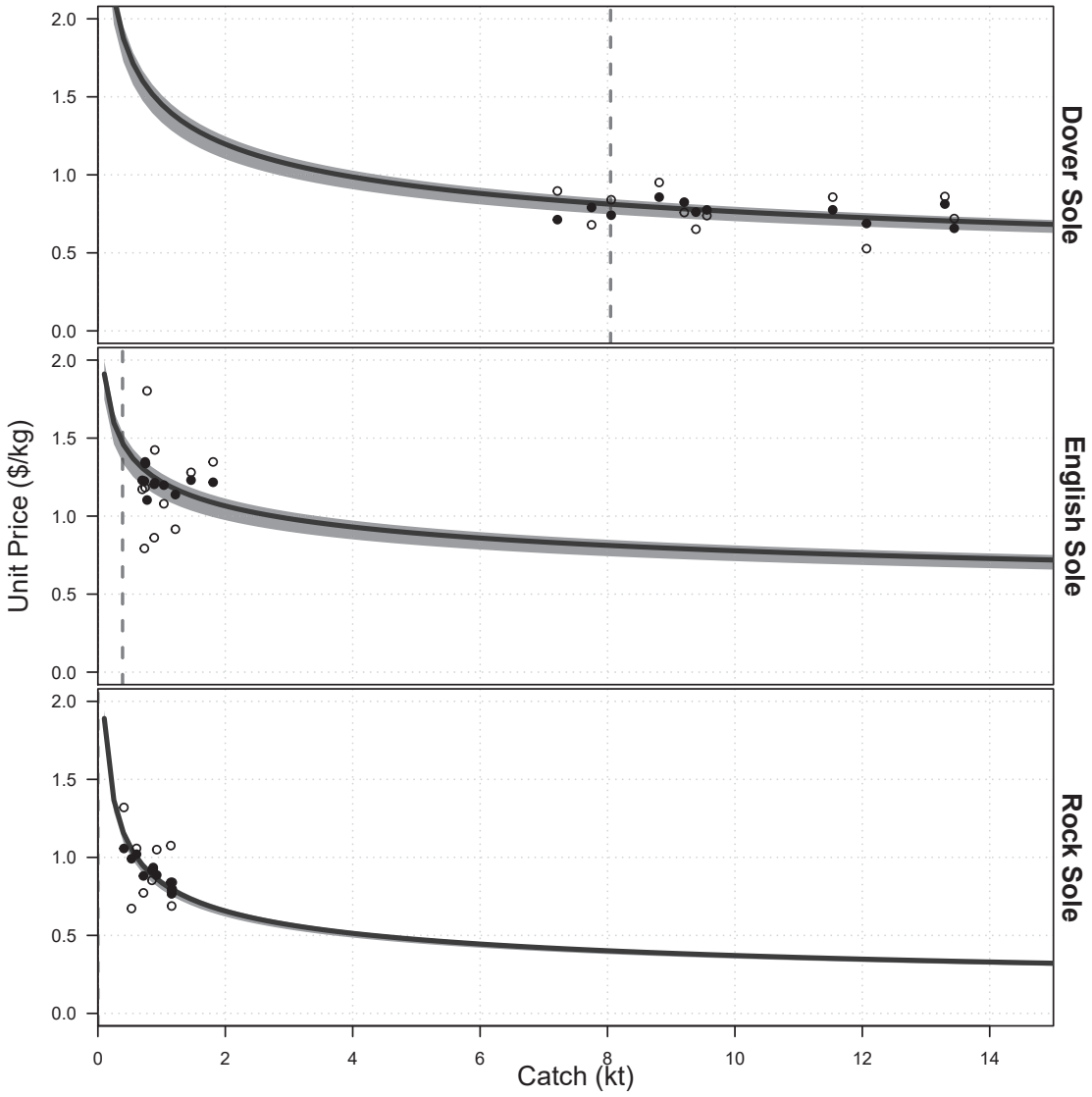

Figure 4.2: Demand curves showing the relationship between unit price ($/kg, vertical axis) and quantity supplied (Catch, kt, horizontal axis) for Dover, English, and Rock soles. Lines show the median over the period 2006 - 2016, while the shaded region shows the central 90% of the demand over the same time period. Points show raw catch and price data (open circles), and the same catch with price estimated with cost of fishing as an instrumental variable (closed circles). The vertical dashed lines show the exogenous US West Coast catch for each species, assumed to be fixed in the projections.

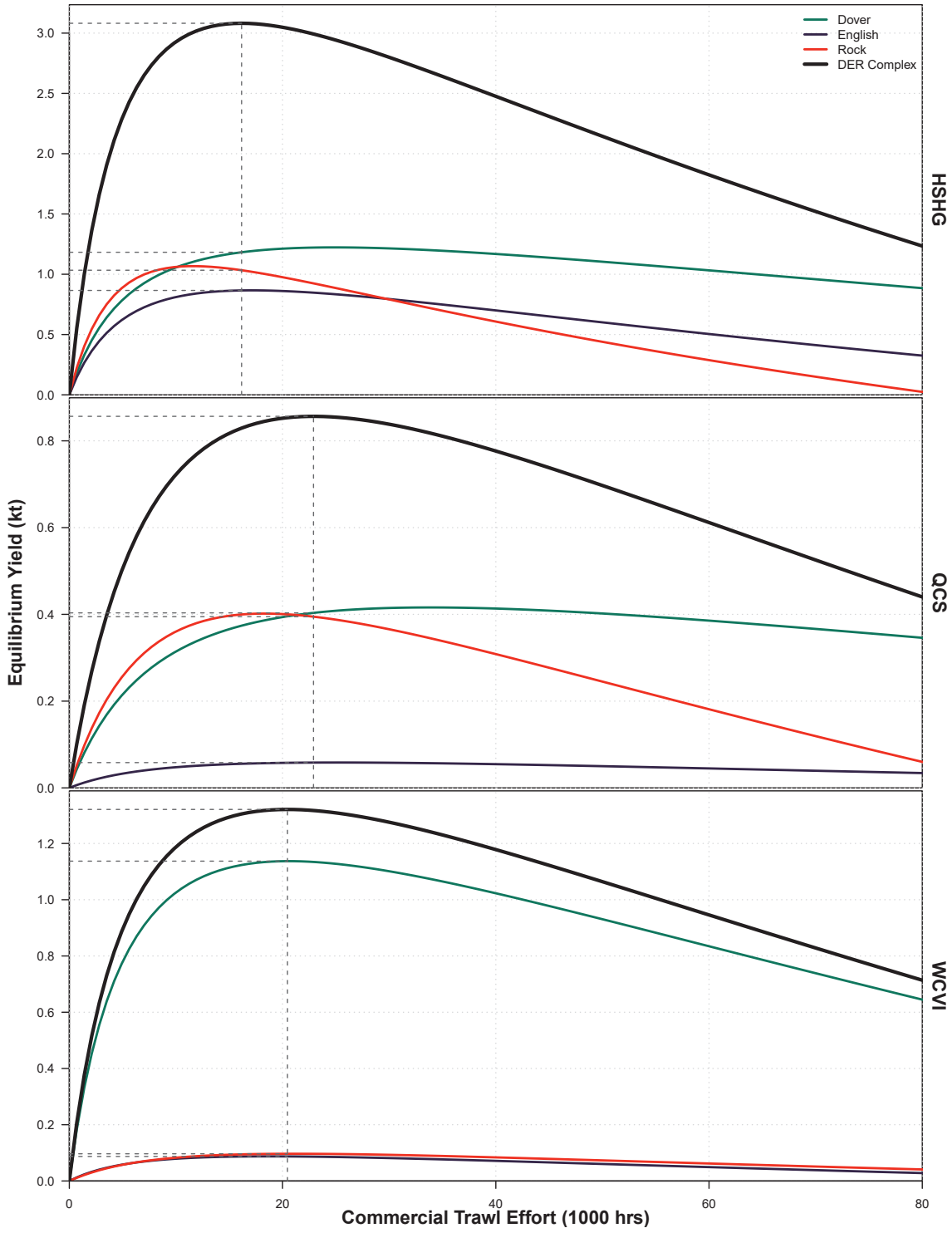

Figure 4.3: Equilibrium catch yield by species (coloured lines) and for the BC flatfish complex (thick black line) as a function of fishing effort in each stock area. Effort $E_{MSY,MS,p}$ maximising the multi-species yield in each area is marked by the vertical dashed line segment, and the multi-species yield $MSY_{MS,p}$ and species yields at that effort are shown by horizontal dashed line segments meeting the $E_{MSY,MS,p}$ segment where it intersects each curve.

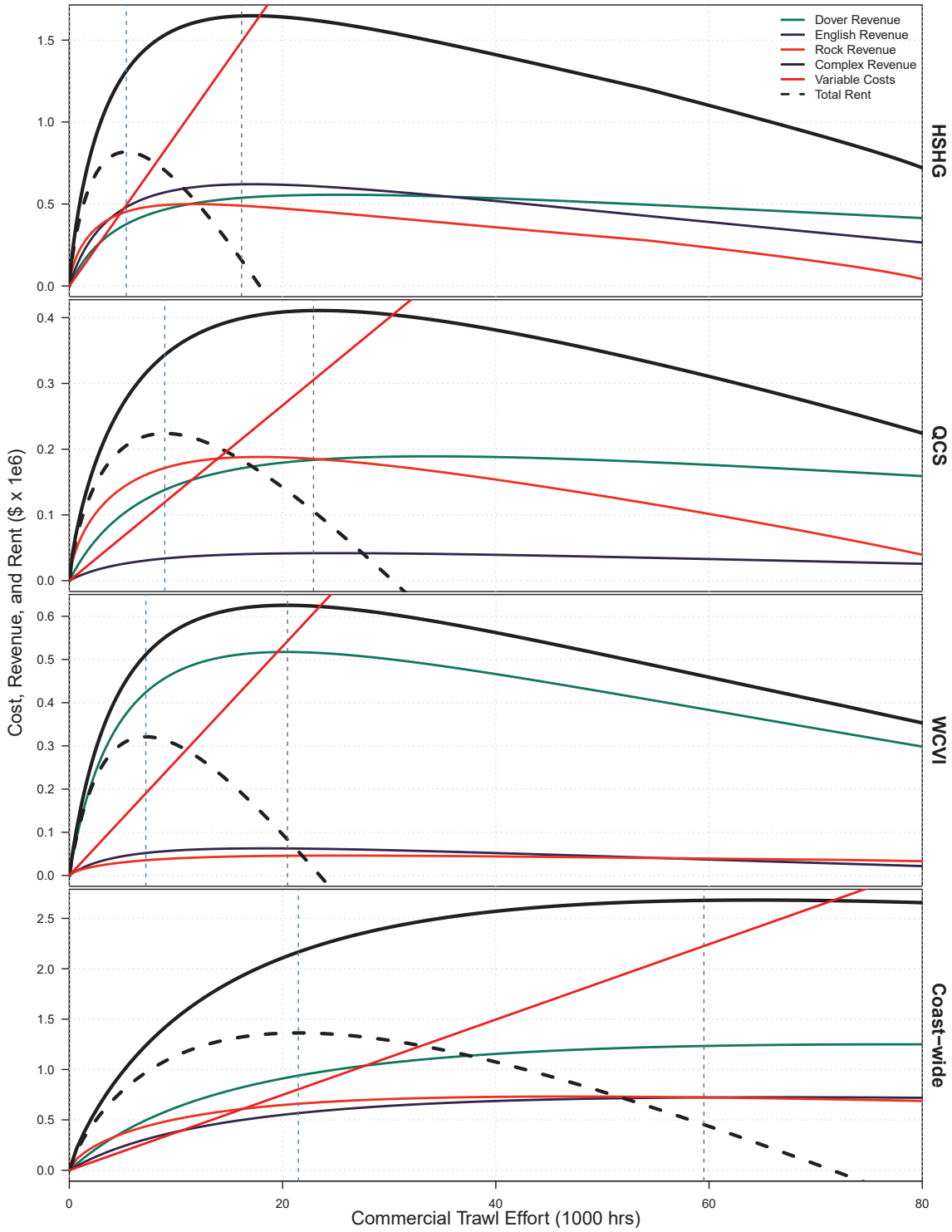

Figure 4.4: Net resource rents (dashed curve), species revenues (coloured curves) and complex revenue (thick black curve), and fuel costs of fishing (red line), all as functions of fishing effort for the BC flatfish complex in each area (first three panels), as well as coast-wide (bottom panel). Vertical dashed lines show $E_{MSY,MS}$ (black, maximising revenue) and E_{MEY} (blue, maximising rent).

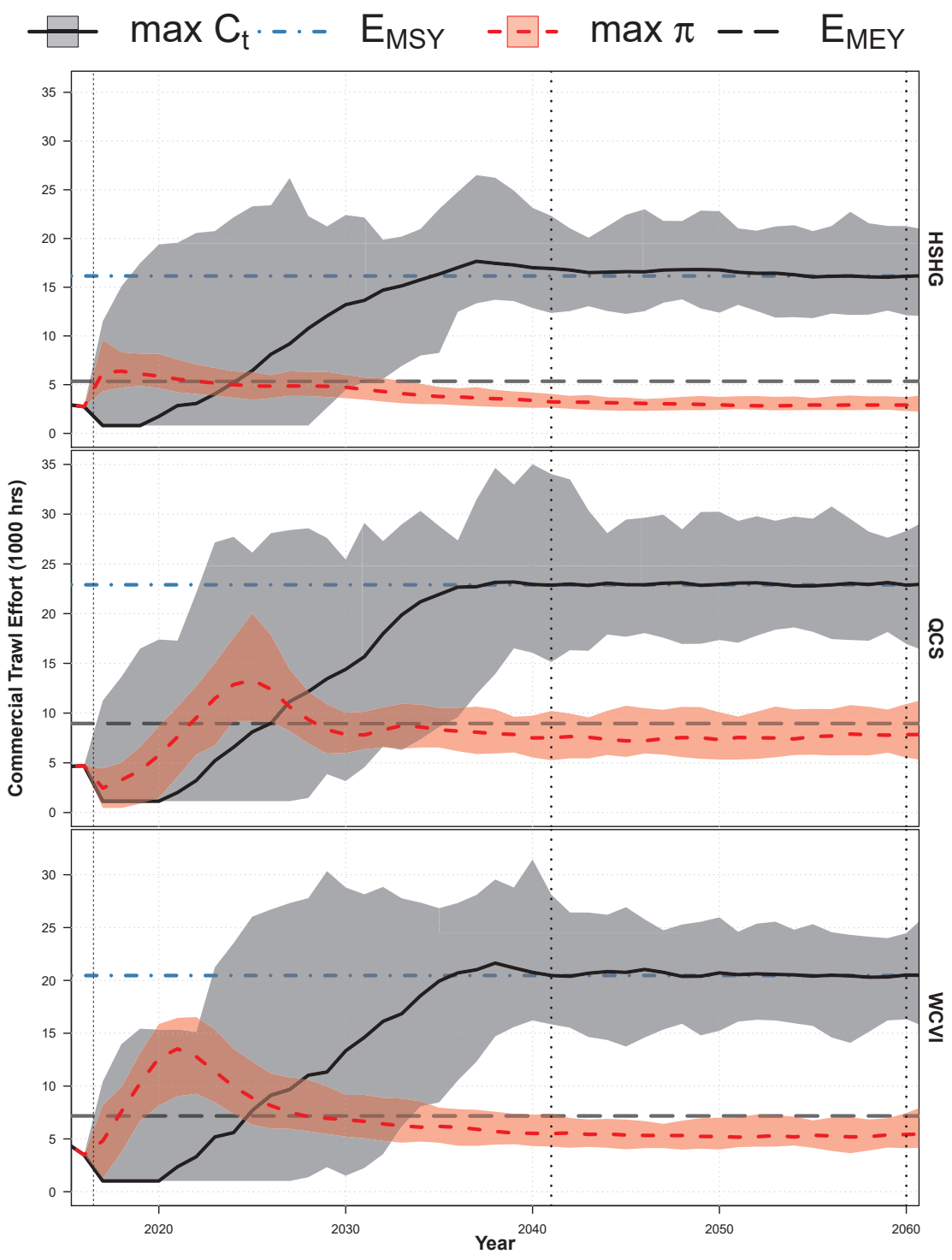

Figure 4.5: Simulation envelopes of commercial trawl effort over the projection period for the stochastic optimisation under the catch maximising (black/grey) and rent maximising (red/pink) objective functions. Horizontal dashed lines show steady-state E_{MSY} (long/short dashed) and E_{MEY} (long dashed), while vertical lines show the beginning of the projection period (thin short dashed in 2017) and the beginning and end of the time period over which dynamic optima are calculated (thicker short dashed lines in 2041 and 2060).

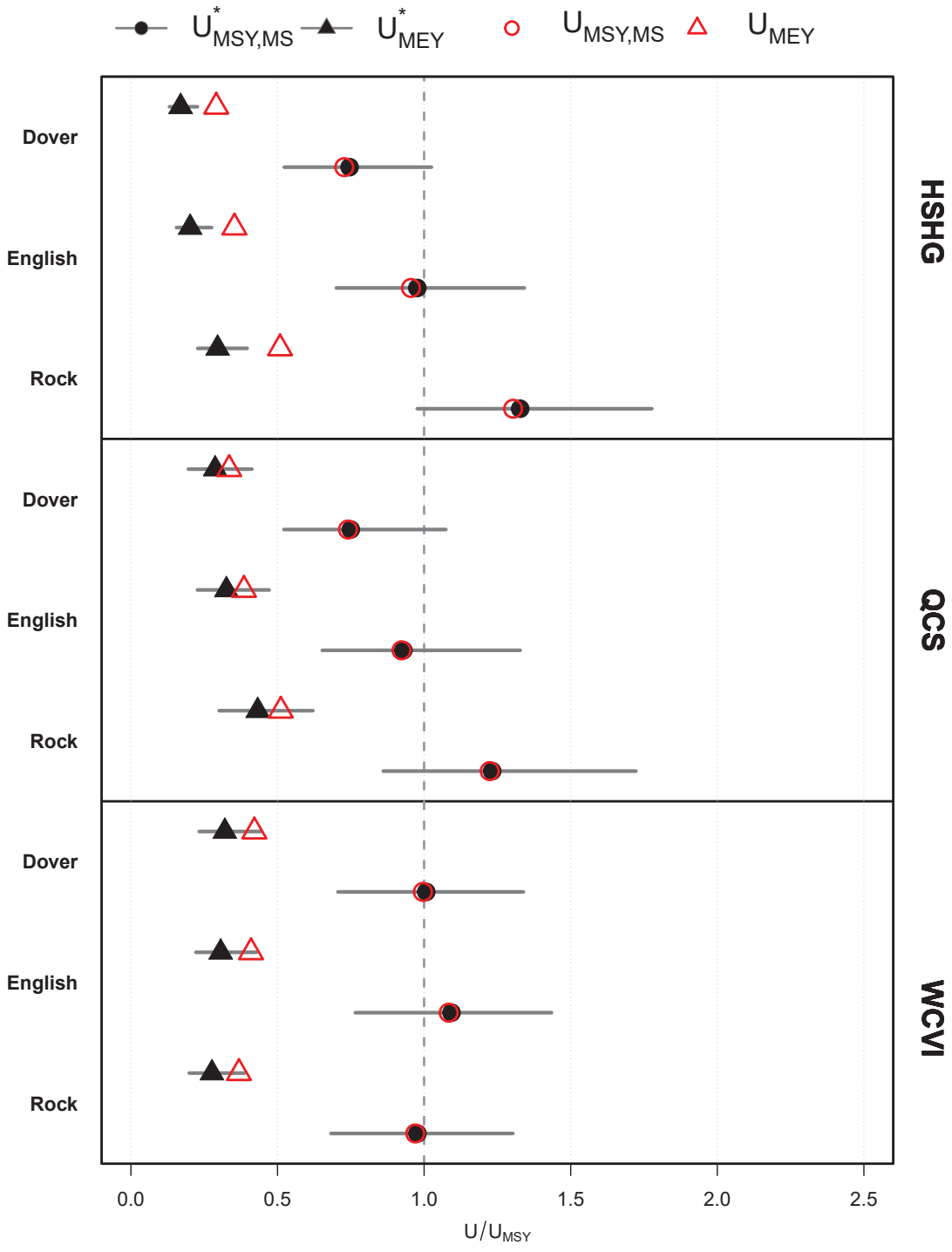

Figure 4.6: Medians (points) and central 95% (grey bars) of simulated harvest rates under the stochastic optimisation between 2041 and 2060 when maximising complex catch (circles) and rent (triangles). The horizontal axis is scaled to single-species equilibrium $U_{MSY,SS}$, indicated by a vertical dashed black line, with steady state optima shown as open red points for comparison.

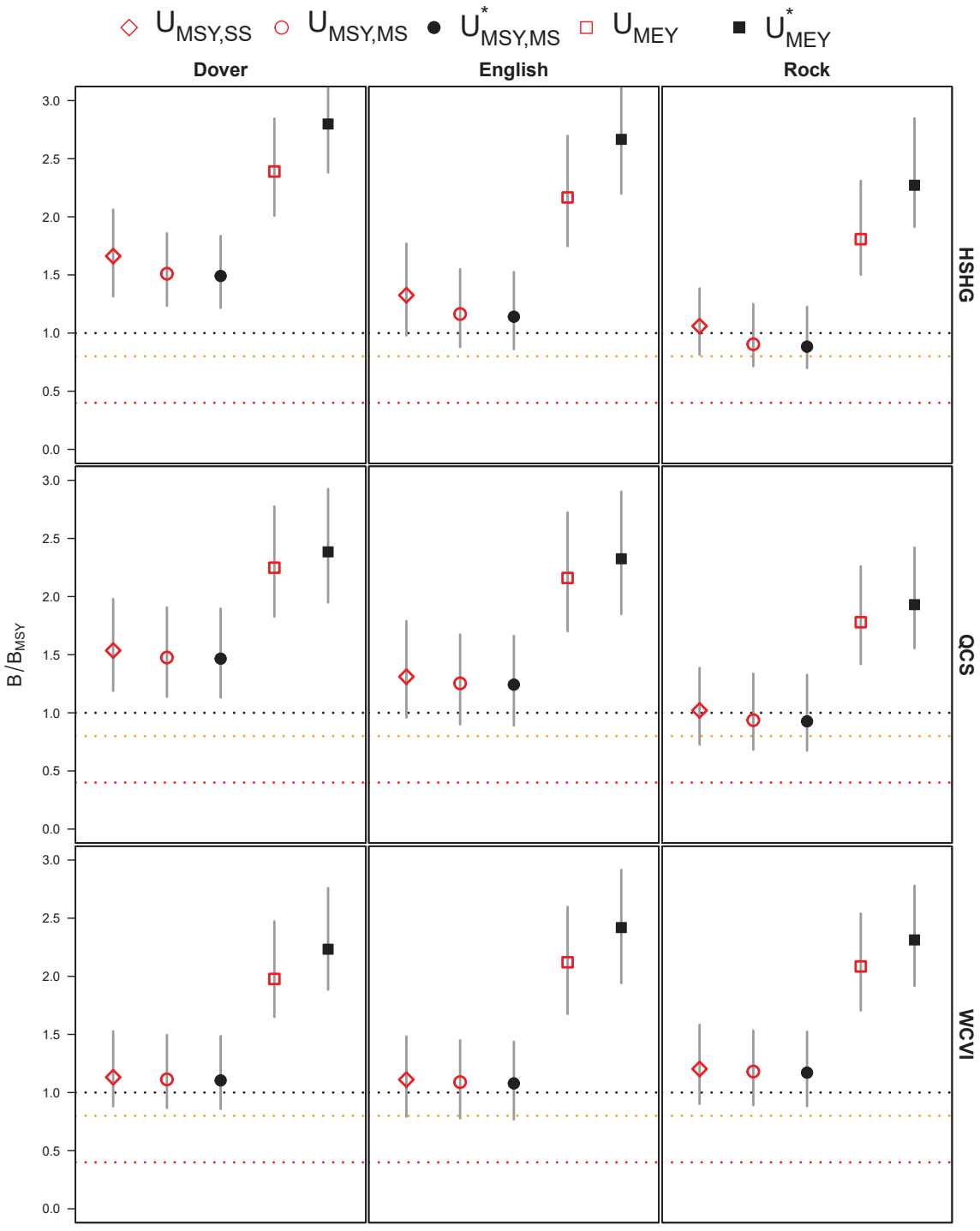

Figure 4.7: Distributions of 2048 biomass relative to single-species B_{MSY} for all BC flatfish complex stocks under the closed loop simulations. Line segments show the central 90% of the simulated biomass distribution and points show median biomas. Line segments differ for the perfect (solid) and imperfect (dashed) information scenarios, while points differ between steady state (open) and dynamic (closed) optimal harvest rates targeting maximum catch (circle) or net-present value of resource rents (square). Horizontal dashed reference lines show B_{MSY} (grey) and $0.8B_{MSY}$ (orange) to indicate conservation risk.

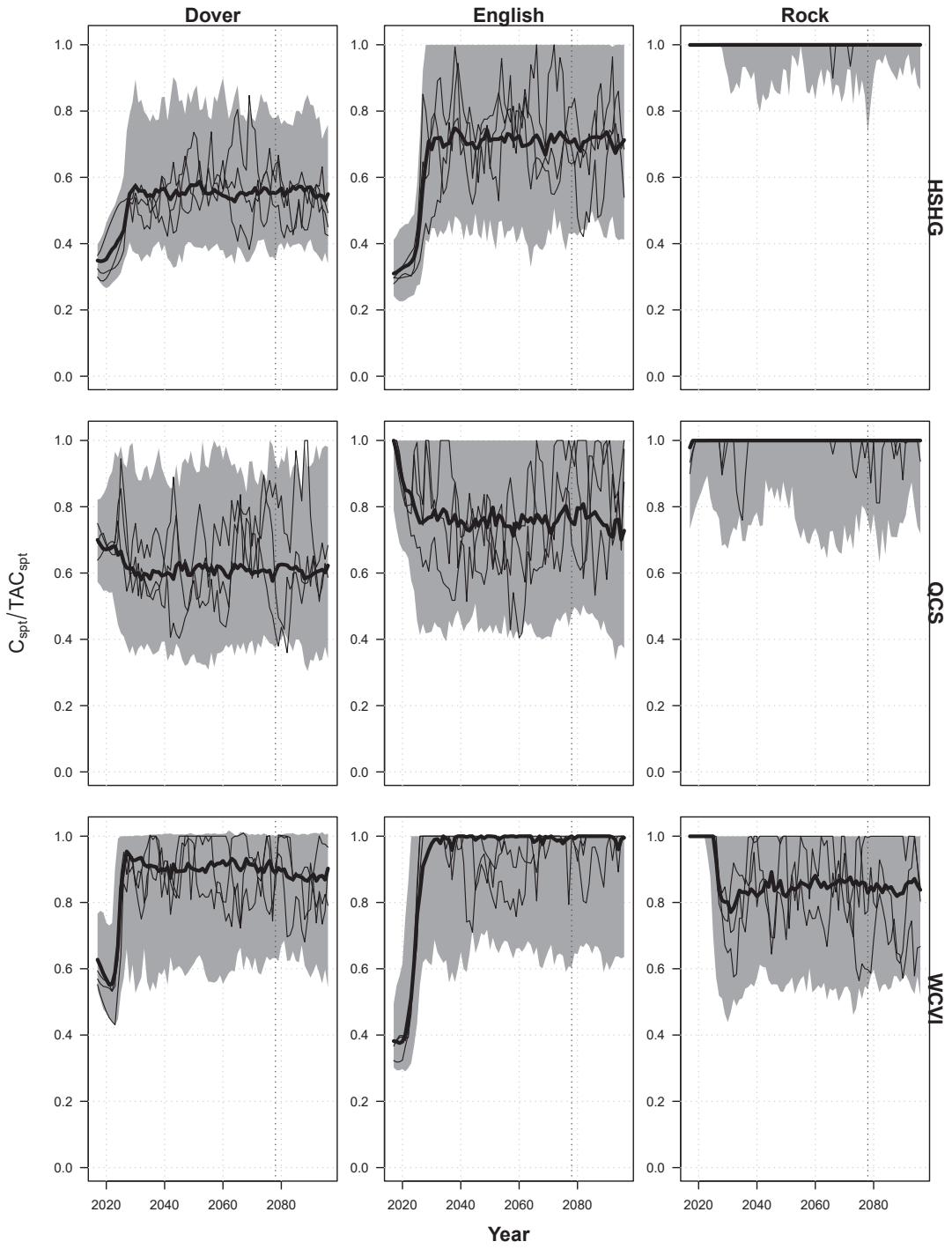

Figure 4.8: Median (thick black line), central 90% (grey region) and single simulation traces (thin lines) for TAC utilisation when targeting single-species $U_{MSY,SS}$ harvest rates in the closed loop feedback simulation.

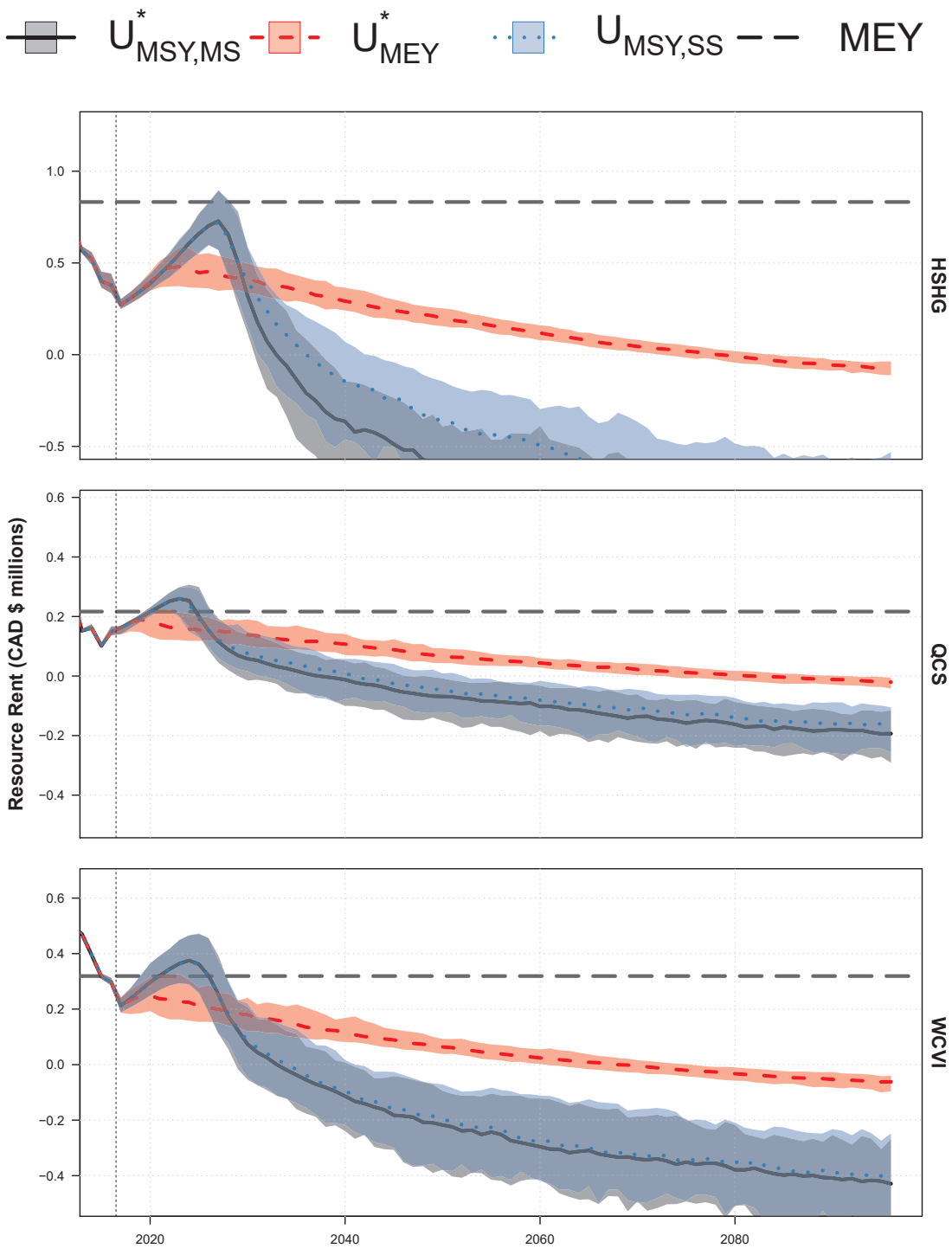

Figure 4.9: Simulated central 90% (shaded regions) and median annual (thick line) resource rents extracted by the trawl fishery over the projection period under the imperfect information scenario with TACs set using dynamic optima U^*_{MSY} (grey region, solid black line) and U^*_{MEY} (pink region, dashed red line). Horizontal dashed lines show steady-state MEY for each stock area, while the vertical dashed line shows the beginning of the projection period in 2017.

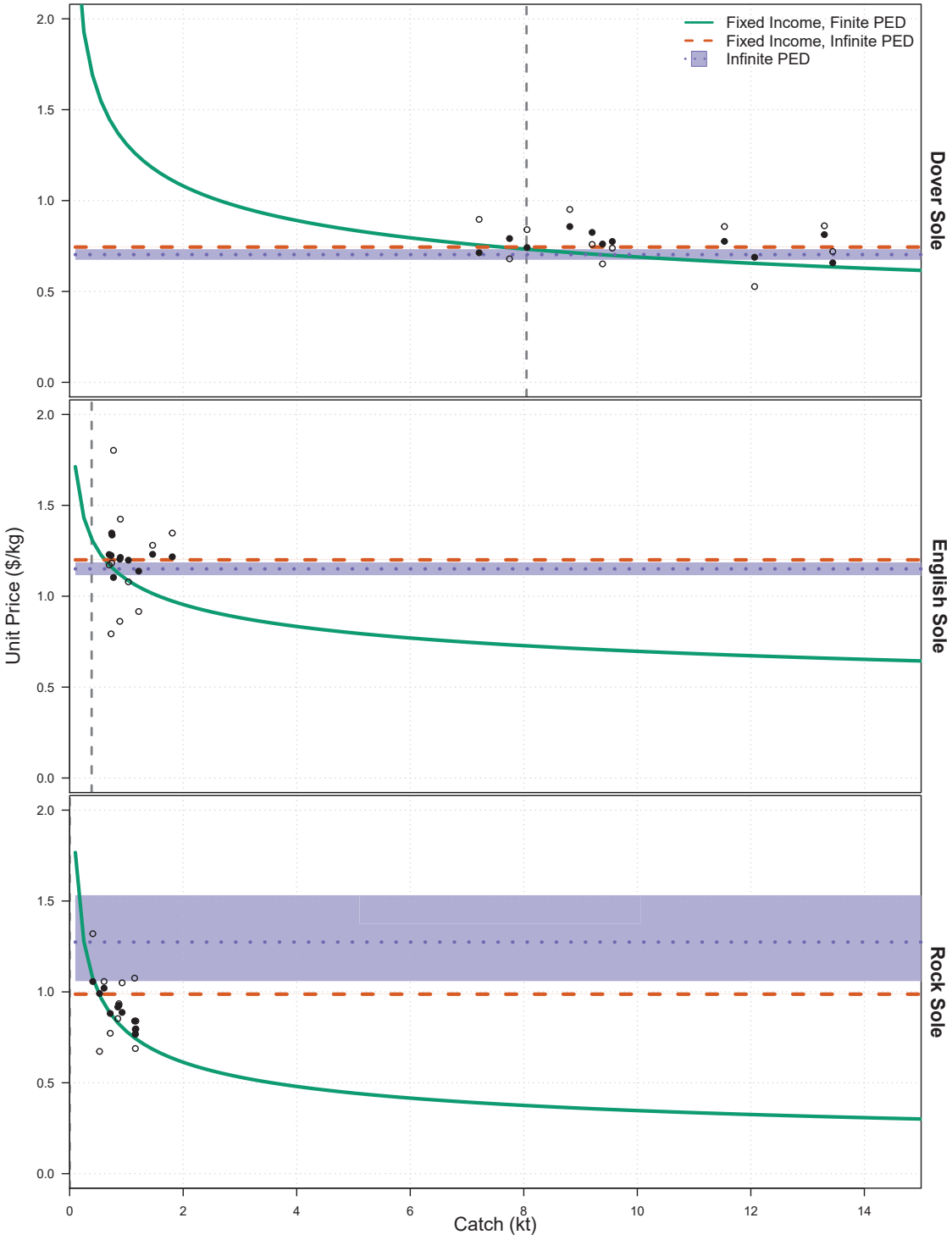

Figure 4.10: Demand curves showing the relationship between unit price ($/kg, vertical axis) and quanitity supplied (Catch, kt, horizontal axis) for Dover, English, and Rock soles for the three demand curve sensitivity analyses. Lines, points, and shaded regions are as in Figure 4.2.

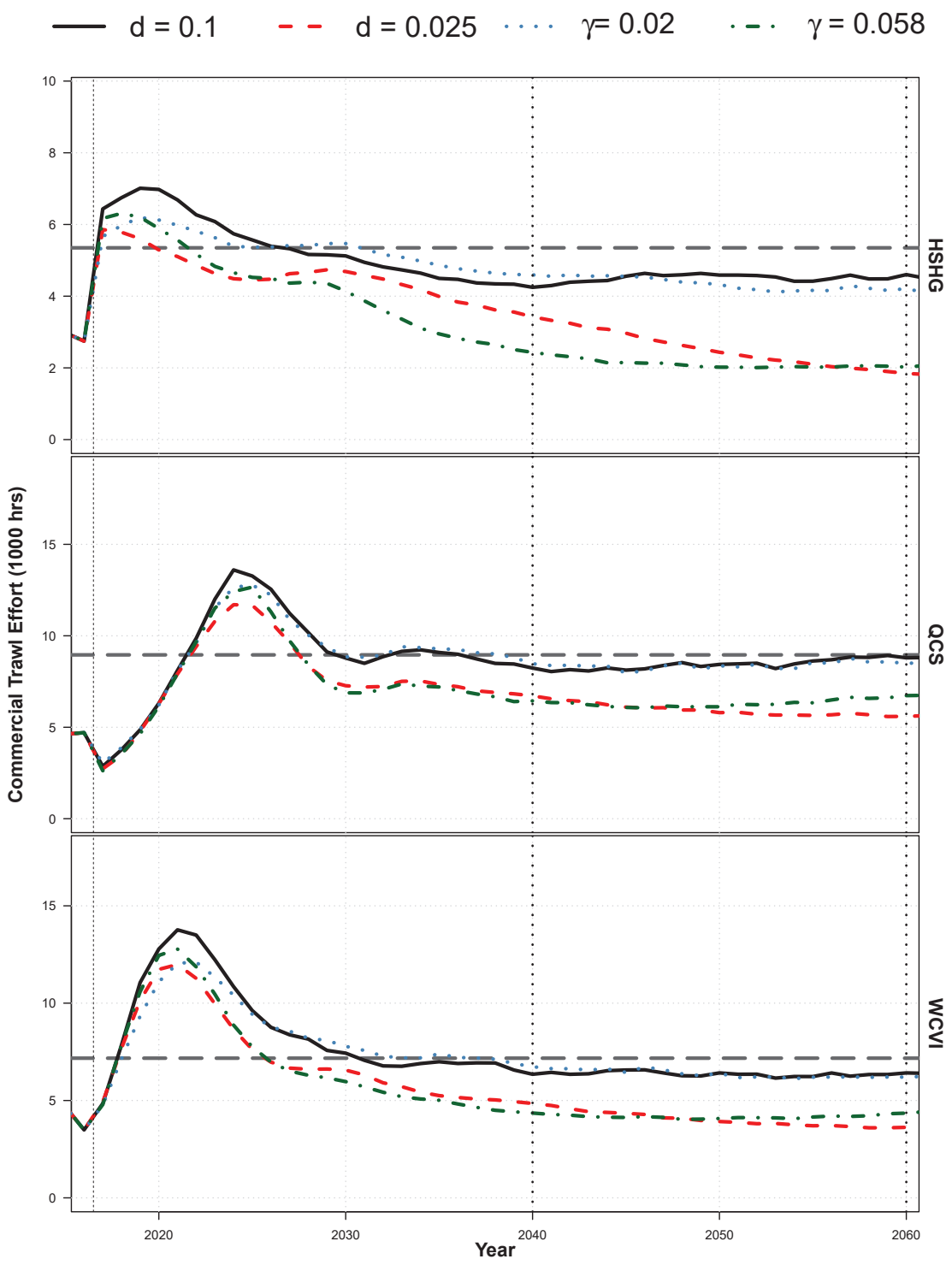

Figure 4.11: Annual median commercial trawl effort maximising NPV under the discount and income growth rate sensitivity analyses. Horizontal dashed lines show YPR E_{MEY} (long dashed), while vertical lines show the beginning of the projection period (thin short dashed in 2017) and the beginning and end of the time period over which dynamic optima are calculated (thicker short dashed lines in 2041 and 2060).

4.7 Appendices

4.A BC flatfish complex demand analysis

The unit price v_s ($/kg) of each BC flatfish complex species s is assumed to follow a downward sloping demand curve for total catch C_s (quantity demanded) with a constant own-price elasticity of demand. Demand curve identifiability was a problem, given the low sample size (11 years) of economic data and the sometimes positive relationship between BC catch and the unit price, indicating either demand or supply shifts that created correlation between explanatory variables (i.e., price) and the residual in the quantity demanded, also known as endogeneity [2].

BC flatfish complex demand curves are fit to landings, ex-vessel prices, and BC household income per capita. Average ex-vessel unit price ($/kg) data for the BC flatfish complex is obtained from dockside fish-slips for the BC trawl fishery from 2006 - 2016, and calculated as total ex-vessel revenues divided by total landings [6]. All ex-vessel prices are specific to the bottom trawl fishery except for English Sole in 2016, which is an outlier with a trawl ex-vessel price almost three times the 2006 - 2016 average, and is therefore replaced with the 2016 ex-vessel price for English sole averaged across all available gear types. All prices are then adjusted to 2016 dollars by using the annual average consumer price index (CPI) from Statistics Canada [7]. BC household income per capita data (Table A1) is also obtained from Statistics Canada [8].

Annual catches of Dover and English soles on the US west coast (i.e., Washington, Oregon, and California, but not Alaska) are added to the BC catch (Table A1) [147, 23]. The BC groundfish trawl fishery is known coloquially as an "I-5 Fishery" (pers comm. Chris Sporer), because, in addition to local BC customers, fish caught in BC are sold to restaurants and retailers along the US west coast via the I-5 highway. Similarly, fish caught on the US west coast are sold in BC. Therefore, the market price for BC trawl catch is dependent on the catch from the US west coast, and demand models fit to catch and price data do not match economic assumptions about market dynamics without the exogenous catch data (i.e., demand curves are upward sloping). Rock sole catch is not assessed on the west coast because there are not significant populations or catches. Further, Rock sole demand curves are downward sloping without the exogenous catch. US west coast catch data are unavailable after 2013 for English sole as there has not been an assessment since

[6]Department of Fisheries and Oceans (2019). Commercial Pacific Landings. Unpublished data Table. Available from Pacific.CatchStats@dfo-mpo.gc.ca

[7]Table 18-10-0256-01 Consumer Price Index (CPI) statistics, measures of core inflation and other related statistics - Bank of Canada definitions, https://doi.org/10.25318/1810025601-eng

[8]Table 36-10-0229-01, Long-run provincial and territorial data, https://doi.org/10.25318/3610022901-eng

that time, so West coast US English sole catch for 2013 - 2016 is assumed to be constant at the 2013 level.

Table 4.A.1: US catches and BC household income per capita data used for estimating the BC flatfish complex inverse demand curves.

Year	Dover	English	GDPpc
2006	5.957	1.0662	34.9156
2007	9.264	0.7894	36.8733
2008	11.203	0.4201	37.6750
2009	11.731	0.4155	37.1238
2010	10.374	0.2581	37.9491
2011	7.261	0.1981	39.4824
2012	6.772	0.2161	40.6199
2013	7.391	0.2241	42.1524
2014	6.011	0.2241	43.2927
2015	5.823	0.2241	45.4112
2016	6.746	0.2241	46.9050

A 2-stage least squares (2SLS) regression was used to reduce endogeneity, predicting the unit price using instrumental variables [1]. The instrument was the unit cost of supplying catch of each species, estimated via assumed fuel costs

$$w_{s,t} = \frac{\sum_p E_p \cdot c_p}{\sum_p C_{s,p}}, \tag{4.4}$$

. where c_p is derived in the Section 2. The first stage then regresses the observed ex-vessel log-price $v_{s,t}$ for species s in year t year on the log-cost of supply

$$\log v_{s,t} = \alpha_{0,s} + \alpha_{1,s} \log w_{s,t}, \tag{4.5}$$

and uses the resulting coefficents $\alpha_{i,s}$ to predict a unit price $\hat{v}_{s,t}$ with less (or no) endogeneity. The predicted price $\hat{v}_{s,t}$ is then used as an explanatory variable in the second stage regression for the demand curve

$$\log C_{s,t} = \beta_{0,s} + \lambda_s \log \hat{v}_{s,t} + \beta_{1,s} y_t, \tag{4.6}$$

where $C_{s,t}$ is the US west coast and BC catch in year t, λ_s is the own-price elasticity for species s [17], $\beta_{1,s}$ is the income elasticity, and y_t is the British Columbia household income per capita in year t. When modeling an inverse demand curve, with price as the response variable, the quantity $\frac{1}{\lambda_s}$ is called price flexibility ƒ [66]. Models were also fit with cross-price elasticity of demand, but no cross-price elasticity parameters were significant, and are therefore not reported or used in projections.

Demand relationships from the 2SLS fit the data well, and are suitable for projections in closed-loop simulations. Coefficients from price regressions on cost of fishing instruments are not significant at the 0.05 level for any species (Table A1), but do reasonably well at reducing endogeneity in the price data (Figure 4.2, main body, compare open points to closed point). Despite the low explanatory power of the fuel cost instrument, IV estimates of unit prices showed that all three species have elastic demand (Table A2), which is within range of previous demand analyses of flatfish [5]. English and Rock sole demand was more elastic than Dover sole demand, which was only slightly elastic (Table A2) and not significant at the $p = 0.1$ level ($p = 0.101$). The higher probability of a spurious estimate for Dover sole is expected, as the catch and price data show a fairly flat relationship, even after the IV regression, that would fit equally well to many different price elasticities of demand (Fig 2, main body). However, the inverse demand relationship is sufficient for the purpose of simulating realistic demand in model projections, as the curve does not contradict the data, and the demand is downward sloping [102]. Further, p values are continuous, and $p = 0.101$ is a low probability of a spurious effect, given the uninformative data. All species have a negative income elasticity parameter β_1, indicating that they are all inferior goods [33], which means that families with more disposable income will switch to more expensive sources of protein, such as higher value fish or beef, rather than the inexpensive flatfishes of the BC flatfish complex.

Table 4.A.2: Two-stage least-squares regression results when modeling inverse demand (price is the response variable). An asterisk denotes that the parameters were significant at the $p = 0.1$ level.

	Price IV model		**Inv demand**			
Species	α_0	α_1	β_0	∱	β_1	λ
Dover	-0.973	-0.632	2.554	-0.279	-0.593	-3.584
English	0.176	0.401	2.537*	-0.195*	-0.636*	-5.128
Rock	0.007	0.477	1.266*	-0.353*	-0.393*	-2.833

To simulate price as a function of catch (i.e., the inverse demand relationship) in model projections, two simplifying assumptions are made. First, US West Coast Dover and English sole catches are assumed to be exogenous and fixed at the average catch for the 2006 - 2016 time period, and second, growth in median BC household income is assumed to be at the fixed rate of 3.9% per year [9].

For sensitivity analyses, inverse demand curves are refit to the price instrument and household incomes assuming own-price elasticity was infinite (i.e., $\frac{1}{\lambda_s} = 0$ for all species).

[9] https://www2.gov.bc.ca/assets/gov/british-columbians-our-governments/government-finances/financial-economic-review/financial-economic-review-2020.pdf

When price-flexibility ƒ is removed from the inverse demand curve, the parameters are very different. Dover and English soles have no significant parameters ($0.4 < p < 0.65$), while Rock sole's inverse demand curve has significant coefficients at the $p = 0.1$ level. Further, while Dover and English soles have a reduction in price with household incomes, as above, Rock sole's relationship with income is reversed, showing that it is no longer an inferior good and demand will increase with household income. This indicates that own-price elasticity and income elasticity are confounded, but is not wholly unexpected, as Rock sole has the smallest magnitude β_1 coefficient in the finite elasticity case, and is sometimes considered a more desirable eating fish.

Table 4.A.3: Two-stage least-squares regression results when modeling inverse demand assuming own-price elasticity is infinite. An asterisk denotes that the parameters were significant at the $p = 0.1$ level.

	Inv demand	
Species	β_0	β_1
Dover	0.255	-0.143
English	0.587	-0.105
Rock	-2.453*	0.634*

4.B Assessment error simulation

Assessment errors for the BC flatfish complex stocks were simulated as an auto-correlated series of multi-variate, log-normal deviations

$$\delta_{j_{s,p},t} = \rho_{j_{s,p}} \delta_{j_{s,p},t-1} + \sqrt{1 - \rho^2_{j_{s,p}}} \cdot \iota_{j_{s,p},t} \tag{4.7}$$

$$\iota_{j_{s,p},t} \; N(\vec{\mu}, \Sigma) \tag{4.8}$$

$$\mu_j = \frac{\sigma^2_{j_{s,p}}}{2} \frac{1 - \rho_{j_{s,p}}}{\sqrt{(1 - \rho^2_{j_{s,p}}}} \tag{4.9}$$

where $j_{s,p} = 1, ..., 9$ is a 1 dimensional index for species s and stock p with the mapping $j_{s,p} = 3(p-1) + s$, $t = 2016, ..., 2095$ is the time step in years, ρ_j is the auto-correlation in the assessment errors δ, and $\vec{\mu}$ is the mean and Σ the covariance matrix (with diagonal entries $\sigma^2_{j_{s,p}}$) for the yearly deviations ι. The mean of the ι distribution is the log-normal bias-correction for auto-correlated deviations, applied so that the median assessment error is unbiased over time. For the remainder of the appendix, the s, p subscripts are excluded from the j index for simplicity. For $t = 2016$, $\delta_{j,t} = \iota_{j,t}$.

The covariance matrix Σ and auto-correlations ρ_j were estimated from the distribution of assessment errors obtained in a previous article [71]. There were 100 replicates of assessment errors from applying a hierarchical, multi-species surplus production stock assessment model to the BC flatfish complex fishery in closed loop simulations for 32 year projection period. Based on the results from the original article where bias in biomass and productivity compensated for each other [71], I adjusted assessment errors so that they were zero mean (i.e., unbiased). To use biased biomass esitmates would have required compensating for the bias by simulating an error for the optimal harvest rate.

Covariance Σ and auto-correlation ρ_j were derived from assessment errors for each individual replicate. The covariance matrix was computed via a diagonal decomposition

$$\Sigma = D \cdot C \cdot D,$$

where $D = [\sigma_j]$ is a diagonal matrix of assessment error standard deviations and C is the correlation matrix of assessment errors, both estimated directly from the distribution of assessment errors. The auto-correlation parameter ρ_j was also estimated directly from the distribution of errors using the acf() function in R [123], which returns a matrix of lag-1 correlations for the series of assessment errors, with auto-correlations along the main diagonal.

4.C Catchability and effort

The optimal allocation of effort in each area p, derived under equilibrium reference point calculations, is very close to the ratio ρ_p of inverse average catchabilty scalars within an area p to the sum of inverse average catchabilities, i.e.,

$$\rho_p = \frac{1/\bar{q}_p^{(F)}}{\sum_p 1/\bar{q}_p^{(F)}},$$
$$\bar{q}_p^{(F)} = \frac{1}{3}\sum_s q_{s,p}^{(F)}.$$

Further, the optimal equilbrium allocation is also very close to the $E_{MSY,MS}$ allocation (Table 4.5), but the pattern does not hold for stochastic simulations under time-varying demand, as explained in the results.

Chapter 5

Conclusion

In this thesis, I ask whether multi-species harvest strategies centred around hierarchical stock assessment models outperform status-quo single-species harvest strategies for data-limited fisheries. Multi-stock and multi-species stock assessment models and harvest strategies were applied to data from a simulated multi-species complex of flatfishes, modeled on the three species complex of Dover sole, English sole, and Rock sole in British Columbia, Canada.

The results of simulation tests of hierarchical stock assessment models suggests that their usage should be increased for stock assessment and management of data-limited, multi-species fisheries. In most tested cases, hierarchical stock assessments and multi-species harvest strategies perform better than assessments/strategies based on a single-species approach, or, in Chapter 3, those based on data-pooling methods. The superior performance of hierarchical models is especially noticeable as data quality and quantity is reduced. For example, in Chapter 2, estimates of key management parameters from hierarchical models are more robust to reductions in quality or quantity of simulated data than single-stock models, maintaining stronger feedback links between harvest decisions and stock status in low statistical power scenarios. However, while estimation performance of hierarchical models is more robust to decreasing statistical power, hierarchical models do not spontaneously create statistical power. Indeed, my results in Chapter 3 show that management performance is better under hierarchical models despite relatively high bias in biomass estimates. Bias in biomass estimates is compensated by bias in productivity estimates, which is in turn affected by the shrinkage priors on productivity parameters. The biomass/productivity biases therefore cancel out, and TACs are better scaled to maximise catch under hierarchical methods.

Results from Chapter 4 indicate that multi-species harvest strategies incorporating technical interactions could increase food security and economic welfare. First, choke effects are less prevalent when harvest rates are based on technical interactions, avoiding higher quota lease costs from under-utilised TACs. Second, under MEY strategies, fishery rents may increase by a factor of two, with only a small reduction in total yield. However, most fisheries

do not supply inferior goods (i.e., species with negative income elasticity of demand), so the relative economic benefits between MEY and MSY strategies will likely be smaller. In any case, the higher biomasses under MEY-based strategies reduce conservation risk, increasing longer-term food security and economic welfare at a small cost of lower average yield.

Expanding existing single-species harvest strategies to include technically interacting species may be an efficient way of increasing adoption of management procedures, while realising potential increases in food security and economic welfare. In 2019, Canadian commercial marine fishery landings totalled 742,634 t, with a landed value of $3.6 billion [1]. Roughly 28% of total Canadian landings are groundfish, accounting for approximately $351 million, or 10% of the total commercial marine fishery landed value. Groundfish are typically captured using non-selective gears that tend to generate technical interactions among target and byproduct species. As of 2016, formal harvest strategies have been adopted for five Canadian groundfish fisheries, namely Greenland Halibut (NAFO Area 3KLMNO) [98], Atlantic Halibut (3NOPs4VWX5) [25], Western Component Pollock [35], British Columbia Sablefish [24], Pacific Halibut [64], and Outside Yelloweye Rockfish [26]. These six fisheries had landings of 25,859 t in 2019, or about 12%, of Canadian groundfish landings, but account for approximately $197 million worth of landed value, or 55% of the total landed value of Canadian groundfish in that year. Despite such high value, all six fisheries are managed under a single-species paradigm in which TACs and realized catches are often mismatched in the presence of bycatch/byproduct species, leading to choke effects, higher fishing costs, and higher quota lease prices [109]. While the marginal economic benefit of adopting formal harvest strategies for the remaining groundfish species in Canada (i.e., the other 88% of landings) is necessarily lower than these high-value examples, it may be offset by improving the economic performance of these high value fisheries. Indeed, it would probably be more straight-forward to adapt existing harvest strategies for high-value species across Canada to include some (or all) technically interacting quota species, than to define new single-species harvest strategies for lower-value species, whose lower rents may not offset the costs of ignoring technical interactions.

Ultimately, making effective harvest decisions under uncertainty in multi-species fisheries requires tools that reflect the complexities and uncertainties in the management system. Outside of harvest strategy objectives (Chapter 4) and simulation performance (Chapters 2 and 3), the hierarchical multi-species modeling framework I developed here has sufficient complexity to answer management questions for most multi-species fisheries dominated by technical interactions. Further, the framework is readily expandable to incorporate extra complexity, such as movement, ecological interactions, and climate change influences on

[1]Statistics Canada. 2017. Table 32-10-0107-01 Aquaculture, production and value; https://doi.org/10.25318/3210010701-eng

recruitment and/or movement. Finally, leveraging technical interactions among species, and their shared evolutionary history, in the form of shrinkage priors adds value by potentially qualifying 'data-limited' fisheries for formal, simulation tested harvest strategies, as required under most fishery policies and eco-certification programs.

Bibliography

[1] Angrist, J. D. and Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives*, 15(4):69–85.

[2] Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

[3] Appeldoorn, R. S. (1996). Model and method in reef fishery assessment. In *Reef fisheries*, pages 219–248. Springer.

[4] Arreguín-Sánchez, F. (1996). Catchability: a key parameter for fish stock assessment. *Reviews in fish biology and fisheries*, 6(2):221–242.

[5] Asche, F., Bjørndal, T., and Gordon, D. V. (2007). Studies in the demand structure for fish and seafood products. In *Handbook of operations research in natural resources*, pages 295–314. Springer.

[6] Baudron, A. R. and Fernandes, P. G. (2015). Adverse consequences of stock recovery: European hake, a new "choke" species under a discard ban? *Fish and Fisheries*, 16(4):563–575.

[7] Benson, A. J., Cox, S. P., and Cleary, J. S. (2015). Evaluating the conservation risks of aggregate harvest management in a spatially-structured herring fishery. *Fisheries Research*, 167:101–113.

[8] Berger, A. M., Goethel, D. R., Lynch, P. D., Quinn, T. J., Mormede, S., McKenzie, J., and Dunn, A. (2017). Space oddity: the mission for spatial integration. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(11):1698–1716.

[9] Beverton, R. J. and Holt, S. J. (1957). *On the dynamics of exploited fish populations*, volume 11. Springer Science & Business Media.

[10] Branch, T. A. and Hilborn, R. (2008). Matching catches to quotas in a multispecies trawl fishery: targeting and avoidance behavior under individual transferable quotas. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(7):1435–1446.

[11] Burnham, K. P. and Anderson, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.

[12] Cadrin, S. X., Kerr, L. A., and Mariani, S. (2013). *Stock identification methods: applications in fishery science*. Academic Press.

[13] Carlin, B. P. and Louis, T. A. (1997). *Bayes and empirical Bayes methods for data analysis*, volume 7. Springer.

[14] Carruthers, T. R. (2018). A multispecies catch-ratio estimator of relative stock depletion. *Fisheries Research*, 197:25–33.

[15] Carruthers, T. R. and Hordyk, A. R. (2018). The data-limited methods toolkit (dlm tool): An r package for informing management of data-limited populations. *Methods in Ecology and Evolution*, 9(12):2388–2395.

[16] Carruthers, T. R., Punt, A. E., Walters, C. J., MacCall, A., McAllister, M. K., Dick, E. J., and Cope, J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research*, 153(0):48 – 68.

[17] Cheng, H.-t. and Capps Jr, O. (1988). Demand analysis of fresh and frozen finfish and shellfish in the united states. *American Journal of Agricultural Economics*, 70(3):533–542.

[18] Clark, C. W. and Munro, G. R. (2002). The problem of overcapacity. *Bulletin of Marine Science*, 70(2):473–483.

[19] Collie, J. S., Beck, M. W., Craig, B., Essington, T. E., Fluharty, D., Rice, J., Sanchirico, J. N., et al. (2013). Marine spatial planning in practice. *Estuarine, Coastal and Shelf Science*, 117:1–11.

[20] Collie, J. S. and Gislason, H. (2001). Biological reference points for fish stocks in a multispecies context. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(11):2167–2176.

[21] Collie, J. S. and Walters, C. J. (1991). Adaptive management of spatially replicated groundfish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 48(7):1273–1284.

[22] Conrad, J. M. and Clark, C. W. (1987). *Natural resource economics: notes and problems*. Cambridge University Press.

[23] Cope, J., Dick, E., MacCall, A., Monk, M., Soper, B., and Wetzel, C. (2013). Data-moderate stock assessments for brown, china, copper, sharpchin, stripetail, and yellowtail rockfishes and english and rex soles in 2013. *National Oceanic and Atmospheric Administration, National Marine Fisheries Service*.

[24] Cox, S., Kronlund, A., and Lacko, L. (2011). Management procedures for the multi-gear Sablefish (*Anoplopoma fimbria*) fishery in British Columbia, Canada. *Can. Sci. Advis. Secret. Res. Doc*, 62.

[25] Cox, S. P., Benson, A., and den Heyer, C. (2016). *Framework for the Assessment of Atlantic Halibut Stocks on Scotian Shelf and Southern Grand Banks*. Fisheries and Oceans Canada, Ecosystems and Oceans Science.

[26] Cox, S. P., Doherty, B., Benson, A. J., Johnson, S. D. N., and Haggarty, D. R. (2020). Evaluation of potential rebuilding strategies for Outside Yelloweye Rockfish in British Columbia. (2020/069):viii + 135 p.

[27] Cox, S. P. and Kronlund, A. R. (2008). Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia, Canada. *Fisheries Research*, 94(3):224–237.

[28] Cox, S. P., Kronlund, A. R., and Benson, A. J. (2013). The roles of biological reference points and operational control points in management procedures for the Sablefish (*Anoplopoma fimbria*) fishery in British Columbia, Canada. *Environmental Conservation*, 40(4):318–328.

[29] Danielsson, J. (2011). *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*, volume 588. John Wiley & Sons.

[30] de la Mare, W. K. (1998). Tidier fisheries management requires a new MOP (management-oriented paradigm). *Reviews in Fish Biology and Fisheries*, 8(3):349–356.

[31] DeMartini, E. E. (2019). Hazards of managing disparate species as a pooled complex: A general problem illustrated by two contrasting examples from hawaii. *Fish and Fisheries*, 20(6):1246–1259.

[32] 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.

[33] Dey, M. M., AlAM, M. F., and Paraguas, F. J. (2011). A multistage budgeting approach to the analysis of demand for fish: An application to inland areas of bangladesh. *Marine Resource Economics*, 26(1):35–58.

[34] DFO (2006). A Harvest Strategy Compliant with the Precautionary Approach. Technical Report 2006/023, DFO Can. Sci. Advis. Rep.

[35] DFO (2011). Western Component (4Xopqrs5) Pollock Management Strategy Evaluation. *Can. Sci. Advis. Sec. Sci. Advis. Rep.*, (054).

[36] Dichmont, C. M., Pascoe, S., Kompas, T., Punt, A. E., and Deng, R. (2010). On implementing maximum economic yield in commercial fisheries. *Proceedings of the National Academy of Sciences*, 107(1):16–21.

[37] Dick, E. and MacCall, A. D. (2011). Depletion-based stock reduction analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. *Fisheries Research*, 110(2):331–341.

[38] FAO (1996). Precautionary approach to capture fisheries and species introductions. Technical report, Food and Agriculture Organization of the United Nations, Rome.

[39] Fargo, J. (1998). Flatfish stock assessments for the west coast of Canada for 1997 and recommended yield options for 1998. Technical Report 97/36, Canadian Stock Assessment Secretariat Research Document.

[40] Fargo, J. (1999). Flatfish stock assessments for the west coast of Canada for 1999 and recommended yield options for 2000. *DFO Can. Stock. Assess. Sec. Res. Doc.*, (1999/199):51.

[41] Fargo, J., Foucher, R., Shields, S., and Ross, D. (1984). English sole tagging in hecate strait, r/v gb reed, june 6-24, 1983. *Can. Data Rep. Fish. Aquat. Sci*, 427:49.

[42] Fargo, J., Kronlund, A. R., Schnute, J. T., and Haigh, R. (2000). Stock assessment of Rock sole and English sole in Hecate Strait for 2000/2001. *DFO Can. Stock. Assess. Res. Doc.*, (2000/166):83.

[43] Fargo, J., Westrheim, S., Station, P. B., of Fisheries, C. D., and Oceans (1987). *Results, Through 1985, of the Rock Sole (Lepidopsetta Bilineata) Tagging Experiments in Hecate Strait (British Columbia) During April-May 1982 with Regard to Stock Delineation*. Canadian manuscript report of fisheries and aquatic sciences. Fisheries and Oceans, Canada.

[44] Fargo, J., Westrheim, S., and Stocker, M. (1985). *Results of the September 1979 Dover sole tagging experiment in northern Hecate Strait, through 1984*. Department of Fisheries and Oceans, Fisheries Research Branch, Pacific ....

[45] Fisheries and Oceans, Canada (2015). Pacific Region Integrated Fisheries Management Plan: Groundfish.

[46] Forrester, C. (1969). Results of english sole tagging in british columbia waters. *Bulletin of Pacific Marine Fisheries Commission*, 7:1–10.

[47] Fournier, D. and Archibald, C. P. (1982). A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1195–1207.

[48] 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(9):2105–2116.

[49] Francis, R. C. (2016). Growth in age-structured stock assessment models. *Fisheries Research*, 180:77–86.

[50] Gaichas, S., Gamble, R., Fogarty, M., Benoît, H., Essington, T., Fu, C., Koen-Alonso, M., and Link, J. (2012). Assembly rules for aggregate-species production models: simulations in support of management strategy evaluation. *Marine Ecology Progress Series*, 459:275–292.

[51] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 3. Taylor & Francis.

[52] Gillis, D. M., Peterman, R. M., and Tyler, A. V. (1993). Movement dynamics in a fishery: application of the ideal free distribution to spatial allocation of effort. *Canadian Journal of Fisheries and Aquatic Sciences*, 50(2):323–333.

[53] Gislason, H. (1999). Single and multispecies reference points for baltic fish stocks. *ICES Journal of Marine Science*, 56(5):571–583.

[54] Goethel, D. R. and Berger, A. M. (2017). Accounting for spatial complexities in the calculation of biological reference points: effects of misdiagnosing population structure for stock status indicators. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(11):1878–1894.

[55] Guillen, J., Macher, C., Merzéréaud, M., Bertignac, M., Fifas, S., and Guyader, O. (2013). Estimating msy and mey in multi-species and multi-fleet fisheries, consequences and limits: an application to the bay of biscay mixed fishery. *Marine Policy*, 40:64–74.

[56] Harling, W., Foucher, R., and Saunders, M. (1982). *Rock sole tagging in Hecate Strait, April 27-May 11, 1982*. Department of Fisheries and Oceans, Fisheries Research Branch, Pacific ....

[57] Hilborn, R. (1979). Comparison of fisheries control systems that utilize catch and effort data. *Journal of the Fisheries Board of Canada*, 36(12):1477–1489.

[58] Hilborn, R. (2018). Measuring fisheries performance using the "goldilocks plot". *ICES Journal of Marine Science*, 76(1):45–49.

[59] Hilborn, R. and Walters, C. J. (1987). A general model for simulation of stock and fleet dynamics in spatially heterogeneous fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 44(7):1366–1369.

[60] Hilborn, R. and Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty/Book and Disk*. Springer Science & Business Media.

[61] Holt, K. R., Starr, P. J., Haigh, R., and Krishka, B. (2016). Stock Assessment and Harvest Advice for Rock Sole (*Lepidopsetta spp.*) in British Columbia. *DFO Can. Sci. Advis. Sec. Res. Doc.*, (2016/009):ix + 256.

[62] Hopf, J. K., Jones, G. P., Williamson, D. H., and Connolly, S. R. (2016). Fishery consequences of marine reserves: Short-term pain for longer-term gain. *Ecological Applications*, 26(3):818–829.

[63] Hoshino, E., Pascoe, S., Hutton, T., Kompas, T., and Yamazaki, S. (2018). Estimating maximum economic yield in multispecies fisheries: a review. *Reviews in Fish Biology and Fisheries*, 28(2):261–276.

[64] IPHC (2019). International pacific halibut commission harvest strategy policy. Technical report, IPHC.

[65] Jacobs, O., Ballance, D., and Horwood, J. (1991). Fishery management as a problem in feedback control. *Automatica*, 27(4):627–639.

[66] Jaffry, S. A., Pascoe, S., and Robinson, C. (1999). Long run price flexibilities for high valued uk fish species: a cointegration systems approach. *Applied Economics*, 31(4):473–481.

[67] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 361–379.

[68] Jiao, Y., Cortés, E., Andrews, K., and Guo, F. (2011). Poor-data and data-poor species stock assessment using a bayesian hierarchical approach. *Ecological Applications*, 21(7):2691–2708.

[69] Jiao, Y., Hayes, C., and Cortés, E. (2009). Hierarchical Bayesian approach for population dynamics modelling of fish complexes without species-specific data. *ICES Journal of Marine Science: Journal du Conseil*, 66(2):367–377.

[70] Johnson, S. D. N. and Cox, S. P. (2018). Evaluating the role of data quality when sharing information in hierarchical multi-stock assessment models, with an application to dover sole. *Can. J. Fish. Aquat. Sci.*

[71] Johnson, S. D. N. and Cox, S. P. (2021). Hierarchical surplus production stock assessment models improve management performance in multi-species, spatially replicated fisheries. *Fisheries Research*.

[72] Kempf, A., Mumford, J., Levontin, P., Leach, A., Hoff, A., Hamon, K. G., Bartelings, H., Vinther, M., Staebler, M., Poos, J. J., et al. (2016). The MSY concept in a multi-objective fisheries environment–Lessons from the North Sea. *Marine Policy*, 69:146–158.

[73] Ketchen, K. (1982). *Stock delineation and growth of rock sole (Lepidopsetta bilineata) as indicated by tagging in British Columbia waters, 1944-66*. Department of Fisheries and Oceans, Fisheries Research Branch, Pacific ....

[74] Kleijnen, J. P. (2008). *Design and analysis of simulation experiments*, volume 20. Springer.

[75] Krigbaum, M. J. and Anderson, C. M. (2021). Increasing value through gear flexibility: A case study of us west coast sablefish. *Canadian Journal of Fisheries and Aquatic Sciences*, (ja).

[76] Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. (2015). Tmb: Automatic differentiation and laplace approximation. *arXiv preprint arXiv:1509.00660*.

[77] MacCall, A. D. (2009). Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. *ICES Journal of Marine Science: Journal du Conseil*, 66(10):2267–2271.

[78] 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.

[79] Magnusson, A. and Hilborn, R. (2007). What makes fisheries data informative? *Fish and Fisheries*, 8(4):337–358.

[80] Malick, M. J., Cox, S. P., Peterman, R. M., Wainwright, T. C., Peterson, W. T., and Krkošek, M. (2015). Accounting for multiple pathways in the connections among climate variability, ocean processes, and coho salmon recruitment in the northern california current. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(10):1552–1564.

[81] Mangel, M. (2000). Irreducible uncertainties, sustainable fisheries and marine reserves. *Evolutionary Ecology Research*, 2(4):547–557.

[82] Mangel, M., Brodziak, J., and DiNardo, G. (2010). Reproductive ecology and scientific inference of steepness: a fundamental metric of population dynamics and strategic fisheries management. *Fish and Fisheries*, 11(1):89–104.

[83] Marchal, P., De Oliveira, J. A., Lorance, P., Baulier, L., and Pawlowski, L. (2013). What is the added value of including fleet dynamics processes in fisheries models? *Canadian Journal of Fisheries and Aquatic Sciences*, 70(7):992–1010.

[84] Marchal, P., Francis, C., Lallemand, P., Lehuta, S., Mahévas, S., Stokes, K., and Vermard, Y. (2009). Catch-quota balancing in mixed-fisheries: a bio-economic modelling approach applied to the new zealand hoki (macruronus novaezelandiae) fishery. *Aquatic living resources*, 22(4):483–498.

[85] Marshall, K. N., Koehn, L. E., Levin, P. S., Essington, T. E., and Jensen, O. P. (2019). Inclusion of ecosystem information in us fish stock assessments suggests progress toward ecosystem-based fisheries management. *ICES Journal of Marine Science*, 76(1):1–9.

[86] Martell, S. J., Walters, C. J., and Hilborn, R. (2008). Retrospective analysis of harvest management performance for bristol bay and fraser river sockeye salmon (oncorhynchus nerka). *Canadian Journal of Fisheries and Aquatic Sciences*, 65(3):409–424.

[87] Maunder, M. N. and Piner, K. R. (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72(1):7–18.

[88] May, R. M., Beddington, J. R., Clark, C. W., Holt, S. J., and Laws, R. M. (1979). Management of multispecies fisheries. *Science*, 205(4403):267–277.

[89] McCullagh, P. (1984). Generalized linear models. *European Journal of Operational Research*, 16(3):285–292.

[90] Melnychuk, M. C., Kurota, H., Mace, P. M., Pons, M., Minto, C., Osio, G. C., Jensen, O. P., de Moor, C. L., Parma, A. M., Little, L. R., et al. (2021). Identifying management actions that promote sustainable fisheries. *Nature Sustainability*, pages 1–10.

[91] Mercer, M. (1982). Multispecies approaches to fisheries management advice. *Can. Spec. Publ. Fish. Aquat. Sci*, 59(169):399–409.

[92] Morgan/Reuters, J. P. (1996). Riskmetrics - technical document, 4th edition. Technical report, Morgan Guaranty Trust Company and Reuters Ltd, New York.

[93] Mortensen, L. O., Ulrich, C., Hansen, J., and Hald, R. (2018). Identifying choke species challenges for an individual demersal trawler in the north sea, lessons from conversations and data analysis. *Marine Policy*, 87:1–11.

[94] Mueter, F. J., Ware, D. M., and Peterman, R. M. (2002). Spatial correlation patterns in coastal environmental variables and survival rates of salmon in the north-east pacific ocean. *Fisheries Oceanography*, 11(4):205–218.

[95] Murawski, S. A. (1984). Mixed-species yield-per-recruitment analyses accounting for technological interactions. *Canadian Journal of Fisheries and Aquatic Sciences*, 41(6):897–916.

[96] Murawski, S. A., Lange, A., and Idoine, J. (1991). An analysis of technological interactions among gulf of maine mixed-species fisheries. In *ICES Marine Science Symposia*, volume 193, pages 237–252.

[97] Nadon, M. O. and Ault, J. S. (2016). A stepwise stochastic simulation approach to estimate life history parameters for data-poor fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(12):1874–1884.

[98] NAFO (2010). Report of the fc working group on greenland halibut management strategy evaluation (wgmse).

[99] Nelson, S. (2011). Pacific commercial fishing fleet: Financial profiles for 2009. *Pacific Commercial Fishing Fleets Financial Profiles Series*, (2011-4):160.

[100] Ono, K., Haynie, A. C., Hollowed, A. B., Ianelli, J. N., McGilliard, C. R., and Punt, A. E. (2017). Management strategy analysis for multispecies fisheries, including technical interactions and human behavior in modelling management decisions and fishing. *Canadian Journal of Fisheries and Aquatic Sciences*, (999):1–18.

[101] Oostdijk, M., Byrne, C., Stefánsson, G., Santos, M. J., and Woods, P. J. (2020). Catch–quota matching allowances balance economic and ecological targets in a fishery managed by individual transferable quota. *Proceedings of the National Academy of Sciences*, 117(40):24771–24777.

[102] Pascoe, S., Hutton, T., and Hoshino, E. (2018). Offsetting externalities in estimating mey in multispecies fisheries. *Ecological Economics*, 146:304–311.

[103] Pascoe, S., Hutton, T., Hoshino, E., Sporcic, M., Yamasaki, S., and Kompas, T. (2020). Effectiveness of harvest strategies in achieving multiple management objectives in a multispecies fishery. *Australian Journal of Agricultural and Resource Economics*.

[104] Pascoe, S., Hutton, T., Thébaud, O., Deng, R., Klaer, N., and Vieira, S. (2015). Setting economic target reference points for multiple species in mixed fisheries. *FRDC Final Report*.

[105] Pella, J. J. and Tomlinson, P. K. (1969). A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin*, 13(3):416–497.

[106] Peterman, R. M. (1990). Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(1):2–15.

[107] Peterman, R. M., Pyper, B. J., Lapointe, M. F., Adkison, M. D., and Walters, C. J. (1998). Patterns of covariation in survival rates of british columbian and alaskan sockeye salmon (oncorhynchus nerka) stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(11):2503–2517.

[108] Pikitch, E. K. (1987). Use of a mixed-species yield-per-recruit model to explore the consequences of various management policies for the oregon flatfish fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, 44(S2):s349–s359.

[109] Pinkerton, E. and Edwards, D. N. (2009). The elephant in the room: the hidden costs of leasing individual transferable fishing quotas. *Marine Policy*, 33(4):707–713.

[110] Punt, A. and Butterworth, D. (1995). The effects of future consumption by the Cape fur seal on catches and catch rates of the Cape hakes. 4. Modelling the biological interaction between Cape fur seals Arctocephalus pusillus pusillus and the Cape hakes Merluccius capensis and M. paradoxus. *South African Journal of Marine Science*, 16(1):255–285.

[111] Punt, A. E. (2003). Extending production models to include process error in the population dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, 60(10):1217–1228.

[112] 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 Modeling*, 19(4):441–464.

[113] Punt, A. E., Butterworth, D. S., Moor, C. L., De Oliveira, J. A. A., and Haddon, M. (2016). Management strategy evaluation: best practices. *Fish and Fisheries*.

[114] Punt, A. E., Dalton, M. G., and Foy, R. J. (2020). Multispecies yield and profit when exploitation rates vary spatially including the impact on mortality of ocean acidification on north pacific crab stocks. *Fisheries Research*, 225:105481.

[115] Punt, A. E., Deng, R., Pascoe, S., Dichmont, C. M., Zhou, S., Plagányi, É. E., Hutton, T., Venables, W. N., Kenyon, R., and Van Der Velde, T. (2011a). Calculating optimal effort and catch trajectories for multiple species modelled using a mix of size-structured, delay-difference and biomass dynamics models. *Fisheries Research*, 109(1):201–211.

[116] Punt, A. E. and Donovan, G. P. (2007). Developing management procedures that are robust to uncertainty: lessons from the international whaling commission. *ICES Journal of Marine Science*, 64(4):603–612.

[117] Punt, A. E., Okamoto, D. K., MacCall, A. D., Shelton, A. O., Armitage, D. R., Cleary, J. S., Davies, I. P., Dressel, S. C., Francis, T. B., Levin, P. S., et al. (2018). When are estimates of spawning stock biomass for small pelagic fishes improved by taking spatial structure into account? *Fisheries Research*, 206:65–78.

[118] Punt, A. E., Smith, A. D., and Cui, G. (2002). Evaluation of management tools for australia's south east fishery. 1. modelling the south east fishery taking account of technical interactions. *Marine and Freshwater Research*, 53(3):615–629.

[119] Punt, A. E., Smith, D. C., and Koopman, M. T. (2005). *Using information for data-rich species to inform assessments of data-poor species through Bayesian stock assessment methods*. Primary Industries Research Victoria.

[120] Punt, A. E., Smith, D. C., and Smith, A. D. (2011b). Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. *ICES Journal of Marine Science: Journal du Conseil*, 68(5):972–981.

[121] Quinn, T. J. (2003). Ruminations on the development and future of population dynamics models in fisheries. *Natural Resource Modeling*, 16(4):341–392.

[122] Quinn, T. J. and Deriso, R. B. (1999). *Quantitative fish dynamics*. oxford university Press.

[123] R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[124] Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, volume 1. Sage.

[125] Russo, T., Parisi, A., Garofalo, G., Gristina, M., Cataudella, S., and Fiorentino, F. (2014). Smart: a spatially explicit bio-economic model for assessing and managing demersal fisheries, with an application to italian trawlers in the strait of sicily. *PloS one*, 9(1):e86222.

[126] Schaefer, M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-American Tropical Tuna Commission Bulletin*, 1(2):23–56.

[127] Schnute, J. (1981). A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(9):1128–1140.

[128] Schnute, J. (1985). A general theory for analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences*, 42(3):414–429.

[129] Schnute, J. T. and Richards, L. J. (1995). The influence of error on population estimates from catch-age models. *Canadian Journal of Fisheries and Aquatic Sciences*, 52(10):2063–2077.

[130] Seber, G. (1997). *Estimation of Animal Abundance*. Oxford University Press, 2 edition.

[131] Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., and Kjesbu, O. S. (2016). Ecosystem processes are rarely included in tactical fisheries management. *Fish and Fisheries*, 17(1):165–175.

[132] Smith, S. J., Hunt, J. J., and Rivard, D. (1993). *Risk Evaluation and Biological Reference Points for Fisheries Management Canadian Special Publication of Fisheries and Aquatic Sciences No. 120*. NRC Research Press.

[133] Srinivasan, B., Biegler, L. T., and Bonvin, D. (2008). Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function. *Computers & Chemical Engineering*, 32(3):572–579.

[134] Starfield, A. M. (1997). A pragmatic approach to modeling for wildlife management. *The Journal of wildlife management*, pages 261–270.

[135] Starr, P. J. (2009). English Sole (*Parophrys vetulus*) in British Columbia, Canada: Stock Assessment for 2006/07 and Advice to Managers for 2007/08. *DFO Can. Sci. Advis. Sec. Res. Doc.*, (2009/069):v + 149.

[136] Stewart, I. J., Hicks, A. C., Taylor, I. G., Thorson, J. T., Wetzel, C., and Kupschus, S. (2013). A comparison of stock assessment uncertainty estimates using maximum likelihood and bayesian methods implemented with the same model framework. *Fisheries Research*, 142:37–46.

[137] Stocker, M., Bennett, D., and Fargo, J. (1980). *Dover Sole Tagging, Eastern Dixon Entrance, September 18-29, 1979*. Government of Canada, Fisheries and Oceans.

[138] Thorson, J. T., Cope, J. M., Kleisner, K. M., Samhouri, J. F., Shelton, A. O., and Ward, E. J. (2015). Giants' shoulders 15 years later: lessons, challenges and guidelines in fisheries meta-analysis. *Fish and Fisheries*, 16(2):342–361.

[139] Thorson, J. T., Cope, J. M., and Patrick, W. S. (2014). Assessing the quality of life history information in publicly available databases. *Ecological Applications*, 24(1):217–226.

[140] Thorson, J. T., Munch, S. B., Cope, J. M., and Gao, J. (2017). Predicting life history parameters for all fishes worldwide. *Ecological Applications*, 27(8):2262–2276.

[141] Tidd, A. N., Hutton, T., Kell, L. T., and Blanchard, J. L. (2012). Dynamic prediction of effort reallocation in mixed fisheries. *Fisheries Research*, 125:243–253.

[142] Wakefield, W. W. (1984). Feeding relationships within assemblages of nearshore and mid-continental shelf benthic fishes off oregon. Master's thesis, Oregon State University.

[143] Walters, C. (1986). Adaptive management of renewable resources.

[144] Walters, C. (1998). Evaluation of quota management policies for developing fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(12):2691–2705.

[145] Walters, C. J. and Bonfil, R. (1999). Multispecies spatial assessment models for the British Columbia groundfish trawl fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(4):601–628.

[146] Westrheim, S. J., Barss, W. H., Pikitch, E. K., and Quirollo, L. F. (1992). Stock delineation of dover sole in the california–british columbia region, based on tagging studies conducted during 1948–1979. *North American Journal of Fisheries Management*, 12(1):172–181.

[147] Wetzel, C. R. and Hamel, O. S. (2019). Catch update for dover sole (microstomus pacificus) along the u.s. west coast in 2019. Technical report, Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

[148] Winker, H., Carvalho, F., and Kapur, M. (2018). JABBA: just another Bayesian biomass assessment. *Fisheries Research*, 204:275–288.