OCEANS & ATMOSPHERE
www.csiro.au

Incorporating the effects of marine spatial closures in risk assessments and fisheries stock assessments.

Geoff Tuck, Malcolm Haddon, Rich Little, André Punt, Neil Klaer, Ross Daley, Jemery Day, Tony Smith, Miriana Sporcic, Sally Wayte, Shijie Zhou

Project 2011/032

May, 2018

Fisheries Research and Development Corporation

FRDC
FISHERIES RESEARCH & DEVELOPMENT CORPORATION

ISBN 978-1-4863-0933-7

Incorporating the effects of marine spatial closures in risk assessments and fisheries stock assessments. FRDC Project 2011/032. Fisheries Research Development Corporation and CSIRO Oceans and Atmosphere, Hobart. By Geoffrey N Tuck, Malcolm Haddon, Rich Little, Andre Punt, Neil Klaer, Ross Daley, Jemery Day, Tony Smith, Miriana Sporcic, Sally Wayte, Shijie Zhou 2017. May, 2018 – ONLINE

FRDC Final Report 2011/032

May 2018

This publication (and any information sourced from it) should be attributed to:

Tuck GN, Haddon M, Little LR, Punt AE, Klaer N, Daley RK, Day JR, Smith ADM, Sporcic M, Wayte S, Zhou S (2018) Incorporating the effects of marine spatial closures in risk assessments and fisheries stock assessments. CSIRO, Australia. FRDC Final Report 2011/032. Hobart

Creative Commons licence

Disclaimer

The Fisheries Research and Development Corporation plans, invests in and manages fisheries research and development throughout Australia. It is a statutory authority within the portfolio of the federal Minister for Agriculture, Fisheries and Forestry, jointly funded by the Australian Government and the fishing industry.

	Researcher Contact Details		FRDC Contact Details
Name	Geoff Tuck		
Address	Oceans and Atmosphere, CSIRO, Castray Esplanade, Hobart, GPO Box 1538, Tasmania, 7001, Australia	Address	25 Geils Court, Deakin ACT, 2600
Phone	03 62325106	Phone	02 62850400
		Fax	02 62850499
E-mail	Geoff.Tuck@csiro.au	E-mail	frdc@frdc.com.au
		Web	www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

Contents

Acknowledgments

The authors would like to thank the FRDC for their financial support of this project. The CSIRO SESSF stock assessment team are thanked for their comments and contributions to much of the work presented. Eva Plaganyi, Dale Kolody (CSIRO), Daniel Goethel (NOAA, SEFSC), and the anonymous reviewers are thanked for their comments on earlier versions of Appendix 5 and 6. This work was funded by FRDC and the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative agreement No. NA10OAR4320148, Contribution No. 2016-01-16. The authors also greatly thank Toni Cracknell (Ultimate Proof Processing) for her efforts in formatting the report.

Executive summary

Background

Within the framework of the Commonwealth Harvest Strategy Policy, the Australian Fisheries Management Authority (AFMA) has adopted clear decision rules to set catch limits for commercially targeted species in Commonwealth fisheries. In several fisheries, AFMA has adopted a tiered harvest strategy framework that specifies both assessment methods and decision rules appropriate to the extent and quality of information available for each target species. For by-product and bycatch species, AFMA has adopted a risk-based and hierarchical Ecological Risk Assessment and Ecological Risk Management (ERA/ERM) framework that identifies higher risk species and prioritizes management responses. With the increase in number and extent of marine spatial closures, whether for conservation or fishery management purposes, there is a need to evaluate the impact of closures on existing assessment methods and rules and, if necessary, modify these methods and rules, or provide new methods that appropriately account for the existence of closures. This project will evaluate and develop assessment methods and a complementary set of meta-rules that can be integrated into the current assessment and management frameworks.

Objectives

1. Develop criteria and procedures for determining whether current methods for incorporating the effects of marine spatial closures in risk assessments and stock assessments are appropriate for all species.
2. Develop a method for incorporating the effects of marine spatial closures in risk assessments and stock assessments for those species where the current approach is not considered effective.
3. Develop a set of rules for determining TACs or catch limits based on the quantity and quality of data available on the species biology, the characteristics of the closure, and the extent of monitoring inside and outside of the closure.

Keywords

Management Strategy Evaluation, Spatial Closure, Harvest Strategy Policy, Coral Trout, Pink Ling, Ecological Risk Assessment.

Introduction

Background

Within the framework of the Commonwealth Harvest Strategy Policy, the Australian Fisheries Management Authority (AFMA) has adopted clear decision rules to set catch limits for commercially targeted species in Commonwealth fisheries. In several fisheries, AFMA has adopted a tiered harvest strategy framework that specifies both assessment methods and decision rules appropriate to the extent and quality of information available for each target species. For by-product and bycatch species, AFMA has adopted a risk-based and hierarchical Ecological Risk Assessment and Ecological Risk Management (ERA/ERM) framework that identifies higher risk species and prioritizes management responses. With the increase in number and extent of marine spatial closures, whether for conservation or fishery management purposes, there is a need to evaluate the impact of closures on existing assessment methods and rules and, if necessary, modify these methods and rules, or provide new methods that appropriately account for the existence of closures. This project evaluated current methods and, where appropriate and possible, suggested alternative assessment methods and complementary sets of meta-rules that can be integrated into the current assessment and management frameworks.

The methods developed under this project were evaluated in the Southern and Eastern Scalefish and Shark Fishery (SESSF) and Queensland Coral Trout Fisheries. The SESSF is a Commonwealth-managed, multi-species and multi-gear fishery that catches over 80 species of commercial value. Catches are taken from both inshore and offshore waters, as well as offshore seamounts. A formal harvest strategy framework (HSF) was adopted in the SESSF for the first time in 2005. This framework includes an agreed process for fishery monitoring, stock assessment, and decision rules for translating stock assessment outputs into clear advice on Recommended Biological Catch (RBC). Methods used to assess SESSF stocks range from simple trend analysis (Tier 4), through use of catch curves (Tier 3) to full quantitative assessments using integrated analysis (Tier 1). There is considerable uncertainty regarding how best to account for marine closures in this fishery, in particular for stocks with considerable proportions of their habitat already closed, such as Deepwater Shark, Silver Trevally (*Pseudocaranx georgianus*) and Pink Ling (*Genypterus blacodes)* fisheries, and when determining if discount factors should be applied between Tier levels, given large proportions of the stock are believed to be secure in a closure.

Coral trout (*Plectropomus leopardus*) is the key target species in the Queensland Coral Reef Fin Fish Fishery (CRFFF). The fishery area spans 14 degrees of latitude, and lies within a world heritage area where approximately one quarter of the available habitat is closed to fishing. The fishery is spatially complex and there is substantial variation in the distribution of coral trout. This makes it challenging to use standard approaches to determining sustainable levels of harvest. The current TAC for coral trout is based on historical commercial catch taken by the fishery, and has changed little since the ITQ system was implemented in 2004. The reliance on historical data to determine the TAC has led to questions regarding the potential profitability and sustainability of the fishery. A stock assessment has been developed for the fishery, but the effects of spatial closures on the biomass available to the fishery, and potential harvest control rules, is not known. Developing ways to deal with these closures in assessments and decision rules is critical for effective management of the resource. From the perspective of target species assessment and management, the key questions that the project addressed relate to closure impacts on (a) data collection (observer, logbook, surveys), (b) stock assessments, given that input data (eg catch, CPUE, length) may only reflect a fraction of the stock, and (c) control rules, in particular if they should be modified (eg through changes in target reference points or discount factors).

For non-target and bycatch species, the project assessed the extent to which the existing assessment methods, Productivity and Susceptibility Analysis (PSA) and Sustainability Assessment for Fishing Effort (SAFE), adequately account for the presence of spatial closures in assessing risk levels to stocks. The project developed rules of thumb for identifying appropriate management responses, including catch limits, depending on assessed levels of risk. This will provide a complement to the assessment / decision rule framework for target species and potentially extend the range of species that can be dealt with formally under the current harvest strategy policy.

Objectives

The objectives of the project were to:

1. Develop criteria and procedures for determining whether current methods for incorporating the effects of marine spatial closures in risk assessments and stock assessments are appropriate for all species.
2. Develop a method for incorporating the effects of marine spatial closures in risk assessments and stock assessments for those species where the current approach is not considered effective.
3. Develop a set of rules for determining TACs or catch limits based on the quantity and quality of data available on the species biology, the characteristics of the closure, and the extent of monitoring inside and outside of the closure.

Methods

The methods and results described here are a summary of the broader work conducted. Greater methodological detail and results can be found in the subsequent Appendices 1 - 8.

Ecological Risk Assessment

Productivity and Susceptibility Analysis (PSA)

The Productivity Susceptibility Analysis (PSA) method is in wide use in fisheries around the world and has been applied to various species and habitats, not just fish (Hobday et al., 2011; Tuck et al, 2011). The PSA analysis characterizes risk as a function of the productivity of a population and its susceptibility to capture. In Appendix 1, the PSA for the otter trawl fishery within the Southern and Eastern Scalefish and Shark Fishery (SESSF) was, firstly, revised to assess the effectiveness of area closures introduced to mitigate the effects of fishing between 2003 and 2013 and, secondly, explored for refinements to the PSA method. The focus was on waters deeper than 700m where substantial areas have been closed to protect Orange Roughy.

The PSA method scores the risk posed by fishing for a given species along two axes using productivity and susceptibility attributes. Productivity attributes are intrinsic to species and biologically fixed, for example, life span, number of eggs etc. Low productivity species, including whales and birds as well as long-lived fish and sharks, tend to have higher overall risk; high productivity species such as prawns and many short-lived teleosts tend to have lower overall risk scores. Because productivity attributes are fixed, they don't change with changes in management.

Susceptibility (exposure) varies according to four attributes: 1. Availability (quantifying the extent of spatial overlap between fishing effort and species range); 2. Encounterability (position in the water column, relating to the likelihood that fishing gear will encounter a species); 3. Selectivity (the likelihood of capture given encounter); and 4. Post Capture Mortality (related to the probability of survival given capture and particularly important for species that are discarded). Any of these four attributes can be influenced by management interventions and regulations. Fishery closures mainly affect Availability. Availability risk for a given species will be reduced by spatial management if the restrictions applied to a given "closure" encompass part of a species range.

In current PSA analysis, availability is measured as the percentage of the species range that overlaps with fishing effort. These availability values are then given categorical risk scores, according to scoring thresholds: <10% overlap = low risk, 10–20% overlap = medium risk; >20% = high risk. We calculated overlap values for seven SESSF species for 2013. The refined PSA uses a continuous value between 0–3 for availability, rather than a categorical score. This provides a higher resolution for smaller changes in spatial closures.

Sustainability Assessment for Fishing Effect (SAFE)

AFMA has adopted the quantitative method Sustainability Assessment for Fishing Effect (SAFE) as a preferred method within the Ecological Risk Assessment for the Effect of Fishing (ERAEF) framework. However, the performance of SAFE in the context of marine closures has not been examined. Appendix 2 investigates the performance of the method with regard to spatial closures. The SAFE method requires a stock indicator, fishing mortality rate in this case, and reference points. Because of a lack of basic data for most non-target species, including time series of catch, abundance index, age composition, etc., SAFE derives estimates of the fishing mortality rate using two alternative methods: Base SAFE and enhanced SAFE.

The Base SAFE method is based on the spatial overlap between species distribution and fishing effort distribution over the jurisdictional area of the fishery (Zhou et al., 2007). The estimate of overlap is fine-tuned by considering habitat and behaviour-dependent encounterability and fishing gear and size-dependent gear selectivity. Therefore, a spatial configuration of fishing effort that includes closures and protected areas is essential in deriving fishing impact and overlap with species range as a proxy for fishing mortality, the indicator of interest. Marine spatial closures within the jurisdictional area reduce the available area to fishing. Closures will result in a decreased fishing impact on stocks that are fully or partially protected by closures if the total fishing effort does not increase, assuming closures do not affect the extent of fish distribution and fish do not move between closed and open areas. For non-target species assessed using the SAFE method, the effect of closures will be reflected in a reduced fishing mortality rate. There are no changes to the basic equations for this method when marine closures are included.

Enhanced SAFE was initially developed to assess hundreds of fish bycatch species in Australia's Northern Prawn Trawl Fishery. Unlike the Base SAFE method, this method allows a heterogeneous density across a fishes' distribution range or between fished and unfished areas. It also uses species-specific gear efficiency. When marine closures are implemented, if fish density in the closures can be assumed to be the same as fish density in the open but unfished area, then no changes are needed to the basic SAFE equations. However, if this assumption is violated, then account needs to be made of the different densities. The enhanced SAFE thus requires more information than the Base SAFE. It is possible to obtain measures of heterogeneous density in different locations from historical surveys or from observer data to predict bycatch species density at locations where no data exist.

This study describes how spatial information is used in the SAFE method and applies it to seven species in the Southern and Eastern Scalefish and Shark Fishery (SESSF) to illustrate how marine closures might mitigate the risk of fishing posed to the sustainability of by-product and bycatch species. The base SAFE method has previously been applied to the SESSF, so this method is again examined in this study. Furthermore, enhanced SAFE requires more information, including fish density in closed areas that is unavailable for bycatch species in the SESSF. However, if it is assumed that fish density is the same between closed areas and open but unfished areas, then closures will have no effect on the result.

The Coral Trout Fishery

Simulation is used in Appendix 3 to evaluate the ability of a two-region age-structured assessment model to provide accurate and precise estimates of stock status, i.e. the ratio of female spawning biomass to unfished female spawning biomass, for coral trout on the Great Barrier Reef (GBR), Australia. The Effects of Line Fishing Simulator (ELFSim) operating model is used to generate the simulated data used by the assessment model. ELFSim is a spatially complex age- and sex-structured population dynamics model that captures the protogynous nature of coral trout. ELFSim models the population dynamics, harvest, and management of coral trout on more than 3,000 individual reefs on the Great Barrier Reef, connected through larval dispersal. It operates stochastically using monthly time steps. Starting from an assumed year in which harvest began (1965), individual reef sub-populations are conditioned on historical commercial, charter, and recreational catches; the selectivity of the gear in catching fish of different sizes; the biological characteristics of coral trout; and the physical characteristics of the individual reefs. Reef sub-populations are then projected from 2012 to 2035 by simulating harvest using an agent-based model of fishing vessel dynamics, subject to management constraints including spatial closures. A simulated line fishing survey vessel was also implemented. The survey vessel collected data from a sample of 30 reefs in September of each projection year. Two scenarios for the survey were simulated based on obtaining samples from: (i) reefs open to fishing; and (ii) reefs open and closed to fishing.

A two-region stock assessment model was used to analyze the simulated data, and applied annually. Two abundance indices were recorded in each year during the projection period. This first was a fleetwide standardized catch-rate index, computed from the aggregate catch and effort data of the commercial fleet from applying a General Linear Model with factors for month, vessel, one degree resolution spatial grid, and year. The second index of abundance was based on the catch rate data collected during the structured line survey. The assessment model aggregated biomass across reefs into two sub-populations (equivalent to regions): an unfished (protected) and a fished portion. The two sub-populations had different time-series of exploitation rate, but were linked through a common stock-recruitment relationship. Assessments were conducted annually on each of 10 projected replicates from 2012 to 2035. Analyses considered the effect on the stock assessment estimates of using the simulated structured line survey to collect data from areas closed or open to fishing. Simulations also examined different levels of larval advection and self-seeding (*ss*) among reefs. The amount of closures in the operating model was varied at two levels corresponding to 21% and 41% of coral trout habitat, measured as reef perimeter, in the marine park. The results of the simulations were summarized by the errors associated with estimation of the ratio of the female spawning biomass to the unfished level in each year of each simulation replicate.

The Southern and Eastern Scalefish and Shark Fishery

Evaluation of current assessment methods

Management strategy evaluation (MSE) is used to evaluate the influence of the inclusion of a no-take marine closure on assessments and control rules currently used to manage harvested stocks from south eastern Australia (Appendix 4). MSE is the examination of alternative options for management via Monte Carlo simulation. A management strategy considers: (1) data collection, to provide relevant information on the state of the system; (2) assessment of stock status using the data collected; and (3) a harvest control rule to translate stock status from an assessment into a management action (a TAC for example). Multiple simulations are conducted for the chosen management strategy. Management strategies are sought that are robust to system uncertainty. This requires stochastic simulations for a chosen scenario and the subsequent evaluation of management performance across multiple scenarios. The simulation process can highlight weaknesses or biases in current or possible management strategies, as well as identify the weaknesses of indicators of the system state.

Marine closures are commonly used as a fishery and biodiversity conservation management tool. The consequent spatial structuring and reduction of data to inform assessments of status is known to bias standard assessments. A spatially explicit operating model is developed to allow the simulation of data from a single harvested area and then following the introduction of a no-take closure. The closure can vary according to size and the mixing rate between open and closed areas. Management strategies varying from data-rich (Tier 1) to data-poor (Tiers 3 and 4) are evaluated in terms of their ability to attain biomass targets and have reasonable biomass risk profiles. Alternative biomass targets are considered, including whether the area open to fishing, or the population as a whole, has a fixed biomass target. The influence on yield and catch rates is also examined according to biological and closure characteristics (eg mixing rates and percentage of the stock closed) and the management strategy chosen (eg the Tier level assessment method, discount factors, use of all data or data only from the open area).

Evaluating alternative assessment methods

Spatial structure in biological characteristics and exploitation rates impact the performance of stock assessment methods used to estimate the status of fish stocks relative to target and limit reference points. Spatially-structured stock assessment methods can reduce the bias and imprecision in the estimates of management-related model outputs. However, their performance has only recently been evaluated formally, in particular when some of the area fished is closed. In order to evaluate the effects of closed areas and spatial variation in growth and exploitation rate when estimating spawning biomass, in Appendix 5 a spatially-explicit operating model was developed to simulate spatial data and five configurations of the stock assessment package Stock Synthesis (three of which were spatially structured) were applied. The evaluation framework was based on the SESSF for Pink Ling off southern Australia. Allowance was made for three spatial zones (nominally zones 10, 20, and 30 of the SESSF). The fish populations in these three zones were assumed to be connected through the distribution of age-0 animals, with animals of age-1 and older being sedentary. Two fleets (trawl and non-trawl) were assumed to operate in each zone, growth could differ among zones, and recruitment was assumed to be stochastic, with spatial variation in the proportion of the total recruitment that settles to each zone, as well as temporal variation in total recruitment. The study examines the consequences, in terms of the ability to estimate time-trajectories of spawning biomass, of closed areas that encompass a large proportion of stock biomass (>15%) as well as the benefits of the availability of surveys in the closed (and open) areas.

The five assessment configurations estimate unfished recruitment, natural mortality, growth by sex, length-specific selectivity parameters, catchability for the CPUE indices, and recruitment for simulated years 1963-2013. The five assessment methods vary in their assumptions regarding spatial structure, namely the way data are aggregated spatially and whether recruitment varies spatially. Assessments commonly used in the SESSF were tested (spatially aggregated and fleets-as-areas) as well as more spatially explicit model structures.

Evaluating harvest control rules

Several studies, including those of Appendix 5, have shown that ignoring spatial structure leads to bias in estimates of management-related quantities from assessment, and this bias is exacerbated by closures. They also show that spatially-structured stock assessments reduce bias in estimates of biomass, at the cost of lower precision but may lead to biased estimates of movement parameters. Ideally, spatially-structured stock assessments should be based on tagging data to inform estimates of movement, but such data are available for very few stocks (exceptions in Australia include Gummy Shark (*Mustelus antarcticus)*, and Toothfish (*Dissostichus eleginoides*)). Consequently, it is common to either ignore spatial structure and instead apply the fleets-as-areas approach, which accounts approximately for spatial structure or use spatial models with no tagging data.

Although the estimates from assessments that ignore spatial structure, or use approximate methods for addressing spatial structure such as the fleets-as-areas approach, are biased (and imprecise), the use of such estimates in harvest control rules may not necessarily lead to an inability to achieve management goals. This is because the feedback nature of harvest control rules mean that errors may be corrected over time. Appendix 6 uses MSE to explore the effect of spatial heterogeneity, including spatial closures, on the ability of feedback-control management strategies to achieve goals relating to conservation and utilization of fishery resources. The operating model underlying the projections is based on Pink Ling off southern Australia and assumes that animals are sedentary following settlement.

The MSE allows for three spatial zones (nominally zones 10, 20, and 30 of the SESSF), one of which (either zone 10 or zone 20) can be assumed to be closed. The stock structure hypothesis is that there is a single biological stock across the three spatial zones, i.e. the recruitment for each zone is determined by the total spawning biomass across all three zones. The fish populations in these three zones are assumed to be mixed through the settlement of age-0 animals, with animals of age-1 and older not moving between zones. Two fleets (trawl and non-trawl) are assumed to operate in each zone, growth can be assumed to differ among zones, and recruitment is assumed to be stochastic, with spatial variation in the proportion of the total recruitment that settles to each zone, as well as temporal variation in total recruitment. Pink Ling is a 'Tier 1' stock and its RBC is based on applying a harvest control rule to the outcomes from stock assessments (separately east and west of 147°E) implemented using Stock Synthesis and CASAL. The management strategies considered are based on the harvest control rule and two Stock Synthesis assessment methods used in actuality for Pink Ling off southern Australia.

The primary aim of the study is to examine whether biases in assessment outcomes caused by spatial heterogeneity in population structure will lead to an inability to achieve management goals.

Empirical evaluation of catch rate standardizations

The work in Appendix 7 constituted an empirical consideration of the effects of closures upon catch-rate (catch-per-unit-effort, CPUE) standardizations through using actual standardizations from the SESSF. It applied three treatments related to the inclusion or exclusion of data from closures and then compared the outcomes of the analyses. The treatments were 1) to ignore the advent of closures (possibly valid if the closure was small or little catch was taken there), 2) treat the closure as a factor in the standardization where all data from the closure area are treated as one level and data from outside a different level of a single categorical factor (possibly valid if the closure has not been present for too many years), and 3) exclude all data ever taken from within the closure region. While very many Commonwealth Marine Protected Areas have been declared, only those in the south-east are currently active so the analyses were restricted to the SESSF. The trawl fisheries off of eastern Australia that were standardized were for Tiger Flathead, Pink Ling, and John Dory. The auto-line fishery for Blue-Eye Trevalla was also analysed in the same manner.

Simulation evaluation of catch rate standardizations

Appendix 8 examined the effect of marine closures on standardized CPUE to determine how well CPUE acts as a relative index of abundance according to different scenarios regarding resource (fish) movement dynamics and fisher behaviours. This was achieved by simulating CPUE data using an agent based model (Resource - Fisher Integrated Model – RESFIM) across selected resource movements based on a generic platycephalid (i.e., tiger flathead) frequently occurring in the SESSF, and selected fisher behaviours that differ in terms of the degree of knowledge of resource dynamics. Generated RESFIM CPUE data were standardized using generalized linear models (GLIMs), and resultant CPUE indices compared with true abundance to examine the effect that marine closures have on the CPUE-abundance relationship. In addition, the estimated bias of the proportionality parameter (which links CPUE to abundance), the degree of improvement of CPUE-abundance linearity following standardisation, and relative errors of annual indices (temporal bias) were also examined in the context of marine closures.

Results and Key Findings

Ecological Risk Assessment

Productivity and Susceptibility Analysis (PSA)

Results obtained using the original PSA method found that there was a substantial 36–50% reduction in overlap between fishing area and the distribution of seven non-target species included in the comparison (Appendix 1). This reduction is attributed to spatial management. However, only one out of seven of the overall species risk scores decreased in response to reduced overlap using the original PSA methods. The method was then refined to use a continuous (percent overlap) rather than categorical (high, medium, low) scoring approach. This refinement was coupled to revised species maps based on new depth contours obtained from the latest seafloor mapping to revise species boundaries. These coupled refinements were used to develop a version of the PSA that is more sensitive to reduced overlap. When this method was applied to the 2013 data, only one of seven species remained at high risk. The revised PSA method is an improvement that can show the effectiveness of management actions.

Sustainability Assessment for Fishing Effect (SAFE)

A Base SAFE was conducted for seven non-target species using trawl fishing effort distributions respectively for 2003 and 2013 (Appendix 2). Two species, Brier Shark (*Deania calcea*) and Plunket's Shark (*Centroscymnus plunketi*), had a fishing mortality rate equal to or slightly greater than the reference point F_{msy} in 2003. Our new results show that fishing mortality is lower than F_{msy} for all species in 2013. A comparison of these two years reveals that fishing mortality in 2013 was reduced by between 35% and 50% from the 2003 level for the evaluated species. The results show that fishing impact on non-target species has decreased substantially since 2003. This decrease was due to a large reduction in total fishing effort rather than spatial closures. Since the SAFE method assumes homogeneous or a random distribution of fish between fished and unfished areas and uses actual fishing effort data, the current method remains valid when marine closures are imposed.

The Coral Trout Fishery

Results from simulations showed that the assessment model estimates stock status at about 50% of initial levels when the simulated line survey collects data only from the reefs in which the fishery operates, whereas stock status in the operating model is around 70-80% (Appendix 3). The estimated stock status in each projection year varied little across the replicate simulations. Stock status is generally under-estimated (negatively biased), with the extent of negative bias related to mis-specification of the breeding strategy of the target fish stock, the impact of the amount of larval connectivity among reefs, including reefs in fished areas and closures, as well as exploitation rates. The estimates of stock status were less negatively biased when fishery-independent index and age- length-composition data were available from closed areas. The amount of closure (21% or 41%) had little effect on the negative bias when self-seeding was low and reefs are highly connected. In contrast, there was a marked effect of closures on the bias of stock status when recruitment to a reef was more independent of that to other reefs (high self-seeding).

The results will inform the development of management strategies for Coral Trout in the GBR, and highlight the importance of basing evaluations of estimation and management performance on operating models that capture ecologically-important processes, such as metapopulation dynamics and protogynous life history.

The Southern and Eastern Scalefish and Shark Fishery

Evaluation of current assessment methods

Results suggest that full quantitative assessments with their associated harvest control rule (Tier 1) generally performed well in terms of meeting target biomass levels for the open area, in particular if all data are used (not just data from the open area) and closure size is small (Appendix 4). The data-poor harvest strategies (Tiers 3 and 4), show increased risk, with the catch curve based Tier 3 having a greater risk profile compared to the catch rate based Tier 4 harvest strategy. If managers choose a stock-wide target (of say 48% of initial biomass levels), then the open area biomass target can be less than the stock–wide target (less than 48%), due to the protection of stock within the closure. A stock-wide biomass target strategy also tends to maintain catches irrespective of mixing level. Alternatively, a harvest strategy based on a target for the open area alone provided higher catch rates and stock-wide biomass for a particular mixing rate, but with potentially less annual catch than a strategy based on maintaining stock-wide biomass. Discount factors reduced stock risk when Tiers 3 and 4 harvest strategies are applied, but do not match the Tier 1 risk profile (across all HCR targets) even with a large discount factor. Using all of the data shows better performance (at meeting the target biomass) across Tiers, but has increased risk compared to only using data from the open area. With low harvest control rule targets, the biomass dynamics for the data-poor harvest strategies become unstable (noting that these low targets are unlikely to be adopted by management). The cycling is dampened by increased mixing rates and discount factors.

Evaluation of alternative assessment methods

Appendix 5 explored the selection of an assessment configuration to apply in the face of spatial variation in exploitation rate, including closed areas. This study, amongst others, has shown that it is desirable to use spatially-structured assessment methods in the face of spatial variation in exploitation rates and biological parameters, and highlighted the consequences of selecting a mis-specified spatial model. Use of such spatial methods in actual stock assessments (rather than in research on stock assessments) is still uncommon, but the number of such assessments is increasing.

The results highlighted that closed areas increase uncertainty (and bias when the assessment is mis-specified). The bias in estimates of spawning stock biomass associated with spatially-aggregated assessment methods increases in the presence of closed areas. These biases can be reduced (or even eliminated) by applying appropriately constructed spatially-structured stock assessments. The performance of spatially-aggregated assessments when estimating spawning stock biomass was found to depend on the interactions among spatial variation in growth, in exploitation rate, and in knowledge of the spatial areas over which growth and exploitation rate are homogeneous.

More specifically, the study showed that estimating spatially-varying growth did not lead to much poorer estimation performance when growth did not vary spatially, but led to markedly improved estimation performance when this was the case. In contrast, ignoring spatial growth could lead to large biases when growth actually varied spatially.

This study also explored the value of survey data in the face of closures. In general, the effects of model mis-specification dominate those of additional data from the survey. However errors for the assessment configuration that is correctly specified were lower when relatively precise survey data were available, particularly when the survey data started before the closed areas were implemented. Unfortunately, the magnitude of improved performance when including survey data was fairly limited even though the assumed survey CV was very small. A key reason for the inability to estimate biomass was the lack of catch-rate and compositional data in the early years of the fishery and this problem was not overcome by collecting survey data for recent years.

The spatial assessment configurations estimate biomass by zone, and in principle, the results from such configurations could form the basis for spatial management (e.g. assessing stock status by zone and setting catch limits spatially based on spatial stock status). Whether basing management advice on the outcomes of such assessments will lead to an inability to achieve management goals with respect to sustainability is explored in the subsequent Appendix.

Evaluating harvest control rules

The results of Appendix 6 confirm those of earlier studies that spatial heterogeneity in abundance and age structure will lead to bias for assessments based on spatially-aggregated population dynamics models. They further confirm that the extent of bias in estimates of total spawning biomass is exacerbated in the presence of spatial closures. Although the simulations involved closures that were in operation for more than 50 years, there was no evidence that biases that arise due to spatial closures (or even spatial heterogeneity in population structure) reduce over time, even when high precision survey data are available for assessment purposes.

However, biased estimates of total spawning biomass do not necessarily lead to a complete inability to achieve management goals. While the time-trajectories of catches and biomass are substantially more variable than would be expected had a single stock been managed, the stock tends to be fairly close to the target level at the end of the projection period, with a low probability of being depleted below the limit reference point, at least in the absence of closures and for closed areas that are up ~25% of the stock area. Larger closed areas lead to lower catches and stock sizes in excess of the target level. The extent to which the stock is above the target level depends on the size of the closed area and the target level.

The management strategies were able to move the resource towards the target level in the absence of spatial closures even though assessment results are biased. The probability of reducing the stock below its limit reference point was higher when growth rates vary spatially, but the effect was small. The probability of the stock being above its target reference point was lower when one of the smaller spatial strata was closed. However, performance was markedly different when a large fraction of the area was closed, with the stock substantially larger than the target level at the end of the projection period.

Empirical evaluation of catch rate standardizations

The outcomes from the three treatments (as measured by the trend of standardized CPUE through the years) barely differed from each other in all the trawl fisheries considered (Appendix 7). This was not surprising as in most cases the closures present only influenced a very small proportion of the area in which catching takes place, and an equivalently small proportion of the catch. In addition to the trawl fisheries, the auto-line fishery for Blue-Eye Trevalla was included and this differed from the trawl fisheries because some of the recent closures, such as the Flinders Research Zone and around the St Helens Hill, together accounted for just over 20% of all catches in Zones 20 and 30. Even in that case, the differences between treatments were minor and mostly occurring up to from 1997 – 2007, but even those were within the bounds of uncertainty of each of the standardizations.

The lack of influence or effect of the current marine protected areas within the south-east on the trawl fisheries should not be surprising. Before the Commonwealth closures were first introduced, a separate research project was initiated tasked with producing alternative closure definitions that attempted to minimize the effect of those closures upon commercial fisheries, which was relatively successful. This is reflected in the minor amounts of catches excluded from the trawl fisheries by those closures, which in turn flows on to the lack of any significant effects upon the standardizations. This could not be an explanation for the lack of effect in the Blue-Eye auto-line fishery because the closures with most influence were introduced later than the Commonwealth closures. Instead, it appears that the Industry vessels and their skippers are capable of rapidly adapting to the advent of even effectively large closures so that any potential effects they might have are masked by the vessels altering their fishing behaviour and developing alternative fishing grounds.

The current south-east closures have only had a minimal effect upon the standardized CPUE trends through a combination of the closures being designed to have minimal effects on fisheries or the fishers themselves adapting to the closure of some of their favoured fishing grounds by developing alternative localities. The optimum strategy for any standardization is to exclude all data taken within the closed area from subsequent analysis, despite closures having a minimal effect when standardizing catch and effort data from a region that contains closures. Then the next best approach would be to treat the inside and outside of a closure as a factor in a standardization, which is akin to declaring two, or more, new areas within the data being standardized if excluding data restricts the amount of data too much for a usable standardization. By definition the assumption would be that there would be no data in the closures after they were established.

Simulation evaluation of catch rate standardizations

Linearity between CPUE and abundance is desired for CPUE to adequately index abundance. Significant improvements (in terms of linearity) over unstandardized indices for most resource/fisher scenarios support the use of standardized CPUE estimates as proxies for abundance with or without closures (Appendix 8). Significant improvements occurred for scenarios where resource movement was non-random, fishers shared information and with perfect fisher behaviour. Standardized CPUE were least effective (i.e. provided minimal improvements towards linearity) at indexing abundance under random resource movement (i.e. fish move randomly in space), irrespective of fisher behaviours with or without marine closures. Bias in standardized CPUE-abundance relationships were greater across each of the resource/fisher scenarios with, than without, closures.

Implications

The refined Productivity Susceptibility Analysis (PSA) methodology developed in this report includes a continuous scoring approach for overlap instead of categorical, and more refined spatial maps of species distribution. This will greatly improve confidence in ERA results for non-quota species (Appendix 1). These refinements provide sensible and appropriate improvements to the methodology and, as shown through the examples, illustrate that substantially different management outcomes may result (a shift from high risk to a lower risk categorization for several of the species considered). The refined SAFE methodology for non-target species that accounts for spatial closures showed that major changes from previous SAFE results were largely due to changes (reductions) in effort than closures. The new methodology of Appendix 2, and the examples provided, conclude that the current SAFE method remains valid and major changes are not required when marine closures are imposed.

This report produced the first stock assessment for Coral Trout that evaluates the biases on stock biomass estimation when there are multiple marine closures (Appendix 3). In the application presented in this report, stock status was generally under-estimated, with the degree of under-estimation related to uncertainties regarding the biology of the species and the larval mixing between reefs. The results from this work, and the assessment methods developed, will greatly improve confidence in stock status estimation and guide management advice on appropriate catches and management strategies in the GBR.

Testing of current assessments and harvest control rules showed that full quantitative assessments (Tier 1) generally performed well at meeting target biomass (Appendix 4). The MSE was able to explore the implications of having a stock-wide biomass target or an open-area only biomass target. This is relevant to current management and the revised harvest strategy policy. This study concluded that if managers chose a stock-wide biomass target, then not surprisingly, the open area target can be less (than 48% of virgin open area biomass say) as stock is protected in the closed area. The degree of adjustment to the open-area target to maintain a stock-wide biomass level is contingent on mixing rates, however. The stock-wide biomass target tends to maintain catches irrespective of mixing level. In contrast, if managers chose to set biomass targets for the open-area only, then catch rates (which can be interpreted as a proxy for economic performance) are higher, as is the stock-wide biomass. However, annual catches may be lower than having a stock-wide biomass target. Other important results include that the Tier 3 risk profile (the probability that the stock biomass is below the limit reference point as a function of mixing and biomass targets) is greater (for a specific mixing rate and target biomass) than Tier 4. This has implications for buffering uncertainty between assessment methods, as currently application of Tier 4 leads to a greater discount (buffer) than Tier 3. Testing discount factors also showed that the current factors are not sufficient to match the risk profile of a Tier 1.

The evaluation of alternative stock assessment methods in the face of spatial variation in growth and marine closures conducted in this report will help guide stock assessment scientists regarding appropriate configurations of stock assessment models (Appendix 5). This will provide greater confidence for managers and industry with regard to model outcomes (TAC setting). Many alternatives are available, from fully aggregated to fully spatial, and which are 'best' is an often debated and uncertain scientific question. The 'best' assessment configuration in the face of spatial variation in growth was FULL whereas SSTVR and SSTVRSEL outperformed FULL when growth was the same spatially. The work concludes that it may be 'safer' to estimate more parameters (through a fully spatially structured model) even when they are not deemed necessary by the data. The study showed that estimating spatially-varying growth did not lead to much poorer estimation performance when growth did not vary spatially, but led to markedly improved estimation performance when this was the case. In contrast, ignoring spatial growth could lead to large biases when growth actually varied spatially. In addition, the value of additional data from closed areas was questionable, in particular when considered across the uncertainties faced with model choice. Mis-specification of the model had a much larger impact on bias than could be overcome by additional data in the closed area. This outcome should be recognized by managers if additional data (and funding) are being considered from closed areas. Not surprisingly, this work highlights that closed areas increase uncertainty. However, basing management advice on the outcomes of such assessments can still achieve management goals with respect to sustainability (Appendix 6). The results of the MSE evaluation of harvest strategies highlighted that effective management may require sampling programs and assessment frameworks designed to support management strategies tailored to there being closed areas (including tagging programs and sampling of age structure in closed areas).

The empirical evaluation of catch rates, with a case study focus of the SESSF, showed that the trend of the standardized catch rate did not vary greatly across the three treatments considered within the trawl fisheries (ignore the closure; data from within the closure are treated as one level of a factor in the GLM; exclude data from the closed area across the full capture history) and the three focal species (flathead, blue-eye and John Dory) (Appendix 7). This was because the current closures represent only a small proportion of the area in which catches occurs and a small proportion of the catch. With the auto-line fishery for Blue-Eye Trevalla the proportion of areas excluded was much greater (up to 30% of catches) but even in that case the differences, which were much more apparent than in the trawl fisheries, remained minor.

The approach adopted in the SESSF for standardizations, namely removing all data taken within the closed area, appears optimal and adequate for catch rates as an input to assessments and management considerations. Simulations of catch rates across multiple factors including fish dynamics and fisher behaviour indicated that standardizations greatly improved the relationship between actual biomass trends and trends estimated from the GLM, whether a marine closure exists within the species habitat or not (Appendix 8).

Recommendations

- Recommendations from the refined Productivity Susceptibility Analysis include that further application use continuous scoring and updated fine-scale spatial mapping of species and effort distributions. Further work is needed to resolve the level of risk associated with different levels of overlap between fishing effort and distribution, particularly for aggregating bycatch and byproduct species. Closures reduce this overlap and consequently reduce availability (exposure) risk. Species that aggregate are more likely to be sensitive to closure size and location. Many bycatch species have poorly known breeding and feeding patterns therefore levels of aggregation are not well known. The PSA method would benefit from future development of some form of statistical method of identifying aggregating species. Patchiness in mapped observer data is one method that could be explored.
- The SAFE methodology was found to be able to quantitatively distinguish whether the change in risk is due to a change in the amount of fishing effort or a change in effort distribution caused by spatial closures. No methodology modification is needed for the base SAFE method when spatial closures are imposed. The enhanced SAFE method assumes a heterogeneous distribution of fish density so will be more accurate in capturing any spatial effect than the base SAFE. The enhanced version requires more fishery data and analytical effort, but is recommended for more rigorous assessment, particularly for species that are assessed as high risk by PSA and base SAFE.
- The development of a stock assessment for coral trout that includes management closures was promising and has been addressed in the latest assessment (Leigh et al. 2014). Evaluation of bias using simulation did show an over-estimation of depletion (estimates of 50% of virgin levels compared to 70-80% in the operating model), depending on biological characteristics such as self-seeding rates. These results should be considered in any further development of the model for management purposes and highlights the importance of understanding population connectivity.
- The MSE evaluation of current tier level assessments showed that generally the catch-rate-based Tier 4 assessment had a lower risk profile than the catch curve based Tier 3 assessment method. This should be considered when setting discount rates for stocks utilising these approaches. Tier 1 assessments performed well at meeting biomass targets in general, and in particular if all data (from closed and open areas) are used in the full assessment. The MSE also showed that targets can be reduced in open areas and still maintain stock-wide targets (eg if a stock-wide target is 48% of virgin levels, depending on mixing rates, the target in the open area can be less than 48%). Having open-area management targets leads to increased catch rates compared to equivalent stock-wide biomass targets, but potentially less catch. Managers should consider which form of biomass target is reasonable given utilisation and conservation objectives. These results will inform managers of the potential outcomes from the decision process.
- Due to the increased uncertainty introduced by marine closures, spatial variation in growth should be explored in stock assessments, and fully spatially-explicit models developed in preference to aggregated models. Additional sampling within closed areas does not greatly improve assessment outcomes (reduce bias) as model mis-specification dominates this bias. As the spatial assessment configurations estimate biomass by zone, the results from such configurations could form the basis for spatial management (e.g. assessing stock status by zone and perhaps setting catch limits spatially based on spatial stock status).

- MSE testing of feedback control rules showed that biased assessments due to unmodelled spatial heterogeneity do not necessarily lead to an inability to achieve management goals. The ability to achieve management goals is affected by closed areas, with the effects greater for large closed areas. Effective management may require sampling programs and assessment frameworks designed to support management strategies tailored to there being closed areas. Such programs could involve tagging programs as well as sampling of age structure in closed areas. Frameworks such as those outlined in this report could be used to evaluate the extent to which alternative management strategies can outperform current approaches.
- Empirical testing of alternative data structures as inputs for catch rate standardizations in the SESSF showed that Commonwealth closures in the South-East have had minimal impact on the resulting time-series of abundance for the stocks examined. The optimal strategy for data inputs when there are marine closures is to exclude all data from within the closed area, and next best was to treat the closure as a factor in the GLM. An exception to these conclusions is the 700m deepwater closure where the principle fishing area for species such as Orange Roughy and the eastern and western Deepwater Sharks (basket TAC species) have been closed. For species where most of the fishable preferred habitat is closed the meaningfulness of any CPUE standardization becomes questionable. Generally the number of available records is greatly reduced and there are repeated reports of fishers altering their fishing behaviour near and around the deepwater closure. Such changes in fishing behaviour would imply that the catches taken by a given amount of effort after imposition of the closure are not necessarily comparable to catches taken by the same amount of effort prior to the closure. Once further marine closures come into effect and exclude fishing, more examples of different degrees of overlap with active fishing areas will become available and further empirical studies of the effects of such closures can be made.
- Simulation testing of impacts on standardisations of marine closures under alternative scenarios of fish and fleet behaviour showed that standardising catch rates improved estimated abundance trends (compared to the 'truth') over nominal catch rates. Statistical standardization analyses should be employed to improve CPUE-abundance relationships and reduce temporal biases in standardized indices either with or without marine closures. At least year (Y), vessel (V), month within year (M), grid-location (G) and interaction term grid-location × month (G × M) should be employed in standardizations to obtain greatest improvements towards linearity and least biased estimates.

Extension and Adoption

The extension of the work and methods presented here is a formal objective of this project and will entail making presentations and explanations to the RAGs and MACs concerned along with other interested stakeholders. Formal explanatory documents describing the methods in detail, based on published literature, will also be presented.

The methods and results of this project have direct relevance to the types of assessments being used to set RBCs and TACs and manage fisheries in Australian and, for that matter, international fisheries. The evaluation of the various assessment methods that include closures will influence RAG decisions regarding the types of assessments used to determine stock status and set TACs in the SESSF. The MSE evaluation of harvest control rules across alternative Tier levels according to the size, mixing level and species characteristics has direct relevance to the recent (2017) revision of the Australian Harvest Strategy Policy. ERA methods such as SAFE and PSA continue to be used to manage non-quota and incidentally caught species. The revisions made given the results of this project, where appropriate, will provide managers with greater confidence in ERA outcomes. With regard to catch rate standardizations, AFMA has adopted the principle of conducting catch rate analyses on records that occur only outside of a closure (for say Deepwater Sharks and Silver Trevally), as confirmed and recommended by this study.

Project Materials Developed

Published peer-reviewed articles:

Punt, A.E., Haddon, M., Little, L.R. and Tuck, G.N. 2016. Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. Fisheries Research. 175: 10-23

Punt, A.E., Haddon, M., Little, L.R. and Tuck, G.N. 2016. The effect of marine closures on a feedback control management strategy used in a spatially aggregated stock assessment: a case study based on pink ling in Australia. Canadian Journal of Fisheries and Aquatic Science. 73: 1–14

Little, L.R., Punt, A.E., Tuck, G.N., and Mapstone, B.D. 2017. Exploring the effect of sampling, protogyny, and larval advection on stock estimates subject to no-take closures in a spatially complex coral reef line fishery on the Great Barrier Reef, Australia. Canadian Journal of Fisheries and Aquatic Science. 74: 1950–1959

References

Abesamis R.A., Alcala, A.C. and Russ, G.R. 2006. How much does the fishery at Apo Island benefit from spillover of adult fish from the adjacent marine reserve? Fishery Bulletin 104, 360-375.

Aburto-Oropeza, O., Erisman, B., Galland, G.R., Mascarenas-Osorio, I., Sala, E., and Ezcurra, E. 2011. Large recovery of fish biomass in a no-take marine reserve. PLoS ONE, 6: e23601. doi:10.1371/journal.pone.0023601

AFMA 2012, Upper-Slope Dogfish Management Strategy AFMA-managed Fisheries October 2012, Australian Fisheries Management Authority. Canberra, Australia. http://www.afma.gov.au/wp-content/uploads/2014/12/Upper-slope-Dogfish-Management-Strategy-14December-2012-FINAL.pdf (last accessed 4 January 2016).

AFMA 2015, Northern Prawn Fishery Directions and Closures 2015, Australian Fisheries Management Authority. Canberra, Australia. http://www.afma.gov.au/wp-content/uploads/2014/08/NPF-Directions-and-Closures-2015.pdf (last accessed 3 January 2016).

Akaike, H. (1974). A new look at statistical model identification. *IEEE Trans. Auto. Control* 19: 716-722.

Anon. 2015. Commonwealth Marine Reserves. http://www.environment.gov.au/topics/marine/marine-reserves Department of the Environment, Australian Government.

Armstrong, C.W. 2007. A note on the ecological–economic modelling of marine reserves in fisheries. Ecological Economics 62, 242-250.

Azarovitz, T.R. (1981). A brief historical review of the Woods Hole laboratory trawl survey time series. *Can. Spec. Publ. Fish. Aquat. Sci*. 58: 62-67.

Babcock, E.A. and MacCall, A. 2011. How useful is the ratio of fish density outside versus inside no-take marine reserves as a metric for fishery management control rules? Canadian Journal of Fisheries and Aquatic Sciences 68, 343-359.

Bull, B., Francis, R.I.C.C., Dunn, A., McKenzie, A., Gilbert, D.J., and Smith, M.H. 2005. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.07-2005/08/21. NIWA Technical Report 127.

Bunnefeld, N., Hoshino, E. and Milner-Gulland, E.J. 2011. Management Strategy Evaluation: A Powerful Tool for Conservation? Trends in Ecology and Evolution, 26(9), 441-447.

Burnham K.P., Anderson, D.R. (1998). *Model Selection and Inference. A Practical Information-Theoretic Approach*. Springer-Verlag, New York. 353 pp.

Burnham, K., Anderson, D.R. (2002). Model selection and multimodel inference. A practical information approach. Second edition. Springer-Verlag New York Inc. 488 pp.

Butterworth, D.S., and Punt, A.E. 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56: 985–998.

Buxton, C., Haddon, M. and Bradshaw, M. 2006. Regional impact assessment for the marine protected areas proposed for the South East region. FRDC Project 2005/083. FRDC Final Report. Tasmanian Aquaculture and Fisheries Institute, University of Tasmania. 204 p.

Cadrin, S.X., and Secor, D.H., 2009. Accounting for spatial population structure in stock assessment: Past, present, and Future. In: The Future of Fisheries Science in North America, Fish & Fisheries Series 31, pp 405-426.

Campbell, R.A. (2004). CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fish. Res*. 70: 209-227.

Chambers, J.M., Hastie, T. (1992). *Statistical Models*. Chapman and Hall, New York. 608pp.

Commonwealth of Australia (2003). *Australia's South-east Marine Region: A user's guide to identifying candidate areas for a regional representative system of Marine Protected Areas.* 85 pp.

Conn, P.B., Williams, E.H., Shertzer, K.W., 2010. When can we reliably estimate the productivity of fish stocks? Can. J. Fish. Aquat. Sci. 67, 511–523.

Cope, J. M., Piner, K., MinteVera, C. V. and Punt, A. E. 2003. Status and Future Prospects for the Cabezon (Scorpaenichthys marmoratus) as Assessed in 2003 Portland, OR: Pacific Fishery Management Council. http://www.pcouncil.org/wp-content/uploads/Cabezon_2003_Assessment.pdf

Cordue, P.L. 2015. Stock assessment results for eastern pink ling: new base model. Post-meeting Draft report to the RAG meeting of 29th October 2015.

DAFF (Australian Government Department of Agriculture, Fisheries and Forestry) 2007. *Commonwealth Fisheries Harvest Strategy: policy and guidelines*, DAFF, Canberra.

DAFF. 2007. Commonwealth Fisheries Harvest Strategy Policy Guidelines. Australian Government Department of Agriculture, Fisheries and Forestry, Canberra, Australia, pp. 55. http://www.agriculture.gov.au/fisheries/domestic/harvest_strategy_policy (last accessed 3 January 2016).

Daley, R.K., Williams, A. Green, M., Barker, B. and Brodie, Pamela. 2015. Can marine reserves conserve vulnerable sharks in the deep sea? A case study of Centrophorus zeehaani (Centrophoridae), examined with acoustic telemetry. Deep Sea Research II. 115: 127-136.

Day, J. 2009. School whiting (*Sillago flindersi*) stock assessment based on data up to 2008. Technical report to the Shelf RAG 23 November 2009.

Day J., Haddon M., and Hillary R. 2015. Stock Assessment of the Macquarie Island fishery for Patagonian toothfish (Dissostichus eleginoides) using data up to and including August 2014. Report to SARAG 51, 24 February 2015. CSIRO, Australia.

Deroba, J.J., and Bence, J.R. 2008. A review of harvest policies: Understanding relative performance of control rules. Fish. Res. 94: 210–223.

Dichmont, C.M., Punt, A.E., Dowling, N., De Oliveira, J.A.A., Little, L.R., Sporcic, M., Fulton, E., Gorton, R., Klaer, N., Haddon, M., and Smith, D.C. In press. Is risk consistent across tier-based harvest control rule management systems? A comparison of four case studies. Fish. And Fish. 00: 00–00.

Fay, G., Punt, A.E., and Smith, A.D.M. 2009. Operating model specifications. In: Wayte, S.E. (Ed) Evaluation of new harvest strategies for SESSF species. CSIRO Marine and Atmospheric Research and Australian Fisheries Management Authority. Report 2006/815. June 2009. 137p.

Fay, G., Punt, A.E., and Smith, A.D.M. 2011. Impacts of spatial uncertainty on performance of age structure-based harvest strategies for blue eye trevalla (Hyperoglyphe antarctica). Fish. Res. 110: 391–407.

Fay, G., Little, L.R., Tuck, G.N., Haddon, M. and Klaer, N. 2013. Maintaining risk equivalency among fishery harvest control rules in Southeast Australia. In Tuck, G.N. (Ed.) Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery: 2012, Part 2. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research. Pp. 1-22

Fernandes, L., et al. 2005. Establishing Representative No-Take Areas in the Great Barrier Reef: Large-Scale Implementation of Theory on Marine Protected Areas. Conservation Biology 19, 1733-1744.

Field, J.C., Punt, A.E., Methot, R.D., and Thomson, C.J. 2006. Does MPA mean "Major Problem for Assessments?" Considering the consequences of place-based management systems. Fish and Fish. 7: 284–302.

Fletcher, W. J., Kearney, R. E., Wise, B. S., and Nash, W. J. 2015. Large-scale expansion of no-take closures 1 within the Great Barrier Reef has not enhanced fishery production. Ecology.

Francis, R.I.C.C., 1992. Use of risk analysis to assess fishery management strategies: A case study using orange roughy (Hoplostethus atlanticus) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49, 922–930.

Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138.

Francis, R.I.C.C., 2014. Replacing the multinomial in stock assessment models: A first step. Fish. Res. 151, 70–84.

Fu, C.H., Fanning, L.P., 2004. Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. N. Am. J. Fish. Manage. 24, 775–784.

Garrison, T.M., Hamel, O.S., Punt, A.E., 2011. Can data collected from marine protected areas improve estimates of life-history parameters? Can. J. Fish. Aquat. Sci. 68, 1761–1777.

Gaertner, D., Dreyfus-Leon, M. (2004). Analysis of non-linear relationships between catch per unit effort and abundance in a tuna purse-seine fishery simulated with artificial neural networks. *ICES J. Mar. Sci*. 61: 812-820.

Greenstreet, S.P.R., Armstrong, E., Mosegaard, H., Jensen, H., Gibb, I.M., Fraser, H.M., Scott, B.E., Holland, G.J., and Sharples, J. 2006. Variation in the abundance of sandeels Ammodytes marinus off southeast Scotland: an evaluation of area-closure fisheries management and stock abundance assessment methods. ICES Journal of Marine Science. 63: 1530 – 1550.

Goethel, D.R., Legault, C.M., Cadrin, S.X., 2015. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (Limanda ferruginea) through simulation analysis. Can. J. Fish. Aquat. Sci. 72, 582-601.

Goodyear, C.P. (2003). Tests of the robustness of habitat-standardized abundance indices using simulated blue marlin catch-effort data. *Aust. J. Mar. Freshwater Research. Res*. 54: 369-381.

Goodyear, C.P. (2006). Simulated Japanese longline CPUE for blue marlin and white marlin. *Col. Vol. Sci. Pap. ICCAT* 59(1): 211-223.

Guan, W., Cao, J., Chen, Y., Cieri, M., 2013. Impacts of population and spatial structures on fishery stock assessment. Can. J. Fish. Aquat. Sci. 70, 1178–1189.

Haddon, M., 2014. Catch rate standardizations for selected species from the SESSF (data 1986 – 2012). p. 57-275 In: G.N. Tuck (Ed). Stock Assessment for the Southern and Eastern Scalefish and Sharkfish 2013, Part 2. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Haddon, M. 2014. Tier 4 Analyses in the SESSF, including Deep Water Species. Data from 1986 – 2012. Pp 352 – 461 *in* Tuck, *G.N. (ed.) Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2013. Part 2.* Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 487 p.

Haddon, M. 2016a. Blue-Eye Auto-Line and Drop-Line CPUE Characterization and Catch-per-Hook 1997 – 2015. Draft Assessment Document presented the SERAG Nov 2016. 52 p.

Haddon, M. 2016b. Blue-Eye (Hyperoglyphe antarctica) Tier 4 Analysis using Catch-per-Hook for Auto-Line and Drop-Line from 1997 - 2015. CSIRO Oceans and mosphere. Hobart. 25p.

Halpern, B.S. 2003. The impact of marine reserves: Do reserves work and does reserve size matter? Ecological Applications 13, S117-S137.

Haltuch, M.A., Ono, K., and Valero, J. 2013. Status of the U.S. petrale sole resource in 2012. Pacific Fishery Management Council, 7700 Ambassador Place NE, Portland, OR 970220. http://www.pcouncil.org/wp-content/uploads/Petrale_2013_Assessment.pdf (last accessed 3 January 2016).

Hanchet, S.M., Blackwell, R.G., Dunn, A. (2005). Development and evaluation of catch per unit effort indices for southern blue whiting (*Micromesistius australis*) on the Campbell Island Rise, New Zealand. *ICES J. Mar. Sci*. 62: 1131-1138.

Harford, W.J., Ton, C., Babcock, E.A., 2015. Simulated mark-recovery for spatial assessment of a spiny lobster (Panulirus argus) fishery. Fish. Res. 165, 42–53.

Harley, S.J., Myers, R.A., Dunn, A. (2001). Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1760-1772.

Hilborn, R. 1985. Fleet dynamics and individual variation: why some people catch more fish than others. Can. J. Fish. Aquat. Sci. 42: 2–13.

Hilborn, R. et al. 2004. When can marine reserves improve fisheries management? Ocean & Coastal Management 47, 197-205.

Hobday, A. J., Smith, a. D. M., Stobutzki, I. C., Bulman, C., Daley, R., Dambacher, J. M., Deng, R. a., et al. 2011. Ecological risk assessment for the effects of fishing. Fisheries Research, 108: 372–384. http://linkinghub.elsevier.com/retrieve/pii/S0165783611000324 (Accessed 19 September 2013).

Hulson, P-J.,F., Miller, S.E., Ianelli, J.N., Quinn, T.J. II., 2011. Including mark-recapture data into a spatial age-structured model: walleye Pollock (Theragra chalcogramma) in the eastern Bering Sea. Can. J. Fish. Aquat. Sci. 68, 1625–1634.

Hulson, P-J.F., Quinn, T.J. II., Hanselman, D.H., Ianelli, J.N., 2013. Spatial modeling of Bering Sea walleye pollock with integrated age-structured assessment models in a changing environment. Can. J. Fish. Aquat. Sci. 70, 1402–1416.

IMCRA. 1998. Interim Marine and Coastal Regionalisation for Australia: ecosystem-based classification for marine and coastal environments. Version 3.3. Environment Australia, Commonwealth Department of the Environment. Canberra. Australia.

International Whaling Commission (IWC) 2012. The Revised Management Procedure (RMP) for Baleen Whales. J. Cet. Res. Manage. 13(Suppl), 485–494.

Klaer, N.L. (2004). Abundance indices for main commercial fish species caught by trawl from the south-eastern Australian continental shelf from 1918-1957. *Aust. J. Mar. Freshw. Res.* 55: 561-571.

Klaer, N. 2010. Tiger flathead (*Neoplatycephalus richardsoni*) stock assessment based on data up to 2009. Technical report presented to the Shelf RAG 8-9 November 2010.

Klaer, N.L., Wayte, S.E., and Fay, G. 2012. An evaluation of the performance of a harvest strategy that uses an average-length-based assessment method. Fish. Res. 134-136: 42–51.

Knuckey, I., Fuller, M., Moore, J., Daley, R. (2009). Empowering Industry R&D: Redefining deepwater closures in the SESSF to reduce the impact on the commercial deepwater fishery and maintain adequate protection of orange roughy. AFMA Project 2008/836. p18.

Kritzer, J.P., and Sale, P.F. 2004. Metapopulation ecology in the sea: from Levins' model to marine ecology and fisheries science. Fish and Fisheries 5, 131-140.

Last, P., Lyne, V., Yearsley, G., Gledhill, D., Gomon, M., Rees, T., and White, W. 2005. Validation of national demersal fish datasets for the regionalisation of the Australian continental slope and outer shelf (>40 m depth). Department of Environment and Heritage and CSIRO Marine Research, Australia.

Lauck, T., Clark, C.W., Mangel, M. and Munro G.M.. 1998. Implementing the precautionary principle in fisheries management through marine reserves. Ecological Applications 8:S72–S78.

Lee, H-H, Maunder, M.N., Piner, K.R., Methot, R.D., 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? Fish. Res. 125-126, 254–261.

Leigh, G.M., Campbell, A.B., Lunow, C. P. and O'Neill, M.F. 2014. Stock assessment of the Queensland east coast common coral trout (*Plectropomus leopardus*) fishery. Technical Report. Queensland Department of Agriculture, Fisheries and Forestry.

Lester, S.E., Halpern, B.S., Grorud-Colvert, K., Lubchenco, J., Ruttenberg, B.I., Gaines, S.D., Airame, S., and Waners, R.R. 2009. Biological effects within no-take marine reserves: a global synthesis. Mar. Ecol. Prog. Ser. 384: 33–46.

Little L.R., Punt, A.E., Mapstone, B.D., Pantus, F., Smith, A.D.M., Davies, C.R., and McDonald, A.D. 2007. ELFSim—A model for evaluating management options for spatially structured reef fish populations: An illustration of the "larval subsidy" effect Ecological Modelling 205, 381-396.

Little L.R., Punt, A.E., Mapstone, B.D., Begg, G.A., Goldman, B., and Ellis, N. 2009a. Different responses to area closures and effort controls for sedentary and migratory harvested species in a multispecies coral reef linefishery. ICES Journal of Marine Science 66, 1931-1941.

Little L.R., Punt, A.E., Mapstone, B.D., Begg, G.A., Goldman, B., and Williams, A.J. 2009b. An agent-based model for simulating trading of multi-species fisheries quota. Ecological Modelling 220, 3404-3412.

Little, L.R., Grafton, R.Q., Kompas, T., and Smith, A.D.M., 2010. Closure strategies as a tool for fisheries management in metapopulations subjected to catastrophic events. Fisheries Management and Ecology 17, 346-355.

Little, L.R., Wayte, S.E., Tuck, G.N., Smith, A.D.M., Klaer, N., Haddon, M., Punt, A.E., Thomson, R., Day, J., and Fuller, M. 2011. Development and evaluation of a cpue-based harvest control rule for the southern and eastern scalefish and shark fishery of Australia. ICES J. Mar. Sci. 68: 1699–1705.

Little, L.R., Kerrigan, B., Thébaud, O., Campbell, A.B., Innes, J., Cameron, D., Mapstone, B.D., Norman-Lopez, A., Punt, A.E., Kung, J.P., Slade, S., Leigh, G.M., O'Neill, M.F., and Tobin, A. 2016. Evaluating candidate monitoring strategies, assessment procedures and harvest control rules in the spatially complex Queensland Coral Reef Fin-fish Fishery. Report to the Fisheries Research and Development Corp. Hobart.

Mangel, M. 1998. No-take areas for sustainability of harvested species and a conservation invariant for marine reserves. Ecology Letters 1, 87-90.

Mangel, M. 2000. Irreducible uncertainties, sustainable fisheries, and marine reserves. Evolutionary Ecology Research 2, 547–557.

Mapstone B.D., Davies C.R., Little L.R., Punt A.E., Smith A.D.M, Pantus F., Lou D.C., Williams A.J., Jones A., Ayling A.M., Russ G.R., McDonald A.D. 2004. The Effects of Line Fishing on the Great Barrier Reef and

Evaluations of Alternative Potential Management Strategies. CRC Reef Research Centre Technical Report No 52. CRC Reef Research Centre, Townsville, Australia.

Mapstone, B.D., Little L.R., Punt A.E., Davies C.R., Smith A.D.M., Pantus F., McDonald A.D., Williams A.J., and Jones A. 2008. Management Strategy Evaluations for line fishing in the Great Barrier Reef: Balancing conservation and multi-sector fishery objectives. Fisheries Research 94, 315-329.

Maunder, M., Punt, A.E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fish. Res.* 70: 141-159.

McAllister, M.K., Ianelli, J.N., 1997. Bayesian stock assessment using catch-age data and the Sampling/ Importance Resampling Algorithm. Can. J. Fish. Aquat. Sci. 54, 284–300.

McCook, L.J., Ayling, T., Cappo, M., Choat, J.H., Evans, R.D., De Freitas, D.M., Heupel, M., Hughes, T.P., Jones, G.P., Mapstone, B., Marsh, H., Mills, M., Molloy, F.J., Pitcher, C.R., Pressey, R.L., Russ, G.R., Sutton, S., Sweatman, H., Tobin, R., Wachenfeld, R.E., and Williamson, D.H. 2010. Adaptive management of the Great Barrier Reef: A globally significant demonstration of the benefits of networks of marine reserves. Proc. Natl. Acad. Sci. U.S.A. 107: 18278–18285.

McCullagh,P., Nelder, J.A. (1989). Generalized linear models. Second edition. Chapman and Hall/CRC Monographs on Statistics & Applied Probability. 532 pp.

McDonald, A.D., Parslow, J.S., Davidson, A.J. (2001). Interpretation of a modified linear model of catch-per-unit-effort data in a spatially-dynamic fishery. *Environ. Model. Softw*. 16: 167-181.

McGilliard, C.R., Hilborn, R., MacCall, A., Punt, A.E., and Field, J.C. 2011. Can information from marine protected areas be used to inform control-rule-based management of small-scale, data-poor stocks? ICES Journal of Marine Science 68, 201-211.

McGilliard, C.R., Punt, A.E., Methot, R.D., Hilborn, R., 2015. Accounting for marine reserves using spatial stock assessments. Can. J. Fish. Aquat. Sci. 72, 262–280.

Methot Jr., R.D., Taylor, I.G., 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Can. J. Fish. Aquat. Sci. 68, 1744–1760.

Methot, R.D., Wetzell, C.R., 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

National Environmental Science Program 2016. http://www.nespmarine.edu.au/improved-species-maps-selected-temperate-sharks-and-rays-australia.

National Marine Fisheries Service Office of Science and Technology (2013). Fisheries Independent Surveys System (FINSS) Public User's Manual (v.4.0). 28 pp.

Ono, Kotaro, Punt, A.E., Hilborn, R. (2015a). How do marine closures affect the analysis of catch and effort data? *Canadian Journal of Fish and Aquatic Sciences*. 72: 1177-1190.

Ono, K., Punt, A.E., Hilborn, R. (2015b). Think outside the grids: An objective approach to define spatial strata for catch and effort analysis. *Fisheries Research*. 170: 89-101.

Ortiz, M. (2012). Standardized CPUE of albacore (*Thunnus Alalunga* Bonnaterre, 1788) caught by the Spanish surface longline fishery in the western Mediterranean, 2004-2010. *Col. Vol. Sci. Pap. ICCAT* 68(2): 569-572.

Pincin, J.S. and Wilberg, M.J. 2012. Surplus production model accuracy in populations affected by a no-take marine protected area. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science, 4, 511-525.

Plagányi, É.E., Rademeyer, R.A., Butterworth, D.S., Cunningham, C.L., and Johnson, S.J. 2007, Making management procedures operational—innovations implemented in South Africa. ICES J. Mar. Sci. 64: 626–632.

Pribac, F., Punt, A.E., Walker, T.I., and Taylor, B.L. 2005. Using length, age and tagging data in a stock assessment of a length selective fishery for gummy shark (Mustelus antarcticus). J. Northw. Atl. Fish. Sci. 35: 267–290.

Punt, A.E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. Fish. Res. 65: 391–409.

Punt, A.E., Methot, R.D., 2004. Effects of Marine Protected Areas on the assessment of marine fisheries. Am. Fish. Soc. Symp. 42, 133–154.

Punt, A.E., and Hobday, D. 2009. Management strategy evaluation for rock lobster, Jasus edwardsii, off Victoria Australia: accounting for uncertainty in stock structure. NZ J. Mar. Freshw. Res. 43: 475-609.

Punt, A.E. and Taylor, B. 2012. Stock assessment of pink ling (Genypterus blacodes) using data up to 2010. Pp 103 – 160 In Tuck, G.N. (ed.) (2012) Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2011. Part 1. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 377 p.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. and Haddon, M. 2014. Management Strategy Evaluation: Best Practices, Fish and Fisheries, doi: 10.1111/faf.12104.

Punt, A.E., Hurtado-Ferro, F., Whitten, A.R., 2014. Model selection for selectivity in fisheries stock assessment. Fish. Res. 158, 124–134.

Punt, A.E., Haddon, M., Tuck, G.N., 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment: A case study based on pink ling in Australia. Fish Res. 168, 85–99.

Punt, A.E., Haddon, M., Little, L.R., and Tuck, G.N. 2016. Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. Fish. Res. 175: 10-23.

Punt, A.E., Haddon, M., Little, L.R., and Tuck, G.N. 2016a. Can a spatially structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. Fisheries Research. 175: 10–23. doi:10.1016/j.fishres.2015.11.008.

Punt, A.E., Haddon, M., Little, L.R., and Tuck, G.N. 2016b. The effect of marine closures on a feedback control management strategy used in a spatially aggregated stock assessment: a case study based on pink ling in Australia. Canadian Journal of Fisheries and Aquatic Sciences. 73: 1-14. dx.doi.org/10.1139/cjfas-2016-0017.

Punt, A.E., Walker, T.I., Taylor, B.L., Pribac, F. (2000). Standardization of catch and effort data in a spatially-structured shark fishery. *Fish. Res*. 45: 129-145.

Quinn, T. J., and Deriso, R. B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Rayns, N. 2007. The Australian government's harvest strategy policy. ICES J. Mar. Sci. 64: 596–598.

Richards, L.J., Schnute, J.T. (1986). An experimental and statistical approach to the question: is CPUE an index of abundance? *Can. J. Fish. Aquat. Sci*. 43: 1214-1227.

Roberts, C. M., Hawkins, J. P., and Gelly, F. R. 2005. The role of marine reserves in achieving sustainable fisheries. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 360: 123–132.

Sainsbury, K.J., Punt, A.E. and Smith, A.D.M. 2000. Design of Operational Management Strategies for Achieving Fishery Ecosystem Objectives, ICES Journal of Marine Science, 57: 731–741.

SEMAC, 2012. Minutes of the South East Management Advisory Committee (SEMAC) 08. 30-31 January, 2012. Australian Fisheries Management Authority, Canberra.

Smedbol, R.K., McPherson, A., Hansen, M.M., and Kenchington, E. 2002. Myths and moderation in marine 'metapopulations'? Fish and Fisheries 3, 20-35.

Smith, A.D.M. 1993. Risks of over- and under-fishing new resources. In Risk Evaluation and Biological Reference Points for Fisheries Management. Edited by. S.J. Smith, J.J. Hunt, and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences 120. pp. 261–267.

Smith, A.D.M. 1994. Management strategy evaluation – The light on the hill. In Population Dynamics for Fisheries Management. Edited by D.A. Hancock. Australian Society for Fish Biology Workshop Proceedings, Perth 24-25 August 1993. Australian Society for Fish Biology, Perth. pp 249-253

Smith, A.D.M., Smith, D.C., Tuck, G.N., Klaer, N., Punt, A.E., Knuckey, I., Prince, J., Morison, A., Kloser, R., Haddon, M., Wayte, S., Day, J., Fay, G., Pribac, F., Fuller, M., Taylor, B., Little, L.R., 2008. Experience in implementing harvest strategies in Australia's south-eastern fisheries. Fish. Res. 94(3), 373–379.

Sporcic, M. (2007). *Fisheries CPUE–Abundance Relationships: Examination Using A Resource Fisher Integrated Model*. Doctor of Philosophy Thesis, University of Tasmania, 256 pp.

Sporcic, M. (2016). Catch rate standardizations for selected SESSF species (data to 2014). pp 91 – 338 *in* Tuck, G.N. (ed.) 2016. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery: 2015. Part 2*. Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere Flagship, Hobart. 493 pp.

Sporcic, M. and Haddon, M. 2015. Catch rate standardizations for selected SESSF species (data to 2013) pp 48 – 276 *in* Tuck, G.N. (ed.) *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2014. Part 2.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere Flagship, Hobart. 432 p.

Sporcic, M., Smith, A.D.M. (2009). Framework for a space-time resource and fisher model to assess CPUE-abundance relationships. 18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009; http://mssanz.org.au/modsim09; pp 2150-2155.

Stewart, I.J. 2009. Status of the U.S. canary rockfish resource in 2009 (Update of 2007 assessment model). Pacific Fishery Management Council, 7700 Ambassador Place NE, Portland, OR 970220. http://www.pcouncil.org/wp-content/uploads/2009_canary_updated_assessment_SAFE_version.pdf (last accessed 3 January 2016).

Stobutzki, I., Vieira, S., Ward, P., and Noriega, R. 2011. Southern and Eastern Scalefish and Shark Fishery overview. In Fishery Status Reports 2010: Status of fish stocks and fisheries managed by the Australian Government Edited by J. Woodhams, I. Stobutzzki, S. Vieier, R. Curtotti and G.A., Begg. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra. Pp 114-124.

Szuwalski, C.S., and Punt, A.E. 2015. Can an aggregate assessment reflect the dynamics of a spatially structured stock? Snow crab in the eastern Bering Sea as a case study. Fish. Res. 164: 135-142.

Thorson, J. T., Stewart, I. J., and Punt, A. E. 2012. Development and application of an agent based model to evaluate methods for estimating relative abundance indices for shoaling fish such as Pacific rockfish (Sebastes spp.). ICES Journal of Marine Science 69:635–647.

Thorson, J.T., Wetzel, C., 2015. The status of canary rockfish (Sebastes pinniger) in the California Current in 2015. Pacific Fishery Management Council. 7700 NE Ambassador Place, Suite 101, Portland, OR 97200-1384, USA. 692pp

Tuck, G.N. 2009 (Ed). Stock assessment and management strategy evaluation for sub-Antarctic fisheries: 2007 - 2009. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 112p.

Tuck, G.N. 2014. Stock assessment of blue grenadier Macroronus novaezelandiae based on data up to 2012. In Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2013. Part 1. Edited by G.N. Tuck. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research. pp 61 – 115

Tuck, G.N., Phillips, R., Small, C., Thomson, R., Klaer, N., Wanless, R.M., et al. 2011. An assessment of seabird-fishery interactions in the Atlantic Ocean. ICES Journal of Marine Science. 68(8): 1628-1637.

Waterhouse, L., Sampson, D.B., Maunder, M., Semmens, B.X., 2014. Using areas-as-fleets selectivity to model spatial fishing: Asymptotic curves are unlikely under equilibrium conditions. Fish. Res. 158, 15–25.

Wayte, S.E. (Ed) Evaluation of new harvest strategies for SESSF species. 2009. CSIRO Marine and Atmospheric Research and Australian Fisheries Management Authority. Report 2006/815. June 2009. 137p.

Wayte S., and Klaer N., 2010. An effective harvest strategy using improved catch-curves. Fish. Res. 106, 310–320.

Whitten, A.E., Punt. A.E., 2014. Stock Assessment of pink ling (Genypterus blacodes) using data up to 2012. p. 116-142. In: G.N. Tuck (Ed). Stock Assessment for the Southern and Eastern Scalefish and Sharkfish 2013, Part 1. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Williams, A.J., Little, L.R., Punt, A.E., Mapstone, B.D., Davies, C.R. and Heupel, M.R. 2010. Exploring movement patterns of an exploited coral reef fish when tagging data are limited. Marine Ecology Progress Series 405, 87-99.

Ying, Y., Chen, Y., Lin, L., Gao, T., 2011. Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. 68, 2101–2120.

Zar, J.H. (1984). *Biostatistical Analysis*. Prentice-Hall Inc. New Jersy. 718 pp.

Zhou, S. 2010. Sustainability assessment of fish species potentially impacted in the Northern Prawn Fishery : 2007-2009: 2007–2009.

Zhou, S., Daley, R., Fuller, M., Bulman, C., Hobday, A., Ryan, P., Courtney, T., et al. 2013. ERA extension to assess cumulative effects of fishing on species.

Zhou, S., Fuller, M., and Smith, T. 2009a. Rapid quantitative risk assessment for fish species in seven Commonwealth fisheries additional seven Commonwealth fisheries.

Zhou, S., and Griffiths, S. P. 2008a. Sustainability Assessment for Fishing Effects (SAFE): A new quantitative ecological risk assessment method and its application to elasmobranch bycatch in an Australian trawl fishery. Fisheries Research, 91: 56–68. http://linkinghub.elsevier.com/retrieve/pii/S0165783607003141 (Accessed 3 November 2013).

Zhou, S., and Griffiths, S. P. 2008b. Sustainability Assessment for Fishing Effects (SAFE): A new quantitative ecological risk assessment method and its application to elasmobranch bycatch in an Australian trawl fishery. Fisheries Research, 91: 56–68. http://linkinghub.elsevier.com/retrieve/pii/S0165783607003141 (Accessed 23 September 2014).

Zhou, S., Griffiths, S. P., and Miller, M. 2009b. Sustainability assessment for fishing effects (SAFE) on highly diverse and data-limited fish bycatch in a tropical prawn trawl fishery. Marine and Freshwater Research, 60: 563. http://www.publish.csiro.au/?paper=MF08207.

Zhou, S., Klaer, N. L., Daley, R. M., Zhu, Z., Fuller, M., and Smith, a. D. M. 2014. Modelling multiple fishing gear efficiencies and abundance for aggregated populations using fishery or survey data. ICES Journal of Marine Science: 1–40. http://icesjms.oxfordjournals.org/cgi/doi/10.1093/icesjms/fsu068.

Zhou, S., Smith, A. D. M., and Fuller, M. 2011. Quantitative ecological risk assessment for fishing effects on diverse data-poor non-target species in a multi-sector and multi-gear fishery. Fisheries Research, 112: 168–178. http://linkinghub.elsevier.com/retrieve/pii/S0165783610002481 (Accessed 18 September 2013).

Zhou, S., Smith, T., and Fuller, M. 2007. Rapid quantitative risk assessment for fish species in selected Commonwealth fisheries Commonwealth fisheries.

Appendices

1 Incorporating marine spatial closures in Productivity Susceptibility Analysis (PSA)

Daley, R.[1], Fuller, M.[2] and Smith, A.D.M.[2]

[1]*Horizon Consultancy, ABN 25259581944 Hobart, Tasmania, Australia*

[2]*CSIRO Oceans & Atmosphere, Hobart, Tasmania, Australia*

1.1 Abstract

The Productivity Susceptibility Analysis (PSA) method is in wide use in fisheries around the world for application to various species and habitats, not just fish. We revised the PSA for the SESSF otter trawl fishery to, firstly, assess the effectiveness of area closures introduced to mitigate the effects of fishing between 2003 and 2013 and, secondly to explore refinements to the PSA method. We focused on waters deeper than 700 m where substantial areas have been closed to protect Orange Roughy. Results obtained using the original PSA methods found that there was a substantial (36–50%) reduction in overlap between fishing area and the distribution of seven non-target species considered in the study. This reduction is attributed to spatial management. Only one out of seven of the overall species risk scores decreased in response to reduced overlap using the original PSA methods. The Method was then refined to use a continuous scoring approach (percent overlap) rather than categorical scoring approach (high, medium, low). This refinement was coupled to revised species maps based on new depth contours obtained from latest seafloor mapping to revise species boundaries. These coupled refinements were used to develop a beta version of the PSA that is more sensitive to reduced overlap. Only one of seven species remained at high risk when this method was applied to the 2013 data. The revised PSA method is an improvement that can show the effectiveness of management actions.

1.2 Introduction

This Appendix focuses on how PSA methods can be used to determine how well AFMA fishery closures mitigate the risks of capture fishing to byproduct and bycatch species. The PSA method is a level 2 (semi-quantitative) method integral to the Ecological Risk Assessment for the Effects of Fishing approach (Hobday *et al*, 2011), and can be applied to all species and habitats, not just fish.

The Southern and Eastern Scalefish and Shark Fishery (SESSF) otter trawl fishery was used as an example because it has undergone substantial changes during 2003 to 2013, including the introduction of many fishery closures (and some environmental closures such as MPAs) on the continental shelf, upper slope and mid-slope, as well as substantial reductions in effort over this period (Knuckey and Upston, 2013).

Here we incorporate an up-to-date effort map (see Appendix 2) and provide a Beta version PSA option with modified scoring for availability, to take better account of spatial management. We also improve the graphical representation of the PSA results. Finally, we briefly discuss the utility of the conventional and Beta PSA options.

1.3 Methods

The PSA method scores the risk posed by fishing for a given species along two axes using productivity and susceptibility attributes. Productivity attributes are intrinsic to species and biologically fixed, for example, life span, number of eggs etc. Low productivity species, including whales and birds as well as long lived fish and sharks, tend to have higher overall risk; high productivity species such as prawns and many short-lived teleosts tend to have lower overall risk scores. Because productivity attributes are fixed, they don't change with changes in management.

Susceptibility (exposure) varies according to four attributes: 1. Availability (measuring the extent of spatial overlap between fishing effort and species range); 2. Encounterability (position in the water column, relating to the likelihood that fishing gear will encounter a species); 3. Selectivity (the likelihood of capture given encounter); and 4. Post Capture Mortality (related to the probability of survival given capture and particularly important for species that are discarded). Any of these four attributes can be influenced by management interventions and regulations. Fishery closures mainly affect Availability. Availability risk for a given species will be reduced by spatial management if the restrictions applied to a given "closure" encompass part of a species range.

1.3.1 Base PSA

In current PSA analysis, availability is measured as the percentage of the species range that overlaps with fishing effort. For the Base PSA, these availability values are then given categorical risk scores, according to scoring thresholds: <10% overlap = low risk, 10–20% overlap = medium risk; >20% = high risk.

In this report, we calculated overlap values for seven SESSF species for 2013. The 2003 overlap values were reduced by the same proportion that the SAFE overlaps decreased (see Appendix 2). The SAFE and PSA methods of calculating effort distribution are very similar. A copy of the 2003 PSA spreadsheet, used to develop the Ecological Risk Assessment reports for the fishery at the time, was then modified to include the revised overlap values. Revised availability overlap scores were then generated automatically in the spreadsheet according to the thresholds above. The corresponding changes to combined susceptibility scores and overall risk scores were then generated automatically. It is important to note that the other susceptibility attributes (encounterability, selectivity, post capture mortality) and the productivity attributes were not changed.

1.3.2 Beta PSA

This project developed more precise mapping methods and applied them to the case study species. Species distribution maps were based on Atlas of Living Australia maps. The inner and outer (depth based) boundaries were refined using the latest available bathymetry data. This improvement is particularly important for species that occur mainly in waters deeper than 700m, outside deep-water closures. To calculate availability, fishing effort was overlaid on the species distributions using a 1km^2 grid. Further precision was added to the analysis by changing the categorical scoring used for availability in previous PSA methods to a continuous value n between 0–3. In the refined scoring the percentage overlap is scored as the overlap fraction (effort/species) x 3. The aim of this refinement is to provide higher resolution for smaller changes in spatial closures and to make the management changes easier to visualize in the PSA plots. Initially these maps were used to calculate overlap between species and effort, but later the analysis was automated in CSIRO's ERA database.

1.3.3 Case Study Species

We selected a set of species that were likely to have been influenced by the introduction of closures. As the deep-water (> 700m) closure is the largest closure to be introduced since the last PSA analysis of the SESSF otter trawl fishery, we included mainly species that occur deeper than 700 m. These include two deep-sea sharks*:* Brier Shark (*Deania calcea*) and Plunket's Shark (*Centroscymnus plunketi*), with a shelf-dwelling species of shark for contrast: Broadnose Sevengill Shark (*Notorhynchus cepedianus*). Teleost fishes are represented by four deep-sea oreo species: Ox-eye Oreo (*Oreosoma atlanticum*), Warty Oreo (*Allocyttus verrucosus*), Black Oreo (*Allocyttus niger*) and Spiky Oreo (*Neocyttus rhomboidalis*).

1.4 Results

1.4.1 Base PSA Results

For each species, there was a substantial reduction in spatial overlap between species distribution and effort distribution, ranging from 36.4% for the Broadnose Sevengill Shark (the shelf species) to 50% for the Spikey Oreo (Table 1.1). The only categorical changes to risk scores were reduced availability and susceptibility scores for two species: For the Broadnose Sevengill Shark, availability was reduced from 2 (med) to 1 (low); availability for Plunket's Shark was reduced from 3 (high) to 2 (medium) (Table 1.2, Table 1.3).

The Broadnose Sevengill Shark was the only species where the overall risk category score changes, falling from high to medium (Figure 1.1). For Plunket's Shark, the overall PSA risk score does not change, even though the reduction in susceptibility is similar to the Broadnose Sevengill Shark. Only Broadnose Sevengill Shark crosses an overall scoring threshold (Figure 1.1).

Table 1.1: Reduction in overlap scores for case study species

Species	SAFE Overlaps		
	2003	2013	Reduction
B.nose sevengill shark	0.022	0.014	36.40%
Brier Shark	0.063	0.034	46.00%
Plunket's shark	0.069	0.038	44.90%
Oxeye Oreo	0.085	0.047	44.70%
Warty Oreo	0.018	0.01	44.40%
Black Oreo	0.036	0.019	47.20%
Spiky Oreo	0.052	0.026	50.00%

Table 1.2: Availability scores and overall risk categories for case study species in 2003 PSA.

Species	Overlap	Availability	Susceptibility	Productivity	Overall score	Overall category
B.nose sevengill shark	19.70%	2	2.33	2.29	3.27	High
Brier Shark	82.30%	3	3.00	2.71	4.05	High
Plunket's shark	61.20%	3	3.00	2.71	4.05	High
Oxeye Oreo	81.40%	3	3.00	2.00	3.61	High
Warty Oreo	74.00%	3	2.33	2.00	3.07	Med
Black Oreo	76.20%	3	3.00	1.86	3.53	High
Spiky Oreo	81.10%	3	3.00	2.00	3.61	High

Table 1.3: Availability scores and overall risk categories for case study species in 2013 PSA. * denotes change in availability score

Species	Overlap	Availability	Susceptibility	Productivity	Overall score	Overall category
B.nose sevengill shark	7%	1*	1.67	2.29	2.83	Med
Brier Shark	38%	3	3.00	2.71	4.05	High
Plunket's shark	27%	2*	2.33	2.71	3.58	High
Oxeye Oreo	36%	3	3.00	2.00	3.61	High
Warty Oreo	33%	3	2.33	2.00	3.07	Med
Black Oreo	36%	3	3.00	1.86	3.53	High
Spiky Oreo	41%	3	3.00	2.00	3.61	High

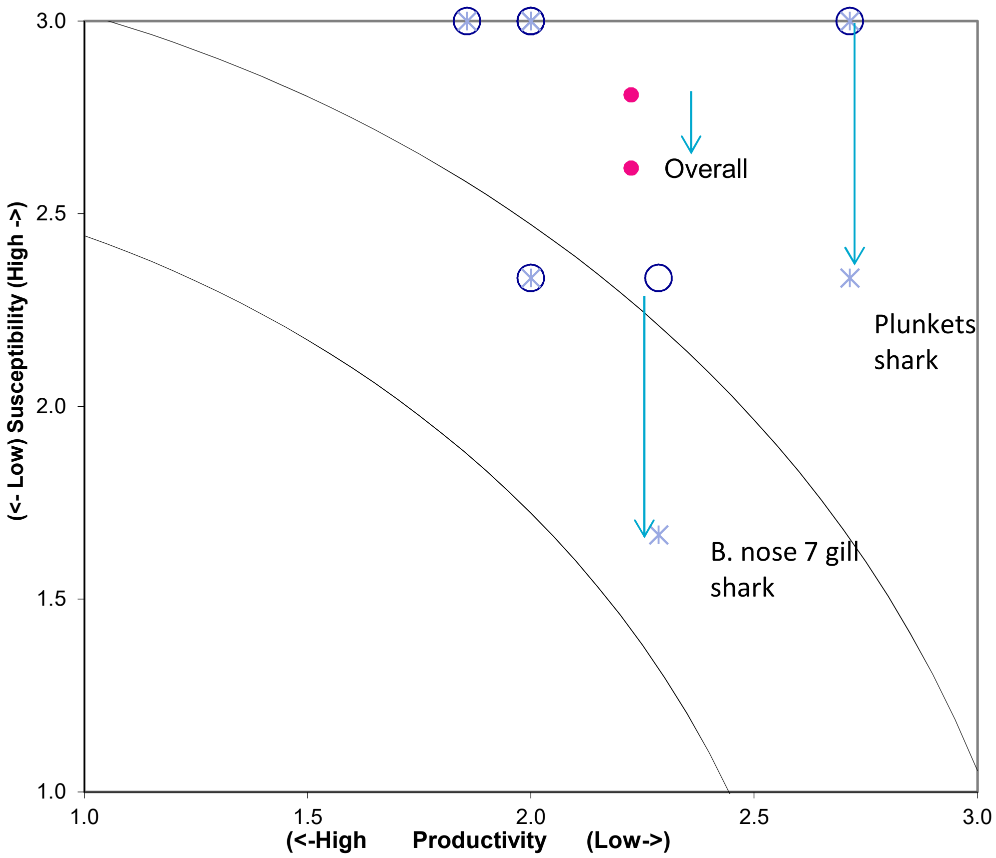


Figure 1.1: Changes in Base PSA results from 2003 to 2013 due to reduced availability scores. O = original values; * = revised values. Curved lines indicate scoring thresholds between risk categories (high risk top right, low risk bottom left, medium risk in between the curved lines).

1.4.2 Beta PSA Results

Using the continuous availability score, the reduction in overlap translates to reduced availability, and resulting susceptibility scores for all case study species (Table 1.4). This results in a reduction in the overall PSA risk for all case study species that is easy to visualize on the PSA graph (Figure 1.2). Using this method only 4 of the 7 species would be high risk using the 2003 data (compared to 6 of the 7 using the Base method), which falls to only one species using the 2013 data (compared to 5 species using the base method). These differences are not directly comparable between methods because PSA uses categorical scoring for availability, whereas Beta PSA uses a continuous measurement (see Discussion).

Table 1.4: Changes in availability and susceptibility scores between 2003 and 2013 using Beta PSA method.

Species	Availability		Susceptibility	
	2003	2013	2003	2013
Broadnose Sevengill	0.59	0.22	1.39	1.14
Brier Shark	2.47	1.14	2.65	1.76
Plunket's shark	1.84	0.82	2.22	1.55
Oxeye Oreo	2.44	1.09	2.63	1.73
Warty Oreo	2.22	0.99	1.99	1.44
Black Oreo	2.29	1.08	2.52	1.72
Spiky Oreo	2.43	1.22	2.62	1.81

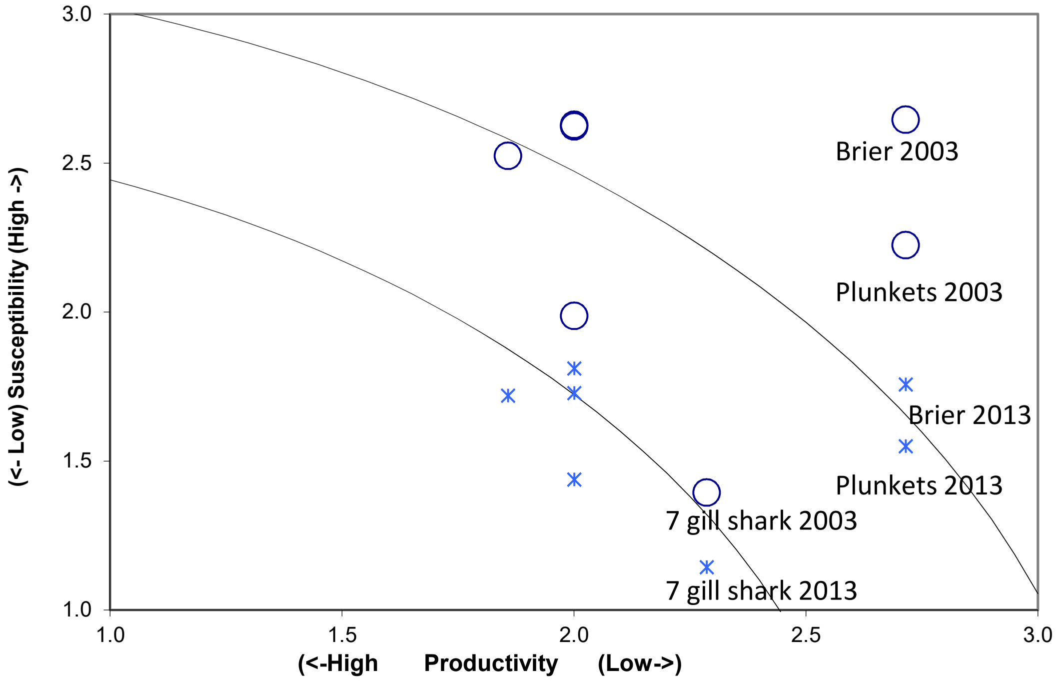


Figure 1.2: Changes in Beta PSA results from 2003 to 2013 due to reduced availability

1.5 Discussion

1.5.1 PSA Methods

The Beta PSA method provides better resolution between species and it is easier to visualise the effects due to management changes. On face value, the Beta PSA results in this report give the appearance that the method provides lower overall risk scores than the original method. This is in fact not necessarily the case. At present the Beta PSA availability scoring is continuous and not weighted for any level of precaution. The availability scores are implicitly cut into thirds to classify as low, medium and high availability. In the original method, categorical thresholds were applied based on the assumption that some species at least can aggregate for breeding and feeding, and the proportion of the overlapped range would need to be less than 10% to pose a low risk to a given species, or 10–20% for a medium risk. These thresholds contain scientific uncertainty and are influenced by the level of precaution a manager might apply.

One limitation of the Beta PSA method is that it required detailed maps and mapping data for each species. Although existing maps for fish and sharks are generally accurate, they had imprecise boundaries and did not show where species aggregate. The improved maps used here address the boundary issue. Calculating availability for aggregating and migratory species (including protected birds and mammals) is an additional problem that will require more observational data and development of more sophisticated mapping methods.

In conclusion, the Beta PSA method offers a better species resolution and easier visualisation than the Base PSA method. Potentially it can be applied to birds, mammals, reptiles and habitats as detailed mapping data become available. However further work is needed to resolve the level of risk associated with different levels of overlap, particularly for aggregating species.

2 Incorporating marine spatial closures in the application of SAFE method

Zhou, S.[1], Fuller, M.[2], Daley, R.[3], Smith, A.D.M.[2] and Tuck, G.N.[2]

[1]*CSIRO Oceans & Atmosphere, Brisbane, Queensland, Australia*

[2]*CSIRO Oceans & Atmosphere, Hobart, Tasmania, Australia*

[3]*Horizon Consultancy, ABN 25259581944 Hobart, Tasmania, Australia*

2.1 Abstract

The Australian Fisheries Management Authority (AFMA) has adopted the quantitative method Sustainability Assessment for Fishing Effect (SAFE) as a preferred method within the Ecological Risk Assessment for the Effect of Fishing (ERAEF) framework. However, the performance of SAFE in the context of marine closures has not been examined. This Appendix investigates the performance of the method with regard to spatial closures. We describe how spatial information is used in the SAFE method and apply the method to seven species in the Southern and Eastern Scalefish and Shark Fishery (SESSF) to illustrate how marine closures might mitigate the risk of fishing posed to the sustainability of by-product and bycatch species. The results show that fishing impacts on non-target species have decreased substantially since 2003. This decrease was due to a large reduction in total fishing effort rather than spatial closures. Since the SAFE method assumes a homogeneous or random distribution of fish between fished and unfished areas and uses actual fishing effort data, the current method remains valid when marine closures are imposed.

2.2 Introduction

Marine spatial closures have become a common tool in fisheries management and biodiversity conservation. Although debates on whether closures can effectively protect biological resources and increase fisheries yield are ongoing (Fletcher *et al.*, 2015; Roberts *et al.*, 2005), there is no doubt that closures will affect the distribution of fishing effort and consequently impacted species. AFMA has adopted a risk-based Ecological Risk Assessment for the Effect of Fishing (ERAEF) framework for managing by-product and bycatch species in Commonwealth fisheries (Hobday *et al.*, 2011). The quantitative method Sustainability Assessment for Fishing Effect (SAFE) has been adopted within the ERAEF framework and applied to most Commonwealth fisheries. During its initial development, SAFE did not specifically take spatial closures within a larger fishing area into consideration (Zhou and Griffiths, 2008a; Zhou *et al.*, 2007, 2009b, 2011). The performance of SAFE in this context is unknown and in this Appendix we examine the method with regard to spatial closures. We describe the SAFE method, focusing on its treatment of spatial information. We apply it to selected species in the Southern and Eastern Scalefish and Shark Fishery (SESSF) to illustrate how marine spatial closures might mitigate the risk of fishing posed to the sustainability of by-product and bycatch species.

2.3 Methods

The SAFE method requires that two major components be defined: stock indicators and reference points. The concept is essentially the same as traditional fishery stock assessment and management (Quinn and

Deriso, 1999). SAFE focuses on a single indictor — fishing mortality rate. Because of a lack of basic data for most non-target species, including time series of catch, abundance index, age composition, etc., SAFE derives estimates of fishing mortality using alternative methods. Currently, two versions of SAFE have been developed and applied to fisheries with different data: the base SAFE and enhanced SAFE. These methods are defined below.

2.3.1 Base SAFE

The Base SAFE method is based on the spatial overlap between species distribution and fishing effort distribution over the jurisdictional area of the fishery (Zhou *et al.*, 2007, 2009a, 2011). This estimate of overlap is fine-tuned by considering habitat and behaviour-dependent encounterability and fishing gear and size-dependent gear selectivity. Therefore, a spatial configuration of fishing effort that includes closures and protected areas is essential in deriving fishing impact and overlap with species range as a proxy for fishing mortality, the indicator of interest (Figure 2.1). For the second component, SAFE derives biological reference points (BRPs) from life-history parameters that are widely available for many species, rather than from time-series of fisheries data. The BRPs have the same meaning as in traditional fishery management, i.e., F_{msy}, F_{limit}, and F_{crash} (Zhou *et al.*, 2011). As these reference points do not depend on spatial information (although less informative reference points based on spatial overlap could also be used), this Appendix will focus on the indicator component of the SAFE method and ignore the reference point component.

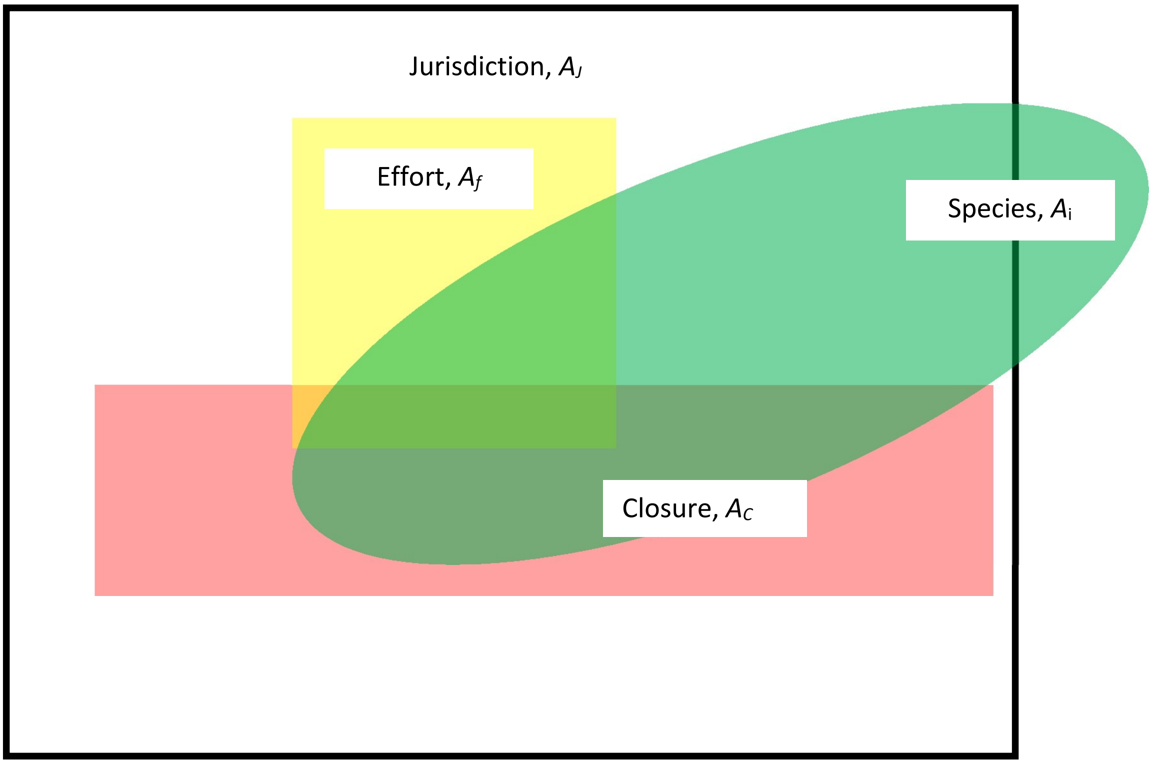


Figure 2.1: Diagram for spatial distribution of species, fishing effort, fishery, and marine closure. Fishing impact is derived from the overlap between effort A_f and species distribution A_i within the jurisdiction A_J.

For fishing gears, such as trawls, that actively sweep the seabed or through the water column, the basic equation for estimating annual fishing mortality can be expressed as

$$
\begin{aligned}
F_i &= \frac{C_i}{\overline{N}_i} \\
&= \frac{A_{i,f|J}}{A_{i,J}} Q_i (1-E_i)(1-S_i) \\
&= \frac{\sum_t L_{i,f|J,t} W}{A_{i,J}} Q_i (1-E_i)(1-S_i)
\end{aligned} \tag{1}
$$

where C_i is the catch of species *i* that is dead after discarding (or retention for by-product species), $\overline{N_i}$ the mean population size over a one year period, $A_{i,J}$ the total area occupied by species *i* within the fishery jurisdiction *J*, $A_{i,f|J}$ the area occupied by species *i* intercepting with the fishing effort distribution *f* within the fishery jurisdiction *J*, Q_i gear-efficiency (probability of a fish entering the net along a track), E_i escapement rate after the fish enters the trawl gear, S_i discard survival rate, *W* the width of trawl wing spread, and $L_{i,f|J,t}$ the trawling distance that intercepts with the species distribution range within fishery jurisdiction *J* at time *t*. This equation means that fishing mortality is the fraction of the species' distribution area swept by fishing gear corrected by catch efficiency and post-capture survival rate if the fish is returned to the sea. For many bycatch species, there may be few data to allow the estimation of fish density. Hence, it is often assumed that fish density is homogeneous or random across the entire distribution range within the jurisdictional area being assessed. On the other hand, a minimum assumption behind this equation is that fish density does not differ between the fished and unfished areas within its distribution range and within the jurisdiction. Clearly, this assumption may result in a serious bias for target species because of fishermen's targeting behaviour, but may be less problematic for non-target species. In this base SAFE, three levels of *Q* are consistent with risk thresholds used in the companion ERAEF PSA tool: low 0.33, medium 0.67, and high 1.0.

The two spatial variables are $A_{i,j}$ and $A_{i,f|J}$, where

$$A_{i,J} = A_i \cap A_J \tag{2}$$

i.e., spatial overlap between species distribution range A_i and fishery jurisdiction A_j, and

$$A_{i,f|J} = A_i \cap A_{f|J} \tag{3}$$

i.e., fishing effort within fishery jurisdiction ($A_{f|J}$) intersecting species distribution (A_i).

Marine spatial closures within the jurisdictional area reduce the available area to fishing. Assuming closures do not affect the extent of fish distribution and fish do not move between closed and open areas, then closures will result in a decreased fishing impact on stocks that are fully or partially protected by closures if the total fishing effort does not increase. The effect of closures will be reflected in a reduced fishing mortality rate for non-target species assessed by the SAFE method. This reduction is through the change in variable $A_{i,f|J}$:

$$A_{i,f|J,O} = A_i \cap A_{f|J} \cap (A_J - A_C) \tag{4}$$

where A_C is the closed area to specific fishing gear, which is often the difference between A_J and open area A_O ($A_O = A_J - A_C$). When SAFE is applied to Commonwealth fisheries, the actual fishing effort from logbooks is used to calculate $A_{i,f|J}$. Effort changes in the open area, either increasing or decreasing, are appropriately accounted for (i.e. the effect of displaced effort is taken into account). Assuming that there is no illegal fishing inside closures, or that effort inside closures is reported in the logbooks, the current method using Eqn (1) remains valid in fisheries where marine closures are imposed. In another words, there is no need to change the existing SAFE method, which has been used for about 17 major Commonwealth fisheries.

2.3.2 Enhanced SAFE

This version of SAFE was initially developed to assess hundreds of fish bycatch species in Australia's Northern Prawn Trawl fishery (Zhou and Griffiths, 2008b; Zhou *et al.*, 2009b), and has been used in that region at the same time as the base SAFE was being used in other regions. The method assumes heterogeneous density across a fishes' distribution range or between fished and unfished area. It also uses species-specific gear efficiency Q_i obtained from field studies, literature, or estimates using survey or fisheries data (Zhou and Griffiths, 2008a; Zhou *et al.*, 2009b, 2013). Without marine closures, fishing mortality rate can be estimated by

$$F_i = \frac{C_i}{\bar{N}_i} = \frac{d_{i,f|J} A_{i,f|J}}{d_{i,f|J} A_{i,f|J} + d_{i,u|J} A_{i,u|J}} Q_i (1 - E_i)(1 - S_i) \tag{5}$$

where d is fish density, and subscript f indicates fished area while u is unfished area. Here fish density is assumed to differ between fished and unfished areas and fish are assumed to remain within these areas (do not move between them), e.g., when the fishing season is short and the species is slow moving. When marine closures are implemented, Eqn (5) remains valid if fish density in the closures can be assumed to be the same as fish density in the open but unfished area. When this assumption is violated, Eqn (5) needs to be modified to partition out the closed area and the open but unfished areas:

$$F_i = \frac{C_i}{\bar{N}_i} = \frac{d_{i,f|J} A_{i,f|J}}{d_{i,f|J} A_{i,f|J} + d_{i,u|O|J} A_{i,u|O|J} + d_{i,C|J} A_{i,C|J}} Q_i (1 - E_i)(1 - S_i) \tag{6}$$

where subscript O indicates open area and C indicates closures. In this equation, sub-populations in closed area, $d_{i,C|J} A_{i,C|J}$ and open but unfished area, $d_{i,u|O|J} A_{i,u|O|J}$ will not be affected by fishing.

The enhanced SAFE thus requires more information than the base SAFE. It is possible to derive heterogeneous density in different locations from historical surveys or from observer data to predict bycatch species density at locations where there are no data. This version of SAFE has been applied to the prawn fishery in the NPF (Zhou and Griffiths, 2008a; Zhou, 2010; Zhou *et al.*, 2009b) and case studies for a few selected species in the SESSF (Zhou *et al.*, 2013). Further, gear efficiency Q for major gear types can be estimated from catch data for some species (Zhou *et al.*, 2013, 2014). The results should presumably be more accurate than assuming Q is constant across species and gear types (i.e., values of 0.33, 0.66, and 1). However, such an analysis has to be carried out species-by-species and requires more time and effort. A model involving varying densities in open and closed areas, species and gear-specific catch efficiency, shot-by-shot fishing effort, as well as species distribution areas inside and outside the closures would be needed to estimate fishing impact for non-target species using this method.

2.3.3 Case study

Since the SESSF has been selected as a case study in this project, we use several species in this fishery to demonstrate how marine closures may affect the SAFE procedure and its results. The base SAFE has previously been applied to the SESSF, so this version is again examined in this Appendix. Seven non-target species are included in the case study: Broadnose Sevengill Shark (*Notorynchus cepedianus)*, Brier Shark (*Deania calcea)*, Plunket's Shark (*Centroscymnus plunketi)*, Ox eyed Oreo (*Oreosoma atlanticum)*, Warty Oreo

(*Allocyttus verrucosus)*, Black Oreo (*Allocyttus niger)*, and Spiky Oreo (*Neocyttus rhomboidalis)*. The first three species are chondrichthyans and the latter four are teleosts. These seven species are the same as those considered in Appendix 1.

Species distribution ranges were obtained from bioregional mapping (IMCRA, 1998; Last *et al.*, 2005). Fishing effort and distribution were from AFMA logbooks. Marine closures have been compiled in several projects and we acquired spatial information from the existing database.

2.4 Results

We carried out base SAFE for seven non-target species using trawl fishing effort distribution respectively for 2003 and 2013. Two species, Brier Shark and Plunket's Shark, had a fishing mortality rate equal to or slightly greater than the reference point F_{msy} in 2003 (Table 2.1 and Figure 2.2). Our new results show that fishing mortality is lower than reference point F_{msy} for all species in 2013. A comparison of these two years reveals that fishing mortality in 2013 was reduced by between 35% and 50% from the 2003 level for the focal species. The average reduction was 44%. To investigate whether this large reduction was a result of marine closures, and how marine closures may have contributed to the difference, we examined the fishing effort pattern and its spatial distribution in relation to spatial closures in the SESSF.

There are about 30 closures where trawl gear has been prohibited since 2004 within SESSF. Overlaying the spatial closures and species distribution maps, we can see that a fraction of each of the species distribution ranges are inside the closures (generally along the edge of the closures, Figure 2.3). These seven species typically reside in a narrow distribution band along the continental slopes or on the continental shelf.

The logbook records show that fishing effort has changed markedly since 2003 (Table 2.2). Total fishing effort gradually increased from the 1980s to a peak in 2001 and then rapidly declined (Figure 2.4). Compared to the base year 2003, effort has reduced by 42% and 75% respectively in the open area and closures (Table 2.2). The overall reduction was 44%. Clearly, much more effort reduction has occurred within the closed areas. However, the fraction of effort within closed areas was not substantial, ranging from 3% to 8% between 1985 and 2014. Hence, the large reduction in total fishing effort since 2003 is the major contributor to the reduced fishing impact on non-target species.

Table 2.1: Comparison of assessments for seven non-target species in 2003 and 2013.

			Reference point F_{msy}				
Class	Science name	Common name	Mean	Min	Max	F_{2003}	F_{2013}
Chondrichthyan	*Notorynchus cepedianus*	Broadnose sevengill shark	0.10	0.06	0.08	0.02	0.01
Chondrichthyan	*Deania calcea*	Brier Shark	0.06	0.02	0.02	0.06	0.03
Chondrichthyan	*Centroscymnus plunketi*	Plunket's shark	0.05	0.02	0.03	0.07	0.04
Teleost	*Oreosoma atlanticum*	Oxeye Oreo	0.25	0.21	0.26	0.02	0.01
Teleost	*Allocyttus verrucosus*	Warty Oreo	0.11	0.08	0.20	0.02	0.01
Teleost	*Allocyttus niger*	Black Oreo	0.13	0.09	0.18	0.05	0.03
Teleost	*Neocyttus rhomboidalis*	Spiky Oreo	0.17	0.13	0.16	0.09	0.05

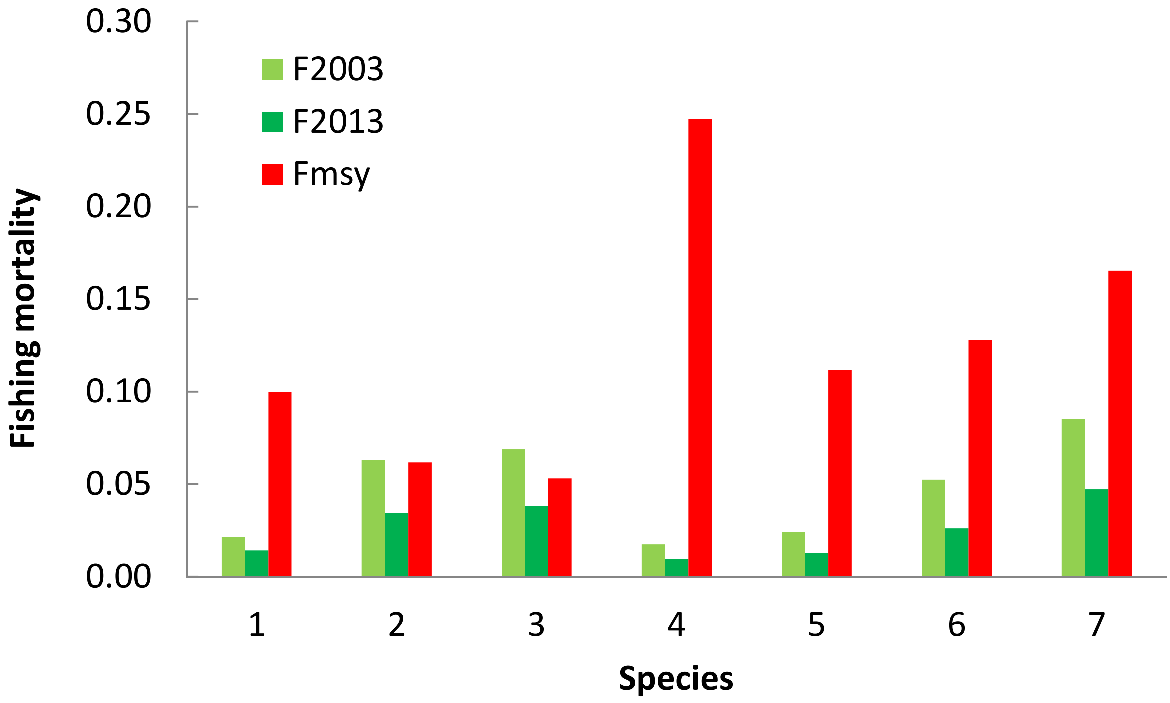


Figure 2.2: Comparison of estimated fishing mortality rates between 2003 and 2013 for seven selected non-target species. Species 1: Broadnose Sevengill Shark; 2: Brier Shark; 3: Plunket's Shark; 4: Ox eyed Oreo; 5: Warty Oreo; 6: Black Oreo; 7: Spiky Oreo.

The effort distribution map shows that fishing activities were wide spread in 2003 (Figure 2.5). The extent of the spatial coverage largely shrank to a narrow band along the continental slope in 2013. The difference between these two years is clear visually.

2.5 Discussion

The SAFE method is a risk-based tool for assessment of fishing effects on non-target species. It is essentially a data-poor stock assessment method. Spatial information is necessary for deriving fishing-induced mortality. Marine closures (and no-take zones, marine reserves etc.) may prohibit particular fishing methods or all fishing gears, and hence reduce the available area to fishing. Such an effect is manifest in fishing effort and its spatial distribution. Because SAFE uses actual fishing effort from fishery data, major modifications to the existing methods are not needed to accommodate spatial closures.

The case study compares the assessment of seven non-target species in the SESSF ten years apart. Estimated fishing mortality has noticeably declined from 2003 to 2013 and this is clearly a result of a reduction in fishing effort. The relative decline of fishing effort in closed areas is more substantial than in the open area, even though the total effort in the closed areas (prior to closure) was lower than in the open areas. For the seven focal species, it appears that the closed areas were not fishing hot spots and had lower fishing intensities even before closures took effect.

Table 2.2: Changes of total fishing effort (trawling hours) in SESSF from 1985 to 2014. The closures are those effective after 2003. Effort reduction is relative to 2003 level.

		Open area		Closures	
Year	Total effort	Effort	Reduction	Effort	Reduction
1985	10414	9910		504	
1986	35815	33249		2566	
1987	36910	34579		2332	
1988	39971	37866		2105	
1989	41379	38438		2941	
1990	39018	37105		1913	
1991	44384	42096		2288	
1992	33768	32116		1653	
1993	43701	42166		1535	
1994	45273	43844		1429	
1995	53725	50106		3619	
1996	58082	54549		3533	
1997	67519	63589		3930	
1998	60742	55759		4983	
1999	65935	60787		5148	
2000	71023	65610		5412	
2001	77746	72994		4752	
2002	73372	69085		4287	
2003	69261	65779	0%	3482	0%
2004	69125	66140	1%	2985	-14%
2005	62004	59719	-9%	2285	-34%
2006	56248	54652	-17%	1596	-54%
2007	43105	42128	-36%	977	-72%
2008	43030	42240	-36%	790	-77%
2009	40861	40001	-39%	860	-75%
2010	43158	42226	-36%	932	-73%
2011	51354	50352	-23%	1002	-71%
2012	42317	41731	-37%	586	-83%
2013	39031	38175	-42%	857	-75%
2014	7715	7492	-89%	223	-94%

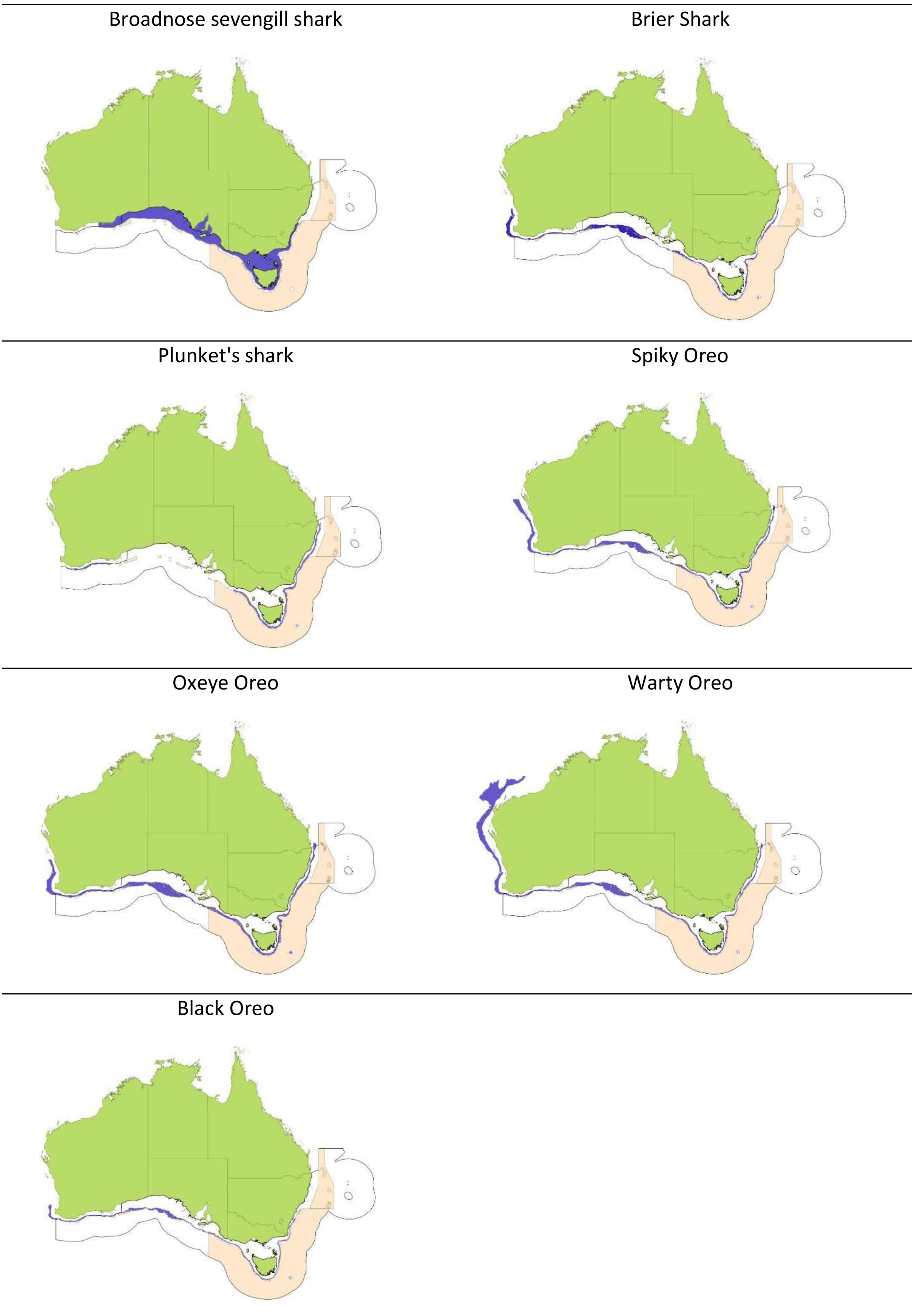


Figure 2.3: Spatial closures (light yellow) to trawling in SESSF in relation to distribution range (blue) of seven fish species. Fish are typically distributed along the edge of the closures.

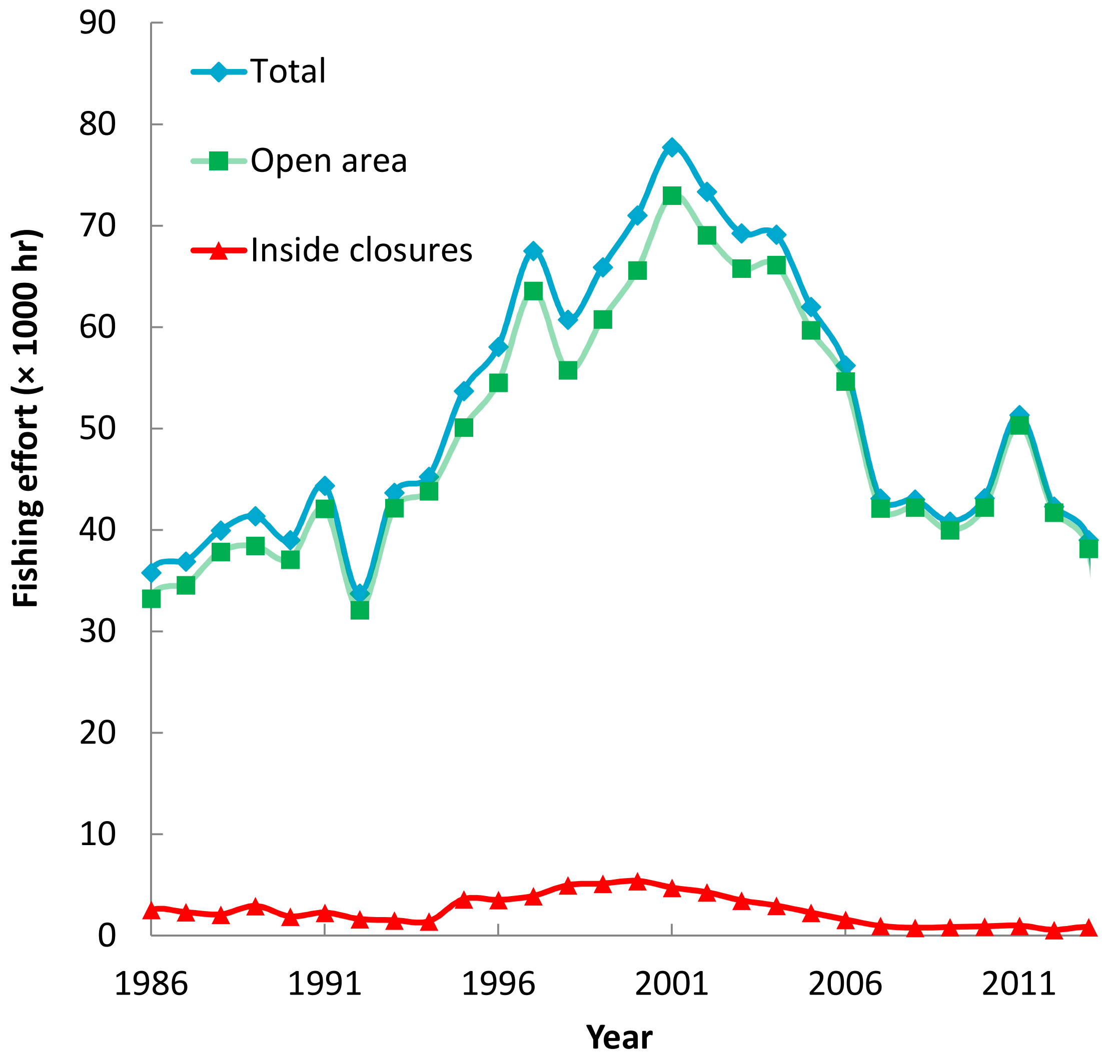


Figure 2.4: Fishing effort (trawling hours) in SESSF from 1986 to 2013. The closures are areas that were closed to trawling since 2004.

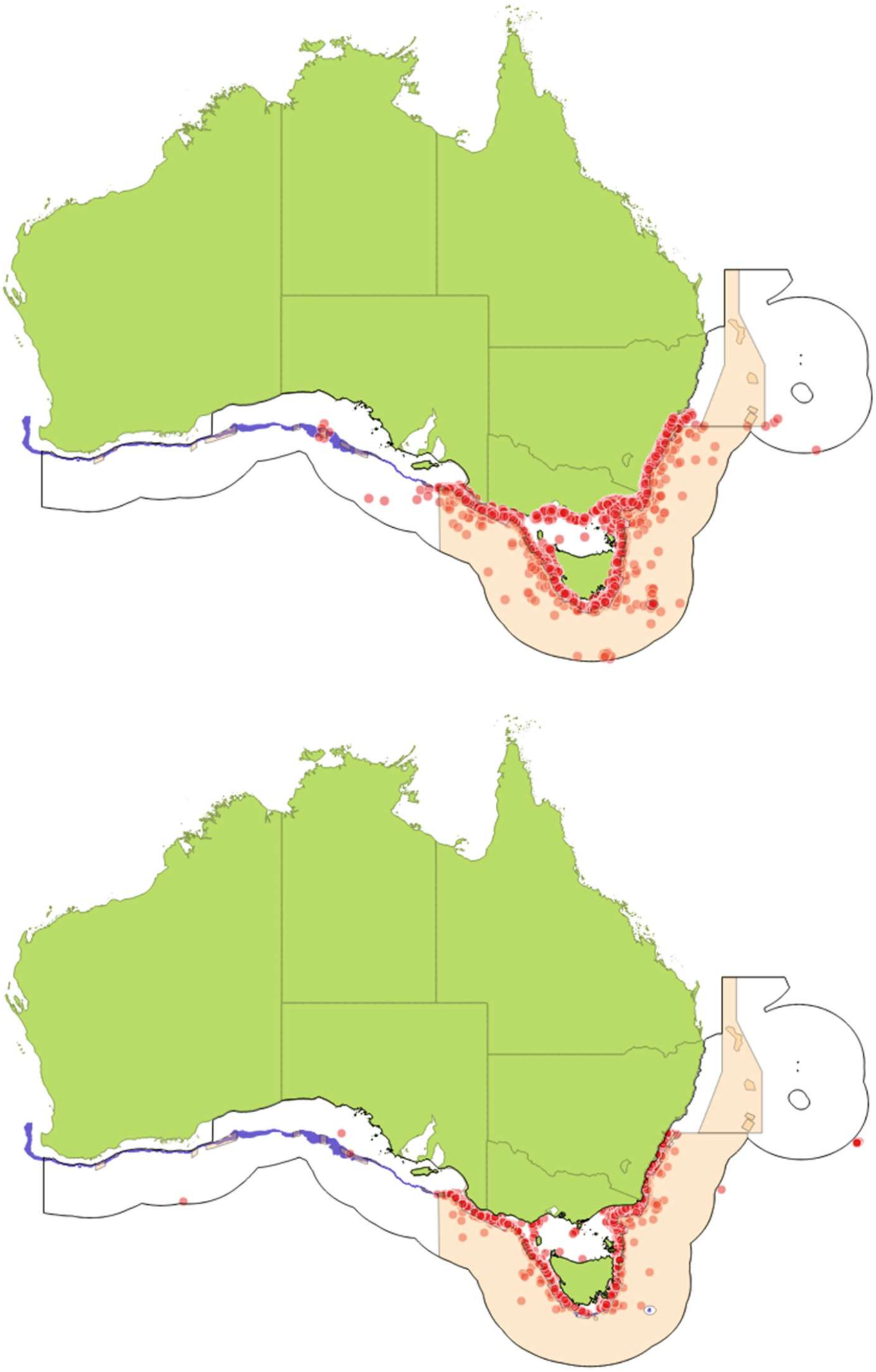


Figure 2.5: Trawling locations in SESSF. Upper panel: 2003; lower panel: 2013. Areas in light pink are closures effective since 2004. Note the clusters of red dots caused by dense effort along the coast and slope.

3 Exploring the effect of sampling, protogyny, and larval advection on stock estimates subject to no-take closures in a spatially complex coral reef line fishery on the Great Barrier Reef, Australia

Little, R.L.[1], Punt, A.E.[1,2], Tuck, G.N.[1] and Mapstone, B.D.[1]

[1]*CSIRO Oceans & Atmosphere, Hobart, Tasmania, Australia*

[2]*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, USA*

This has been published as L. Richard Little, André E. Punt, Geoffrey N. Tuck, Mapstone, B.D. 2017. Exploring the effect of sampling, protogyny, and larval advection on stock estimates subject to no-take closures in a spatially complex coral reef line fishery on the Great Barrier Reef, Australia. Canadian Journal of Fisheries and Aquatic Sciences 74, 1950-1959.

3.1 Abstract

Simulation is used to evaluate the ability of a two-region age-structured assessment model to provide accurate and precise estimates of stock status, i.e. the ratio of female spawning biomass to unfished female spawning biomass, for coral trout, *Plectropomus leopardus*, on the Great Barrier Reef (GBR), Australia. The model used to generate the simulated data used by the assessment model is a spatially complex age- and sex-structured population dynamics model that captures the protogynous nature of coral trout. Stock status is under-estimated (negatively biased), with the extent of negative bias related to mis-specification of the breeding strategy of the target fish stock, the impact of the amount of larval connectivity among reefs, including reefs in fished areas and closures, as well as exploitation rates. The estimates of stock status were less negatively biased when fishery-independent index and age- length-composition data were available from closed areas. The results will inform the development of management strategies for coral trout in the GBR, and highlight the importance of basing evaluations of estimation and management performance on operating models that capture ecologically-important processes such as metapopulation dynamics and protogynous life history.

Keywords: coral trout, fishery-independent sampling, simulation, closed areas, two-region assessment model

3.2 Introduction

The effectiveness of marine spatial closures as a conservation management tool is well documented (e.g., Lauck et al. 1998; Mangel 2000; Halpern, 2003; McCook et al. 2010). Many claims have been made that closing parts of the range of a fish stock would benefit the part that remains open to fishing, through adult ‘spill-over’ (Abesamis et al. 2006) or larval ‘subsidy’ (Hilborn et al. 2004). However, the conditions for such benefits can be quite specific (Armstrong 2007). The reason for this is the porousness or viscosity of the stock between the areas open and closed to fishing. In fact, spatial closures are often implemented for management purposes with the understanding that there is at least some impediment to movement in the species being managed (Field et al. 2006).

This impediment or "viscosity" in movement may present a problem for assessing a stock that is partially found in a marine closure, because most assessments assume the entire stock is subject to the same fishing mortality rate (Punt et al. 2015), and the consequences of violating this assumption can be biased or highly imprecise estimates of biomass. This is especially true if the assessment relies strongly on fishery-dependent data (Field et al. 2006).

In general, assessment of stocks that are partially found in spatial closures can be addressed relatively easily when the species is sedentary, or there are very high levels of mixing. For the former case, the assessment relates only to the biomass in the area open to fishing, whereas in the latter case, it relates to the biomass in areas that are open and closed combined. The difficulty comes when a stock has intermediate amounts of mixing, or mixing that is density-dependent, such as for populations structured as a metapopulation (Smedbol et al. 2002; Kritzer and Sale 2004; Little et al. 2010). Metapopulations characterize many coral reef fish species (Mapstone et al. 2008), where migration occurs during the larval stages, but not for the established or settled adult component. Such dynamics could allow spatial closures to subsidize areas open to fishing (Little et al. 2007).

Metapopulation dynamics are starting to be considered in stock assessment (Cadrin and Secor 2009; Punt et al. 2015), and spatial closures present a particular challenge to estimating stock status (Punt et al. 2015). However, spatial closures offer the opportunity as a potential source of data and information on the unexploited component of the stock (McGilliard et al. 2011; Babcock and MacCall 2011). More accurate estimates of biomass might be obtained by using data from within an area closed to fishing (Pincin and Wilberg 2012).

We show results from a simulation study of fisheries stock assessment performed on a coral reef metapopulation subject to a system of spatial no-take areas. The species of interest is coral trout (*Plectropomus leopardus*), the principal target species of the Queensland (Australia) Coral Reef Fin Fish Fishery (CRFFF), which operates within the Australian (federally) managed Great Barrier Reef Marine Park. The Park is managed for biodiversity conservation using a series no-take zones. We simulated data collection from a structured line survey that samples catch rates, and age- and length-frequencies from a subset of reefs, either inside the no-take closed areas, or outside of them using an operating model developed for the CRFFF. These simulated data were then analyzed to provide estimates of stock status, the ratio of female spawning biomass to unfished female spawning biomass, across all areas. We explored estimation performance under a range of conditions, including various configurations of spatial closures, and levels of larval advection from locally spawned reefs.

3.3 Materials and Methods

3.3.1 Operating model

ELFSim (the Effects of Line Fishing Simulator: Mapstone et al. 2004, 2008; Little et al. 2007, 2009a) was used as the operating model to generate data for input into the assessment model. ELFSim models the population dynamics, harvest, and management of coral trout on more than 3,000 individual reefs on the Great Barrier Reef, connected through larval dispersal. It operates stochastically using monthly time steps, with each simulation consisting of two parts: initialization and projection.

Initialization operates historically, and is used to derive a credible distribution of age-structure and biomass of coral trout on individual reefs by 2012. Starting from an assumed year in which harvest began (1965), individual reef sub-populations are conditioned on historical commercial, charter, and recreational catches; the selectivity of the gear in catching fish of different sizes; the biological characteristics of coral trout; and the physical characteristics of the individual reefs. Reef sub-populations are then projected from 2012 to 2035 by simulating harvest using an agent-based model of fishing vessel dynamics, subject to management constraints including spatial closures, minimum fish size, gear limits, and catch controls implemented with individual transferable quotas (Little et al. 2009b). In the simulations reported here, we implemented three scenarios involving different levels of TAC: (a) 1,300 t, which represented the actual TAC for coral trout that has been in place since 2004, (b) a TAC that is 50% larger than this (1,950t), and (c) a TAC that is 50% smaller than this (650t).

A simulated line fishing survey vessel was implemented using this agent-based model (Little et al. 2016). The survey vessel was chosen randomly at the start of a projection from the commercial vessels characterized in ELFSim, to reflect that scientific surveys for the GBR are based off commercial vessels. That vessel collected data from a sample of 30 reefs in September of each projection year in an attempt to mimic the patterns of previous surveys (Little et al. 2016). These reefs were selected randomly, but stratified regionally (Mapstone et al. 2004), based on a probability proportional to the historical commercial catch rates experienced on the reef prior to 2012, when the projections began, but with a 10% chance of selecting a reef that had never experienced commercial fishing. Two scenarios for this survey were simulated based on obtaining samples from: (i) reefs open to fishing; and (ii) reefs open and closed to fishing (Table 3.1).

Table 3.1: Factors considered in the operating model scenarios. Simulations were undertaken for all combinations of factors.

Factor	Level
Location of line fishing survey	Open area only; open and closed areas
Extent of larval self-seeding	ss = 1; ss = 0.1
TAC	1,300t; 1,950t; 650t
Level of closure	Pre-RAP (21% of the GBR); RAP (42% of the GBR)

A two-region stock assessment model was used to analyze the simulated data, and applied annually. Two abundance indices were recorded in each year during the projection period. This first was a fleetwide standardized catch-rate index, computed from the aggregate catch and effort data of the commercial fleet from applying a General Linear Model with factors for month, vessel, one degree resolution spatial grid, and year. The second index of abundance was based on the catch rate data collected during the structured line survey. Additionally, on each sample reef, the simulated survey sampled lengths and ages from 100 fish taken from the selectivity-weighted age distribution. However, variability of length length-at at-age was added to the associated length measurement, in the form of a normal deviate, $N(0, \sigma_l^2)$, where the variability in the length measurement, σ_l was set to 6.17 cm (Little et al. 2007, page 228). This had the effect of smoothing the frequencies across length bins, because each age class in ELFSim has only a single associated length.

3.3.2 Assessment model

The assessment model used in this study was based on the stock assessment model developed for cabezon *Scorpaenichthys marmoratus* (Cope et al. 2003). It aggregated biomass across reefs into two sub-populations (equivalent to regions): an unfished (protected) and a fished portion, and represents a potentially a greater simplification of the operating model than a recent assessment model developed for coral trout in 2014 (Leigh et al. 2014).

The assessment model was applied annually. Two abundance indices were simulated in each year during the projection period to represent how fisheries data could be collected from a fishery with no-take areas. The first index was a fleet-wide standardized catch-rate index, computed from the aggregate catch and effort data of the commercial fleet by applying a General Linear Model with factors for month, vessel, one one-degree resolution spatial grid, and year. The second index of abundance was based on the catch rate data collected during the structured line survey.

The two sub-populations in the assessment model had different time-series of exploitation rate, but were linked through a common stock-recruitment relationship, which allocates a constant proportion (r_p) of the recruitment to each sub-population p such that recruitment for each sub-population in each year is:

$$R_{p,t} = \frac{r_p \, 4hR_0S_t}{0.5S_0(1-h)+(5h-1)S_t} e^{x_t-\sigma_r^2/2} \tag{1}$$

where x_t is the estimated recruitment deviation for year t, ($x_t \sim N(0;\sigma_r^2)$), r_p is the proportion of recruitment that is allocated to sub-population *p*, R_0 is the total number of age-0 animals at unfished equilibrium, S_0 is the number of spawners at unfished equilibrium in both sub-populations combined, S_t is the number of spawners in both populations combined at time *t*, h is the steepness parameter that controls the degree of compensation, and σ_r is the (assumed) standard deviation of log-recruitment.

The differences between the operating and the assessment models are summarized in Table 3.2. One of the main differences is that the operating model represents the coral trout population as protogynous (Figure 1a), i.e. all fish are born females, but change sex as they grow, which is represented in the operating model using a logistic equation (Mapstone et al. 2004, Little et al. 2007). In contrast, half of each age-class is assumed to be female in the assessment model. An implication of this difference is that total spawner weight per recruit at unfished equilibrium is 19% greater in the operating model than the assessment model (Figure 3.1b). This difference was considered in the results by deriving a comparable summary spawning stock biomass from the assessment model based on the age distribution, and the proportion of females at age that occurs in the operating model.

Assessments were conducted annually on each of 10 projected replicates from 2012 to 2035, which was thought to balance the time it takes for the operating model to run with the expected replication needed to demonstrate differences among the scenarios. Most models converged with a negative log-likelihood of about 1,000. Model results (<10%) were omitted for assessments with a negative log-likelihood greater than 10,000.

Assessment output was not used to change the TAC in the operating model because we sought to determine the accuracy with which the assessment model could estimate stock status and changing exploitation rates as a function of the outputs from the assessment model, was deemed likely to confound or complicate the results, and also no control rule existed in the fishery that could use the estimates from the assessment.

We were interested in the effect on the stock assessment estimates of using the simulated structured line survey to collect data from areas closed to fishing. Consequently, we conducted analyses in which the line survey collected catch-rate, age and length information from reefs open to fishing as well as analyses where the line survey collected data from reefs closed to fishing only (Table 3.1).

Table 3.2: Key differences between operating and assessment models.

Model feature	Operating model	Assessment model
Spawning	Larvae are spawned on a reef based on the amount of habitat it contains.	Spawning and recruitment are combined by sub-population.
Movement	Based on the self-seeding (*ss*) parameter, larvae from each reef are pooled and then advected according to a connectivity or migration matrix.	No movement between sub-populations
Recruitment	The number of 1-year old recruits is determined for each reef by a Ricker stock recruitment function based on the number of larvae self-seeded, and advected from other reefs. $h = 0.5$	A single Beverton-Holt stock-recruitment relationship combining the spawning biomass from both sub-populations. Resulting recruitment is partitioned to sub-populations according to a fixed proportion. $h = 0.5$
Spawning biomass definition	Mature females. Protogynous: Fish are born female, but proportion declines with age.	Mature females. Females comprise 50% of population.

Simulations also examined different operating model conditions. First, the amount of larval connectivity, or self-seeing (*ss*) among reefs, was varied at two levels in the operating model: 1.0, and 0.1. Self-seeding (*ss*) equal to 1.0 meant that all of the larvae spawned on a reef remained there, whereas self-seeding equal to 0.1 meant that 10% of the larvae spawned on a reef remained there, with the remaining being advected based on a migration matrix (Note however, that with 10% self-seeding, some of the remaining 90% of larvae advected return to their native reef due to the hydrodynamics associated with that reef captured in the migration matrix. Each reef however is technically different; Mapstone et al. 2004).

Second, the amount of closures in the operating model was varied at two levels corresponding to 21% and 41% of coral trout habitat, measured as reef perimeter, in the marine park (Table 3.1). This represented roughly the amount of closed coral trout habitat before and after the Representative Areas Program (RAP) was implemented in 2002, which increased the amount of no-take marine reserve coverage (Fernandes et al. 2005). The amount of area assumed to be closed in the assessment model was set to correspond to level in the operating model.

The results of the simulations were summarized by the errors associated with estimation of the ratio of the female spawning biomass to the unfished level in each year of each simulation replicate.

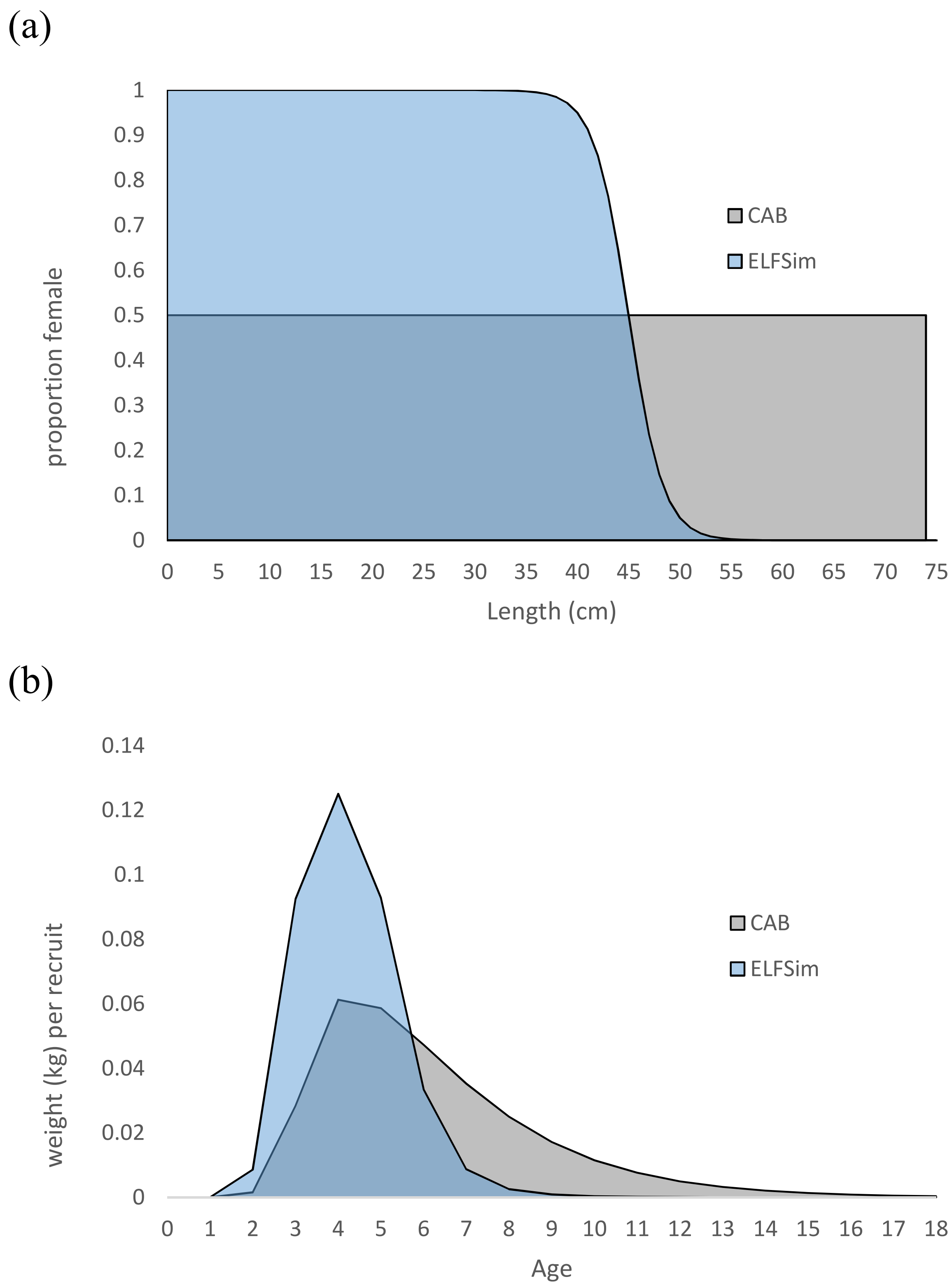


Figure 3.1: (a) The constant proportion of females at length assumed in the assessment model (CAB) and the protogynous assumption for the proportion of females in the operating model (ELFSim). (b) Weight (kg) at age of mature female-per-recruit of coral trout at equilibrium under zero fishing mortality defined by the assessment model (CAB) and the operating model (ELFSim).

3.4 Results

3.4.1 The effect of a fishing survey in area closures on stock assessment

Figure 3.2 shows that the assessment model estimates stock status at about 50%, whereas stock status in the operating model is around 70-80% when the simulated line survey collects data only from the reefs in which the fishery operates. The estimated stock status in each projection year varied little across the replicate simulations, except at the start of the projections when relatively low amounts of simulated data resulted in highly variable estimates of spawning biomass. The estimates of stock status are more accurate, but more variable across the replicates, when the line survey collects data from reefs closed to fishing (Figure 3.3).

3.4.2 Sensitivity of estimation performance to the structure of the operating model

The data points in Figure 3.4 represent stock status as a function of operating model biomass for each projection year and replicate, and across a range of TACs and assuming survey sampling in closed areas only. Perfect estimation in Figure 3.4 would be points on the 1:1 line. In general, there was a strong correlation between the operating model biomass and the estimated biomass across all larval advection and closure scenarios.

The estimates of stock status were generally negatively biased (i.e. they are below the 1:1 line; Figure 3.4). However, the bias differed among larval advection and closure scenarios, as represented by the slopes of the lines through the respective panels in Figure 3.4. Specifically, under the lesser self-seeding scenario ss = 0.1, when larvae were shared among reefs, and the metapopulation behaved more closely to a single stock, the bias was proportional to the underlying actual stock status, and largest at high actual stock size (bottom panels, Figure 3.4). For example, the estimated stock status was approximately 81% of the actual stock status (0.70 / 0.87 and 0.71 / 0.88) when the operating model stock status was high (large blue points, bottom panels, Figure 3.4), and declined to 73% (0.52 / 0.71 and 0.53 / 0.72) when the operating model stock status was low (large grey points, bottom panels, Figure 3.4).

In contrast, the pattern was reversed, and the estimates of stock status were more accurate when self-seeding was such that reefs are independent of each other, ss = 1.0, (and hence the larvae stay on the reef on which they were spawned), (top panels, Figure 3.4), especially when the operating model stock status was high. For example, the estimated stock status was between 92% and 98% of the actual stock status (0.81 / 0.82 and 0.75 / 0.81) when the operating model stock status was high (large blue points, top panels, Figure 3.4). However, as the operating model stock status declined, the negative bias increased, such that the estimated stock status ranged between 72% and 74% (0.48 / 0.67 and 0.49 / 0.66) of the actual stock status (large grey points, top panels, Figure 3.4).

The issue was complicated by the fact that the various scenarios represent different portions of the metapopulation that are closed to fishing. In the left panels of Figure 3.4, 21% of the coral trout habitat was assumed to be closed to fishing, while on the right panels, 41% was assumed to be closed. The amount of closure had little effect on the negative bias when ss=0.1 and reefs are highly connected as seen in comparing the two lower panels of Figure 3.4. In contrast, there was a marked effect of closures on the bias of stock status when recruitment to a reef was more independent of that to other reefs (ss = 1.0; top panels, Figure 3.4). In this case, the bias became increasingly negative as stock status in the operating model declined under the lower closure scenario of 21% (top left panel, Figure 3.4). The effect was reduced when the amount of closures was increased so that the estimated bias was more constant with respect to underlying actual stock status, but still negative (top right panel, Figure 3.4).

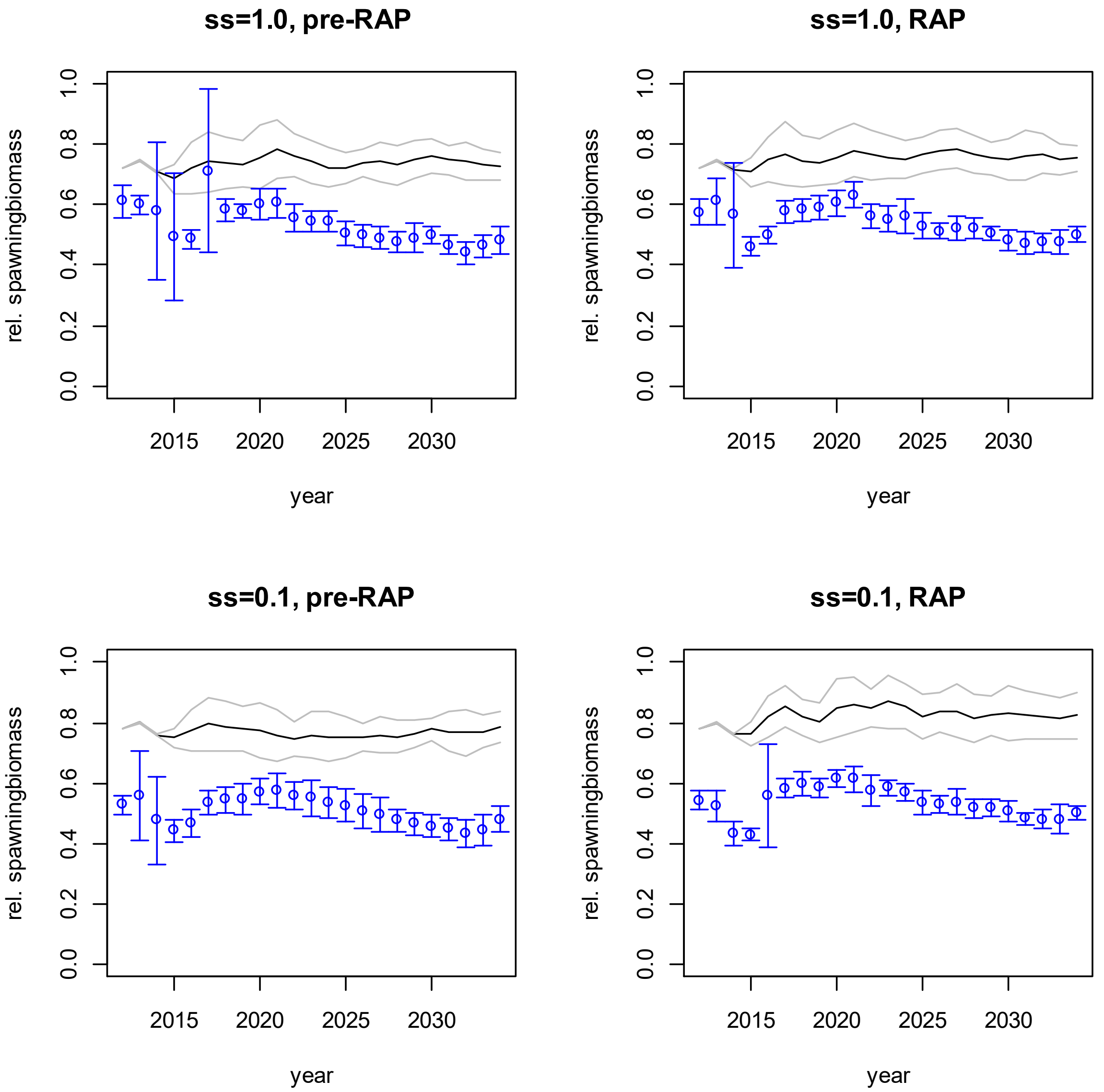


Figure 3.2: Average spawning biomass (mature females) across project replicates, relative to pre-exploitation equilibrium values, for the operating model (black) ± SD (grey) and the associated estimates (blue ± SD) when the annual survey only samples in the areas open to fishing. Top panels represent scenarios where self-seeding (ss) is 1.0, and bottom panels where it is 0.1. Left panels represent scenarios where the no-take areas totaled 21% of the coral trout habitat (pre-RAP closures), and right panels 41% of coral trout habitat (RAP closures).

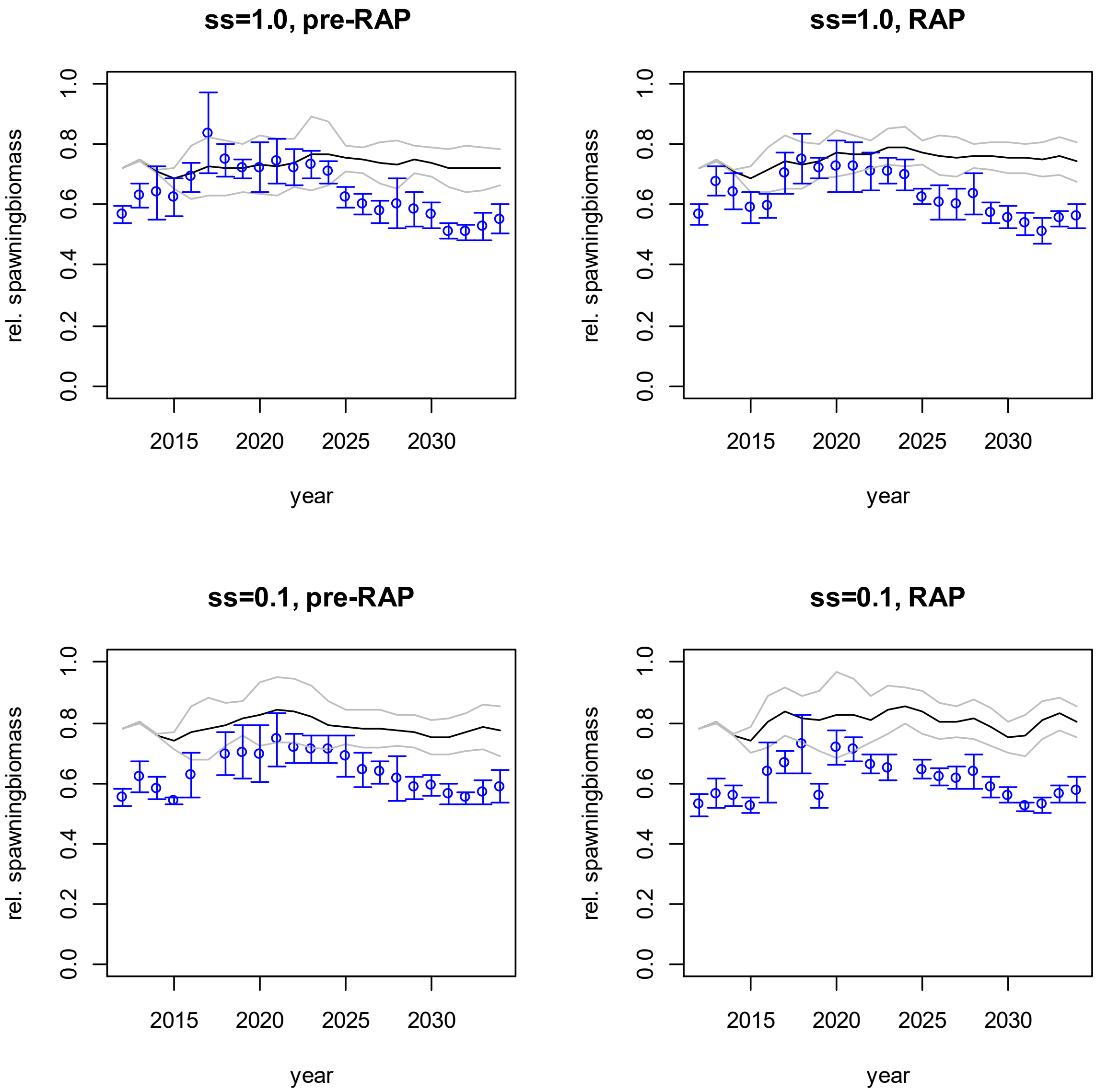


Figure 3.3: Average spawning biomass (mature females) relative to pre-exploitation equilibrium values for the operating model (black) ± SD (grey) and the associated estimates (blue ± SD) when the annual survey only samples in the areas closed to fishing. Top panels represent scenarios where self-seeding (ss) is 1.0, and bottom panels where it is 0.1. Left panels represent scenarios where the no-take areas totaled 21% of the coral trout habitat (pre-RAP closures), and right panels 41% of coral trout habitat (RAP closures).

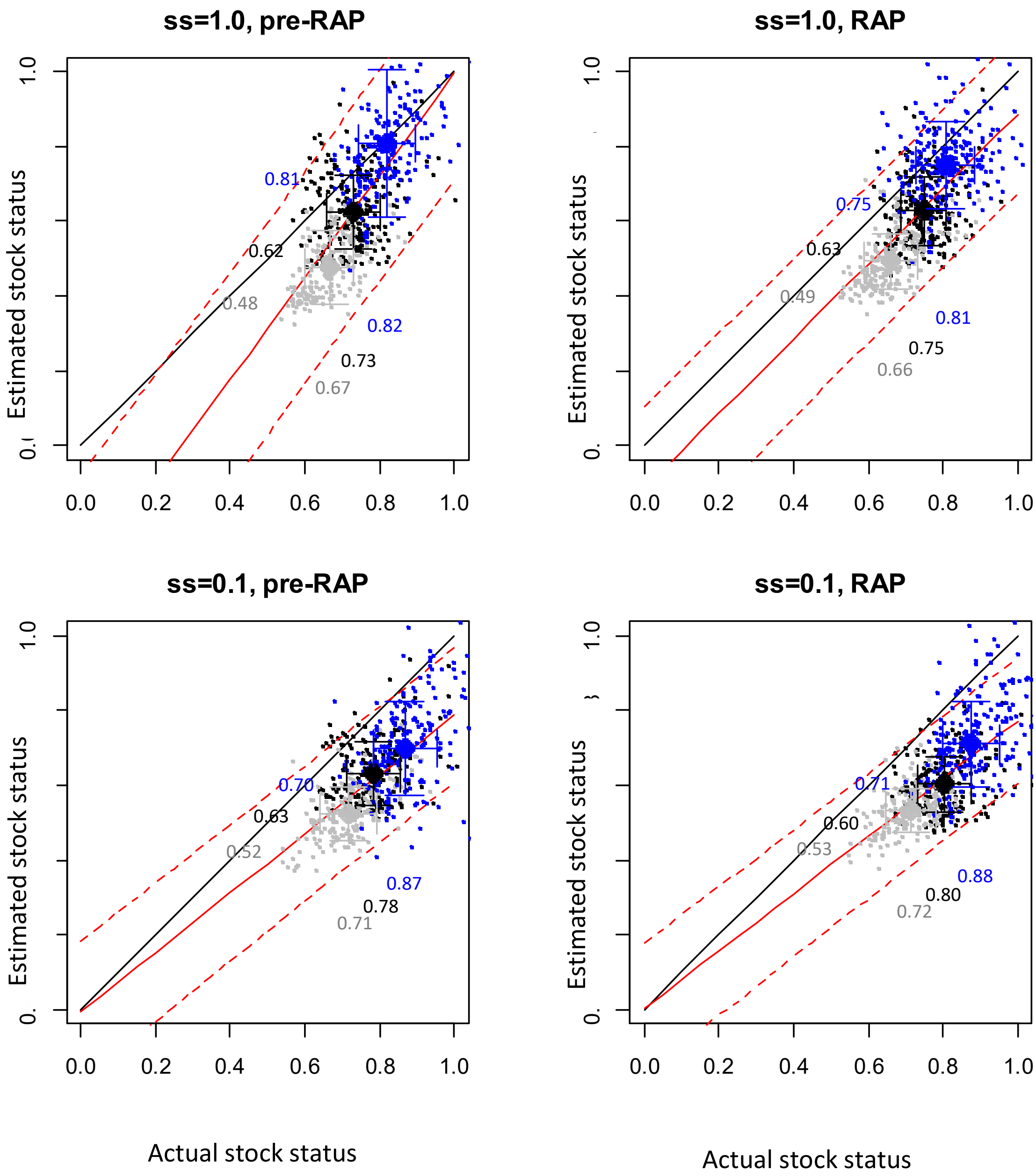


Figure 3.4: Estimated stock status from the assessment model, corrected for the difference in definition of spawning biomass, as a function of stock status in the operating model, when there are annual survey only samples in the areas closed to fishing, for two levels of self-seeding in the operating model (ss = 0.1; ss = 1.0) and two levels of closures (pre-RAP, RAP). Black points: TAC = 1,300t; blue points: TAC = 650t; grey points: TAC = 1,950t. Large points represent the mean, with their corresponding values (± st. dev). Red lines represent the fitted regression lines to the data points ± 95% confidence interval. Top panels represent scenarios where self-seeding (ss) is 1.0, and bottom panels where it is 0.1. Left panels represent scenarios where the no-take areas totaled 21% of the coral trout habitat (pre-RAP closures), and right panels 41% of coral trout habitat (RAP closures).

3.5 Discussion

Closures that restrict fishing and other extractive activities within their boundaries have been proposed for a range of purposes including conserving the habitat and species within them, as well as increasing the biomass and thus yield from fishing activities outside of them. While closures have been seen to achieve conservation objectives inside closure boundaries (e.g. Lester et al. 2009), the ability to achieve fishery objectives outside of them is less apparent (e.g. Mangel 1998) because fish move.

Movement complicates fisheries management because animals may cross jurisdictional and spatial management boundaries. It also poses a problem for conservation management, because putatively protected stocks become vulnerable when animals move outside of the closure. Even if animals do not move between open and closed areas, implementing marine reserves challenges fisheries management partly because it changes the portion of the fishery that is open, and partly because it does not reduce fishing effort. Thus, marine closures will often displace effort, and increase fishing mortality elsewhere, as in the case in the operating model (Mapstone et al. 2004).

When there is no mixing, the population in the open and closed areas can be considered separately. This is represented by the scenario in which ss = 1.0 in the operating model. However, the assessment model produced biased estimates of stock status owing in part to the fact that spawning biomass used to calculate recruitment in the stock recruitment function is scaled-up from individual reefs and pooled, with the extent of bias dependent on the amount of area closed to fishing (pre-RAP vs RAP closures).

In addition, another main cause of bias was the difference in how female mature biomass was defined in the assessment and operating models. Specifically, the assessment model assumed more of the large, older fish were female, while the operating model assumed fish were born female but became male as they grew (protogyny). Even though we corrected for any apparent differences in the output statistic, this difference would still result in fishing mortality having a greater impact on the number of females in the assessment model, and contribute to the negative bias. When more of the stock was exposed to fishing (Figure 3.4, top left panel), and under lower fishing mortality (large blue point), the assessment model would have estimated more females survived than in reality (in the operating model), resulting in greater productivity, and relatively higher estimated stock status. Increasing the amount of closures, and protecting more of the larger older part of the population, increased the negative bias of the assessment model as the assessment model estimated much fewer females were affected under reduced fishing pressure (Figure 3.4, top right panel, large blue point).

When the reefs are connected through larval mixing (e.g. ss = 0.1), larval subsidy, the process of larvae spilling over from the closed areas to the open areas, would be expected to increase the actual (operating model) biomass across both open and closed areas (Little et al. 2007), thus potentially leading to higher actual stock status in the lower panels of Figure 3.4 than in the upper panels (Little et al. 2007). At the same time, the assessment model, which did not consider this, estimated relatively lower stock estimates from the assessment model (Figure 3.4, bottom panels) compared to more restricted larval mixing (Figure 3.4 top panels), especially when the fishing pressure was low (blue points).

The assessment model used to account for closures is a robust, but biased way of dealing with spatial heterogeneity. Similar assessment models have also shown to be biased when they are mis-matched compared to the underlying dynamics (e.g. Punt et al. 2016). Our results indicate similar effects when life-history characteristics are mis-matched. Mixing has also been shown to bias estimates of management-related quantities from assessment models in previous studies. For example, McGillard et al. (2015) showed that a mixed stock subject to closures resulted in negatively biased estimates from a single stock perspective, and positively biased estimates when separate assessments for the open and closed areas are combined. Mixing can confound life history parameters such as natural mortality (Williams et al. 2010), with movement rates (Garrison et al. 2011; McGilliard et al. 2015).

Sampling from the closed areas improved estimation performance. This was not surprising given that this provided the assessment model, which included two regions/populations, data from trend and demographics on one in which data was limited. Nevertheless, the exercise shows how fishing data from no-take areas can be used to manage the stocks outside of them.

Overall, the results of this study contribute to further understanding of assessment models applied to support fisheries management. The ideal would be that the estimates of stock status are unbiased and precise, or at least that any bias was independent of the features of the operating model such as the extent of mixing, exploitation rate, and closures. However, our results suggest that the extent of bias in estimates of stock status depend on all of these factors. Moreover, the estimates of stock status are still biased when the assessment model is close to capturing the structure of the operating model (i.e. ss=1.0).

The cause of this bias was confounded between reproductive strategies, scaling-up from individual reefs, and life history (protogyny) assumptions that varied between the operating model and assessment model. As a first attempt at estimating the coral trout population using a simple off-the-shelf assessment model, the results stress the potential implications of trying to estimate the state of a stock without explicitly using an assessment model developed for such stocks. However, even a highly-specific coral trout assessment model (Leigh et al. 2014) still resulted in bias estimates (Little et al. 2016), although in this case with positive, over-estimated results. With an operating model that includes over 3,000 individual populations, each of which may be connected to the others and subject to different levels of exploitation, it is infeasible to conduct assessments by reef or even collect monitoring data from each reef. However, it is still necessary to provide management advice, recognizing that some reefs will be more depleted than intended and others less so, no matter what management strategy is applied. The lessons learned as part of this study will help to assist managers as they identify candidate management strategies, including harvest control rules, for testing and eventual adoption for this complex fishery.

4 MSE testing of tier level performance with the inclusion of a marine closure

Tuck, G.N.[1], Day, J.R. [1], Klaer, N.L. [1], Wayte, S.E. [1] and L.R. Little[1]

[1]*CSIRO Oceans & Atmosphere, Hobart, Tasmania, Australia*

4.1 Abstract

Management strategy evaluation (MSE) is used to evaluate the influence of the inclusion of a no-take marine closure on assessments and control rules currently used to manage harvested stocks off south eastern Australia. Marine closures are commonly used as a fishery and biodiversity conservation management tool. The consequent spatial structuring and reduction of data to inform assessments of status is known to bias standard assessments. A spatially explicit operating model is developed to allow the simulation of data from a single harvested area and following the introduction of a no-take closure. The closure can vary according to size and account cam be taken of the mixing rate between open and closed areas. Management strategies varying from data-rich (Tier 1) to data-poor (Tier 3 and 4) are evaluated for their ability to attain biomass targets and have reasonable biomass risk profiles (namely, the probability of the biomass falling below the limit reference point). We also examine how yield and catch rates are influenced by closure characteristics and the management strategy chosen. Our results suggest that full quantitative assessments with their associated harvest control rule (Tier 1) generally performed well in terms of meeting target biomass levels for the open area, in particular if all data are used (not just data from the open area) and closure size is small. The data-poor harvest strategies (Tiers 3 and 4), show increased risk, with the catch curve-based Tier 3 having a greater risk profile compared to the catch-rate-based Tier 4 harvest strategy, particularly when a suitable reference period was chosen for Tier 4. I. If managers choose a stock-wide target (of say 48% of initial biomass levels), then the open area biomass target can be less than the stock–wide target (less than 48%), due to the protection of stock within the closure. A stock-wide biomass target strategy also tends to maintain catches irrespective of mixing level. Alternatively, a harvest strategy based on a biomass target for the open area alone led to higher catch rates and stock-wide biomass for a particular mixing rate, but with potentially less annual catch than a strategy based on maintaining stock-wide biomass.

4.2 Introduction

The examination of alternative options for management using Monte Carlo simulation is called Management Strategy Evaluation (MSE). In this Appendix, we use MSE to evaluate the effect of a marine closure on stock assessments and harvest strategies (HS) commonly used in the Southern and Eastern Scalefish and Shark Fishery (SESSF). MSE is a computer simulation tool that allows various alternative biological and management scenarios to form the basis for evaluations (Sainsbury et al., 2000; Bunnefeld et al., 2011; Punt et al., 2014). This approach formally recognizes that the quality of management decision-making is contingent upon the uncertainty of the system and the ability to obtain data of sufficient quality and quantity to inform managers (Polacheck et al., 1999). The key features of MSE are:

a) uncertainty is specified and assessed for its influence on management outcomes.

b) it allows for experimentation across multiple management strategies (current and alternatives) and system structures in circumstances where real-world manipulation is problematic (financially, ethically).

c) optimal solutions are not necessarily sought, rather the aim is the identification of feasible options with an explicit outline of the inherent trade-offs among competing objectives.

d) the framework should promote learning of the system dynamics, for both decision-makers and other stakeholders providing input to the formulation of the simulation model, the management objectives, performance indicators, target points and performance measures (Bunnefeld et al., 2011).

Trade-offs should become clear when framing the objectives, performance indicators, reference points and measures in a natural resource system. For the managers of natural resource systems, choices must be made between utilization and conservation of that resource. MSE is widely used to evaluate management strategies in fisheries, and more recently in marine and terrestrial biodiversity conservation and sport (Bunnefeld et al., 2011; Tuck, 2011; Punt et al., 2014; Tuck et al., 2015).

The first step in the MSE process requires that a scenario is chosen for evaluation; that is, a particular parameterization of the operating model. A management strategy is then chosen from among a set of identified options. We define a management strategy to include (1) data collection, to provide relevant information on the state of the system; (2) assessment of stock status using the data collected; and (3) a harvest control rule to translate stock status from an assessment into a management action (a TAC for example). Multiple projections are then conducted for the chosen management strategy.

The performance measures and evaluation statistics generated by each replicate projection are then recorded for the chosen management strategy. Evaluation statistics include descriptive results such as the maximum, minimum, mean and variance of each indicator across the period simulated (see Section 4.3.10). A further important evaluation statistic is the probability of an indicator moving above or below chosen target or limit reference points. Once all management strategies of interest have been simulated, comparison of these alternative management strategies may proceed.

Due to the uncertain dynamics of a complex system, multiple future realities are possible. As such, management strategies are sought that are robust to system uncertainty. This requires stochastic simulations for a chosen scenario and the subsequent evaluation of management performance across additional scenarios. The simulation process can highlight weaknesses or biases in current or alternative management strategies.

MSE has previously been used to evaluate alternative metrics and assessment methods for fishery management when stocks include a marine closure. Punt and Methot (2004) used MSE to test alternative assessment methods in terms of estimating key quantities of interest for management. They found that impact on performance is slight if account can be made of the spatial structure of the population in the assessment model (see also Punt et al. 2016 and this report). Due to the complexity of the assessment methods and the additional data needed to deal with spatial heterogeneity (which is true of fisheries with or without a marine closure), indicators attached to control rules have been proposed to manage fisheries with closures (Field et al., 2006; Punt et al., 2015). Babcock and MacCall (2011) considered the ratio of fish density outside to inside a no-take reserve as a measure of stock depletion for use in a control rule similar to the depletion-fishing mortality based control rules applied in the SESSF. For example, if the ratio is small, then catch should be reduced (and if below 0.20, then the fishery is closed) and if the ratio is on target (say 0.60) then effort levels remain unchanged. Using MSE, the authors found that the metric and associated controls were effective at maintaining spawning biomass and yield for five Californian stocks. However, they also found that the metric was less effective if there was fish movement between open and closed areas, and so concluded that the density-ratio control rule is most effective for fish stocks that do not migrate or that move distances within the scale of the marine reserve. Wilson et al. (2010) also used MSE to test indicators from outside and inside closed areas as an indicator, namely the proportion of older aged fish in the population.

They argue that their management strategy is better suited to managing data-poor, spatially structured stocks than management strategies based on traditional stock assessment methods.

A number of authors have reviewed the impact and design of marine closures (Guenette et al., 1998; Botsford et al., 2003; Gerber et al., 2003; Halpern 2003; Field et al., 2006). The perceived advantages of marine closures include protection against over-fishing and biodiversity conservation (Field et al., 2006). Design criteria to be considered include the size, number of and position of the closures, the relative mixing between areas, and the species/habitat protected. Marine closures can impact the yield received by fishers and raise questions about the ability of methods that have traditionally been used to assess status, which assume homogeneous stocks (Field et al., 2006; McGilliard et al., 2015). Many authors have shown that yield can be maintained or reduce only slightly when a marine closure exists (Mangel, 1998; Botsford et al., 2003; Tuck and Possingham, 2000; Field et al., 2006; Barnes and Sidhu, 2013). More recently, attempts have been made to show the impacts of including marine closures on alternative configurations of stock assessments (Punt and Methot, 2004; McGilliard et al., 2013; Punt et al., 2015; Appendix 3). These authors show that large biases in estimation can occur, and while improvements can be made by including spatial structure, the uncertainties created by spatial structure and marine closures (in data sampling, surveys, appropriate model structure, and assumptions regarding mixing) complicate decisions regarding the choice of harvest strategy. Similarly, there are questions regarding how the ‘preserved’ stock in the closure is included in measures of stock status for comparison against targets for management purposes. Namely, should biomass targets focus only on the open area or on stock-wide biomass? This question was raised by Field et al. (2006) and considered by Barnes and Sidhu (2013). Barnes and Sidhu (2013) developed a deterministic model that explored the impact of a closure on yield and how this might depend upon stock mobility, closure size, and the management strategy applied. They found that stock-wide levels are generally higher management targets are computed relative to the open area only, with similar yields realised. As the model is deterministic, no measure of risk is considered and assessment models that sample data from the population (such as those typically used in the SESSF) are not considered.

The MSE software developed for the SESSF has established mechanisms that allow the SESSF assessment and TAC decisions rules to be evaluated. These include the full quantitative assessments of Tier 1, the catch curves of Tier 3 and the catch-rate-based assessment of Tier 4 (Smith et al. 2008; Wayte and Klaer, 2010; Little et al., 2011). The operating model on which the MSE is based can accommodate several alternative SESSF species. These stocks can then be considered under various levels of current depletion, mixing rates between spatial zones and closure sizes.

In this Appendix, the harvest strategies of the SESSF are evaluated using MSE to ascertain their ability to meet biomass and yield objectives, as well as having risk profiles that are acceptable to management under the HSP (Smith et al., 2008). The impacts on current harvest strategies of a closure are evaluated by the inclusion of a single no-take area that can have various levels of mixing between open and closed areas, and varying size. Typical data for assessment purposes are simulated, and include catch, catch per unit effort (catch rate, CPUE), lengths and ages. Two SESSF species are considered, Tiger Flathead and School Whiting. These are key SESSF stocks showing contrasting life-histories. The operating model (reality) was conditioned upon quantitative stock assessments (Tier 1) of these stocks, and choices can be made regarding the type of assessment used to estimate biomass for use in harvest control rules (HCRs) to set annual catch quotas, the type of data used for assessment (all or only from the open area), discounts when computing RBCs, and alternative harvest control rule targets (relating to the whole stock or only the open area). The three tier level assessments are known to vary in their ability to estimate biomass (Smith et al., 2008), but to this point, the respective ability of the harvest strategies to provide robust and acceptable management outcomes when marine closures exist is unknown.

4.3 Methods

A factorial design was implemented when setting the operating model and the subsequent population assessment (Table 4.1). The MSE included two population areas, with a specified proportion of the population assigned to each area at the start of the projection period, a constant percentage of mixing among areas, and with annual recruitment from the combined stock apportioned to the areas according to the percentage of the stock closed. The operating model specifications and equations are not repeated here and can instead be found in Fay et al. (2009). Likewise, details of the assessment methods can be found in Smith and Smith (2005), Smith et al. (2008), Wayte (2009), Wayte and Klaer (2010), and Little et al. (2011).

Table 4.1: The factors to be examined and their alternate levels in the MSE. Each combination of factors is a scenario.

Factor	Levels
Species	Tiger Flathead (FLT), School Whiting (WHS)
Starting stock status	Low
% stock closed	0, 25, 50, 75
Mixing rates (% per year)	0, 5, 10, 25, 50,75, 100
Harvest strategy	Perfect information, Tier 1, Tier 3 , Tier 4
Harvest control rule target (%B_0)	48,44,39,35,30,25,20
Discount factor (% reduced from original TAC)	0, 8, 15, 30
Future data	Open area only (no sampling from closure)
Catch history	Use all data, or proportional to open area

4.3.1 Species Considered

The specifications for the current stock status and population dynamics used to parameterize the operating model were based on recent full quantitative stock assessments for two major SESSF species: Tiger Flathead (Klaer, 2010), and School Whiting (Day, 2009); thus they are consistent with the available historical and biological information on these stocks. These species were chosen to represent contrasting life-history strategies. Tiger Flathead are a moderately long-lived species, with relatively constant recruitment, while School Whiting are short-lived with highly variable recruitment.

4.3.2 Starting Stock Status

The stock status at the start of the period in which the harvest strategies were applied was set to be below the target stock status, as it is the behaviour of this situation that is of most interest to fishery managers and industry. A lower stock status was obtained by manipulating the initial stock size in the operating model, so that true depletion at the end of the historical period is 0.35.

4.3.3 Harvest Strategies

The guidelines for implementation of the Commonwealth Harvest Strategy Policy (CHSP) encourage a tiered approach to cater for varying levels of knowledge about a stock (DAFF, 2007). In the SESSF, the Tier 1 harvest strategy uses a fully-integrated quantitative stock assessment, implemented in Stock Synthesis (Methot and Wetzell, 2013), to estimate the current biomass level, which is input into a target- and limit-based harvest control rule (HCR; Smith et al. 2008). The Tier 3 harvest strategy (Wayte and Klaer, 2010) uses information on the age frequencies of annual catches, annual total catch, and basic biological parameters to estimate current fishing mortality, which is then used in a HCR to calculate the subsequent year's intended fishing mortality. Tier 4 stocks are assessed using an empirical rule based on trends in standardized catch rates combined with target catches (Little et al., 2011). To remove the potential for bias brought about by an assessment, a perfect information case is also considered, where the spawning stock biomass (SSB) used for the HCR is taken directly from the operating model rather than being estimated.

In reality, in the SESSF there is a delay of approximately 17 months from the end of data used in stock assessments to the generation of a Recommended Biological Catch (RBC) and Total Allowable Catch (TAC) for the fishery. In simulations, to better account for this delay, there is a two year delay from the end of data used in assessment to the application of a fishery TAC (e.g. the assessment performed in year *n* uses data to the end of year *n-1* for estimating the TAC in year *n+1*).

4.3.4 Harvest Control Rule Target

The CHSP specifies B_{MEY}, the biomass that should lead to maximum economic yield, as the target biomass reference point, and half of B_{MSY}, the biomass corresponding to maximum sustainable yield, as the limit biomass reference point. The use of proxies of B_{40} (40% of unfished SSB) for B_{MSY}, and $1.2B_{\mathrm{MSY}}$ for B_{MEY} result in a limit SSB reference point of B_{20}, and a target SSB reference point of B_{48}. Harvest strategies can also be specified in terms of the corresponding fishing mortality rates (i.e. F_{48} is the fishing mortality rate that on average leads to the spawning stock equilibrating at B_{48}).

The harvest control rule used in the Tier 1 harvest strategy and perfect information scenarios for calculating the following year's fishing mortality is:

$$\begin{cases} 0 & if\ d < 0.2 \\ F_{targ}(d-0.2)/(break-0.2) & if\ 0.2 \le d < break \\ F_{targ} & if\ d \ge break \end{cases}$$

where *d* is the current depletion, and *break* is the HCR breakpoint (e.g. 0.35).

The form of the rule is shown in Figure 4.1. This HCR incorporates break-point and target fishing mortality levels of F_p – the *F* values that will reduce the spawning biomass to *p*% of the unexploited level. The rule is referred to by its limit:break-point:target, namely 20:35:48. Alternative control rule targets were considered that had lower target biomass reference points, because closing a substantial proportion of a stock may imply that greater exploitation in the open area may be feasible, while still maintaining overall conservation objectives for the total stock. The control rules targets are: 48,44,39,35,30,25,20 with the breakpoint being the corresponding linear point between F_{20} and F_{48} (Figure 4.1).

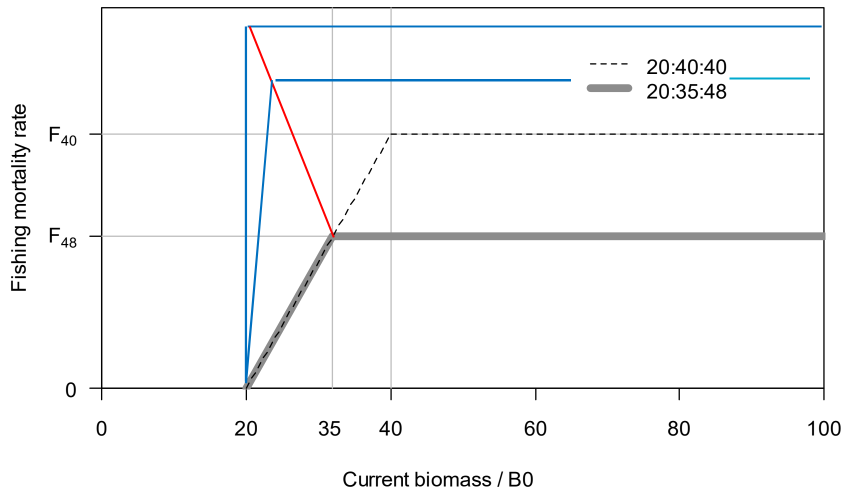


Figure 4.1: The Tier 1 harvest control rule variants. The solid grey line is the current Tier 1 control rule specification (20:35:48). Shown in blue are two alternative control rules with a HCR target of 20 and 30 and a break-point that follows the linear path from 35% biomass to 20% biomass (red line).

4.3.5 Percentage of Stock Closed

At the nominated year at which a marine closure is established a specified proportion of the population is closed, being 25%, 50% or 75% of the biomass, and there is no longer any catch taken from the closed area.

4.3.6 Mixing Rate

In the operating model, the historical population projection is performed with only one region until the year at which a marine closure is established. At this point, the numbers of fish at each age are split into regions according to the percentage of the stock closed. The projection then continues with two regions, with no catch in one of the regions. At the start of each projected year *M*% of the numbers in each age class are added together (where *M* is the mixing rate), and then distributed back to the regions according to the percentage of stock closed or open.

Mixing can be thought of as two cups of water, with the amount of water in each cup determined initially by the proportion of the stock in the open and closed areas. For example, with a 25% closure, 25% of the water is in the closed cup and 75% is in the open cup in the year of closure. With a mixing rate of, for example 10%, 10% of the water from each cup is poured into a third cup, mixed, and then poured back in the proportion of 25% to the closed area and 75% to the open area. As projections progress, catches and population dynamics lead to changing amounts of water in each cup, but mixing is always done in the same way with 10% of each to the third cup, and then poured back 25% to the closed area and 75% to the open area. This uses the ecological notion of carrying capacity in the assumption that the proportion poured back into each area remains fixed through time, meaning that a fully mobile section of the population will distribute itself according to environmental suitability only. It also allows for a net movement of fish out of the closed area due to mixing as the open area is exposed to fishing. Mixing was implemented in this way because it can be naturally described, and can be determined by a single number, limiting the number of dimensions required for the overall analysis.

Recruitment in each year is calculated from the whole stock spawning biomass and then also distributed to regions according to the percentage of stock closed or open.

At 100% mixing we assume that the stock will behave as a single stock regardless of any closed areas - i.e. catch rates and catches will translate smoothly across the period when the fishery closure is introduced. The concept of effective available biomass was developed to allow smooth translation of catch rates across the closure period for the 100% mixing case, while also correctly accounting for expected behaviour at 0% mixing (the populations in each area are independent) and also intermediate values. It makes sense to understand that the fish actually available to fishing gear are those in the open area as well as the proportion mixed from the closed area. For the special case of 100% mixing, all of the stock from open and closed areas will actually be available to the fishing gear.

The effective available biomass contributing to the CPUE is the total amount of available biomass (from open and closed areas) that was mixed and distributed among areas plus the remaining unmixed available biomass in the open area.

For historical years by age:

$$N_e = N_{all}\,(O + M.C)$$

= Numbers in the open area + (Numbers in the closed area)M

In the projections :

$$N_e = N_O + M.N_C$$

where N_e is effective numbers, N_{all} is numbers in the stock for each historical year, N_o is numbers in the open area, N_C is numbers in the closed area, O is the proportion open, C is the proportion closed, and M is the mixing rate. If mixing is zero, then the effective numbers reduces to numbers in the open area only. Likewise, if mixing is 100% (full mixing) then the effective numbers are all fish of all age classes.

N_O and N_C in the projections are not simply fractions of N_{all}, as N_O and N_C are modelled independently, taking into account proportion closed, mixing rate, and fishing.

For the perfect information scenarios, depletion is the ratio of current and initial effective available biomass, and yield is calculated by applying F_{RBC} to effective N. For CPUE generation, the CPUE for each fleet is generated from the retained vulnerable biomass calculated using effective N. Using this method, as expected, the trajectories of effective biomass and CPUE are smooth over the transition from non-closure to closure.

4.3.7 Catch History

In the stock assessment, the historical catch used can include data from both areas, or alternatively, only data that has come from the area that has become the open area since establishment of the marine closure. This factor affects each harvest strategy in a different way. For the perfect information scenario it makes no difference as catch is not used. For Tier 1, it changes the historical catch series used in the integrated assessment. For Tier 3 it affects the reference catch used in calculating the RBC, and will have a much larger effect on long-lived species as they will be using a longer period of historical catches. The period used for the reference catch is relative to the present time, so eventually past catches will not be used in this HS. For Tier 4, the target catch is always calculated over a fixed reference period in the past.

4.3.8 Tier 3 and Tier 4 Discount Factors

For the data-poor harvest strategies, Tier 3 and Tier 4, a discount factor (or buffer) is used in the SESSF in an attempt to balance risk between these assessments and the more data rich Tier 1 assessment (Fay et al., 2013). Tier 1 assessments would be expected to produce more robust results than the data poor assessments and so a discount factor, being 5% for Tier 3 and 15% for Tier 4, acts to reduce the RBC from the raw Tier 3 and Tier 4 analyses. The limit and target reference points of the HCRs are equivalent for all Tiers and so do not provide any additional precaution themselves. The application of discount factors can be removed under certain circumstances, namely if sufficient precaution already exists for the stock. This can occur if stock indicators (catch and catch rates) have been stable for a sufficiently long time, or if a closure exists that is believed to capture sufficient quantity of the stock biomass.

4.3.9 Simulations

One hundred replicates were conducted for each scenario with differences among replicates due to observation error in the generated data, and process error in the population dynamics (future recruitment deviations). Each scenario was projected into the future for at least 30 years for each simulation. Summary statistics were combined over all simulations to provide a set of performance measures for comparing results between scenarios.

4.3.10 Performance Measures

The stated objective of the Commonwealth Harvest Strategy Policy introduced by the Australian Government in 2007 is "the sustainable and profitable utilisation of Australia's Commonwealth fisheries in perpetuity through the implementation of harvest strategies that maintain key commercial stocks at ecologically sustainable levels and within this context, maximise economic returns to the Australian community" (DAFF, 2007). To achieve this objective, harvest strategies will seek to:

- Maintain fish stocks at a **target** biomass point equal to the stock size required to produce maximum economic yield
- Ensure fish stocks will remain above a **limit** biomass level where the risk to the stock is regarded as too high
- Ensure that the stock stays above the limit biomass level at least 90% of the time

The performance of the full Tier 1 HS was evaluated by plots of the trajectory of relative spawning biomass and catch over time, and the comparative performance for each harvest strategy was evaluated by summary plots of the following six performance measures relating to stock level, catch, and variability in catch:

1. average annual catch over the projection period;
2. spawning stock biomass (SSB) in the final year relative to unfished SSB;
3. spawning stock biomass (SSB) in the final year relative to unfished SSB in the open area;
4. spawning stock biomass (SSB) in the final year relative to unfished SSB in the closed area;
5. catch variability: average absolute percentage inter-annual change in catch (%AAV) over the projection period:

$$\%AAV = 100 \sum_{t=y_2}^{y_f} \left| C_t - C_{t-1} \right| \Big/ \sum_{t=y_1}^{y_f} C_t$$

   where y_1, y_2, and y_f are the first, second and final years of the projection period, respectively, and C_t is the catch in year t; and

6. probability of the spawning biomass going below the limit reference point (B_{20}) during the projection period in any region (the proportion of the projected years in which the depletion in at least one area is <0.2).

4.3.11 Tier 4 specifications

The Tier 4 HCR requires the identification of a historical period that is a desirable target in terms of CPUE, catches and status of the fishery. The RBC is calculated as

$$RBC = C_{targ} \frac{(CPUE_{av} - CPUE_{lim})}{(CPUE_{targ} - CPUE_{lim})}$$

where $CPUE_{targ}$ is the average cpue over the reference period, C_{targ} is the average catch over the reference period, $CPUE_{lim}$ is 0.4*$CPUE_{targ}$ and $CPUE_{av}$ is average cpue over the last 4 years.

If the fish stock is not fully-fished by the reference period, then the target catch-rate is halved. The target catch is also halved if a fish-down is considered to have happened in the reference period (i.e. catches are high).

There are three factors which need to be set when running Tier 4 scenarios in the MSE: reference period, catch rate multiplier, and fleet.

1. Reference period

   The success of Tier 4 is highly dependent on the reference period chosen. Tier 4 will get to the specified target, but that may not be the correct target. It is necessary to specify how the reference period for each species is chosen. For the Marine Closures runs, the reference period was chosen so that the average depletion in this period was on target, for a scenario where the starting point for projections was on target.

2. Catch rate multiplier

   Where a fishery is not considered to be fully developed by the start of the reference period, the target catch rate, $CPUE_{targ}$, is divided by two as a proxy for expected changes to catch rates as the fishery develops and the resource stock size declines towards the target of 48% unfished biomass.

   Flathead and Whiting are considered to be fully fished by the start of the reference period, so the catch rate multiplier is 1. In some cases for Tier 4, the reference catch is also halved (if a fish-down was occurring the reference catch is halved).

3. Fleet

   The third input is which fleet to use for Tier 4 cpue. In the MSE, it is possible to either set it to use the fleet with the highest proportion of catches in the last five historical years, or specify a particular fleet. For Tiger Flathead and School Whiting the former option was used (which leads to trawl for whiting and trawl (excl Tasmania) for flathead.

species	Ref period	CR multiplier	Fleet used
FLT	1992-2001	1	Trawl (excl Tasmania)
WHS	1997-2002	1	Trawl

4.4 Results

4.4.1 Outputs

Box and whisker plots of the performance measures showing the median, 50 percentile and 95 percentile are used to summarise each scenario. A scenario being a particular combination of stock, closure size, harvest control rule (HCR) target, stock status assessment method, mixing rate, discount factor and data used for the assessment (Table 4.1). These plots provide an indication of the variation experienced among each of the 100 replicates. Contour plots of the median of risk (the probability of being below the limit reference point calculated across all projection years), the final biomass depletion level, final catch rate and final catch are also used to make more general conclusions about the influence of each factor.

4.4.2 Simulations

For each scenario, 100 replicates were conducted across at least a 30 year projection period (some scenarios were extended to 100 years to ascertain long-term equilibrium behaviour). The operating model factors and resultant numbers of scenarios are described in Table 4.2. The total number of scenarios was: 2595 × 2 = 5,190.

Table 4.2: Number of simulation scenarios.

FLT & WHS	% stock closed	Mixing rate	Historical catch	HCR target	Discount factor	# scenarios
T3	3	7	2	7	4	1176
T4	3	7	2	7	4	1176
T1	3	4	2	4	1	96
Tier P	3	7	1	7	1	147
						2595

4.4.3 Model Outputs

Figures are categorised according to species, assessment method, mixing rate and percent closure. Other factors include alternative harvest control rules (HCRs) and discount factors applied. Key outcomes to note are that populations with 100% mixing show, depending on the robustness of the assessment method, both open and closed populations on the target biomass reference point regardless of the percent closed. Likewise, with no mixing the open population should be on target. Where this does not occur, biases have been introduced through the harvest strategies used, such as the uncertainties inherent in the assessment methods (e.g. data-poor methods), selection of assessment parameters (e.g. reference periods), or the data collected (e.g. large reserves reducing fishery-dependent data for assessment).

4.4.4 Utility of the Tier Estimators with Closures

The performance measures of the perfect information (Tier P) cases illustrated in the figures provide an indication of the potential biases of the Tier level estimators of biomass (Tier 1, 3 and 4) and the consequent effect on stock status when the biomass values, or proxies, are used within the associated HCR. While Tier P uses perfect information of stock biomass for the HCR it is, however, not deterministic, as its performance is influenced by stochasticity in recruitment. Note that with perfect information of stock biomass, data are not needed to estimate biomass (using an assessment), and so the input data scenarios that use either all of the data, Ha, or only use data from the open area, Ho, do not apply.

The Tier 1 harvest strategy generally performs well in comparison to the perfect information case (Figure 4.3 compared to Figure 4.4; and Figure 4.19, Figure 4.21 for tiger flathead, and Figure 4.7 compared to Figure 4.8; and Figure 4.11, Figure 4.13 for School Whiting). With a fixed HCR target of $0.48B_0$, using all the data (Ha) in a Tier 1 harvest strategy appears to provide more robust results with respect to achieving the target in the open area compared to only using data from the open area (Ho) (Figure 4.4 and Figure 4.8). However, when considered across HCR targets and mixing rates, using only the open area data appears to better approximate the stock status results of the Tier P case (Figure 4.11 compared to Figure 4.13 for School Whiting; Figure 4.19 compared to Figure 4.21 for tiger flathead; see Section 4.4.6).

The Tier 3 harvest strategy generally performs well in terms of its ability to reach the HCR target of $0.48B_0$ in the open area for tiger flathead (Figure 4.5) with exceptions being for low mixing rates for both the Ha and Ho input data scenarios. The Tier 3 harvest strategy for School Whiting is less robust, with increased risk, and also a reduced ability to reach the HCR target across most mixing and input data scenarios (Figure 4.9). Tier 3 is known to not perform well for School Whiting as the stock is more recruitment-driven and short-lived (Wayte and Klaer, 2010). Contour plots of risk and final stock-wide biomass for School Whiting when using Tier 3 show reduced biomass for a given mixing level and HCR in comparison to Tier P (Figure 4.11 and Figure 4.12). For Tiger Flathead contours show, surprisingly, an increase in stock status with decreasing HCR targets for Tier 3 (Figure 4.20 and Figure 4.22). This is because of an increased propensity for the stock dynamics to dramatically cycle as the HCR target decreases toward 20 (a very low and unrealistic target). Therefore, performance measures that use the value at the final year of projections are likely to be heavily influenced by whether the cycle is at a peak or trough coincident with the final year. This effect, which implies the Tier 3 harvest strategy for Tiger Flathead is unreliable, is discussed further in Section 4.4.9.

The Tier 4 harvest strategy consistently leads to final stock status that is under the target for the open area (except for large closure sizes when using only data from the open area) for both school whiting (Figure 4.10, Figure 4.12) and Tiger Flathead (Figure 4.6, Figure 4.20). This is because the Tier 4 estimator requires a pre-determined reference period that is assumed to approximate the desired long-term catch rate target. In this case, the Tier 4 reference periods have simulated an inadvertent use of an historical reference period with lower than target biomass to determine the target catch rate, and so the equilibrium stock-wide depletions are as a consequence lower than the desired target (compare the contours with the perfect information case). Determining a suitable reference period is a known issue for the catch-rate-based Tier 4 assessment method (Smith et al., 2008; Wayte 2009; Little et al., 2011).

For the perfect information case, and the Tier 1 harvest strategy, there is minimal risk of the stock falling below the limit reference point of 20% of pre-exploitation biomass levels. School whiting and Tiger Flathead (more so) have higher risk under Tier 1 if all data are used (Ha) and the closure size is large (Figure 4.4 and Figure 4.8). As expected, the data-poor assessment methods, Tier 3 and Tier 4, often have substantially increased risk in comparison to Tier P and Tier 1. This is observed in the lower plots of Figure 4.5, Figure 4.6, Figure 4.9 and Figure 4.10 (note however that flathead using only the data from the open area does not appear to show an increased risk; Figure 4.6). In addition, application of lower HCR targets in the open area increases risk (Figure 4.11 and Figure 4.12 for school whiting), with Tier 3 having a greater risk profile for a particular HCR and mixing rate than Tier 4. Other analyses (Fay et al. 2013; Little et al. 2014) have also concluded that risk may be greater for Tier 3 than for Tier 4, which raises a question regarding the greater RBC discount currently being applied for Tier 4 in comparison to Tier 3. Interestingly, the mixing rate does not appear to influence risk greatly (the contours of risk are largely vertical).

4.4.5 Catch Rates, Catch, Depletion and Management Targets

The establishment of a marine closure influences outcomes from traditional assessment methods (Punt et al., 2016a,b), and leads to questions about the appropriate HCR target to use, i.e. should targets apply to the biomass of the whole stock or only that in the open area? Specifically, there are two options: (i) apply a stock-wide HCR target of x_a, or (ii) apply a HCR target of x_o to the open area only. An argument for the former would be that if sufficient stock is contained within the closed area, then the open area should be able to sustain a lower HCR target than would be the case without a closure (in order to meet conservation imperatives, but also maintain reasonable catch from the open area). An argument for the latter would be that the economic target of MEY should be maintained for the open area stock only, irrespective of the closed area (assuming conservation imperatives are satisfied).

The conservation implications of the two HCR objectives with x_a = 48 and x_o = 48 can be considered by looking at (i) the stock-wide status contour curves showing 0.48, and (ii) the stock status values produced by a HCR of 48 (vertical from the x-axis HCR of 48). This is illustrated below for school whiting, where for a particular mixing value (in this case 40%), the HCR in the open area that leads to a stock-wide biomass of 48%Bo is approximately x_o = 39. This implies, as expected, that a target in the open area which is lower than the target if there were no closure can maintain a stock-wide biomass of 48%Bo. The HCR target in the open area to maintain a particular stock-wide biomass varies according to the mixing level (Figure 4.2). Alternatively, if the management HCR target of 48 is to be maintained in the open area alone, irrespective of the closure, then the vertical dashed line at HCR of 48 shows the resultant stock-wide biomass. Again, biomass outcomes vary with mixing levels, with less mixing leading to a greater (than 48%Bo) stock-wide biomass level (see Barnes and Sidhu (2013) for a similar result). Full mixing essentially implies the closure will make no difference to stock management, although, the closure may have purposes beyond single species stock management.

Comparisons of the catch and catch rates (as a proxy for economic performance) under the two HCR target alternatives can be made by considering the contour plots of Figure 4.11 and Figure 4.26. Taking the 25% closure, school whiting, perfect information of biomass scenario as an example, the stock-wide final depletion is shown in Figure 4.11 (top right). The corresponding contours of final catch and catch rates are in Figure 4.15 (top). Interestingly, the contours of catch generally match the contours of final stock-wide biomass. This implies, for example, that maintaining a stock-wide biomass of x_a = 48 by adjusting the HCR target in the open area will also maintain a catch of approximately 1700t (Figure 4.15, top right). The catch rate with x_a = 48, however, declines from approximately 1.0 units at full mixing, to 0.45 units at zero mixing (Figure 4.15, top left). Alternatively, maintaining a HCR target of x_o = 48 in the open area leads to catches that vary from 1700t at full mixing to 1300t at zero mixing. Catch rates also vary from 1.0 units at full mixing to 0.8 units at zero mixing. Stock-wide biomass varies from approximately 50%Bo at full mixing to 60%Bo at zero mixing. Barnes and Sidhu (2013) find with a harvest strategy of x_o = 50 that stock-wide biomass can increase with only marginal loss in catch. Here we find a similar increase in biomass but that catches decline substantially, depending on the mixing rate.

The example and results discussed above were contingent on the use of perfect information of biomass (Tier P) within the HCR to set annual catches. Clearly, many of the relationships assumed in this ideal case break down with the data-poor (or poorer) assessments and associated harvest control rules. For example, assuming that the HCR target in the open area will actually be achieved becomes less certain. With school whiting, a comparison of the open area depletion for Tier P (all median values are on 48%Bo with a HCR target of 48, as expected) compared to the other tiers shows how this assumption can fail (Figure 4.7 to Figure 4.10, third row of the figures).

4.4.6 Input Data – use all data (Ha) or only that in the open area (Ho)

The establishment of a no-take marine closure immediately removes a source of fishery-dependent data from the closed area that has, or may have, been used in previous assessments of the stock. The data stream is in a sense broken and a decision needs to be made about whether it is appropriate to continue using historical data that has come from the now closed area, or restrict analyses to use data only from the open area. This is particularly relevant for catch rate analyses, where capture rates may differ between local areas. If growth rates are spatially heterogeneous (not examined here) then length data for a particular age may also vary.

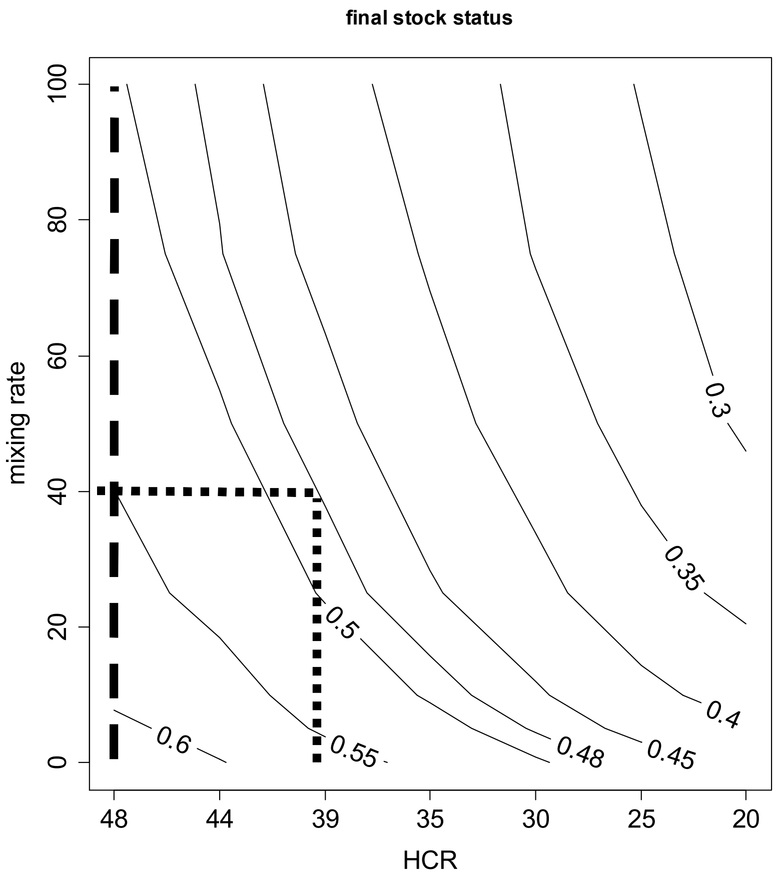


Figure 4.2: An example contour plot of final stock-wide status as it relates to the mixing rate and harvest control rule (HCR) target in the open area. Dotted lines indicate, for 40% mixing, the HCR target in the open area required to maintain a stock-wide biomass of 48% of initial levels. The vertical dashed line shows the resultant stock-wide biomass if the HCR target is 48 in the open area alone.

Results from using all of the data (Ha) or only historical data from the open area (Ho) in assessments of biomass are illustrated in several of the figures. The perfect information case (Tier P) uses the actual biomass in the HCR and so Ha and Ho are not relevant in this instance. For Tiger Flathead and school whiting, Figure 4.4 to Figure 4.6, and Figure 4.8 to Figure 4.10 show box plots of the performance metrics for each of the assessment methods (Tier 1, 3 and 4) according to mixing rates, percentage closed and input data assumptions. A comparison across Tier levels for Tiger Flathead shows that some bias is introduced if only data from the open area is used and the closure size is large (75%). This is evident in the median values for biomass depletion of the open area, which theoretically should meet the HCR target of 48. Tier 1 appears to over-shoot the target when using open area data only, Tier 3 under-shoots the target and Tier 4 over-shoots the target, in particular for highly mixed stocks. Of course, when interpreting the results from Tier 4, one needs to be cognisant of the biases introduced by an inappropriate choice of reference period. Using all the data, Ha, more consistently reaches the HCR target. However, for Tiers 1 and 4 using all the data increases the risk of falling below the limit reference point and increases catch variability. Similar results are found for school whiting. Tier 3 shows similar performance with respect to meeting the HCR target of 48 and risk regardless of input data assumptions. As with tiger flathead, there appears to be a slightly elevated risk to the stock if all data are used compared to only the open area for all tier level assessments.

The contour plots of risk, depletion, catch and catch rates as a function of mixing and HCR targets also provide an ability to compare results for Ha and Ho for both stocks (Figure 4.11 to Figure 4.26). These figures re-iterate the results found from the box and whisker figures discussed previously. For example, a comparison of Figure 4.11 and Figure 4.13 (bottom) for school whiting assessed with a Tier 1 harvest strategy shows that the risk contours have increased for a particular mixing rate and HCR target when using Ha compared to Ho.

4.4.7 Closure Size

As the closure size increases from 25% to 75%, contours of final stock-wide status show, for a particular HCR target and mixing level, an increase in the equilibrium proportion of total spawning stock biomass compared to virgin levels (Figure 4.27). This is expected as more of the stock has the potential to be contained in the closed area. However, this result is dependent on mixing rates between the open and closed areas. For example, having a 75% area closure does not guarantee that at least 75% of the stock will be protected in the closure. With increasing mixing, the preservation influence of the closure reduces, until at full mixing there is no influence on stock status of the closure at all. For the perfect information case (Tier P), Figure 4.27 shows contours of stock-wide status for school whiting and tiger flathead. The contours are similar for a particular closure size for both stocks. This is because the contours should follow certain paths irrespective of the biology of the stock. As mentioned, at full mixing, the contours of stock status *s* should approximate the HCR target, *x*, so $s_{100} = x$. Likewise, at zero mixing, with a closure size of *y* and a HCR target of *x*, the stock status will be approximately $s_0 = y + (1-y)*x$, irrespective of stock biology. The contours between full mixing and zero mixing connect the points between s_0 and s_{100}. As the robustness of the assessment method used to estimate biomass deteriorates, this relationship may also dissipate (see Figure 4.12 as an example).

The ability of the harvest strategies to meet HCR targets weakens as assessments of biomass become less robust. For Tier 1 (Figure 4.4 and Figure 4.8) while generally meeting the HCR target of 48 in the open area across scenarios, if only data in the open area are used (Ho) with 100% mixing and the closure is large (75%) then the target is substantially over-shot. This is not evident if all data are used (Ha) from the historical period in the Tier 1 assessment. However, if all data are used, mixing is less than 100% and closure size is large, then the final stock status of the open area is generally below the desired target and the risk increases (i.e the probability of the stock declining below the limit reference point increases, and is greater than 0.1 for tiger flathead; Figure 4.4). These results are also observed and more evident for Tier 4 (Figure 4.6 and Figure 4.10) but less evident for Tier 3 (Figure 4.5 and Figure 4.9).

4.4.8 Discount Factors

The Tier 3 and 4 harvest strategies require the implementation of a discount factor to the resultant RBC, so as to account for the risk induced by the application of a data-poor assessment method. For the Tier 3 and Tier 4 harvest strategies including a discount on the RBC not surprisingly reduces the equilibrium catch and catch rate, reduces risk and increases stock status compared to not having a discount (Figure 4.28 to Figure 4.35). In the contour figures, a comparison can be made between no discount and a large discount of 30%. The discount of 30% used here is larger than the current discount factor of 15% for Tier 4 to illustrate the resultant behaviour. As has been noted earlier, results for flathead become unreliable due to cyclic behaviour as the HCR target approaches 20 (see Section 4.4.9).

Interestingly, if the intent of the discount factor is to achieve risk equivalence between the Tier 1 harvest strategy and the data-poor strategies, then the risk profiles when including the 30% discount factor should be compared against those of the Tier 1 harvest strategy. For example, the school whiting Tier 1 risk profile of Figure 4.11 (bottom left) can be compared against the resultant risk profiles with a discount seen in Figure 4.28 for Tier 3 and Figure 4.29 for Tier 4. While a HCR target of 48 leads to minimal risk of breaching the limit reference point irrespective of Tier (except for Tier 3 perhaps), the risk resulting from use of Tiers 3 and 4 with a 30% discount remains greater than that of Tier 1. However, the increased risk profile for Tiers 3 and 4 with a 30% discount is largely restricted to the lower HCR targets, which are not used within the SESSF and, as has been noted, produce unreliable results for some scenarios with Tiers 3 and 4.

4.4.9 Cyclic behaviour in resource dynamics

Comparing the perfect information case (Tier P), where the true biomass is used in the HCR, with each of the tier level assessments that estimate biomass can provide an indication of the level of bias introduced by the estimators and provide insights into the expected dynamic behaviour under ideal conditions and data-poor conditions. Not surprisingly, as HCR targets decrease from 48 to 20, the Tier P contour plots show that the stock-wide depletion also declines. As the assessment used becomes less robust, the ability to mimic the perfect information case reduces. This can usually be explained by a poor selection of reference period (Tier 4 for example), fewer data for assessment (large closures limiting fishery dependent data), or known high variance in estimation ability (Tier 3 for example). However, it would still seem logical, or desirable for a harvest strategy, that as the HCR target decreases, the proportion of the stock remaining should also decrease. This does not happen for Tiger Flathead when using Tier 3 and for Ha with Tier 4 (Figure 4.20 and Figure 4.22). Under these circumstances, the stock-wide status improves with decreasing HCR targets, in particular for low HCR targets.

To further explore this phenomenon, the original time horizon was extended for all Tier 3 and Tier 4 scenarios (Figure 4.36 to Figure 4.43). Results illustrated that substantial cycling (Figure 4.38) and non-equilibrium behaviour (Figure 4.41) is evident in the Tiger Flathead biomass with low HCR targets; with some projections showing cycles that dampen, while others do not. As such, the depletion can vary substantially depending upon when the final year is chosen. Namely, some years may have high depletion while only a few years later depletion may be low. This cycling has implications for management. In particular, it raises questions about the stability of the assessment method (Tiers 3 and 4) under the particular spatial closure scenario parameters chosen. Cycling is most evident for the low HCR targets because the target becomes close to, or equal to (in the case of the HCR target being 20), the limit. This is not a sensible HCR, as the fishery varies between being shut (zero catch) and open (see Tuck (2009) for an example fishery where the target biomass reference point was equal to the limit). Interestingly, as mixing increases, the cycling behaviour dampens (Figure 4.37). Cycling is potentially enhanced by flows of recruitment from the closed to open areas that are then subject to cyclical fishing pressure when the species is relatively long-lived. In such species, a sequence of lightly or heavily fished cohorts will influence the overall biomass for a considerable period. School whiting does not exhibit cycling behaviour, even for low HCR targets (Figure 4.42 and Figure 4.43). This may be because its short-lived nature does not allow biomass to be retained for many years, and allows it to be more responsive to management rules.

4.5 Discussion

The interpretation of outcomes from the MSE testing of the impacts on assessments and harvest control rules of marine closures is complicated by the multi-factorial nature of the problem. Namely, there are several estimators of biomass (Tiers), varying mixing levels, closure sizes, discount factors, control rule targets, input data and biology that combine to influence management outcomes. These have largely been discussed in the results section. A brief general summary of the results is provided below:

- Tier 1 assessments generally meet the desired HCR target for the open area stock and meet Harvest Strategy Policy risk objectives.
- Tiers 3 and 4 show increased risk, and varying levels of bias, compared to Tier 1 and Tier P, depending upon:
  - Closure size
  - HCR targets
  - Input data used
  - Reference periods chosen

- Appropriate HCR targets can be considered depending on management objectives. Namely, (i) a stock-wide HCR target or (ii) a HCR target for the open area only.
- As stock is protected by closure, stock-wide biomass status and catches can be maintained (at say 48% of initial levels) by appropriately (depending on mixing) reducing the HCR target in the open area.
- There is a greater reduction in realised catch rates when maintaining a stock-wide HCR target (of 48% of initial levels say) than if adopting the same HCR target in the open area only. By contrast, there is a smaller reduction in catches.
- Discount factors reduce the stock risk when Tiers 3 and 4 harvest strategies are applied, but do not match the Tier 1 risk profile (across all HCR targets) even with a large discount factor.
- The Tier 4 harvest strategy shows less risk than the Tier 3 harvest strategy (counter to the current SESSF discount application of 5% for Tier 3 and 15% for Tier 4).
- Using all of the data, Ha, shows better performance (at meeting the target biomass) across Tiers, but has increased risk compared to only using data from the open area, Ho.
- With low HCR targets, the biomass dynamics for the data-poor harvest strategies become unstable (noting that these targets are unlikely to be adopted by management).
- Cycling at low HCR targets is dampened by increased mixing and discount factors.
- While increased closure size has the potential to capture more of the stock, the stock-wide influence of fishery management measures is dependent on mixing rates.
- The larger the closure size, the less stock-wide fishery dependent data are available for assessment, which can increase stock risk when applying the data-poor assessments.

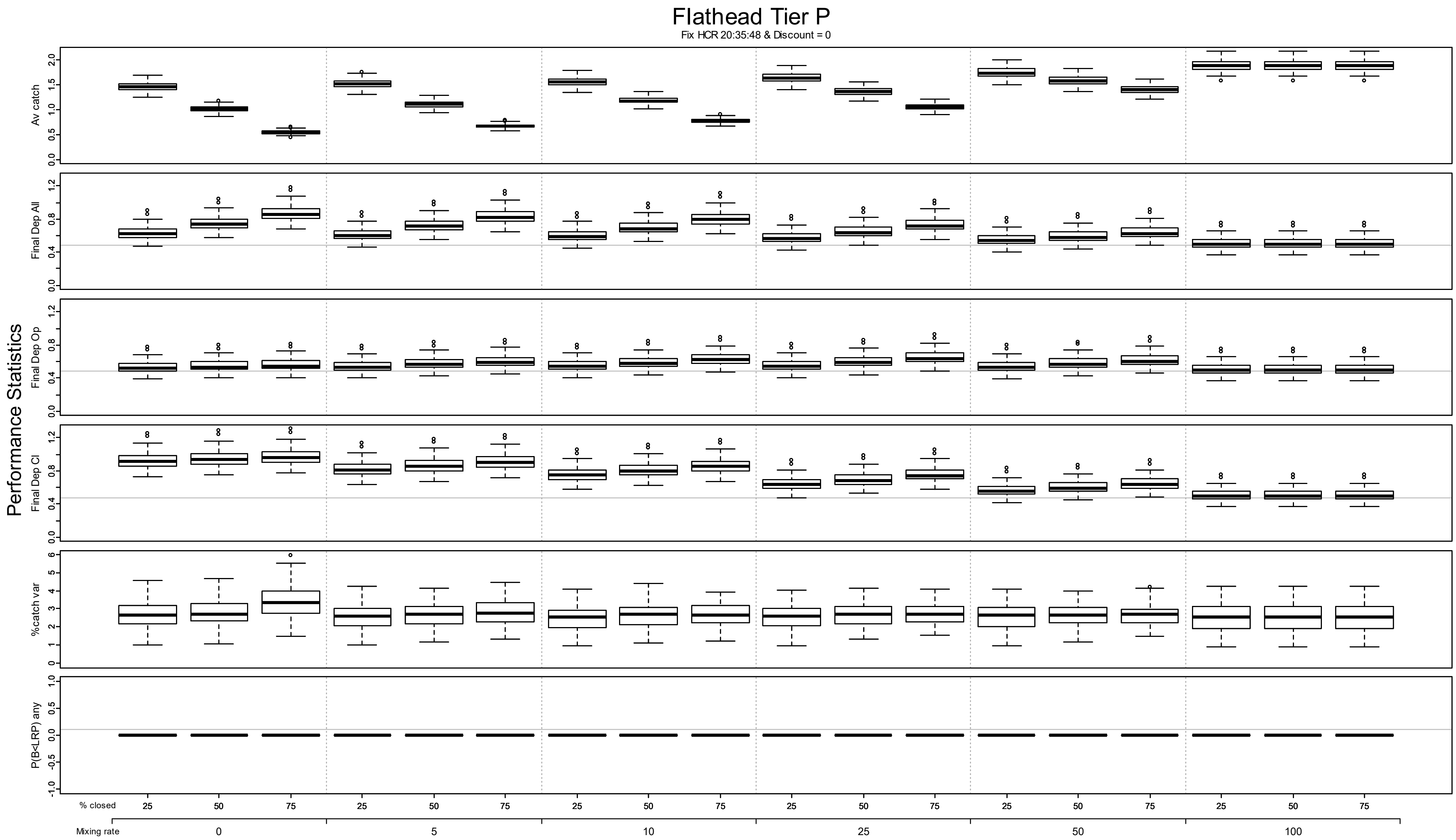


Figure 4.3: For Tiger Flathead with perfect information of biomass (Tier P) and a target of 48%Bo, shown are whisker plots of each of five performance measures under differing closure size (25%, 50%, 75%) and mixing rate (0, 5%, 10%, 25%, 50%, 100%). The grey horizontal line represents the target biomass level (48%Bo) in the three relative biomass plots and is the 0.10 probability in the plot showing the probability of being below the limit reference point.

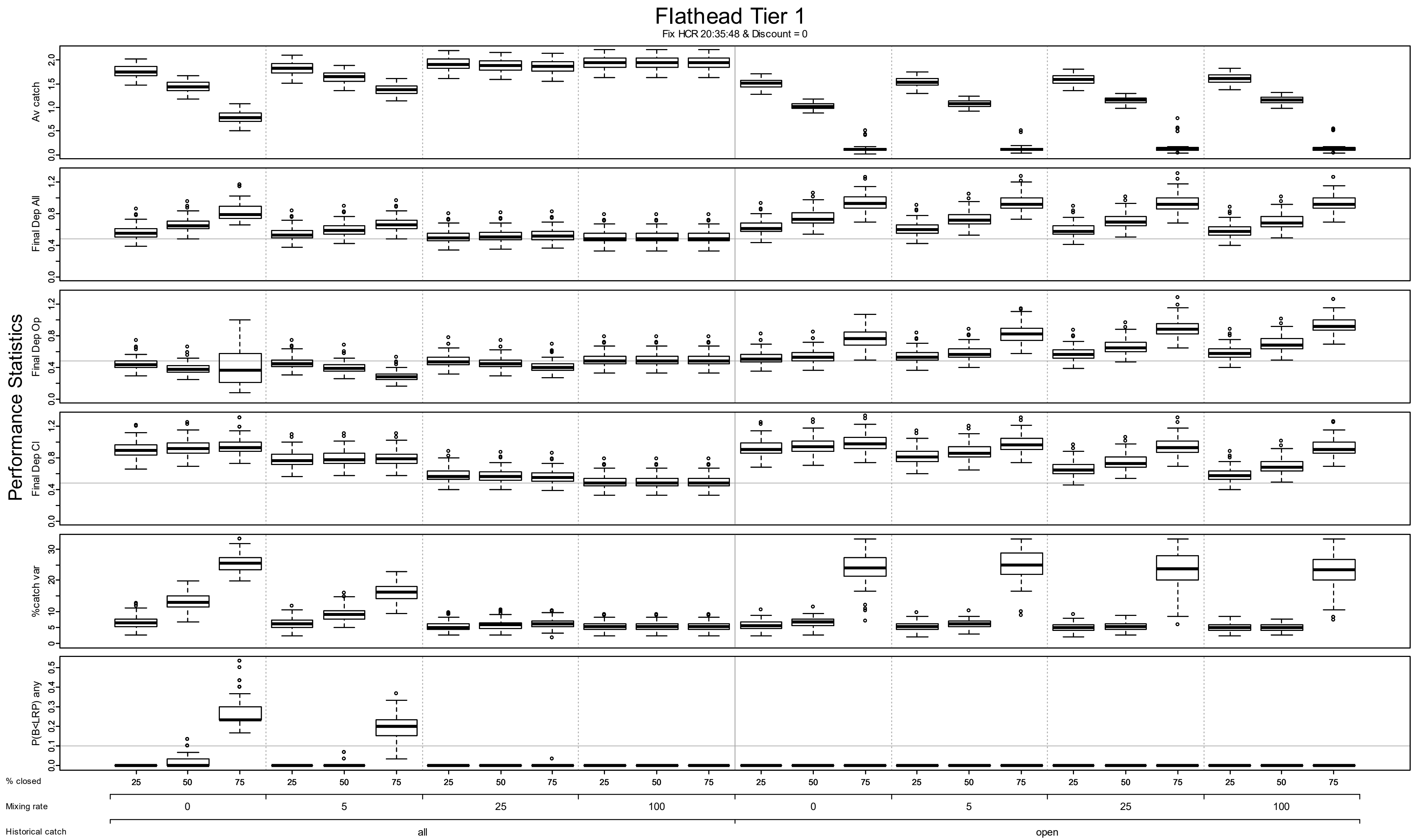


Figure 4.4: For Tiger Flathead assessed using the Tier 1 harvest strategy and a target of 48%Bo, shown are whisker plots of each of fix performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).

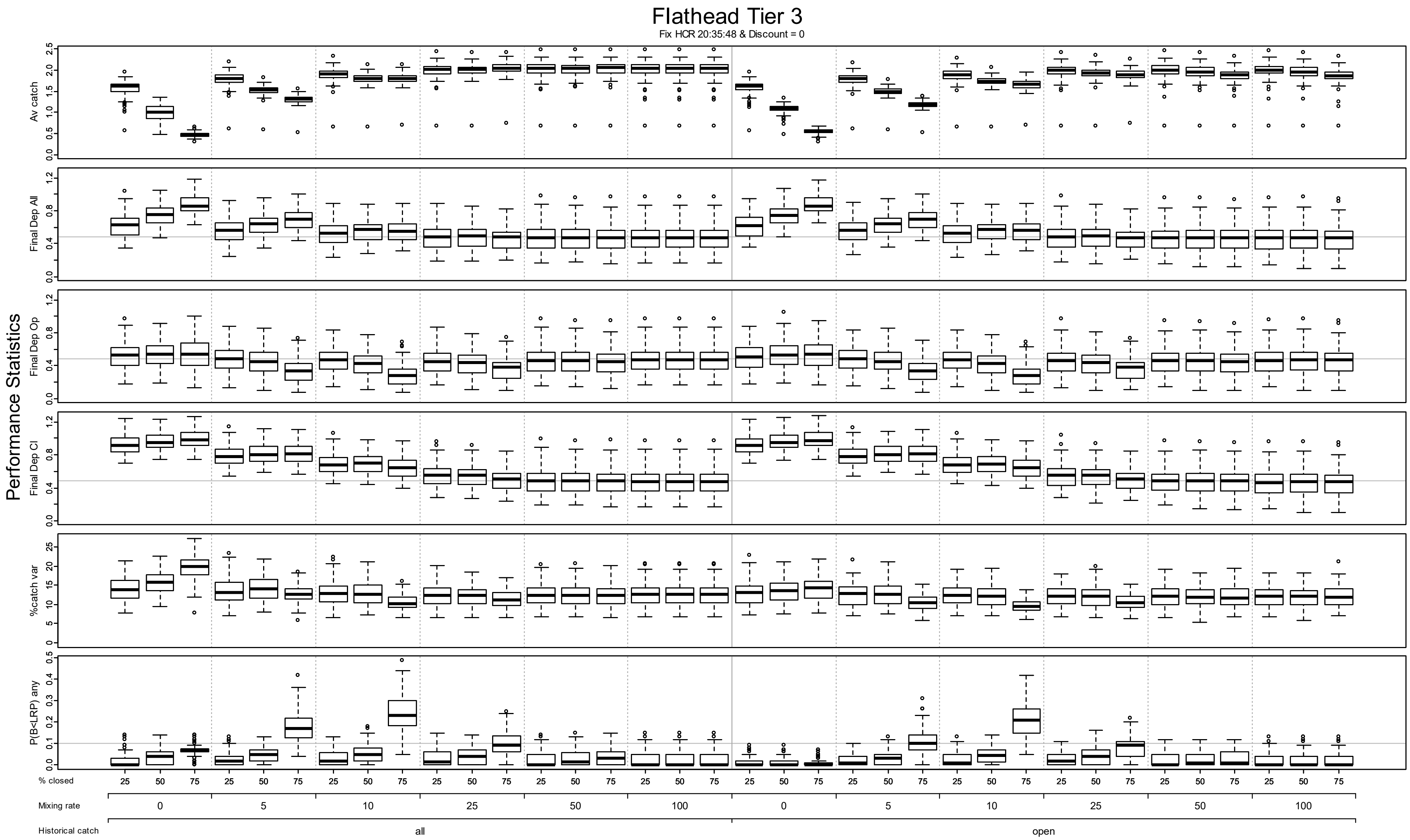


Figure 4.5: For Tiger Flathead assessed using the Tier 3 harvest strategy and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).

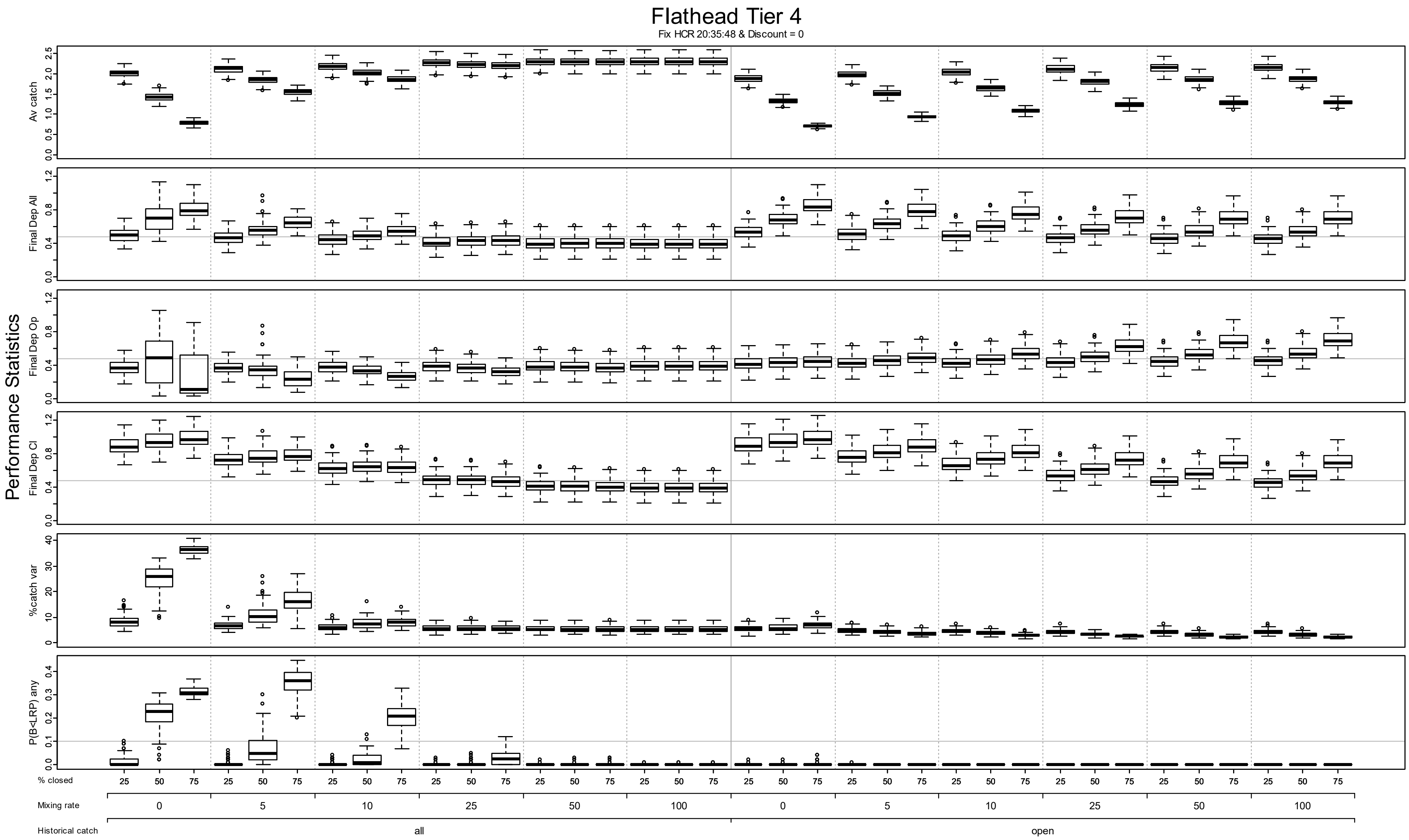


Figure 4.6: For Tiger Flathead assessed using the Tier 4 harvest strategy and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).

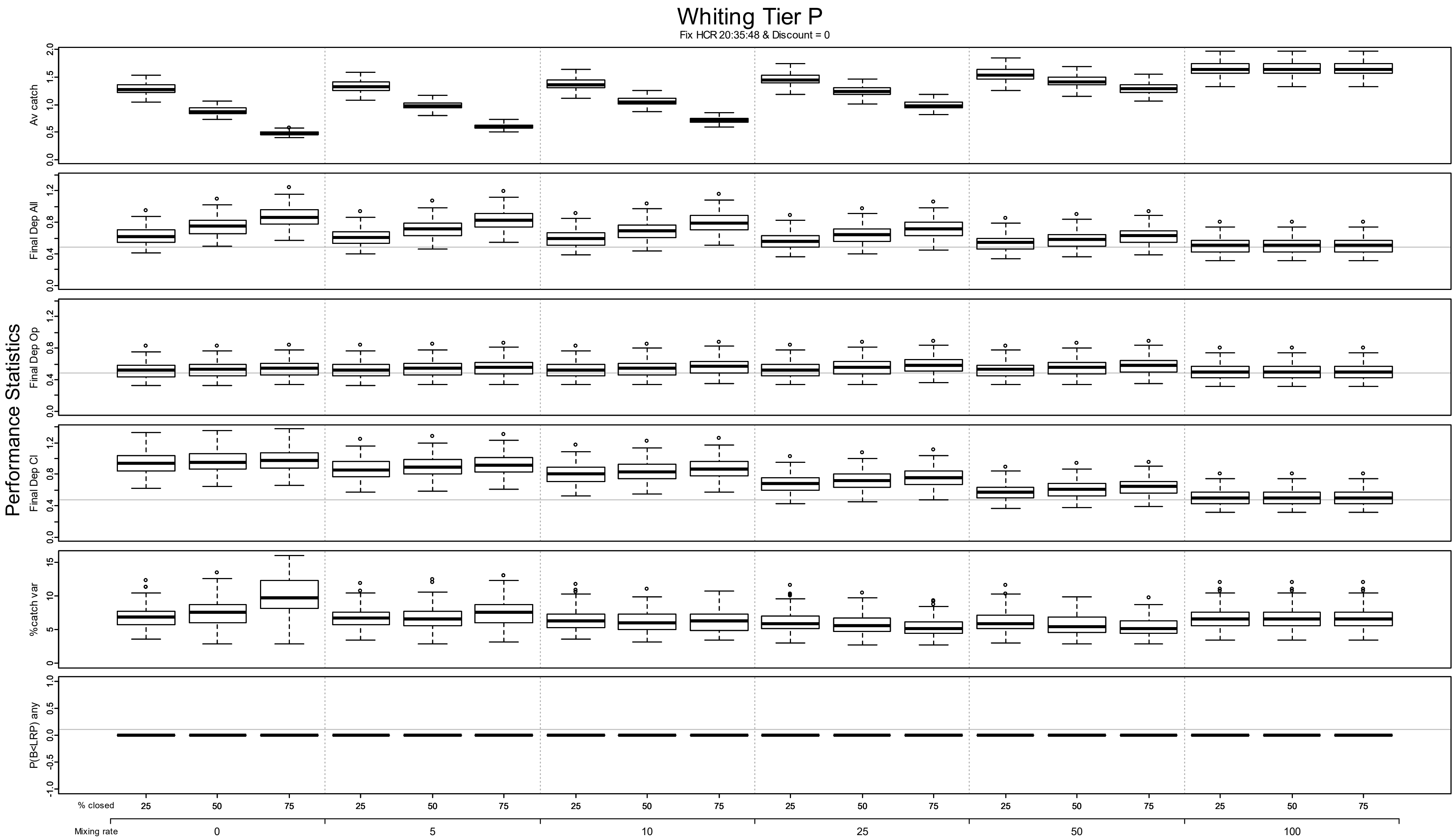


Figure 4.7: For school whiting with perfect information of biomass (Tier P) and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%) and mixing rate (0, 5%, 10%, 25%, 50%, 100%). The grey horizontal line represents the target biomass level (48%Bo) in the three relative biomass plots and is the 0.10 probability in the plot showing the probability of being below the limit reference point.

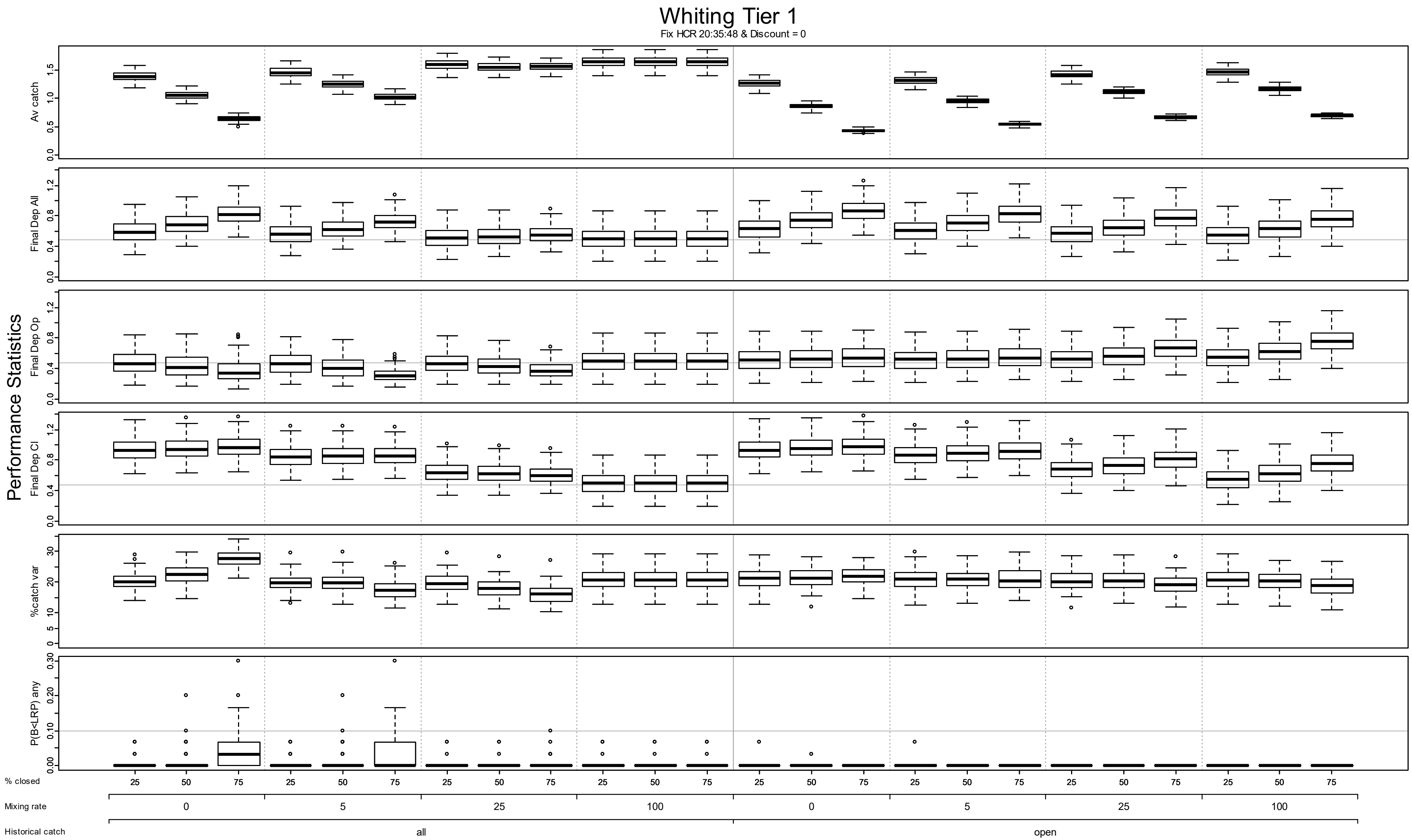


Figure 4.8: For school whiting assessed using the Tier 1 harvest strategy and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).

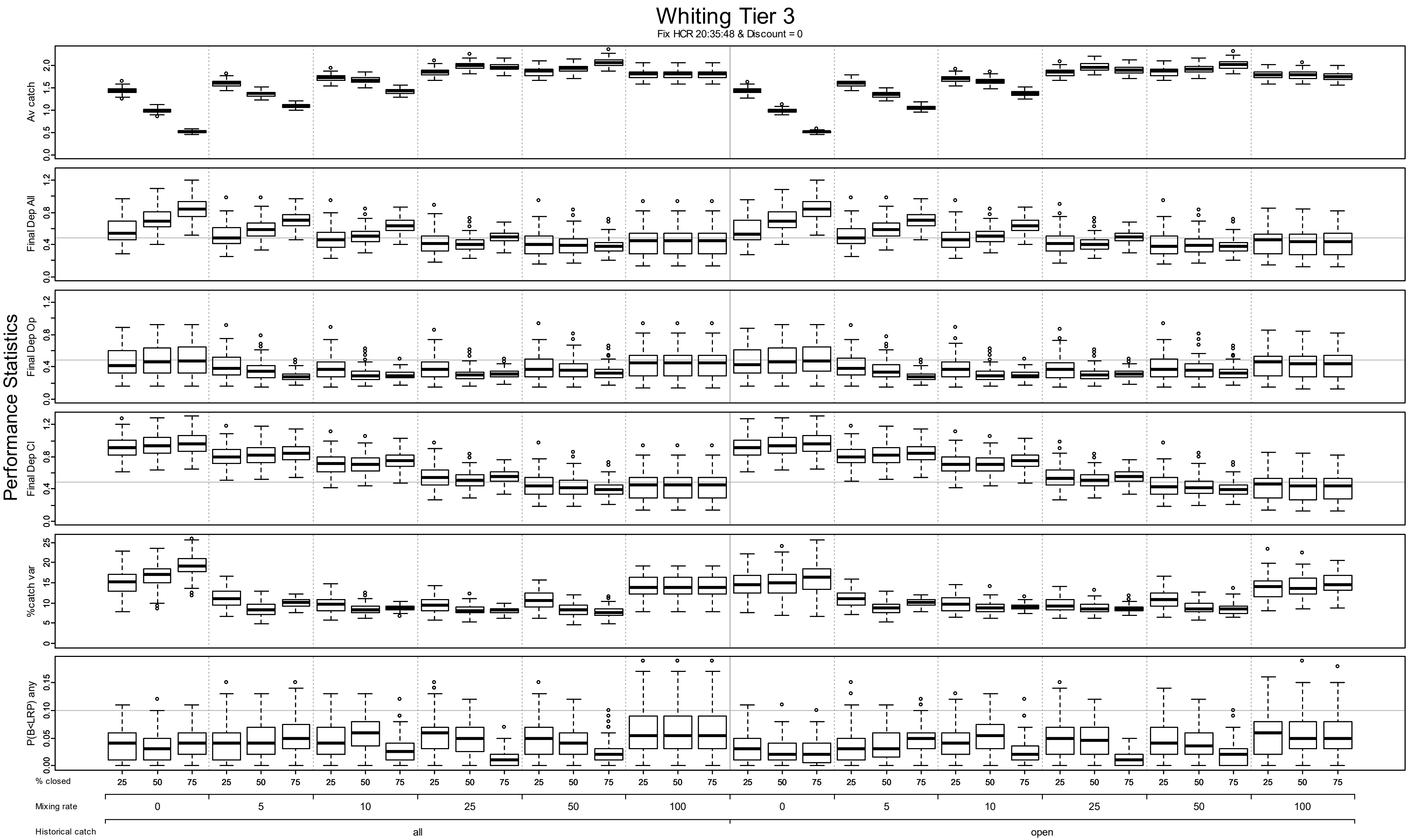


Figure 4.9: For school whiting assessed using the Tier 3 harvest strategy and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).

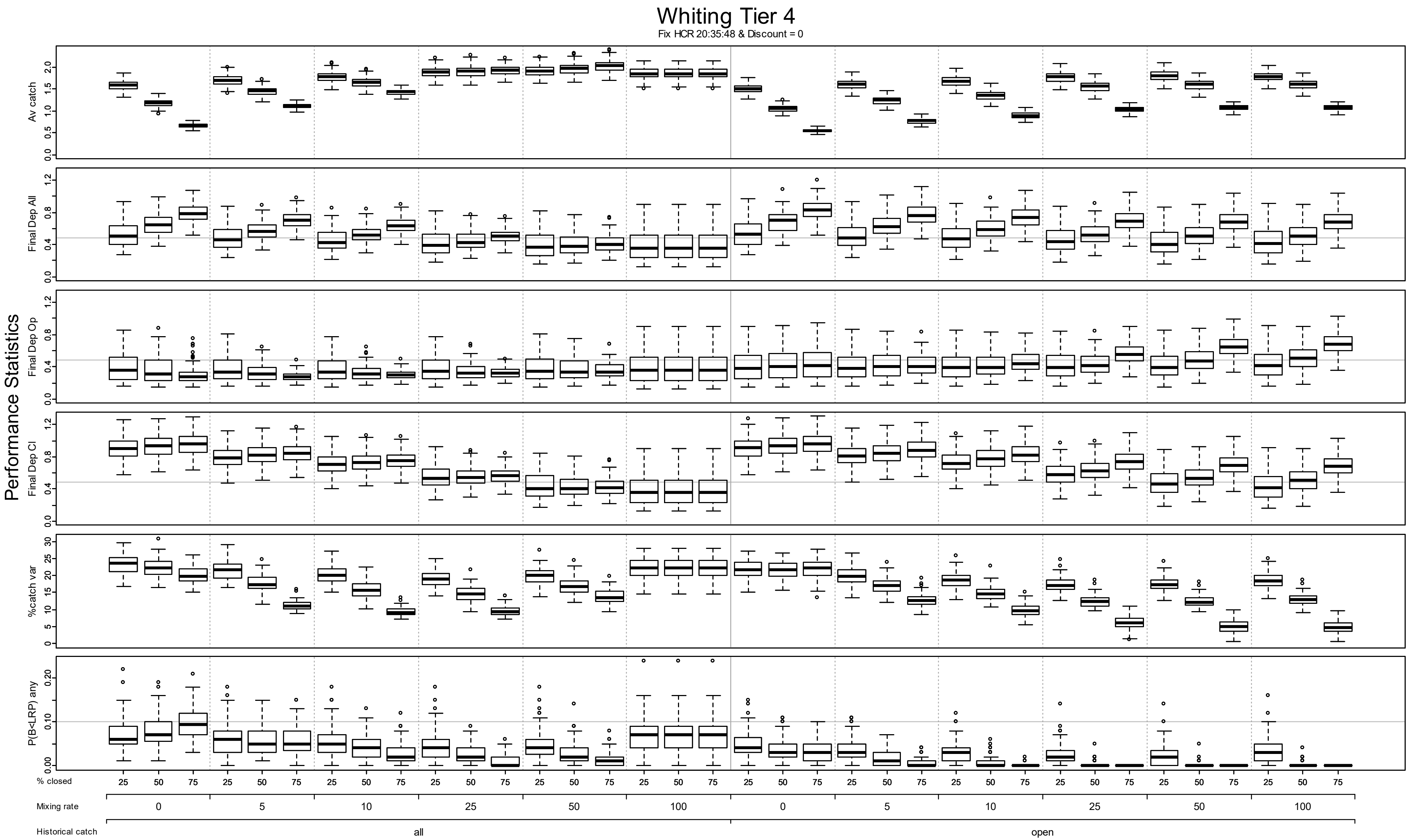
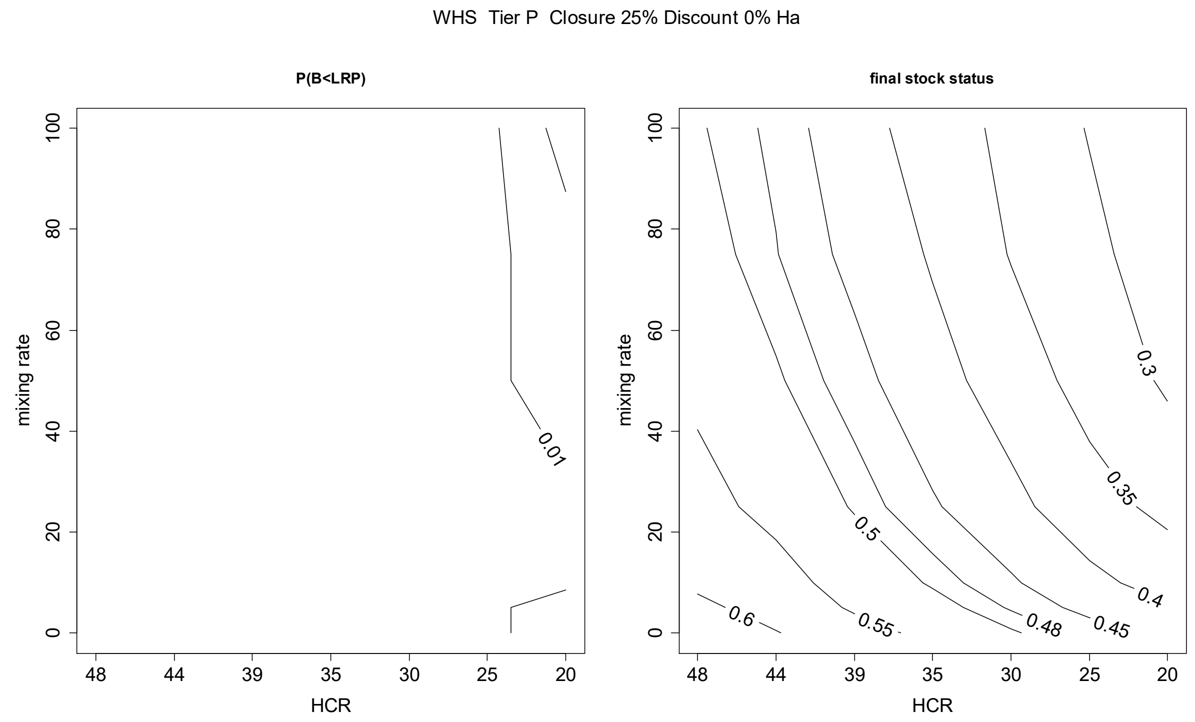


Figure 4.10: For school whiting assessed using the Tier 4 harvest strategy and a target of 48%Bo, shown are whisker plots of each of 5 performance measures under differing closure size (25%, 50%, 75%), mixing rate (0, 5%, 10%, 25%, 50%, 100%) and data used (all data or only from the open area).



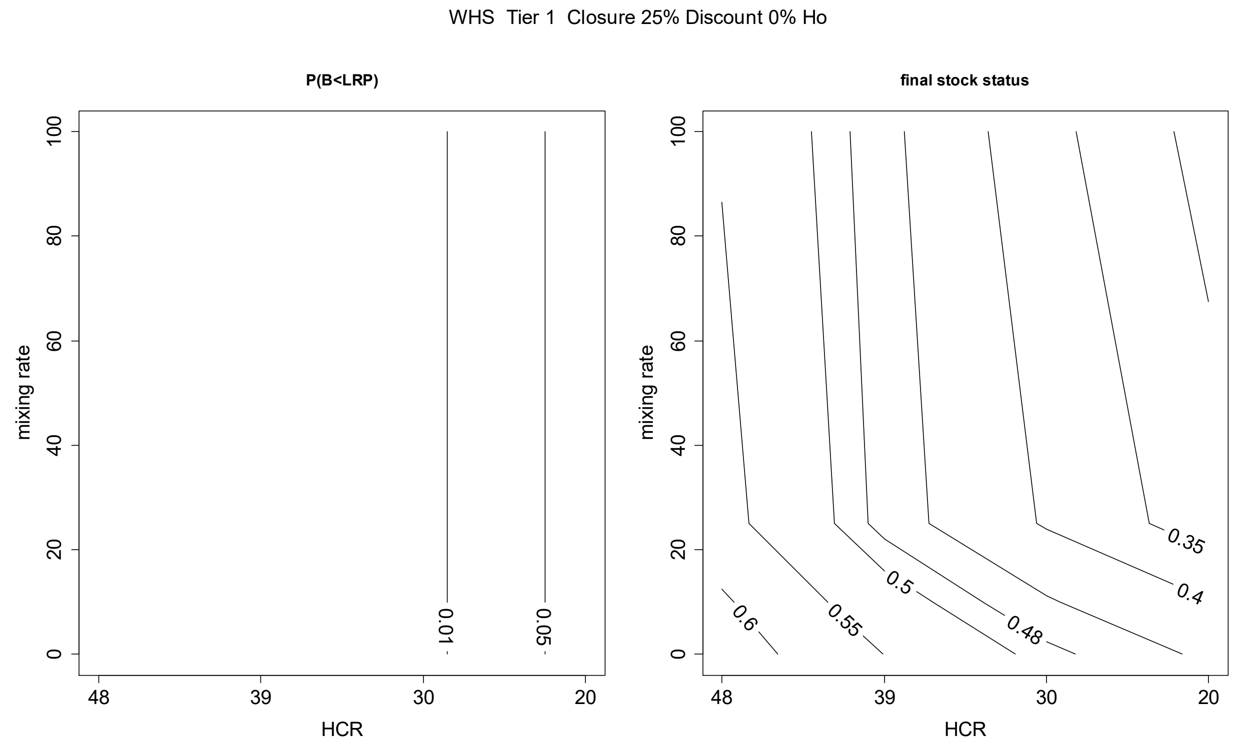


Figure 4.11: The probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

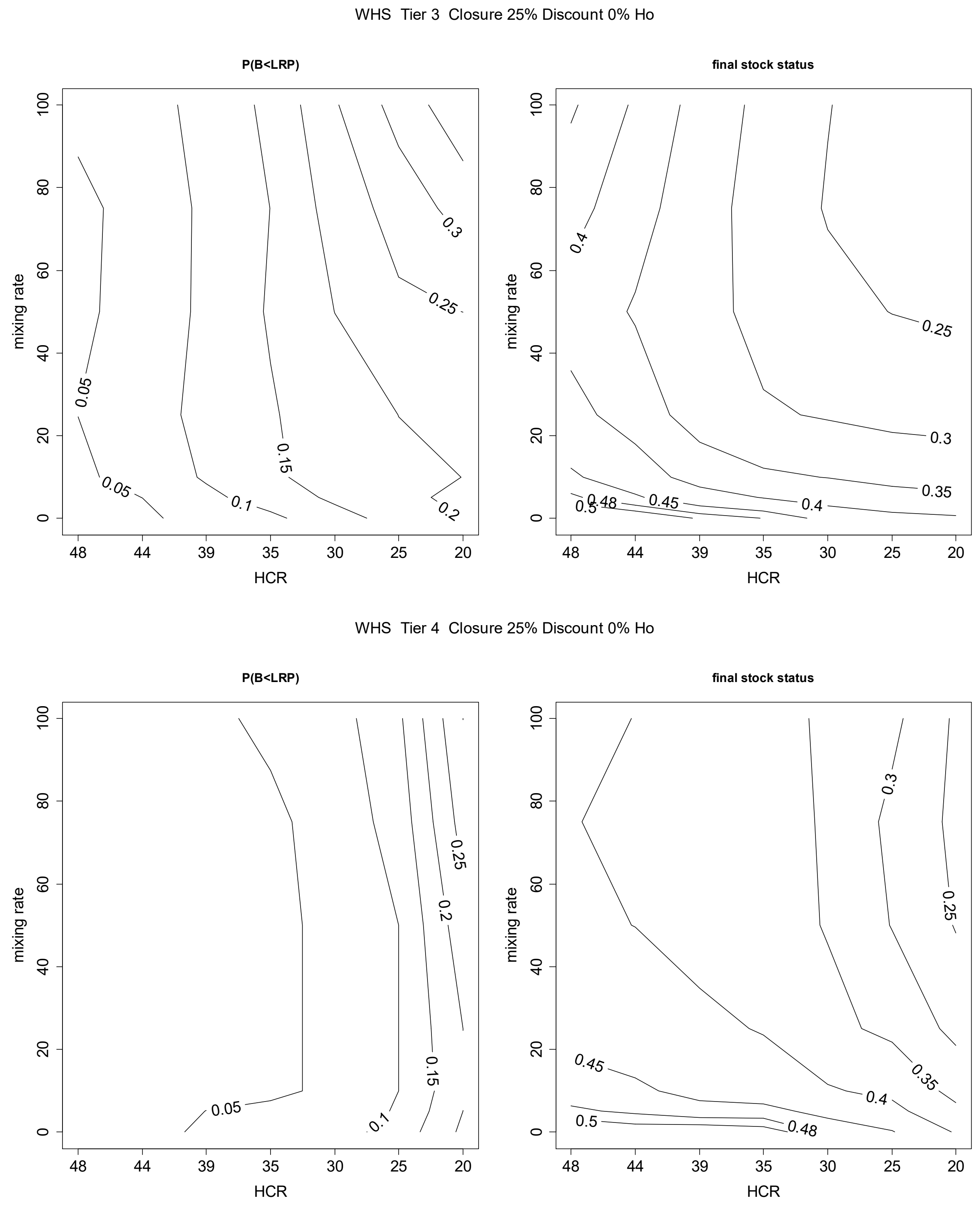


Figure 4.12: The probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting assessed with Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

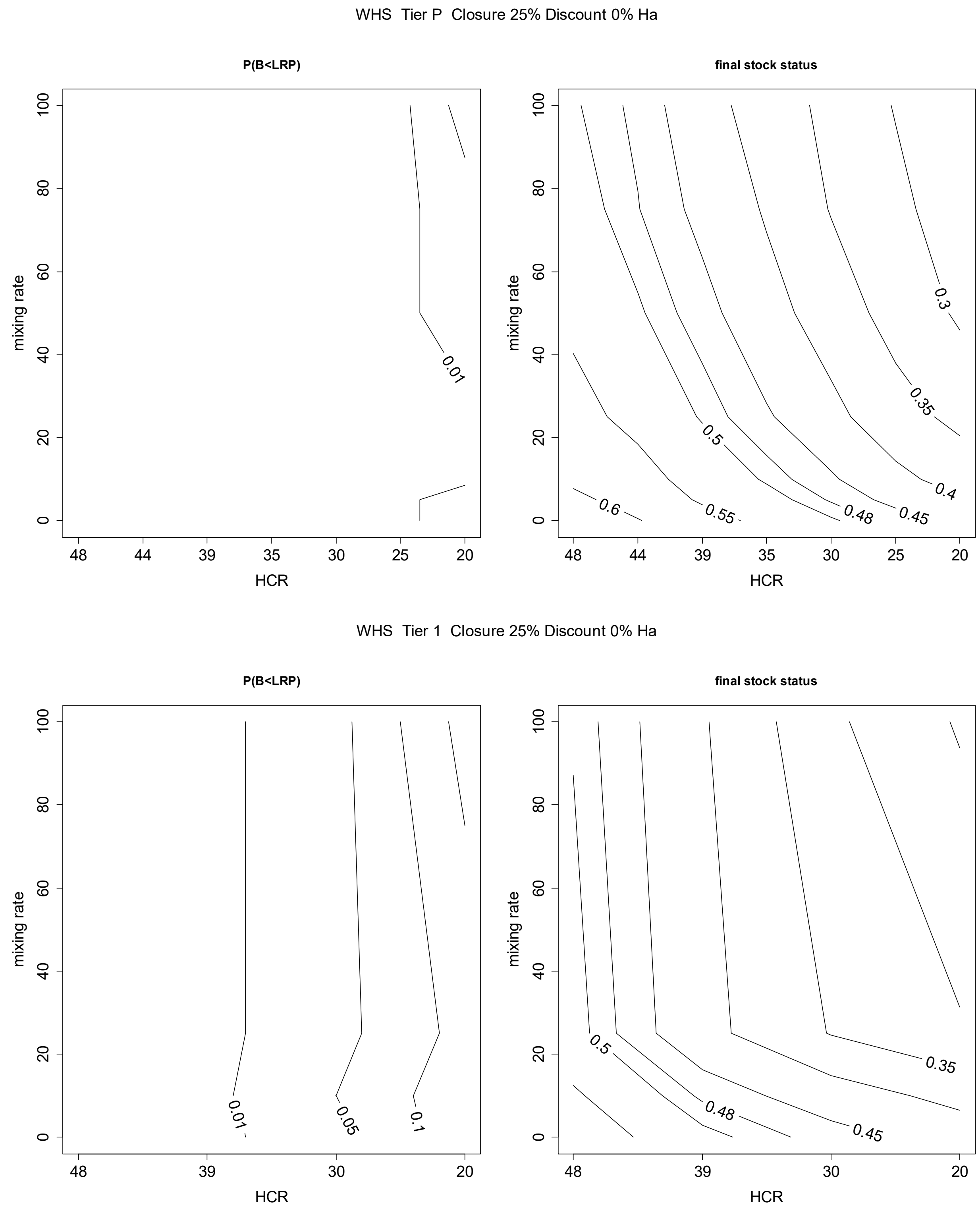


Figure 4.13: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).

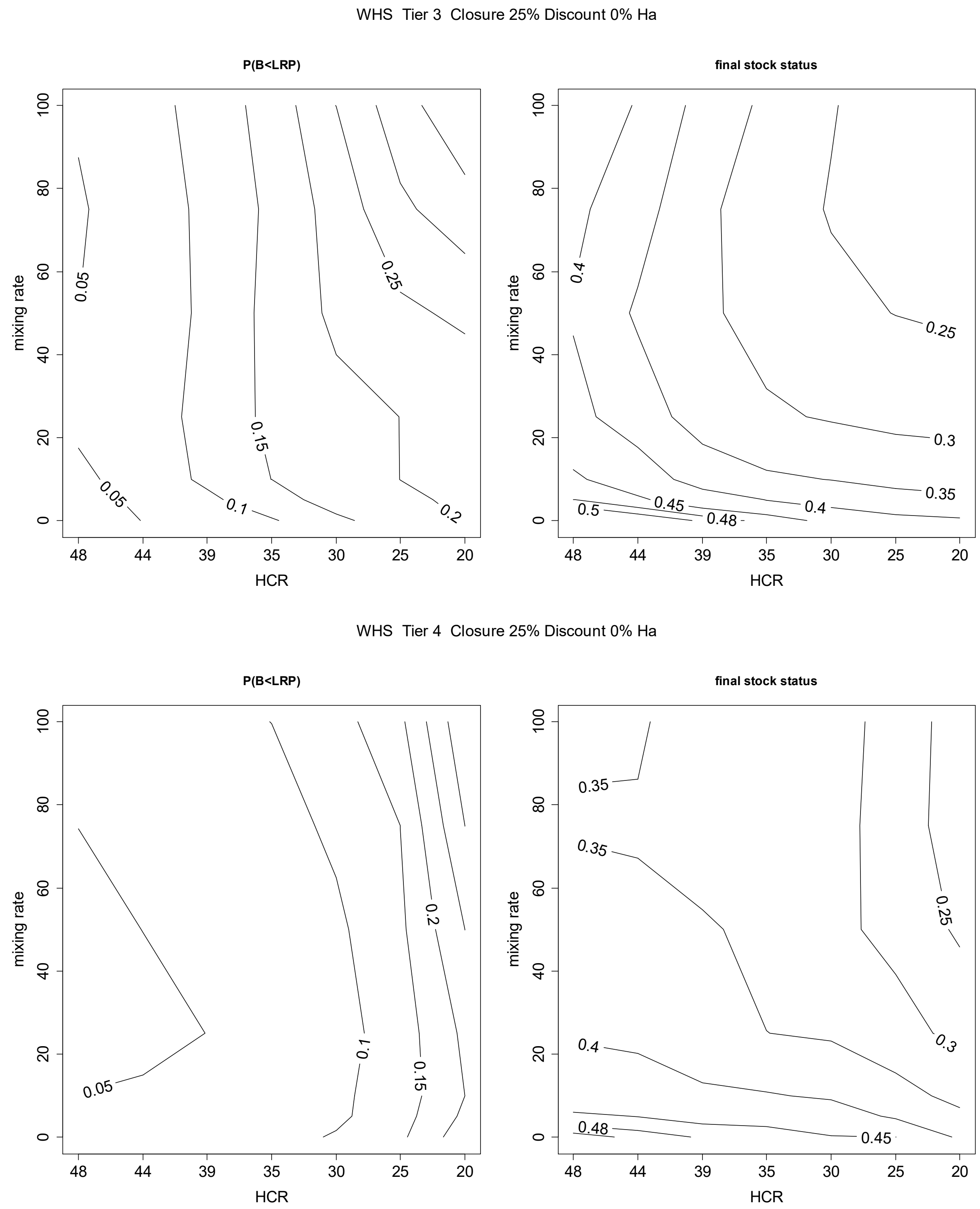
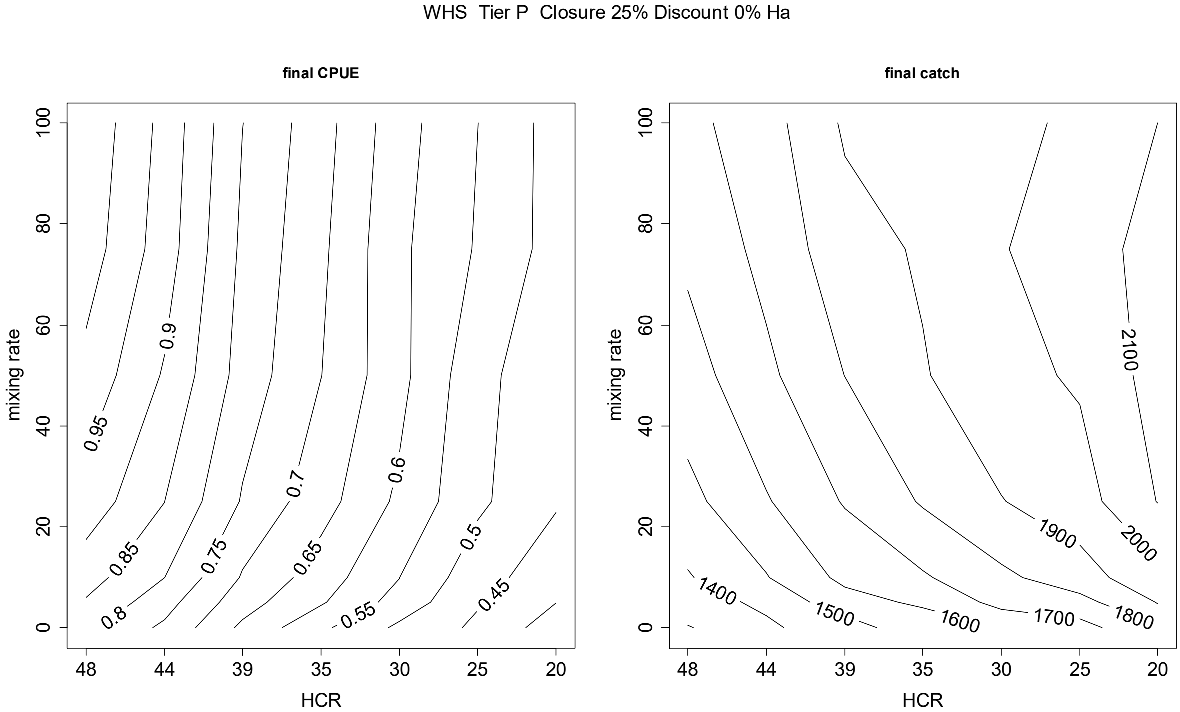


Figure 4.14: The probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting assessed with Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).



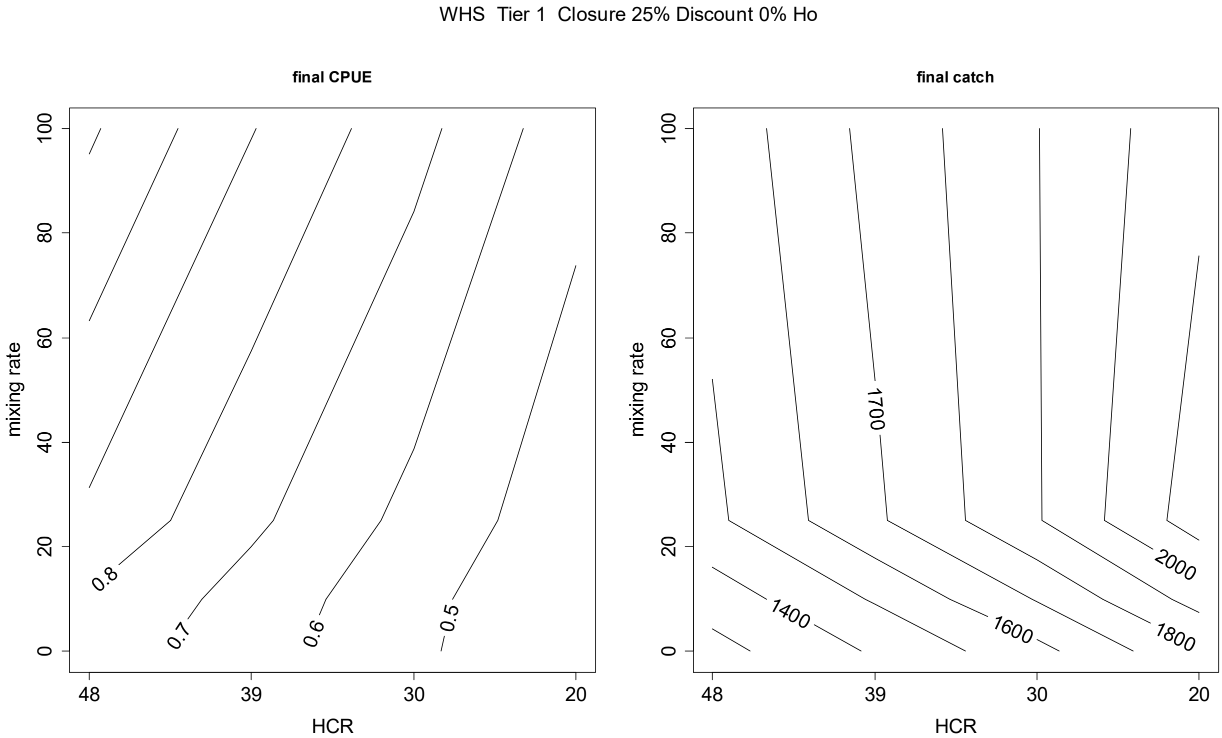


Figure 4.15: Final CPUE (left) and catch (right) contours for school whiting assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

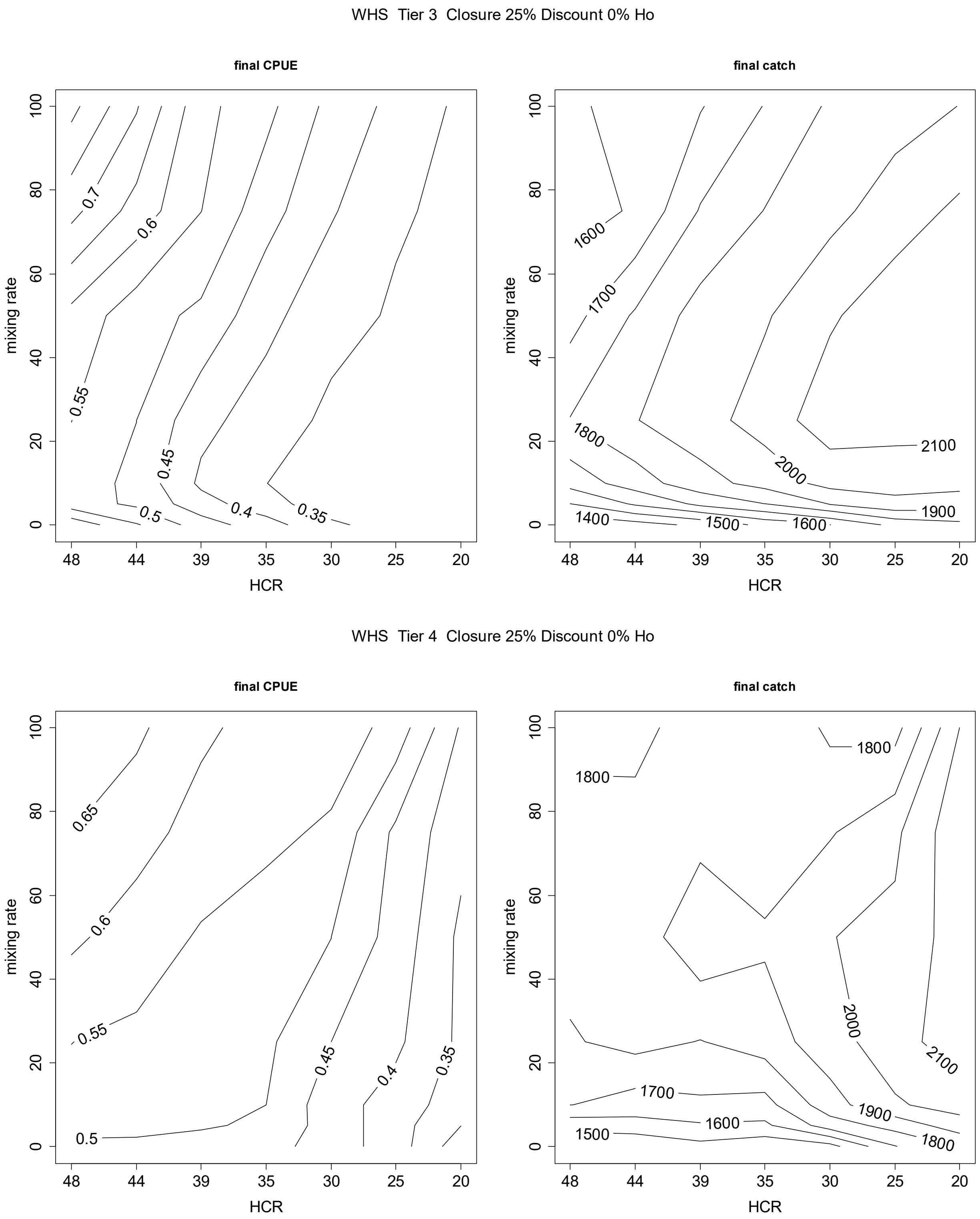


Figure 4.16: Final CPUE (left) and catch (right) contours for school whiting assessed with the Tier 3(top) and Tier 4 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

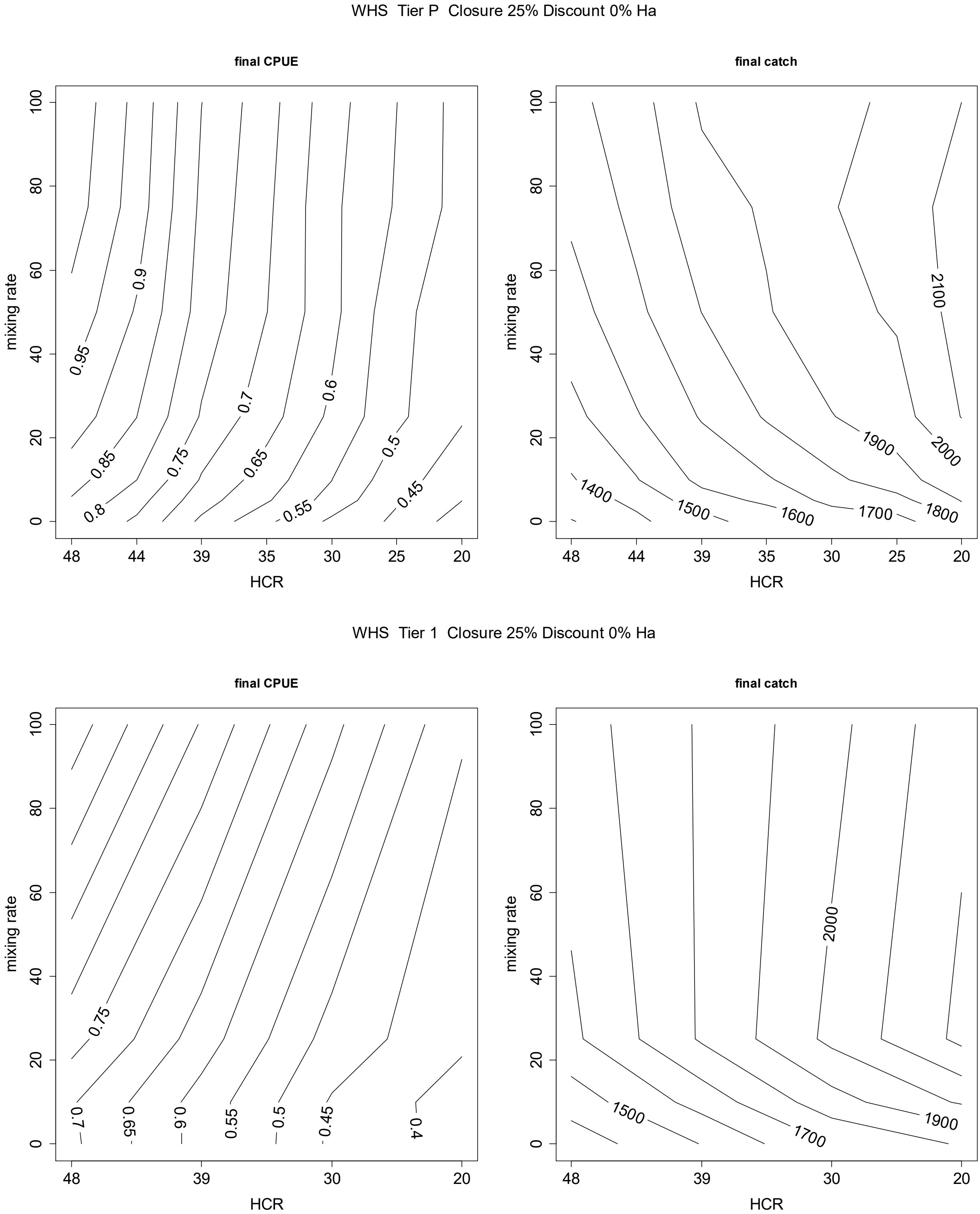


Figure 4.17: Final CPUE (left) and catch (right) contours for school whiting assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).

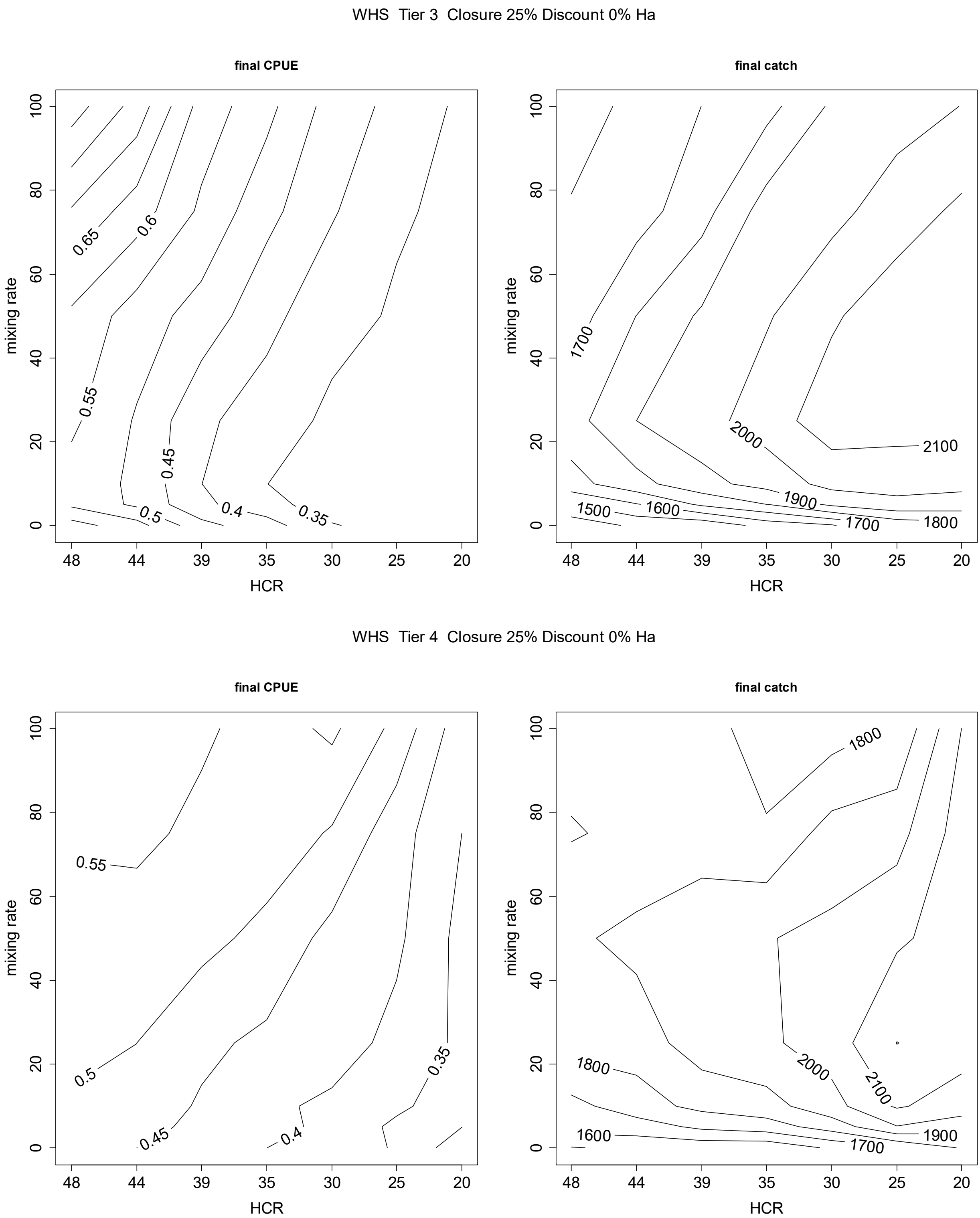


Figure 4.18: Final CPUE (left) and catch (right) contours for school whiting assessed with the Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).

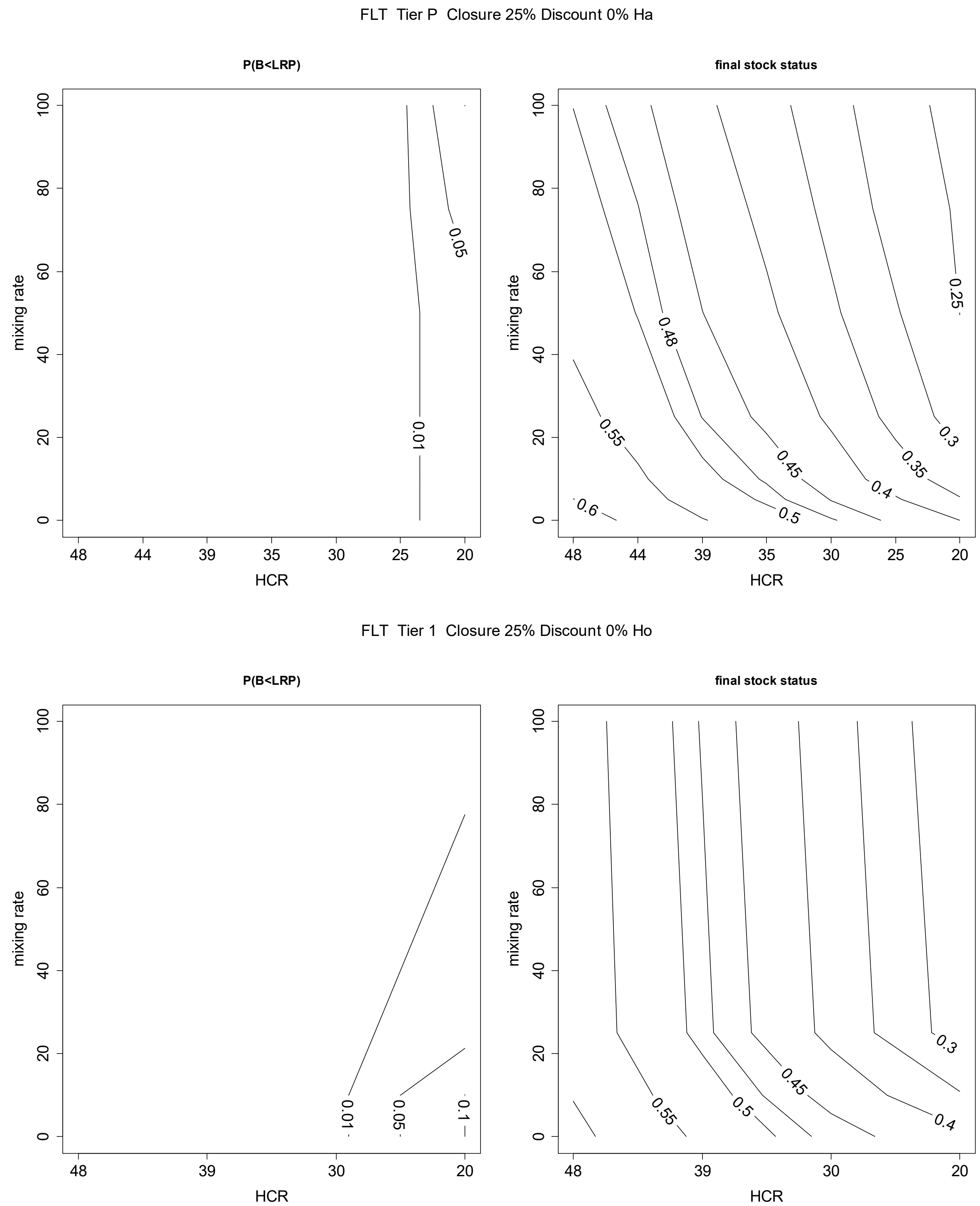


Figure 4.19: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

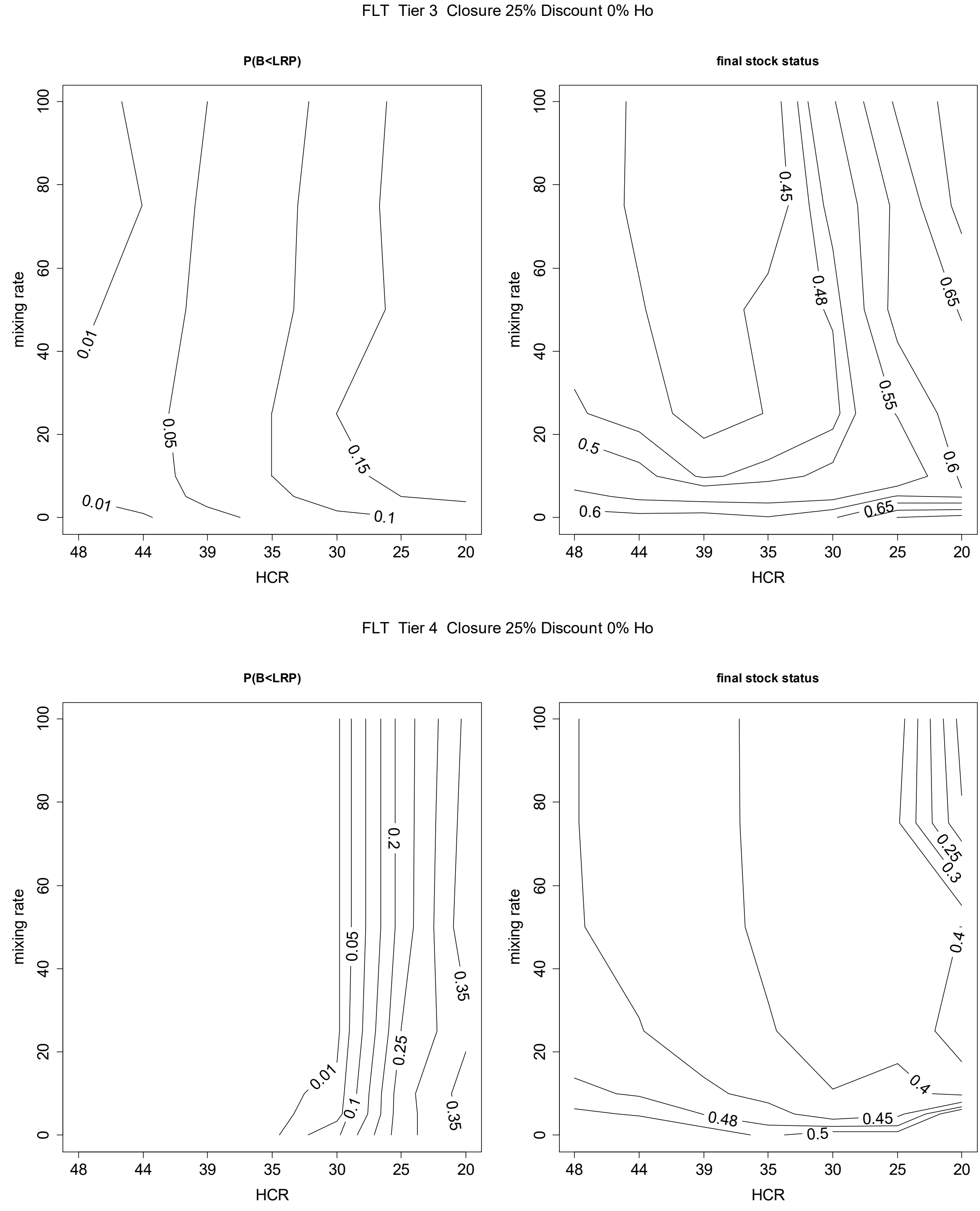


Figure 4.20: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead assessed with Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

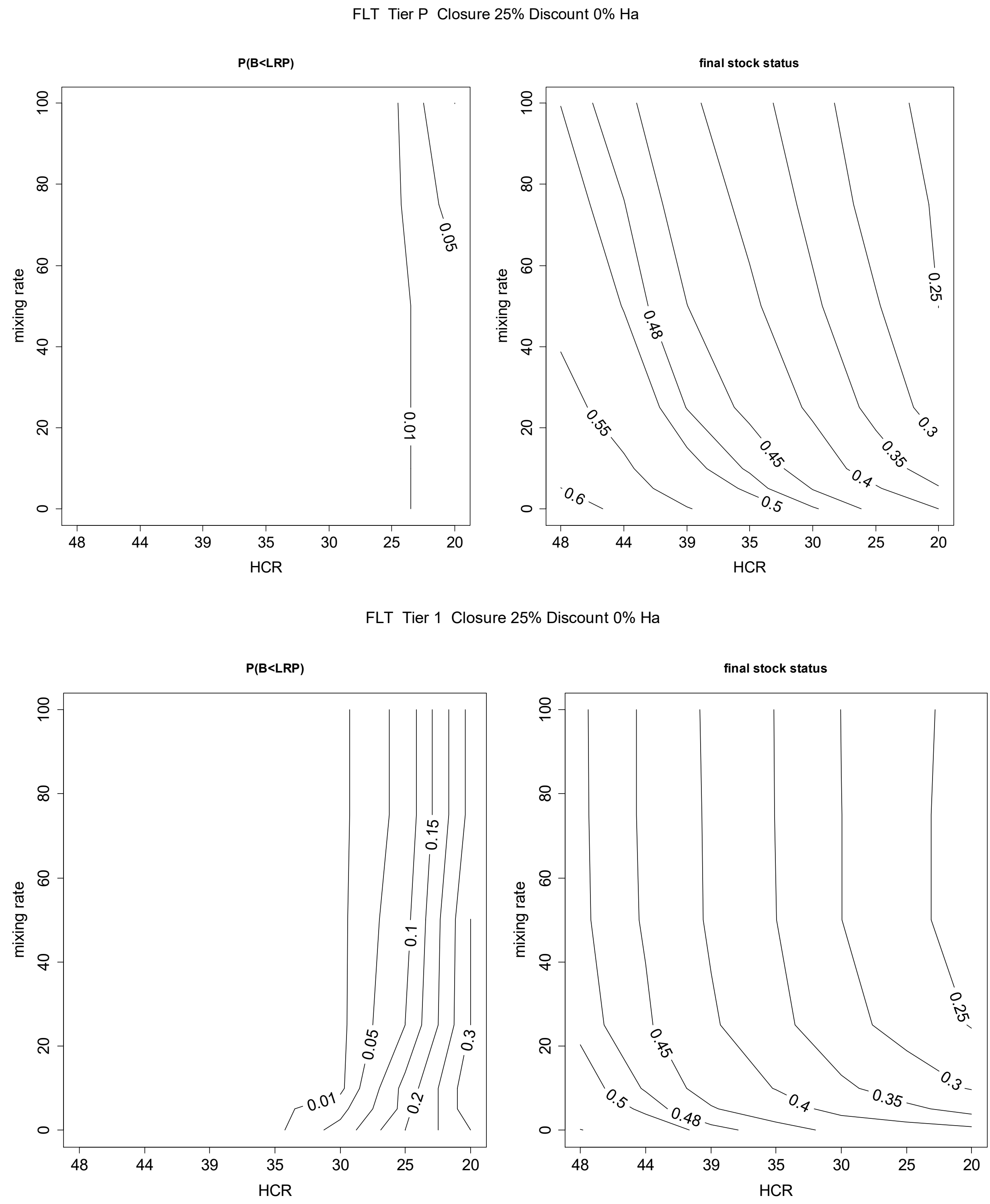
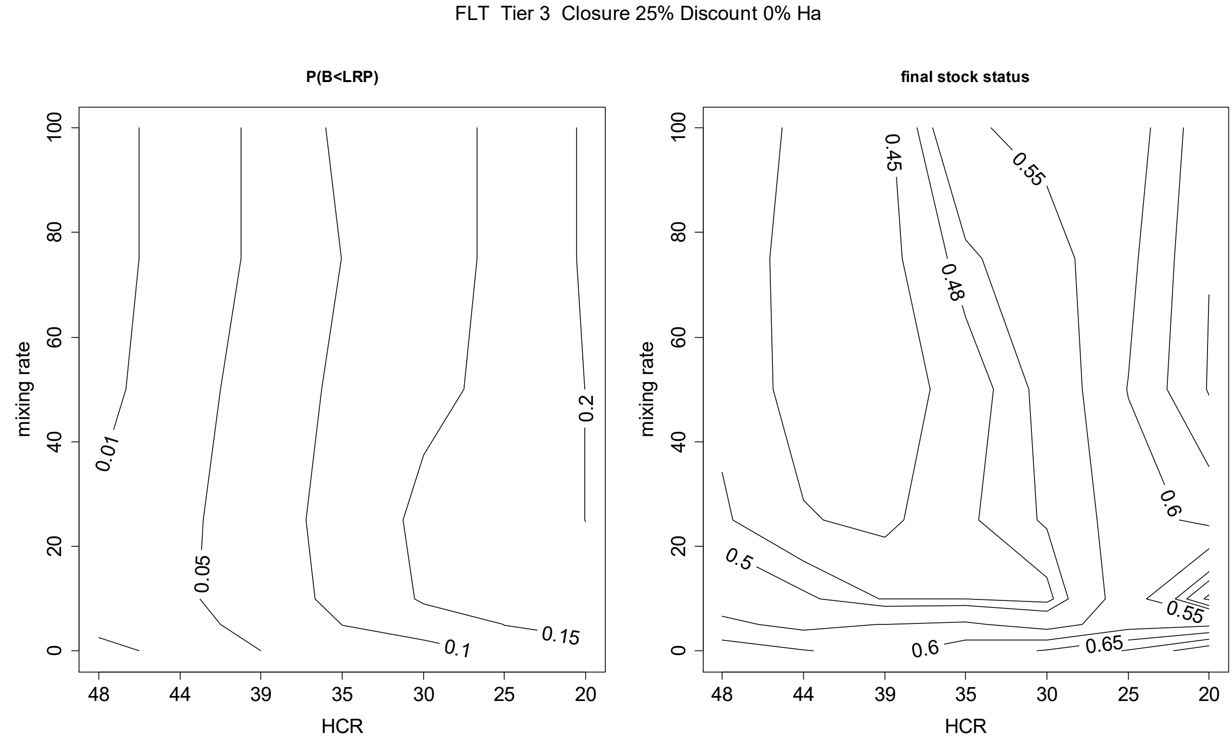


Figure 4.21: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Historical data from the closed and open areas used in the assessment (Ha).



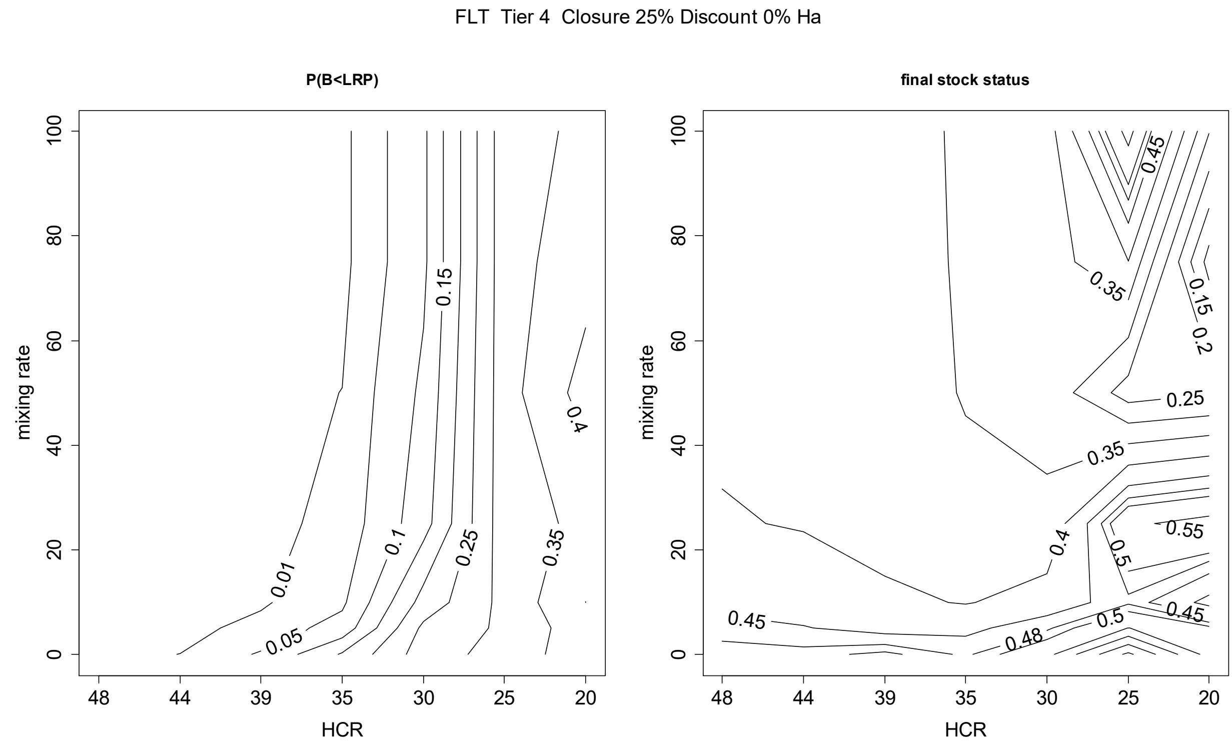


Figure 4.22: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead assessed with Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Data from the closed and open areas used in the assessment (Ha).

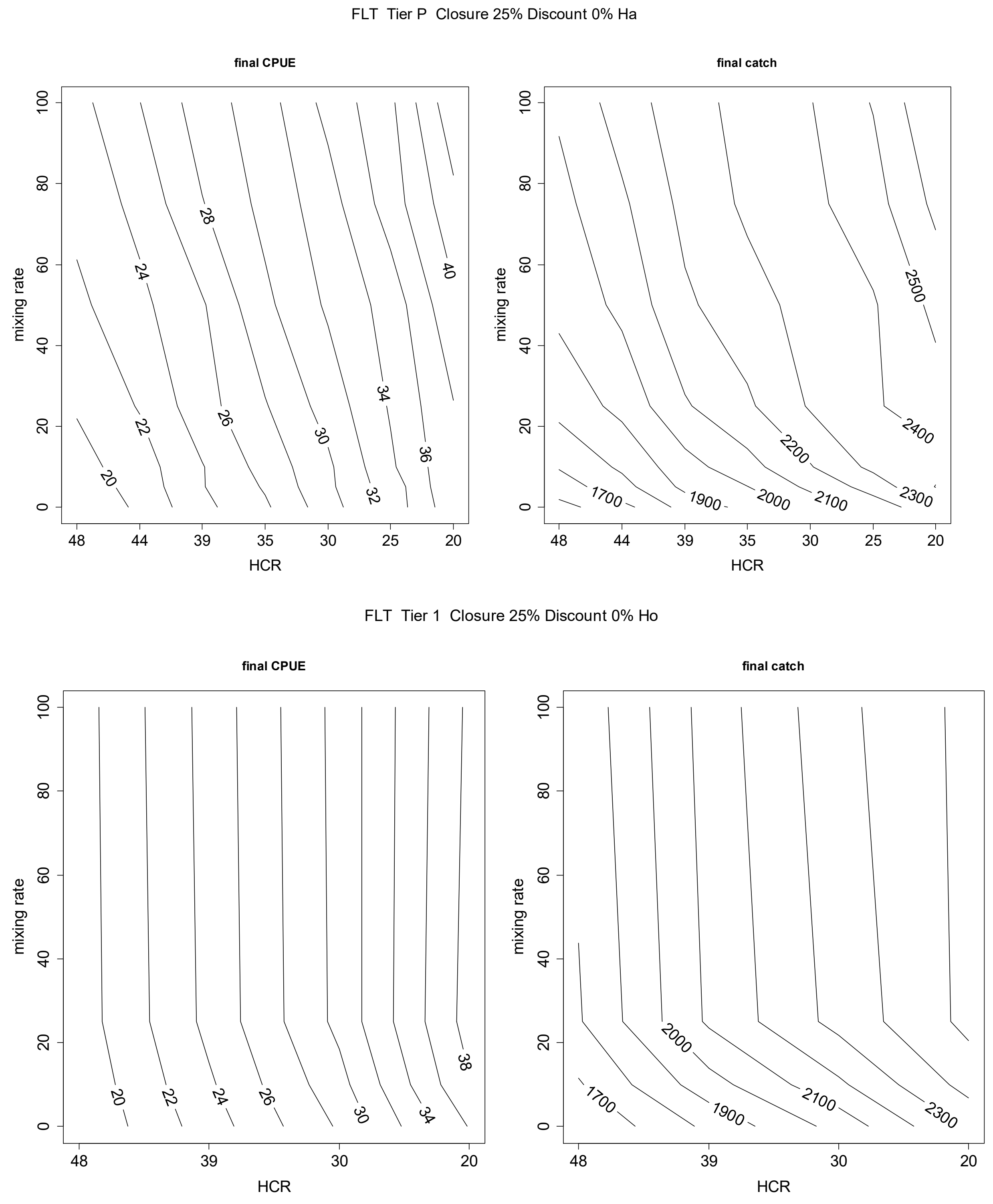


Figure 4.23: Final CPUE (left) and catch (right) contours for Tiger Flathead assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

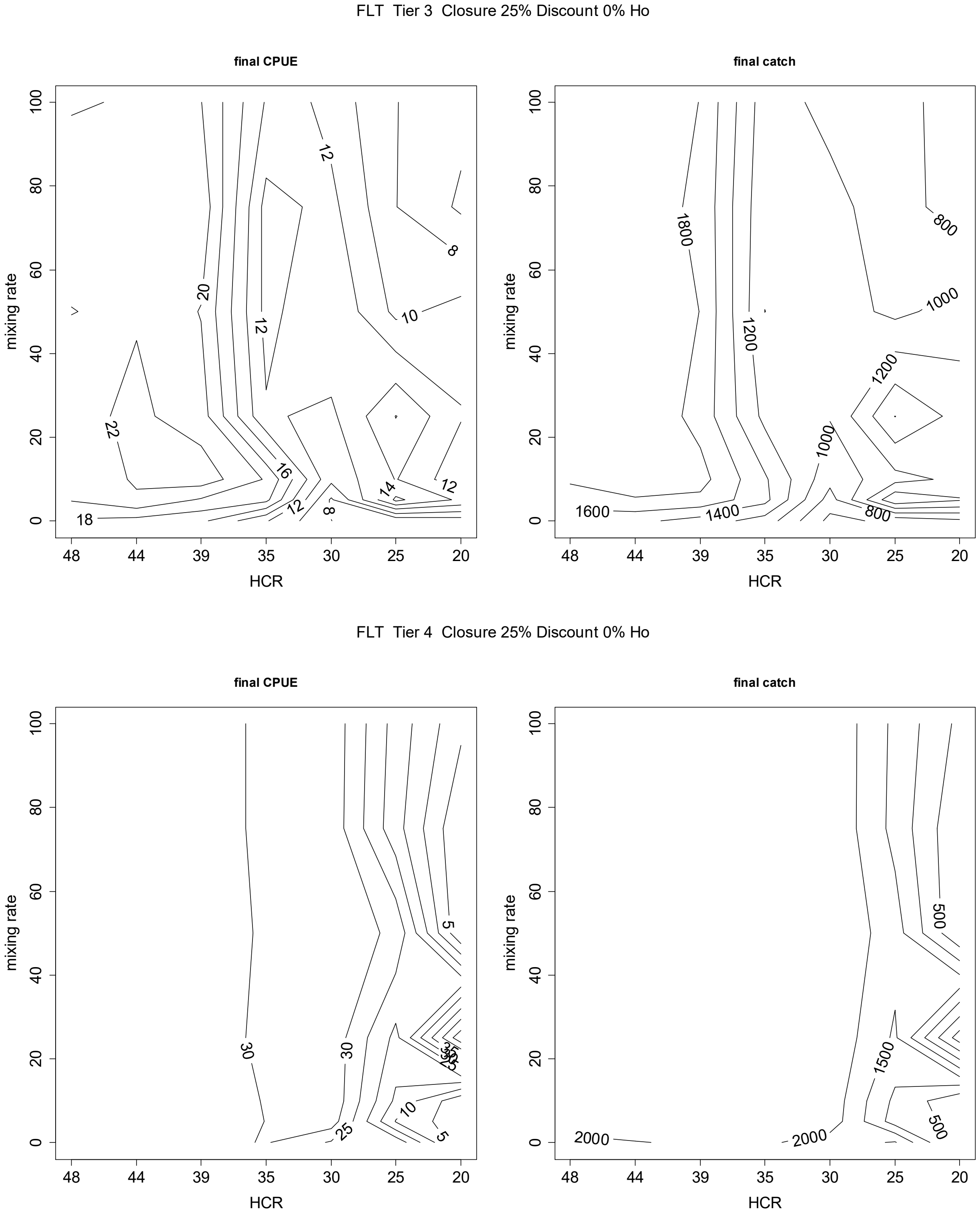


Figure 4.24: Final CPUE (left) and catch (right) contours for Tiger Flathead assessed with the Tier 3(top) and Tier 4 (bottom) harvest strategies with a 25% closure. Only data from the open area used in the assessment (Ho).

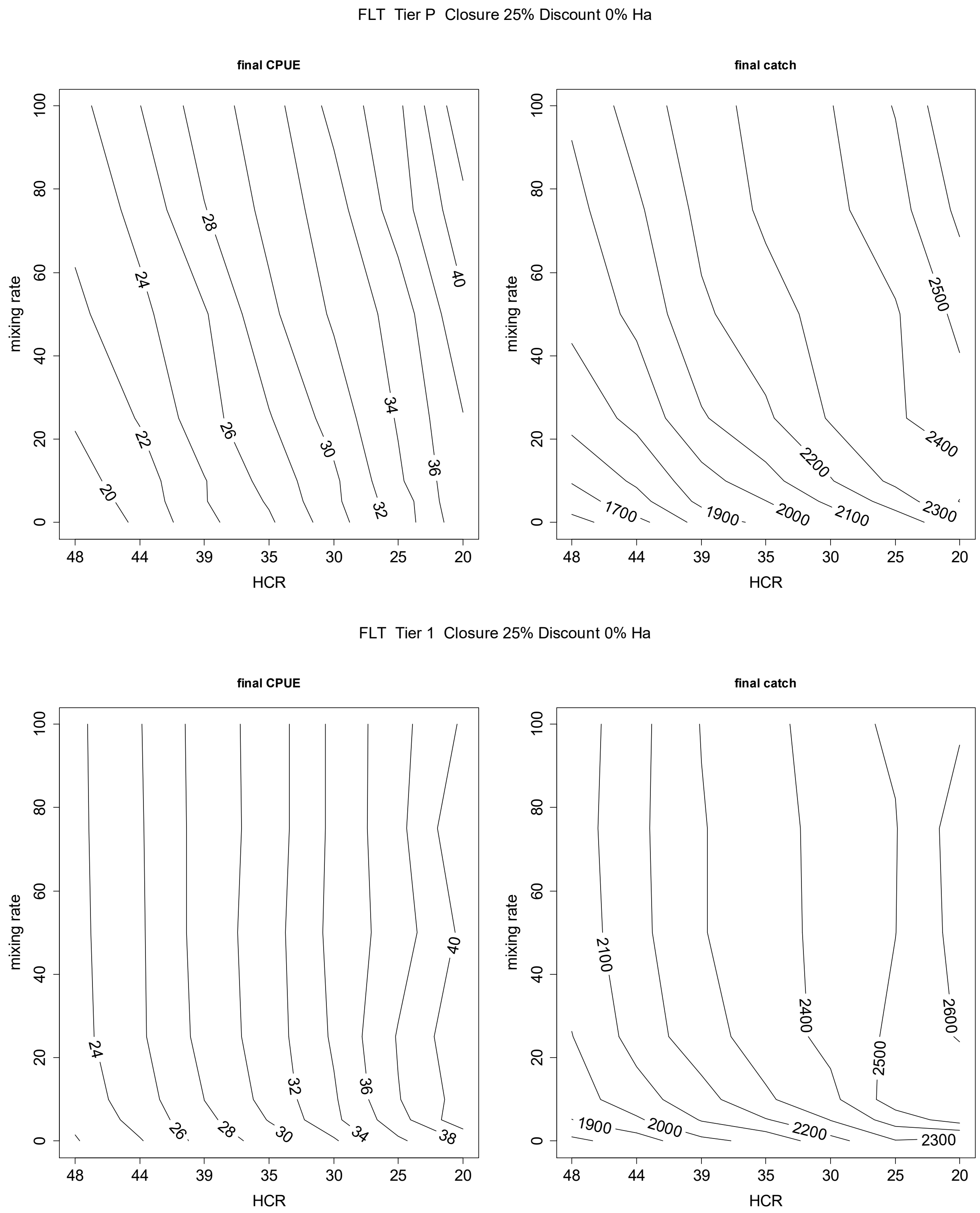
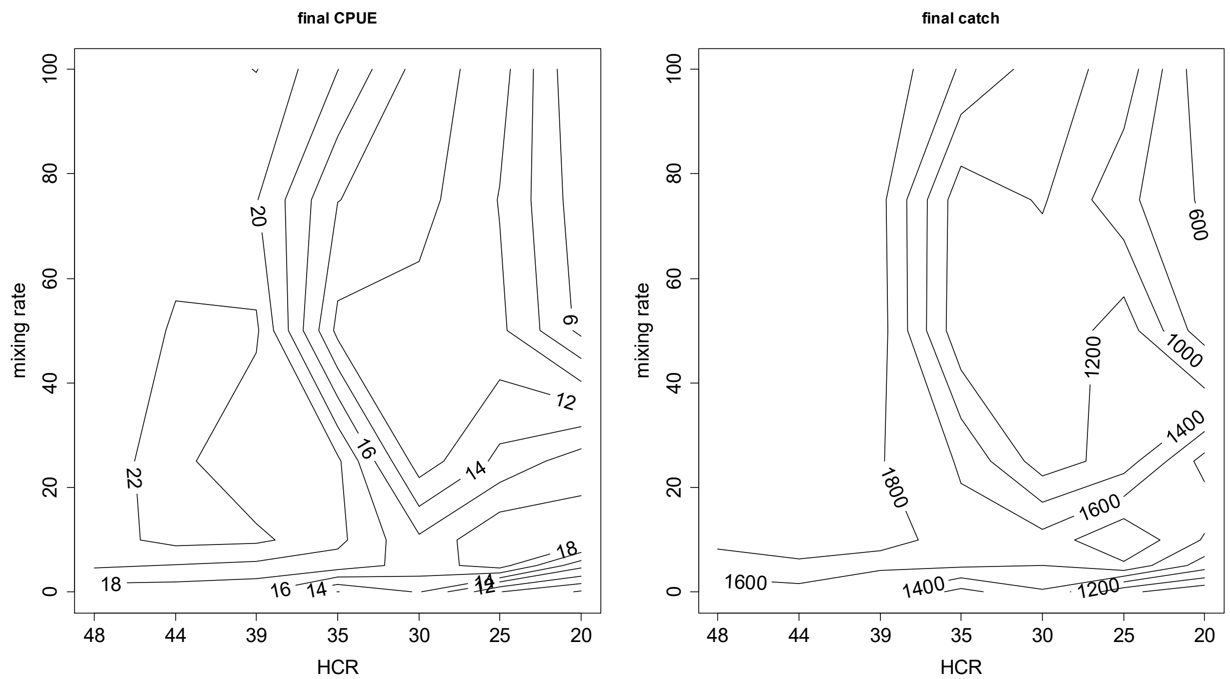


Figure 4.25: Final CPUE (left) and catch (right) contours for Tiger Flathead assessed with perfect information (Tier P; top) and Tier 1 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).

FLT Tier 3 Closure 25% Discount 0% Ha



FLT Tier 4 Closure 25% Discount 0% Ha

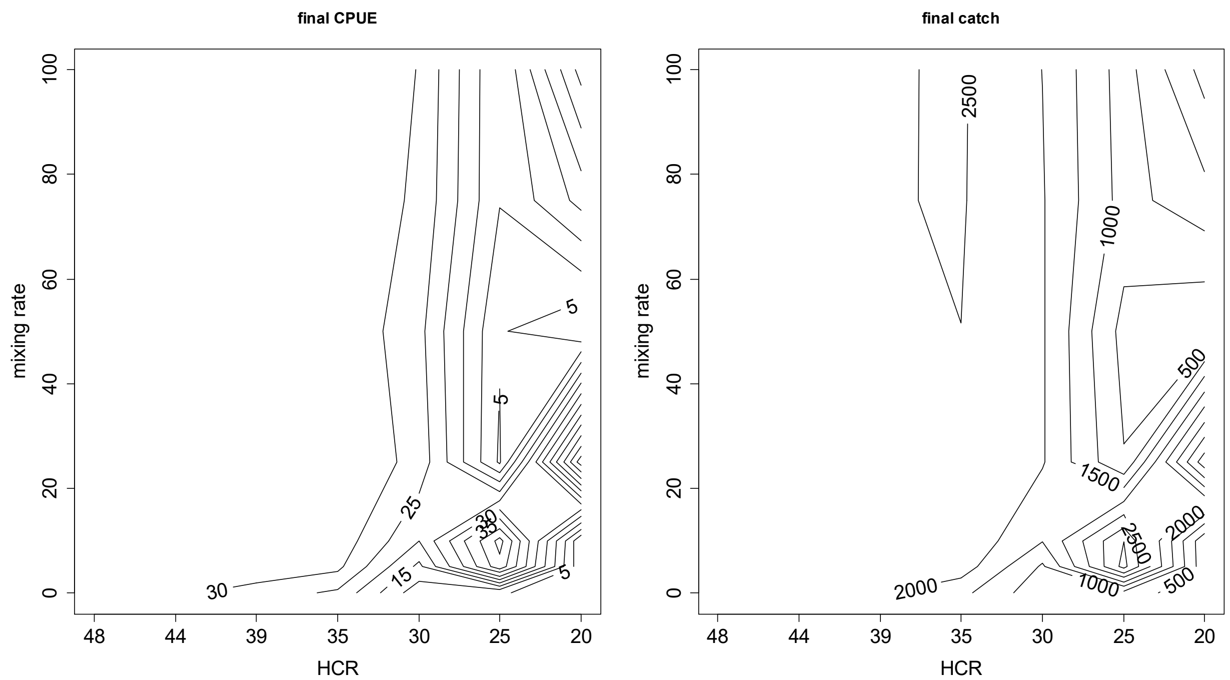


Figure 4.26: Final CPUE (left) and catch (right) contours for Tiger Flathead assessed with the Tier 3 (top) and Tier 4 (bottom) harvest strategies with a 25% closure. Data from open and closed areas used in the assessment (Ha).

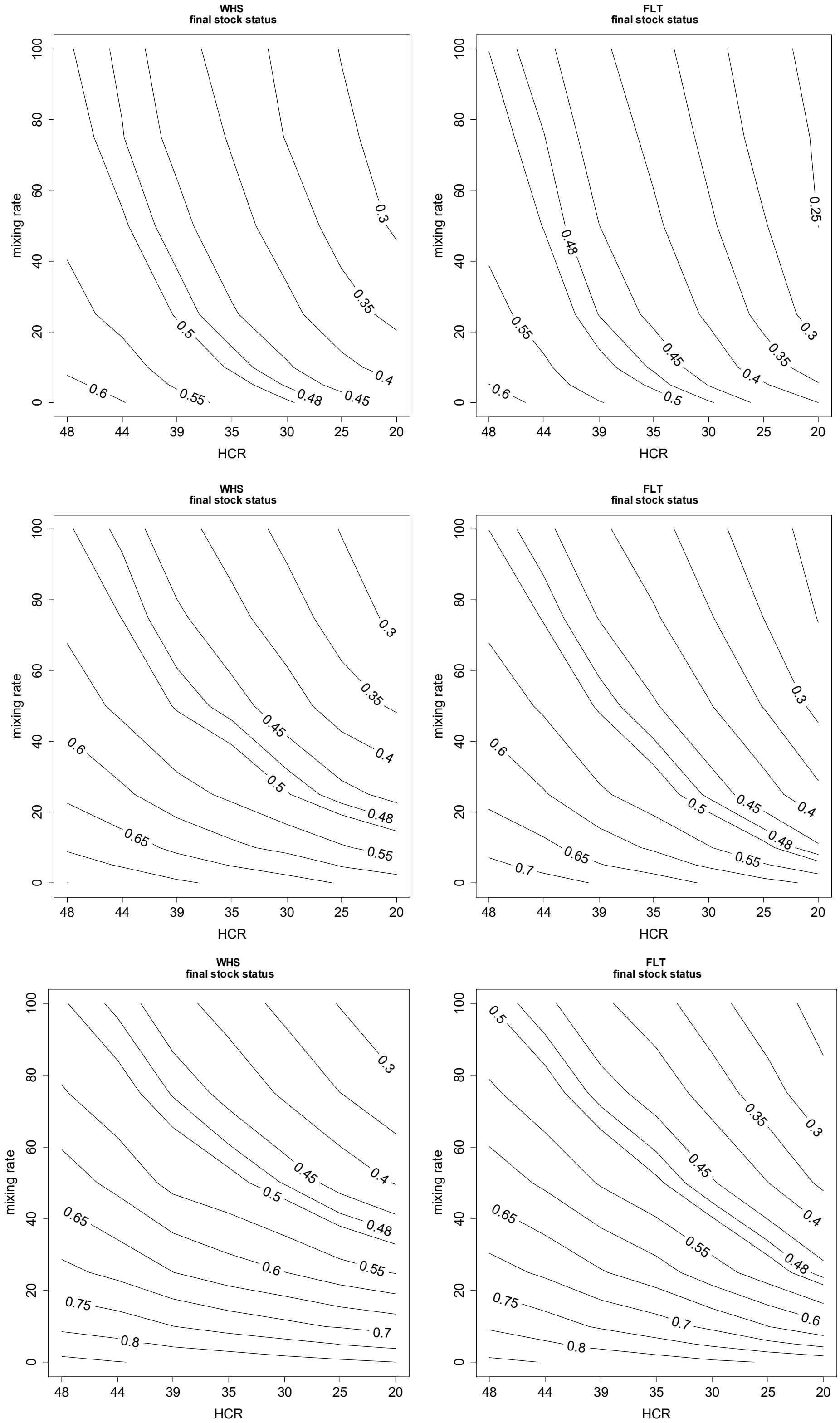
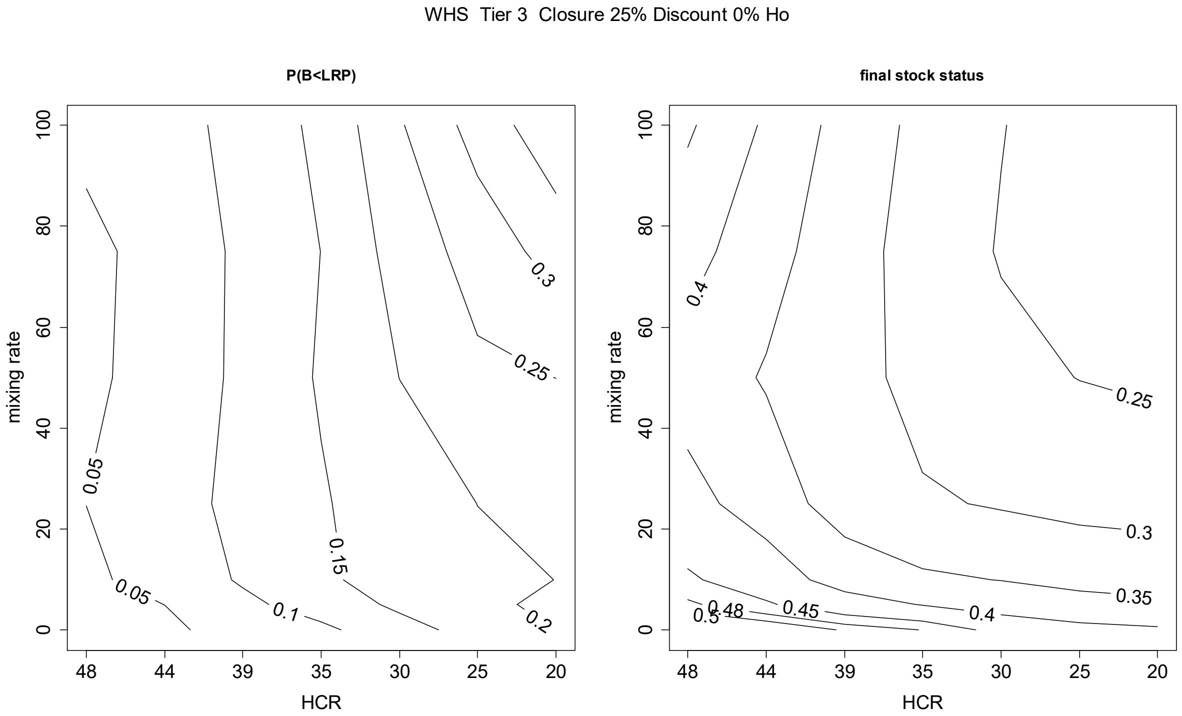


Figure 4.27: Final stock-wide status for school whiting (left) and Tiger Flathead (right) for 25%, 50% and 75% closures (top to bottom respectively) for the perfect information case (Tier P).



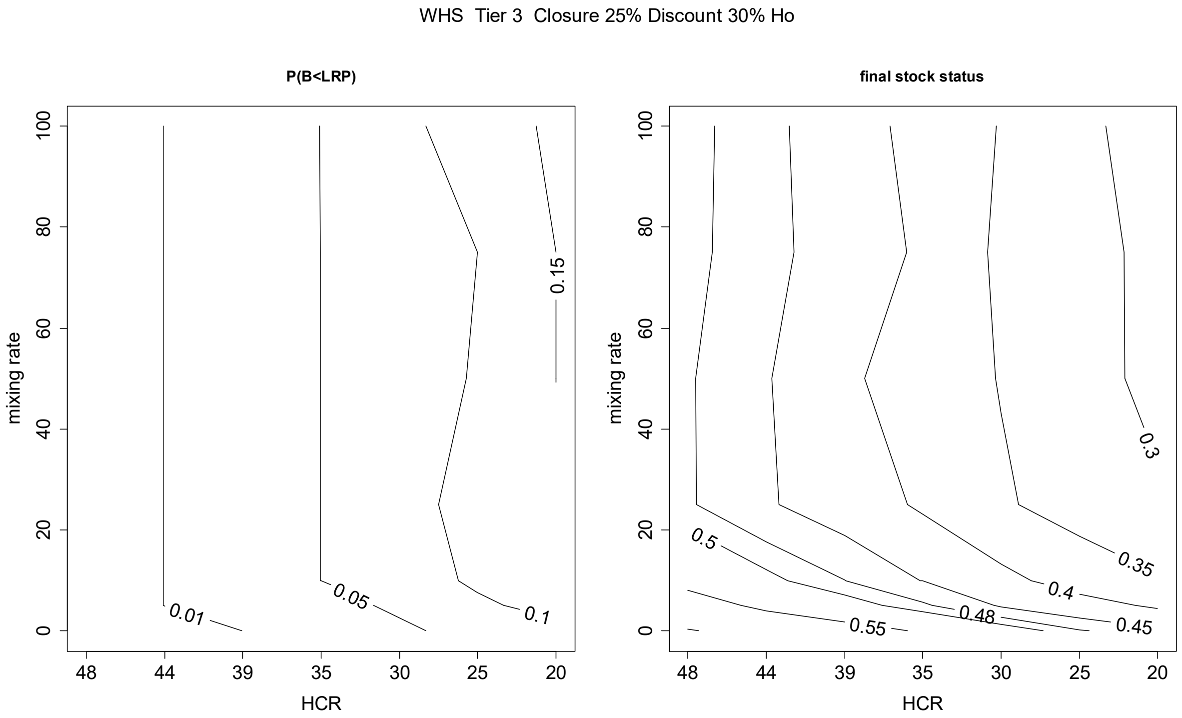


Figure 4.28: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting with a 25% closure and assessed with the Tier 3 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).

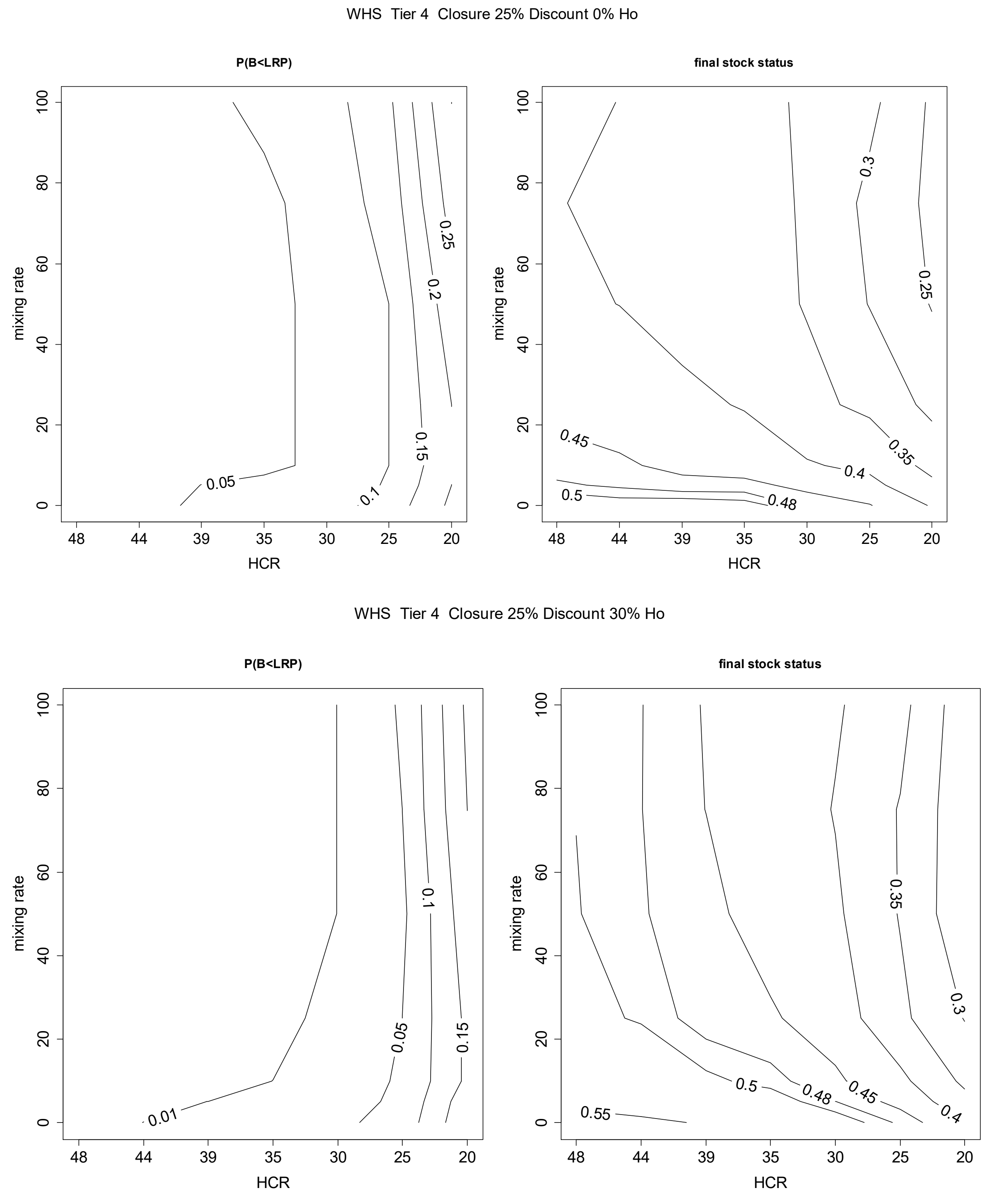


Figure 4.29: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for school whiting with a 25% closure and assessed with the Tier 4 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).

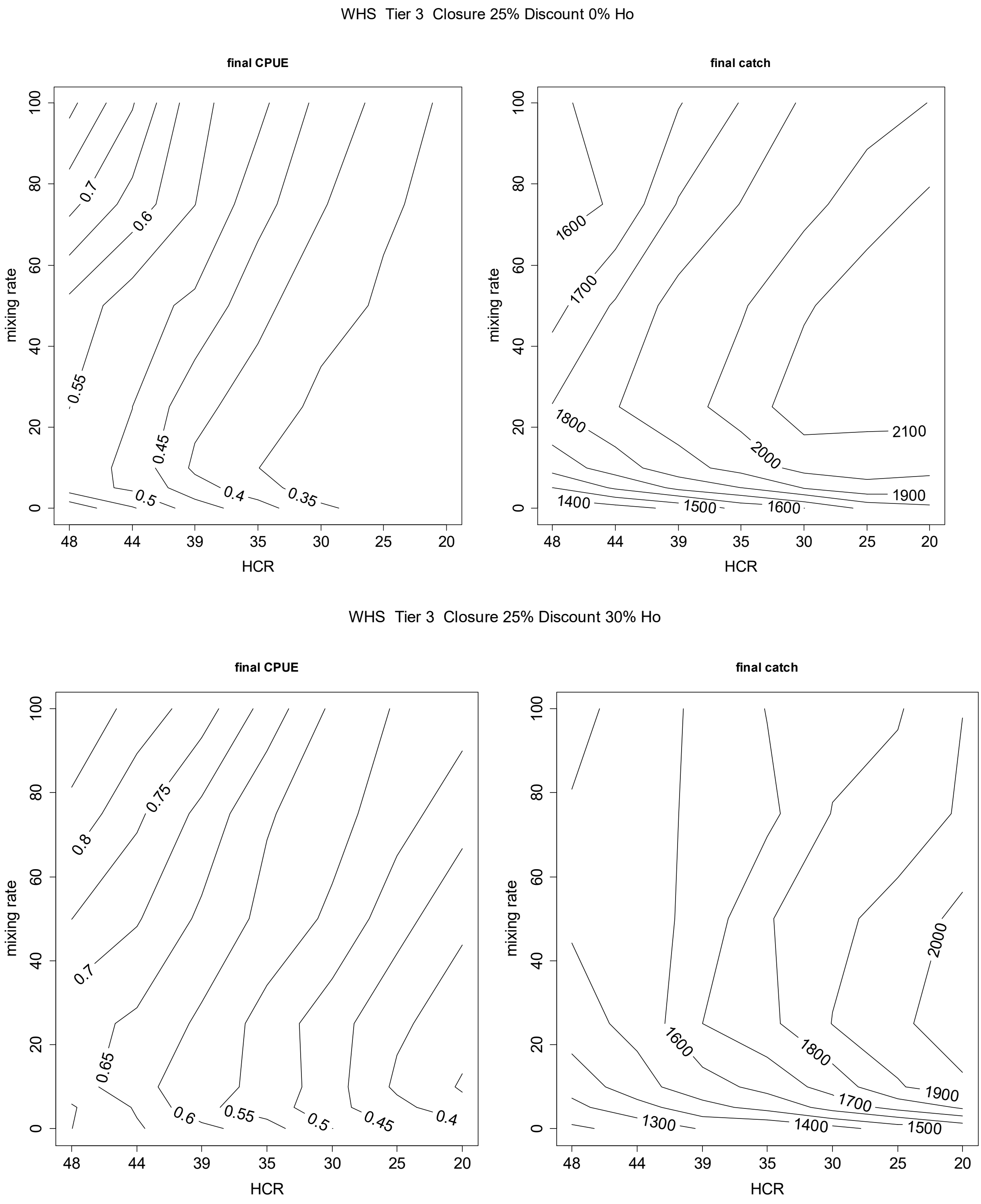


Figure 4.30: Final CPUE (left) and catch (right) contours for school whiting with a 25% closure and assessed with the Tier 3 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).

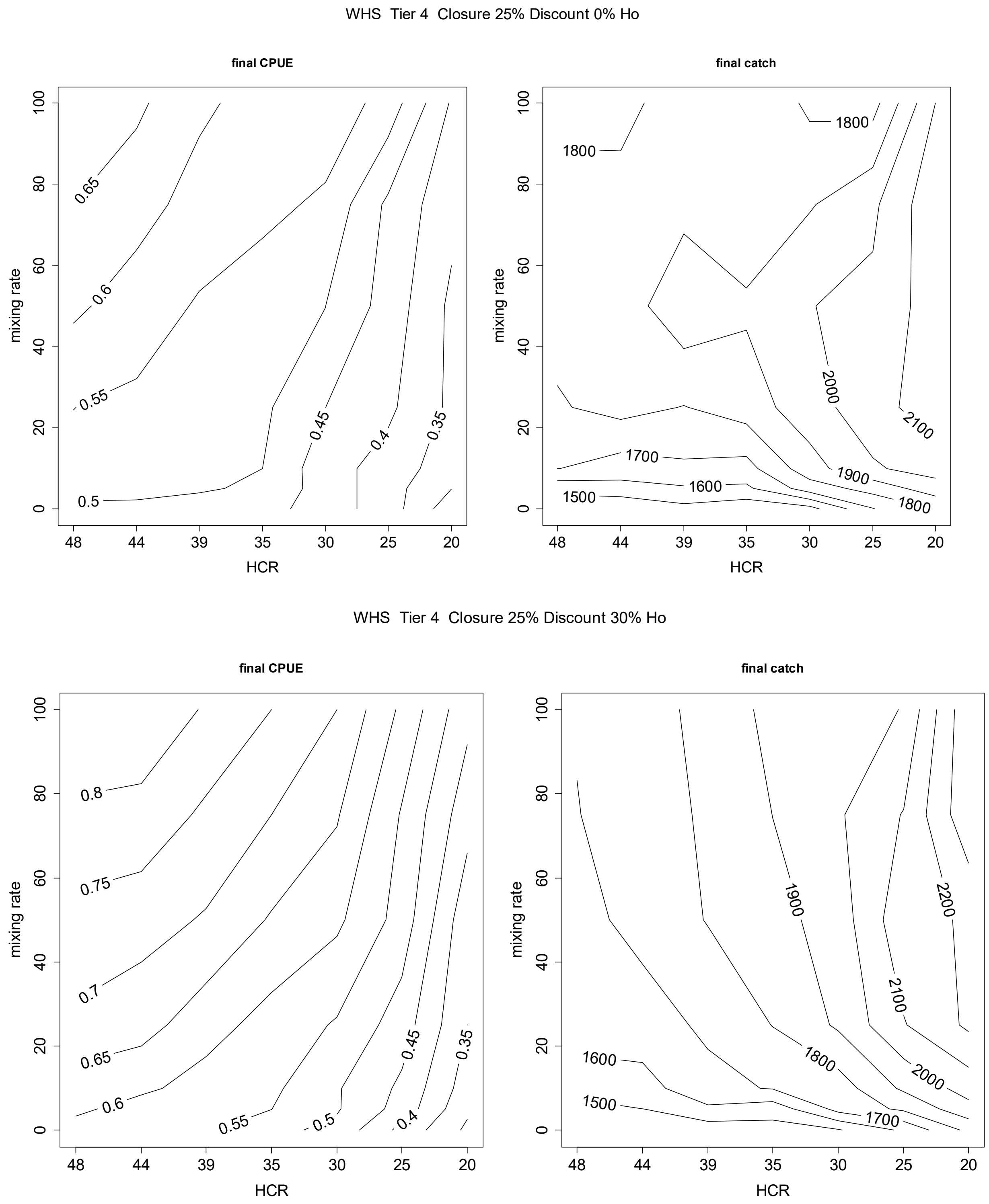
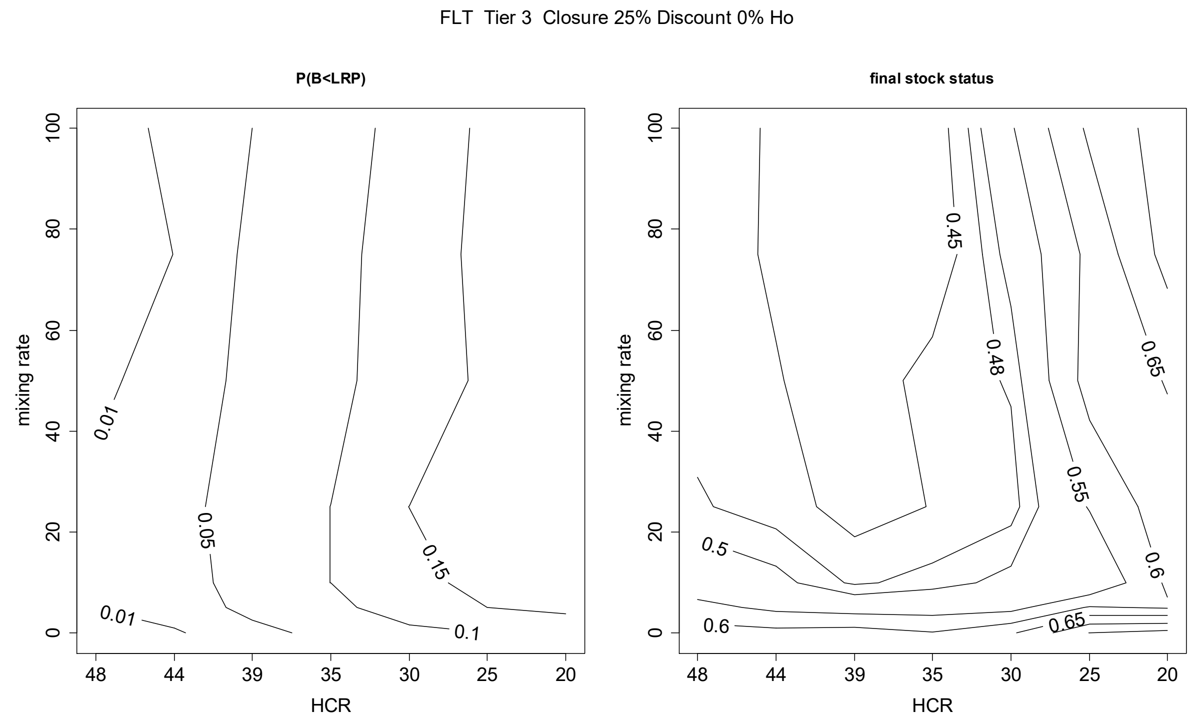


Figure 4.31: Final CPUE (left) and catch (right) contours for school whiting with a 25% closure and assessed with the Tier 4 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).



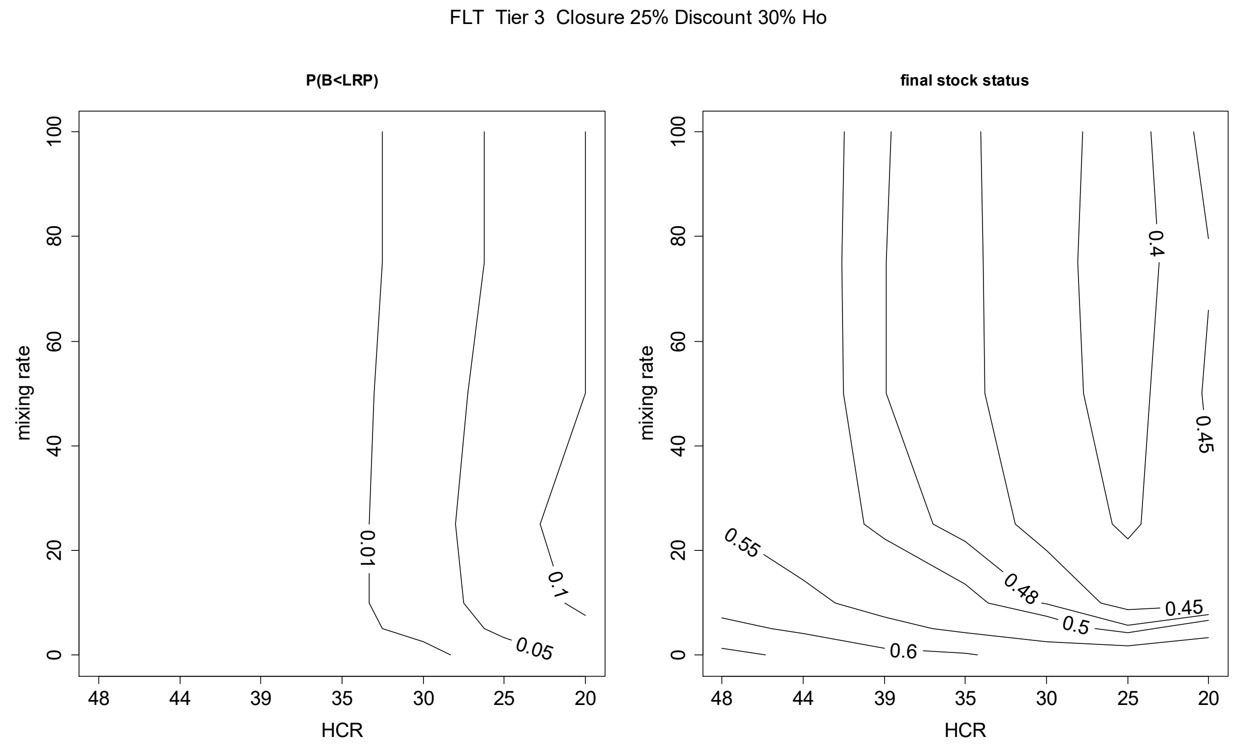


Figure 4.32: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead with a 25% closure and assessed with the Tier 3 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).

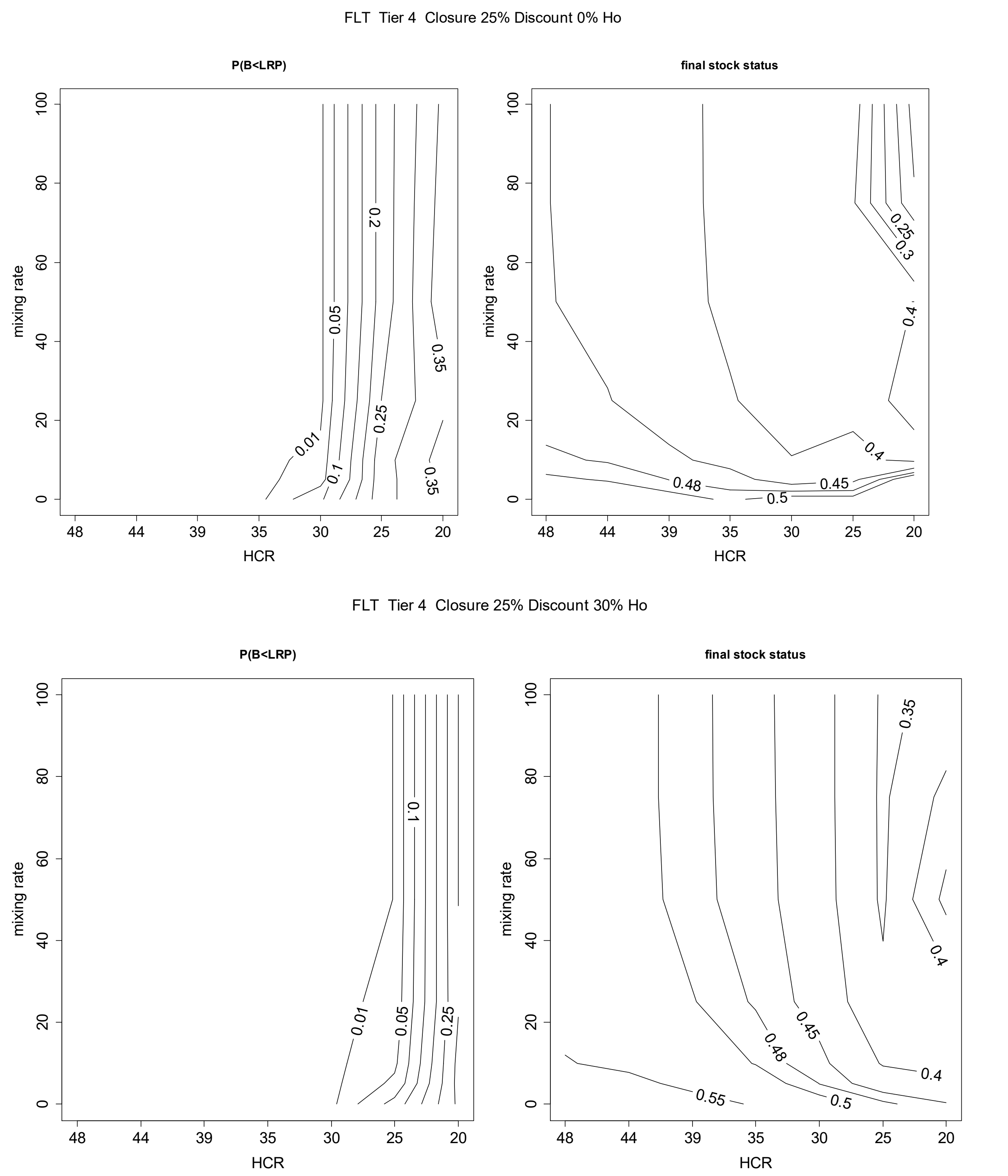


Figure 4.33: Probability of falling below the limit reference point (left) and final stock-wide status (right) contours for Tiger Flathead with a 25% closure and assessed with the Tier 4 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).

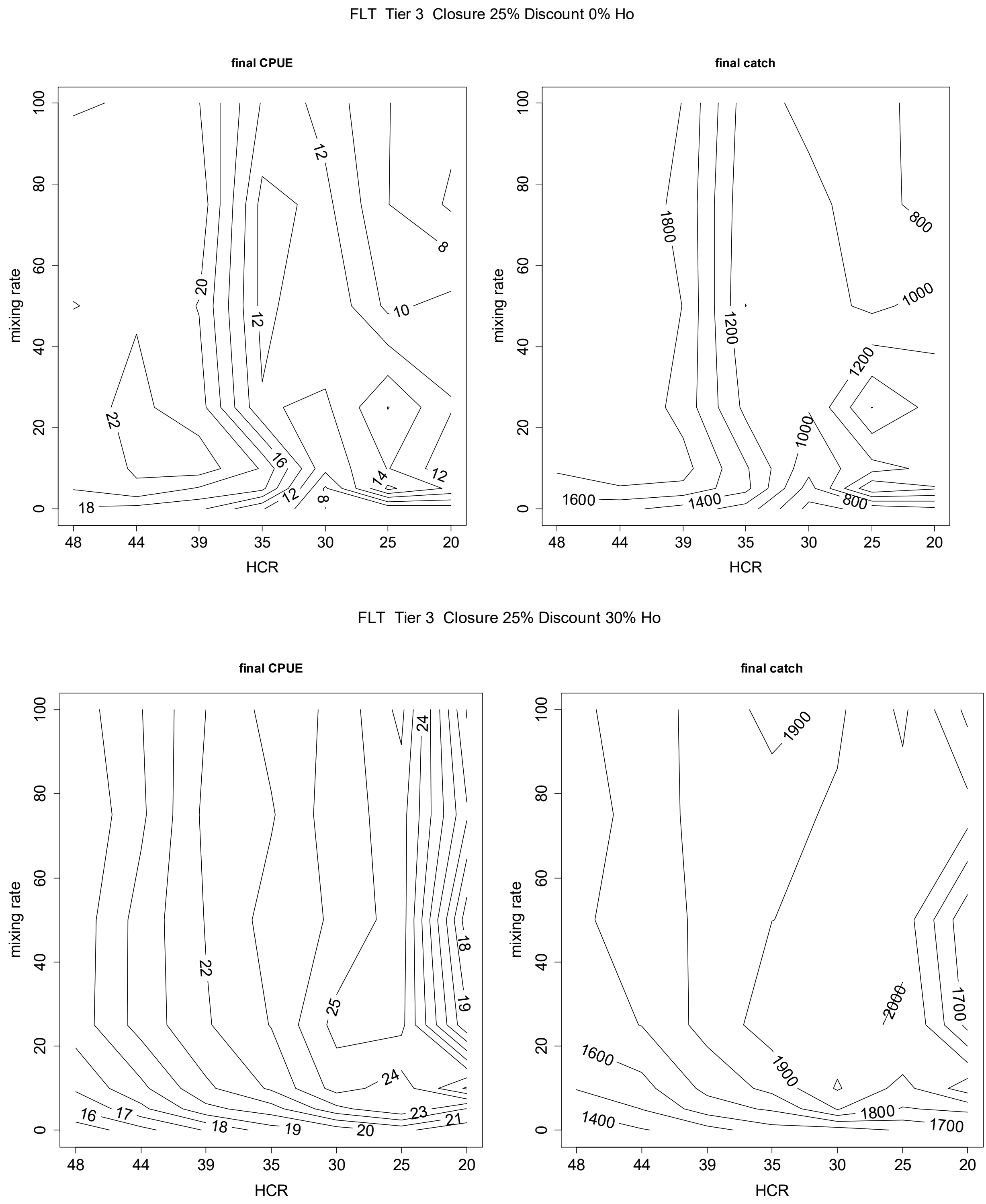
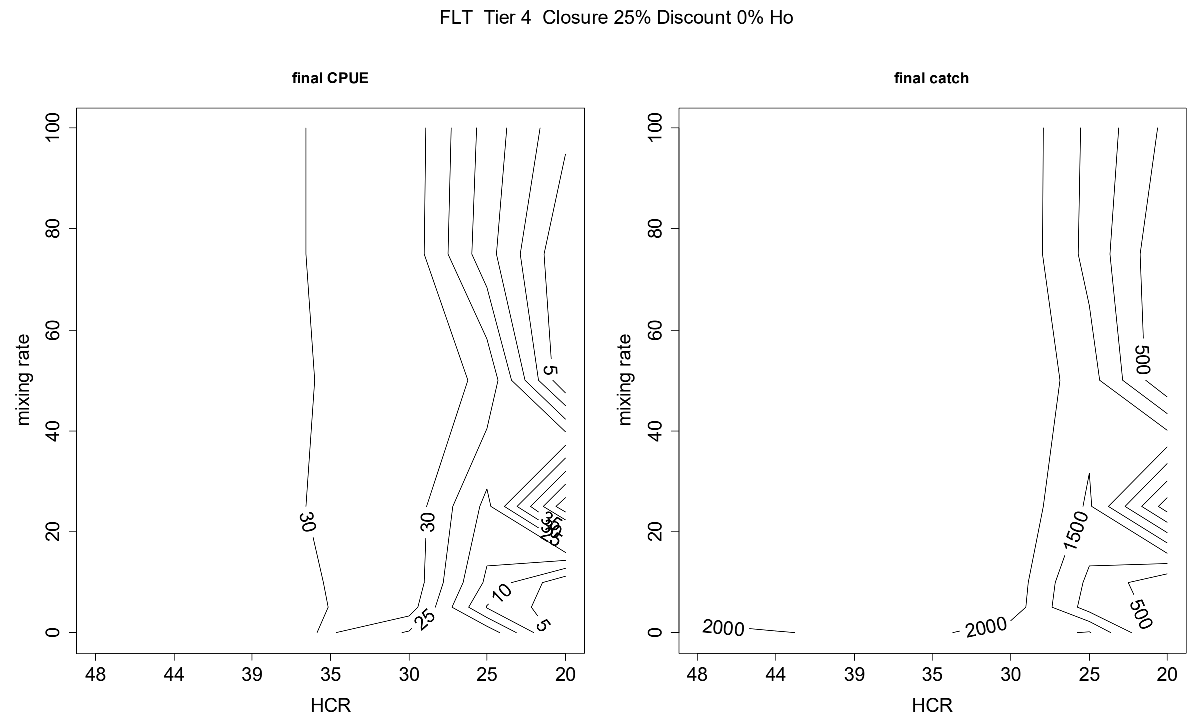


Figure 4.34: Final CPUE (left) and catch (right) contours for Tiger Flathead with a 25% closure and assessed with the Tier 3 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).



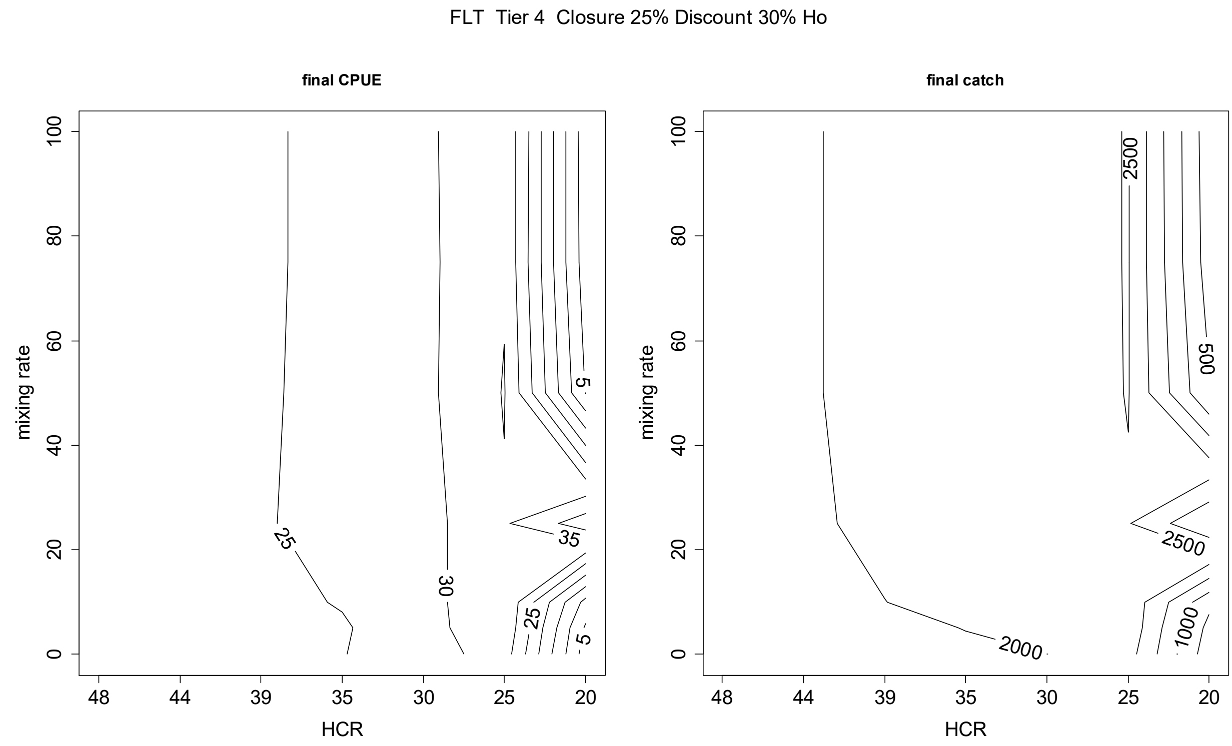
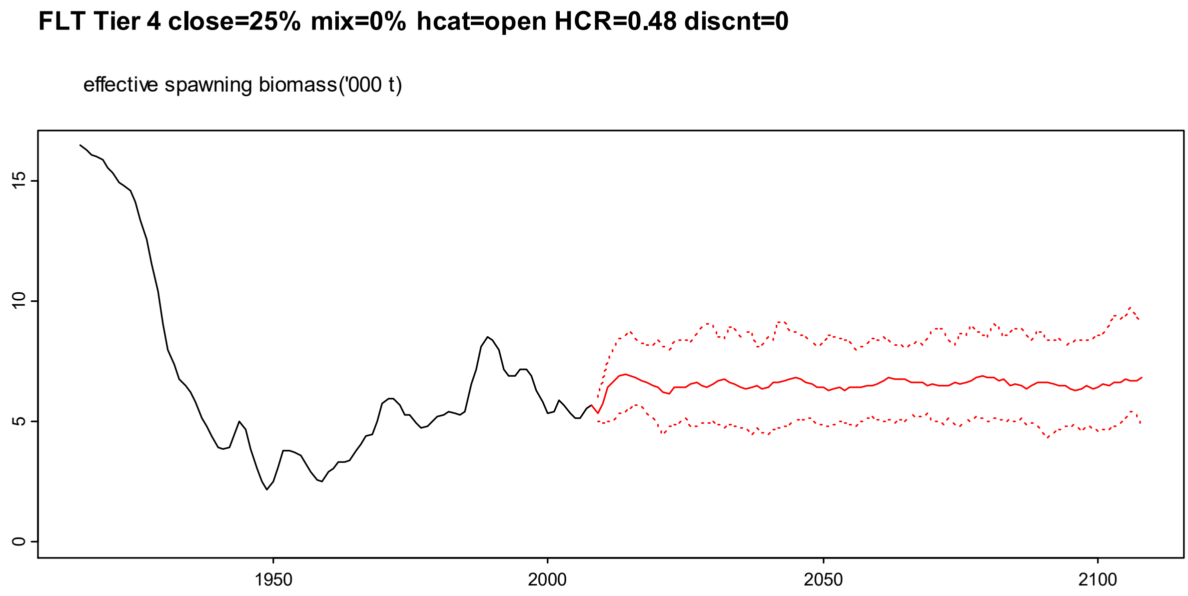


Figure 4.35: Final CPUE (left) and catch (right) contours for Tiger Flathead with a 25% closure and assessed with the Tier 4 harvest strategy and no discount (top) or a 30% discount (bottom). Only data from the open area used in the assessment (Ho).



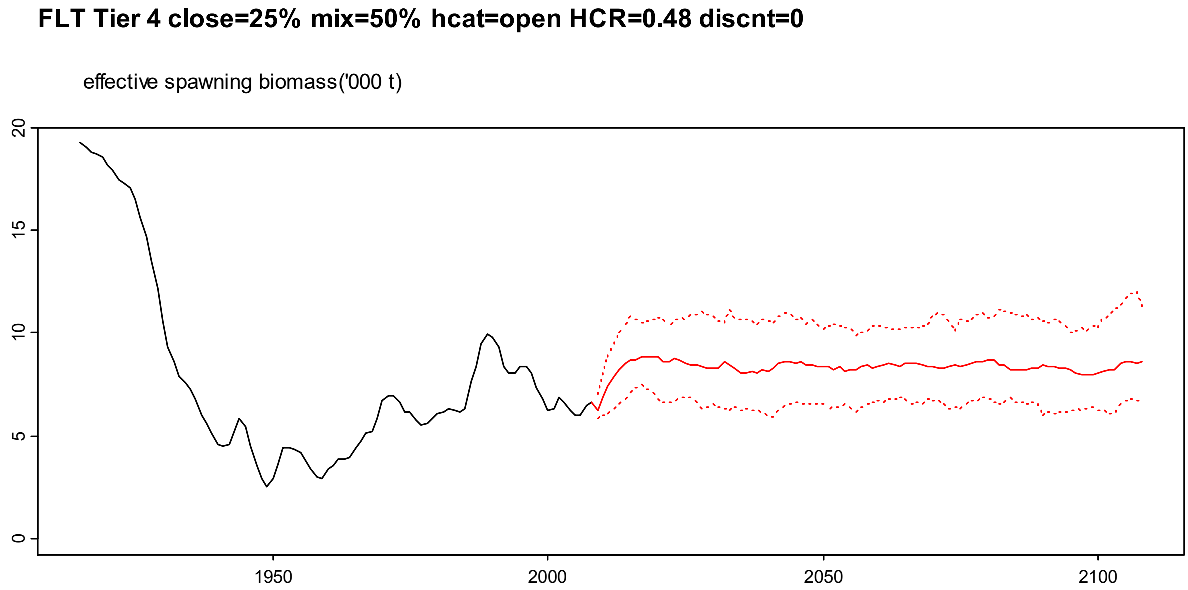
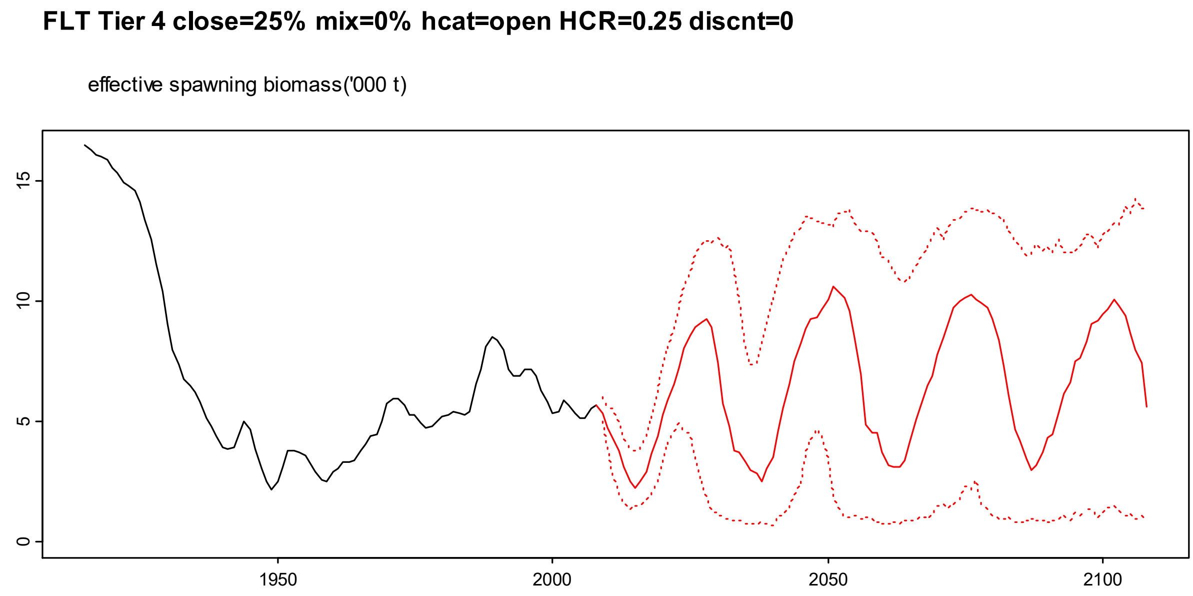


Figure 4.36: Effective biomass for tiger flathead: Tier 4; 25% closure; HCR= 48%; mixing rate =0% (upper) and 50% (lower).



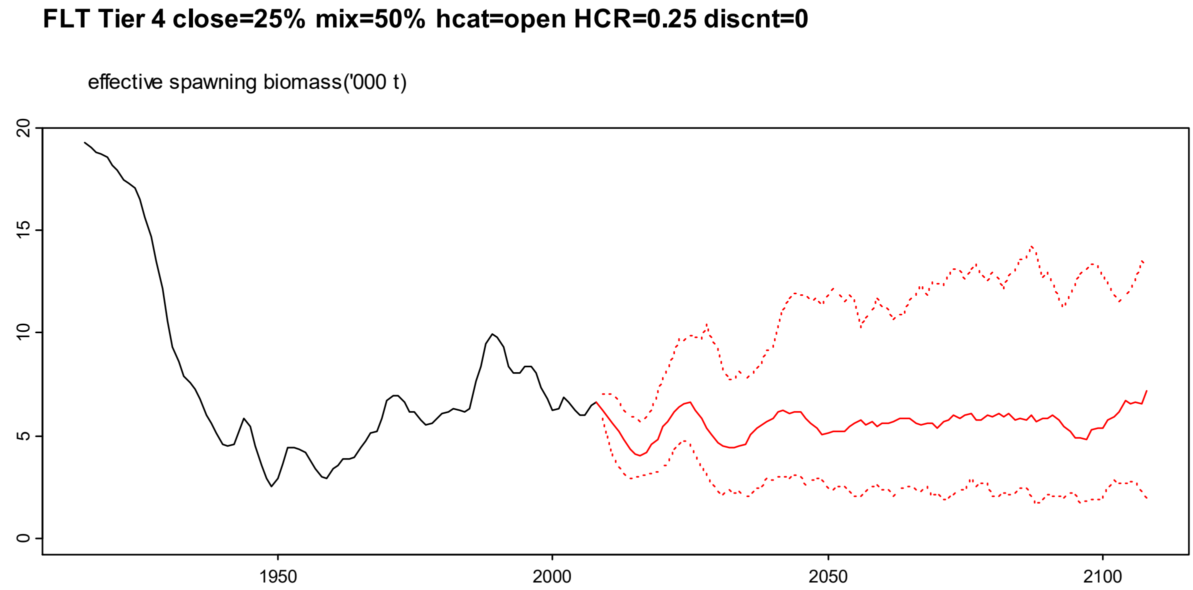
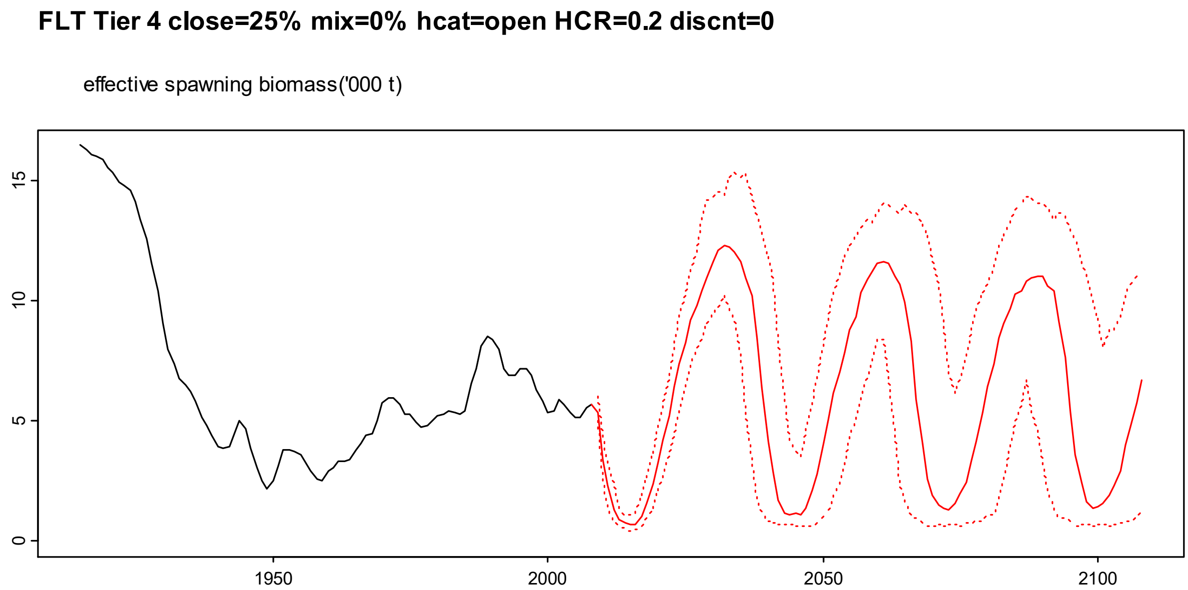


Figure 4.37: Effective biomass for tiger flathead: Tier 4; 25% closure; HCR= 25%; mixing rate =0% (upper) and 50% (lower).



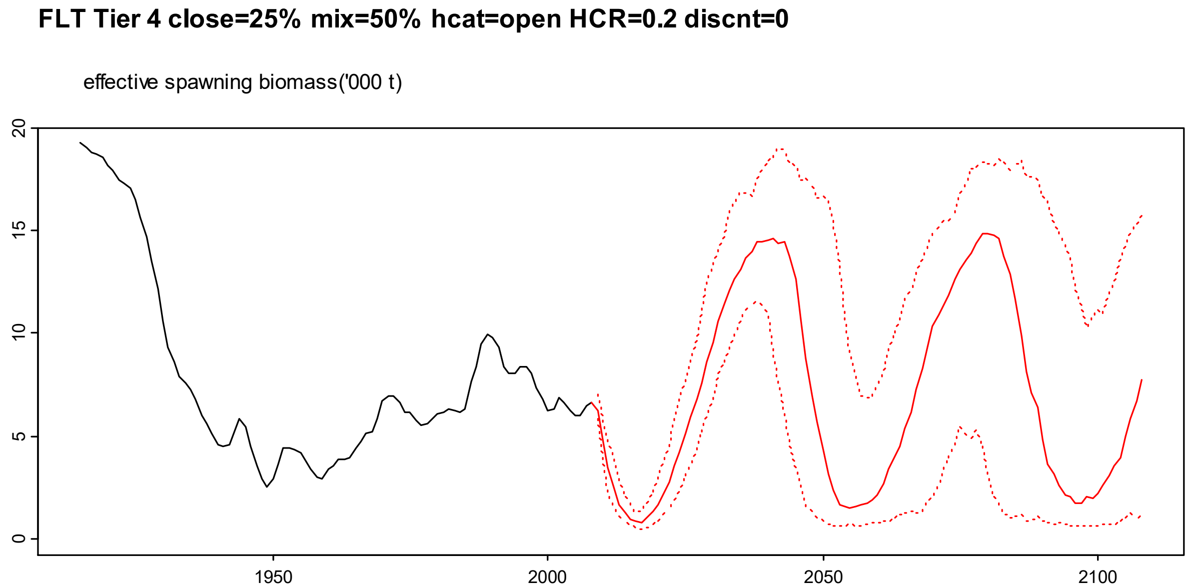


Figure 4.38: Effective biomass for tiger flathead: Tier 4; 25% closure; HCR= 20%; mixing rate =0% (upper) and 50% (lower).

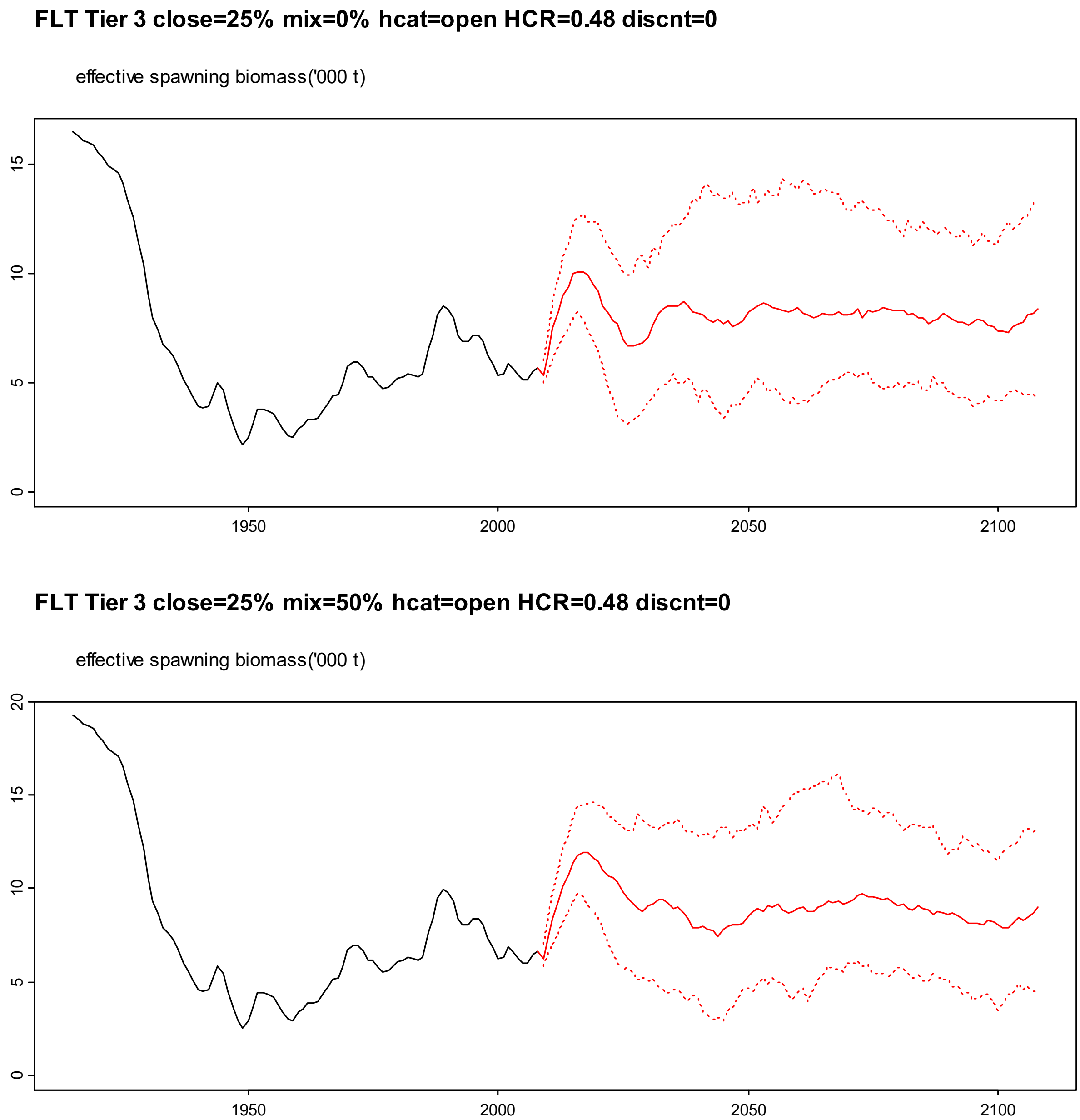


Figure 4.39: Effective biomass for tiger flathead: Tier 3; 25% closure; HCR= 48%; mixing rate =0% (upper) and 50% (lower).

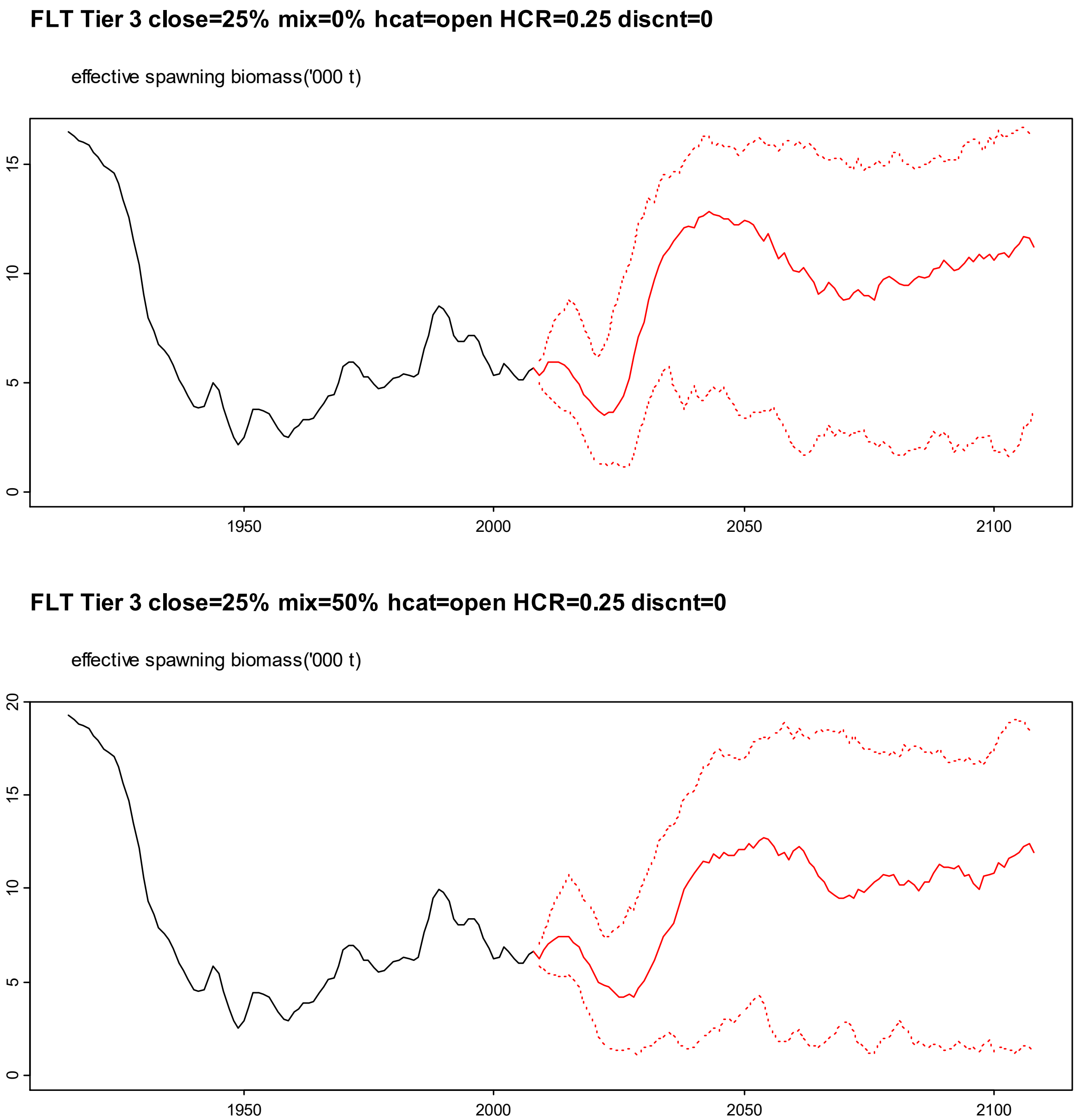
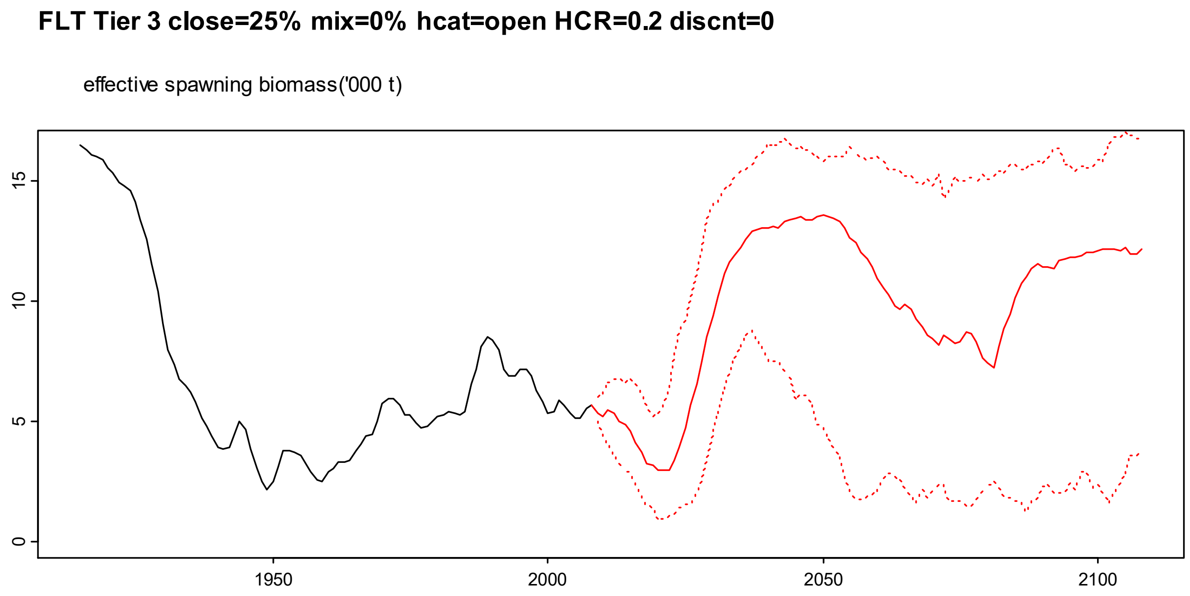


Figure 4.40: Effective biomass for tiger flathead: Tier 3; 25% closure; HCR= 25%; mixing rate =0% (upper) and 50% (lower).



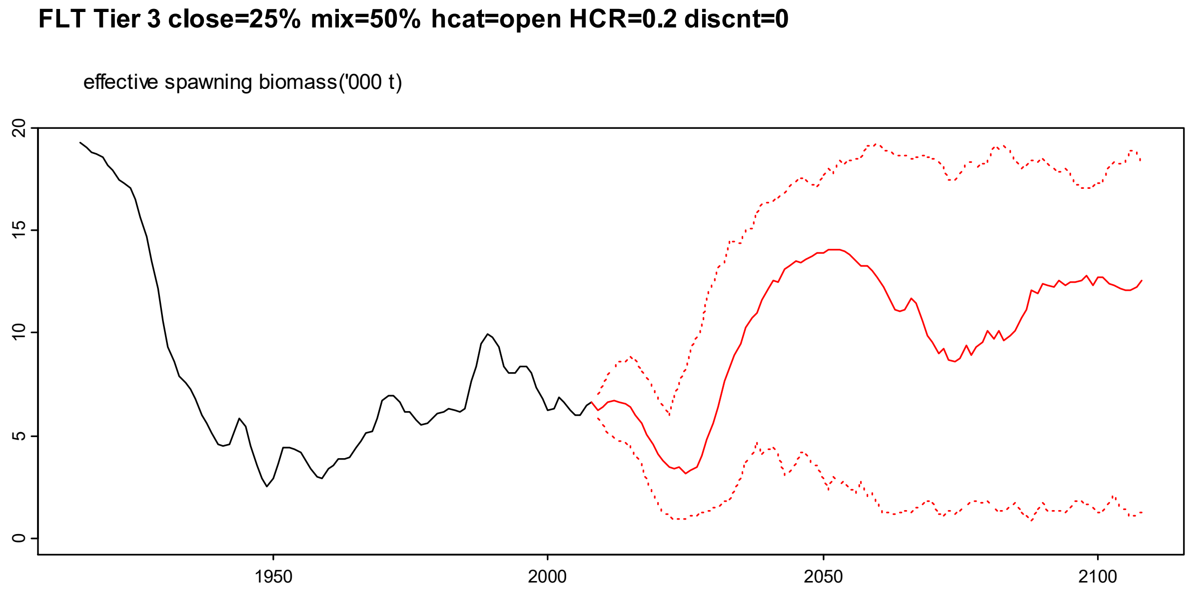
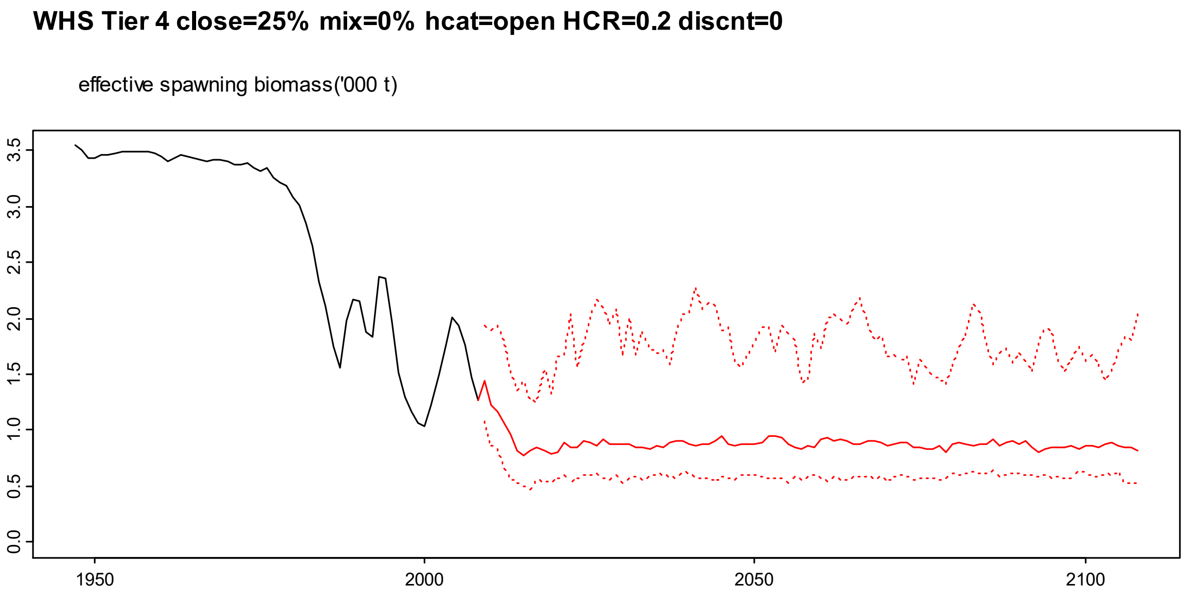


Figure 4.41: Effective biomass for tiger flathead: Tier 3; 25% closure; HCR= 20%; mixing rate =0% (upper) and 50% (lower).



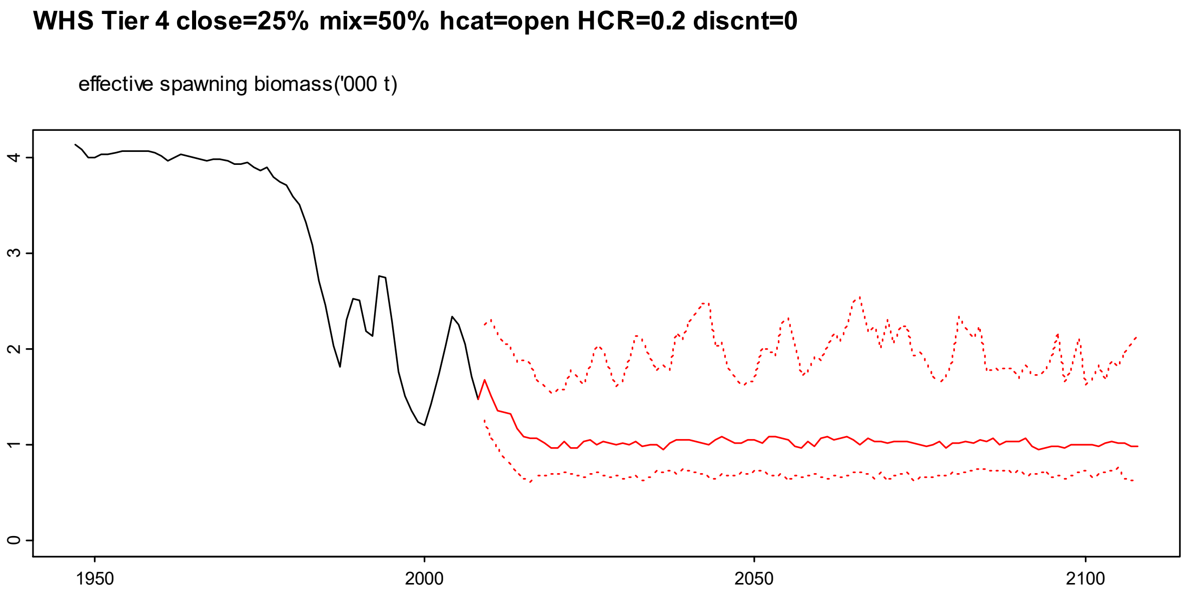
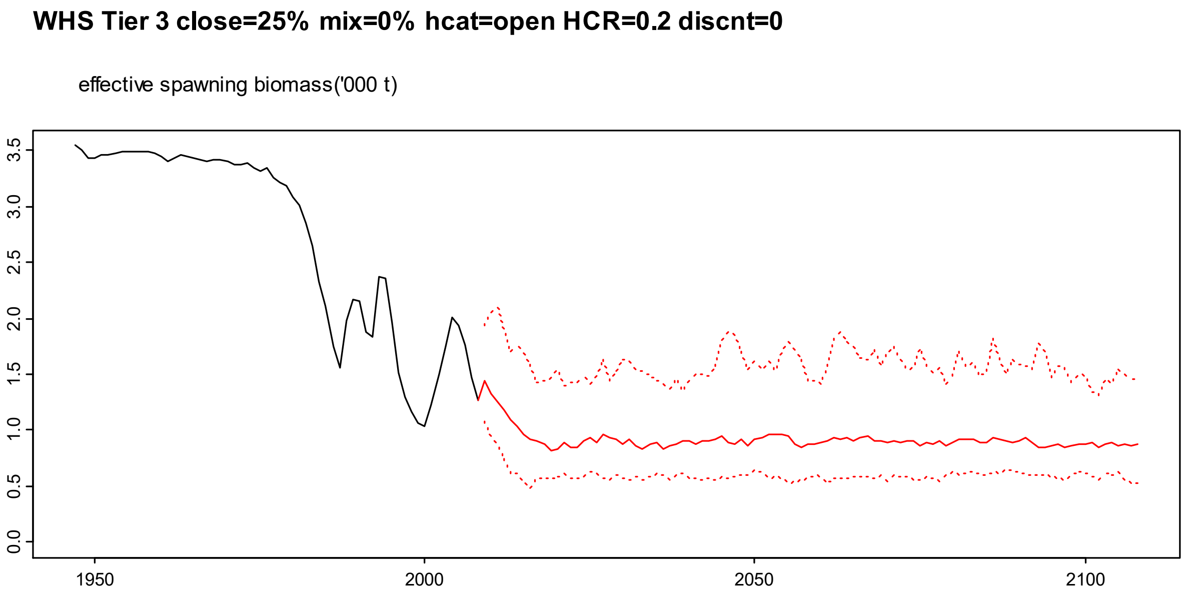


Figure 4.42: Effective biomass for school whiting: Tier 4; 25% closure; HCR= 20%; mixing rate =0% (upper) and 50% (lower).



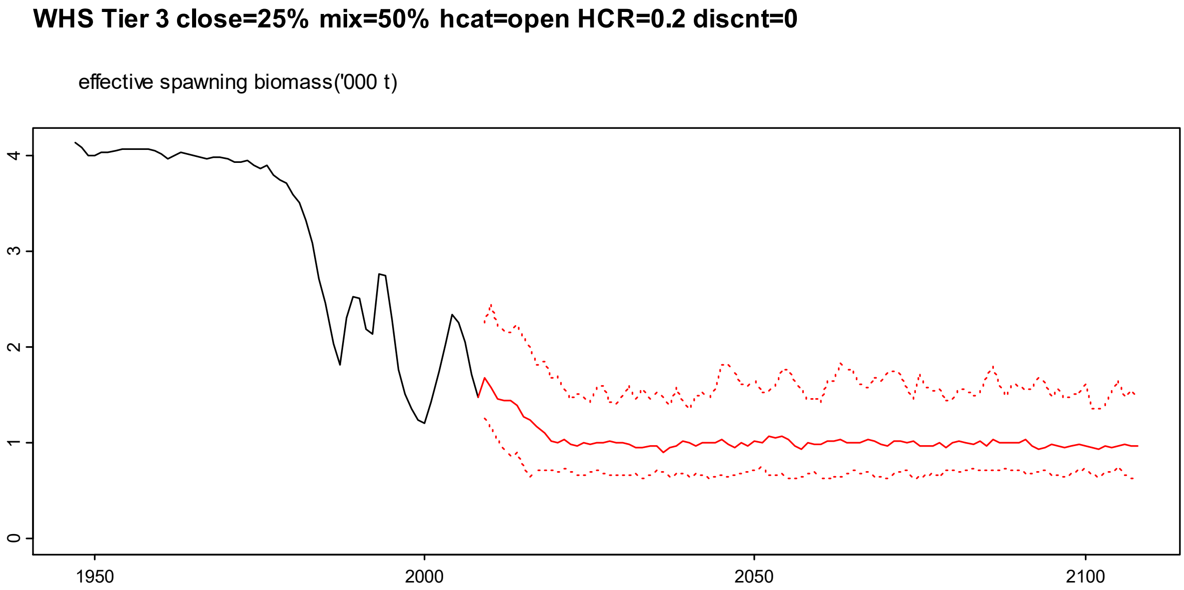


Figure 4.43: Effective biomass for school whiting: Tier 3; 25% closure; HCR= 20%; mixing rate =0% (upper) and 50% (lower).

5 Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on Pink Ling in Australia

Punt, A.E.[1,2], Haddon, M.[2], Little, L.R.[2], Tuck, G.N.[2]

[1]*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, USA*

[2]*CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, Tasmania, Australia;*

This has been published as André E. Punt, Malcolm Haddon, L. Richard Little, Geoffrey N. Tuck 2016. Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. Fisheries Research 175, 10-23.

5.1 Abstract

Spatial structure in biological characteristics and exploitation rates impact the performance of stock assessment methods used to estimate the status of fish stocks relative to target and limit reference points. Spatially-structured stock assessment methods can reduce the bias and imprecision in the estimates of management-related model outputs. However, their performance has only recently been evaluated formally, in particular when some of the area fished is closed. In order to evaluate the effects of closed areas and spatial variation in growth and exploitation rate when estimating spawning biomass, a spatially-explicit operating model was developed to simulate spatial data and five configurations of the stock assessment package Stock Synthesis (three of which were spatially structured) were applied. The bias in estimates of spawning stock biomass associated with spatially-aggregated assessment methods increases in the presence of closed areas while these biases can be reduced (or even eliminated) by applying appropriately constructed spatially-structured stock assessments. The performance of spatially-aggregated assessments when estimating spawning stock biomass is found to depend on the interactions among spatial variation in growth, in exploitation rate, and in knowledge of the spatial areas over which growth and exploitation rate are homogeneous.

Keywords: age-structured stock assessment methods, closed areas, simulation, spatial structure.

Highlights

- Simulation modelling explores estimation performance when the assessed area contains closed areas.
- Accounting for spatial structure can reduce estimation bias markedly.
- Estimation performance depends on the interaction between spatial trends in growth rates and those in exploitation rate.

5.2 Introduction

Punt et al. (2015) developed a simulation framework to evaluate the performance of various stock assessment methods implemented using Stock Synthesis (Methot and Wetzel, 2013) in the face of spatial structuring of fished populations. The assessment configurations considered by Punt et al. (2015) ranged from models that aggregated catch, length-frequency and conditional age-at-length data over space, to treating spatial regions as "fleets", and to spatially-explicit models. The testing framework was based on the Southern and Eastern Scalefish and Shark Fishery (SESSF) for pink ling, *Genypterus blacodes*, off southern Australia (Smith et al., 2008). The SESSF covers the region from southern Queensland, around Tasmania, to Cape Leeuwin in Western Australia. Allowance was made for three spatial zones (nominally zones 10, 20, and 30 of the SESSF; Figure 5.1). The fish populations in these three zones were assumed to be connected through the distribution of age-0 animals, with animals of age-1 and older being sedentary. Two fleets (essentially trawl and non-trawl) were assumed to operate in each zone, growth could differ among zones, and recruitment was assumed to be stochastic, with spatial variation in the proportion of the total recruitment that settles to each zone, as well as temporal variation in total recruitment.

Numerous small marine closures exist in south-east Australia, both for biodiversity conservation (under an Australian federal government initiative to establish a National Representative System of Marine Protected Areas (NRSMPA); Anon, 2015) and for fisheries management (declared under the Commonwealth Fisheries Management Act, 1991). Pink Ling is assessed as two separate stocks, separated east and west at 147˚E, but with a single total allowable catch for management purposes. Since 2005, four seasonal closures (from approximately September to November) have been in place to protect the spawning stock and reduce fishing mortality of Pink Ling at Maria Island, Seiner's Horseshoe and Everard Horseshoe in the east, and the Ling Hole in the west (Figure 5.1). Closures within these areas have been both voluntary and legislated (SEMAC, 2012). These particular closures are relatively small in area, but are considered among the most productive and previously favoured fishing grounds.

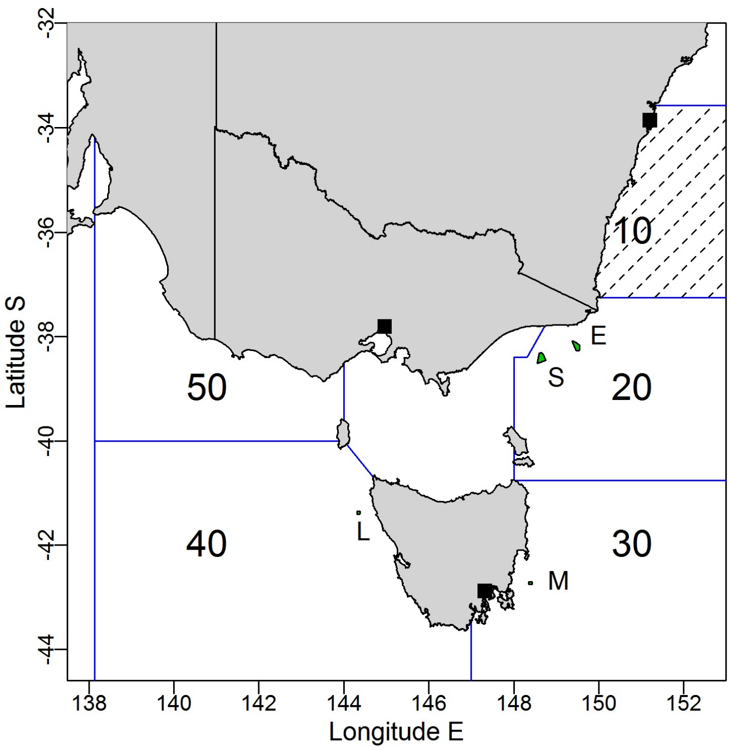


Figure 5.1: Schematic map of SESSF reporting blocks 10 – 50, with the fine blue lines representing block boundaries. The locations of Sydney, Melbourne, and Hobart are indicated by black squares from top to bottom. The east stock of pink link is found in zones 10, 20 and 30; the line between zones 30 and 40 is at 147^{0}E. The cross-hatched zone is the area closed to fishing in the bulk of the simulations; the real world closures include M, Maria Island; S, Seiners Horseshoe; E, Everard Horseshoe, and L, the Ling Hole.

The simulations conducted by Punt et al. (2015) showed that non-spatial assessment configurations that aggregate data spatially provided more precise, but biased estimates of initial and final spawning biomass, as well as of the ratio between final and initial spawning biomass. In contrast, assessments that allowed for spatial structure generally provided imprecise and highly biased estimates, although performance could be improved by changing the relative weighting applied to different data types. The exception to this general conclusion was when the population dynamics model underlying the assessment matched the model used to generate the pseudo data sets. Punt et al. (2015) recommended conducting sensitivity analyses based on several model configurations to select the most appropriate structure for an assessment based on, for example, residual patterns.

The analyses conducted by Punt et al. (2015) did not account for the possibilities of closed areas. Rather, they assumed that the modelled fisheries operated over entire zones that are homogenous with respect to age- and size-structure, as well as relative density. In addition, the analyses conducted by Punt et al. (2015) ignored the possibility of the availability of survey data. This paper therefore extends the analyses of Punt et al. (2015) to examine the consequences, in terms of the ability to estimate time-trajectories of spawning biomass, of closed areas that encompass a large proportion of stock biomass (>15%) as well as the benefits of the availability of surveys in the closed (and open) areas.

Several studies have considered the performance of stock assessment methods in the face of spatial heterogeneity in exploitation rates (e.g., Fu and Fanning, 2004; Hulson et al., 2011, 2013; Goethel et al., 2015; Guan et al., 2013; Harford et al., 2015). In addition, several previous studies have evaluated the impact of closed areas on the performance of stock assessment methods. Punt and Methot (2004) showed that stock assessment methods that assume that the stock is distributed homogeneously across space will lead to biased results when this assumption is violated to a substantial extent, with the magnitude and direction of bias depending on the extent to which the assumptions underlying the stock assessment are violated. Garrison et al. (2011) and McGilliard et al. (2015) found that applying spatially-structured stock assessment methods reduced or eliminated the bias when there are large area closures. This present study extends these earlier studies by considering the possibility that growth and trends in fishing mortality may also vary spatially.

5.3 Methods and Materials

The evaluation of alternative scenarios using simulation is based on specifying a model of the population dynamics. This ('operating') model is assumed to represent the truth for the simulations and is used to generate pseudo data sets. The pseudo data sets are then analysed using each of five configurations of the stock assessment package Stock Synthesis, and results are summarized to determine the overall performance of each configuration.

5.3.1 The operating model

The simulation evaluation involves an operating model that models a single population with spatial variation in age structure and a single stock-recruitment relationship. It includes spatial variation in growth and in the proportion of the total recruitment that settles by zone and can implement spatial closures. The operating model covers a 43-year period (nominally 1970 to 2012). The three zones are assumed to receive different proportions of the total recruitment in an unfished state (0.28, 0.49, and 0.23 respectively for zones 10, 20 and 30, which reflect roughly the relative amount of habitat for Pink Ling off southeastern Australia), with the extent of variation in spatial distribution, σ_ϕ, set to 0.7. Given a Beverton-Holt stock-recruitment relationship, the recruitment (at age-0) to zone z at the start of year y, R_y^z, is given by:

$$R_y^z = \frac{e^{\phi^z + \eta_y^z}}{\sum_z e^{\phi^z + \eta_y^z}} \frac{4\,h\,R_0\,\tilde{S}_y / \tilde{S}_0}{(1-h) + (5h-1)\tilde{S}_y / \tilde{S}_0} e^{\varepsilon_y - \sigma_R^2/2} \quad ; \quad \varepsilon_y \sim N(0;\sigma_R^2) \quad ; \quad \eta_y^z \sim N(0;\sigma_\phi^2) \qquad (1)$$

where h is the "steepness" of the stock-recruitment relationship (Francis, 1992), R_0 is the unfished equilibrium recruitment, $S_y^{♀}$ is total (over zones) spawning biomass, $S_0^{♀}$ is the unfished total spawning biomass, ϕ^z defines the expected proportion of the total recruitment that settles to zone z, σ_ϕ determines the variation about the expected proportion recruiting by zone across years, and σ_R is the standard deviation among recruitment deviations in log space. Spawning biomass is defined as:

$$S_y^{♀} = \sum_z \sum_a O_a^z \, N_{y,a}^{\text{fem},z} \tag{2}$$

where $N_{y,a}^{s,z}$ is the number of animals of sex s and age a in zone z at the start of year y, O_a^z is the product of maturity-at-age and weight-at-age (see Methot and Wetzel [2013] for details of how O_a^z is calculated) based on current stock assessment parameters.

The value for *h* is set to 0.75 and that for σ_R to 0.7 (Whitten and Punt, 2014). Punt et al. (2015) outline the selectivity patterns by gear (assumed to be the same among zones). Figure 5.2 shows the spatial variation in relative fishing mortality in the absence of closed areas. The fishery is assumed to start in zone 10 and then increase progressively southward over time – this reflects the fact that the fisheries off southeast Australia started in the mainland ports (and within zone 10). As in Punt et al. (2015), the maximum level of fishing mortality is assumed to be the same spatially, while the fully-selected fishing mortality for the non-trawl fleet is assumed to be half that for the trawl fleet. For consistency with the actual assessment for Pink Ling (Whitten and Punt, 2014), selectivity for the non-trawl fleet is assumed to be a monotonic logistic function of length, while that for the trawl fleet is modelled using a unimodal double-normal selectivity function.

As in Punt et al. (2015), catch-rate data are only assumed to be available from 1986 (Haddon, 2014), while collection of age- and length-composition data is assumed to start in 1975. Between 1975 and 1985, length data are assumed to be available for 5% of the combinations of years, gears and zones. The years with length data are selected at random, and gear-zone combinations for the years with samples for data on length selected in proportion to the size of the catch in weight. This generation process reflects that catches are sampled both by port samplers and onboard vessels. Length data are assumed to be available for 20% of year-gear-zone combinations from 1986 to 1997 and for 70% of these combinations thereafter. The observed catch length-composition data for a year-gear-zone combination is a Dirichlet sample from the true catch length-composition, with an effective sample size of 100. This is lower than the actual sample sizes for pink ling, but the length composition data for Pink Ling are known to be over-dispersed, and the length data are considerably down-weighted to reflect this when assessments are conducted (Whitten and Punt, 2014). The length-composition data are assumed to be unsexed, as is the case for Pink Ling and most fish stocks. The age-length keys are assumed to be obtained from a subset of the length-frequencies (56%), with a sample size of 500 for each sex. Given a year-gear-zone, the age data are assumed to be a simple random sample from the catches by age and length, as is the intent of the sampling program. Following Punt et al. (2015), the age-estimates are assumed not to be subject to age-reading error.

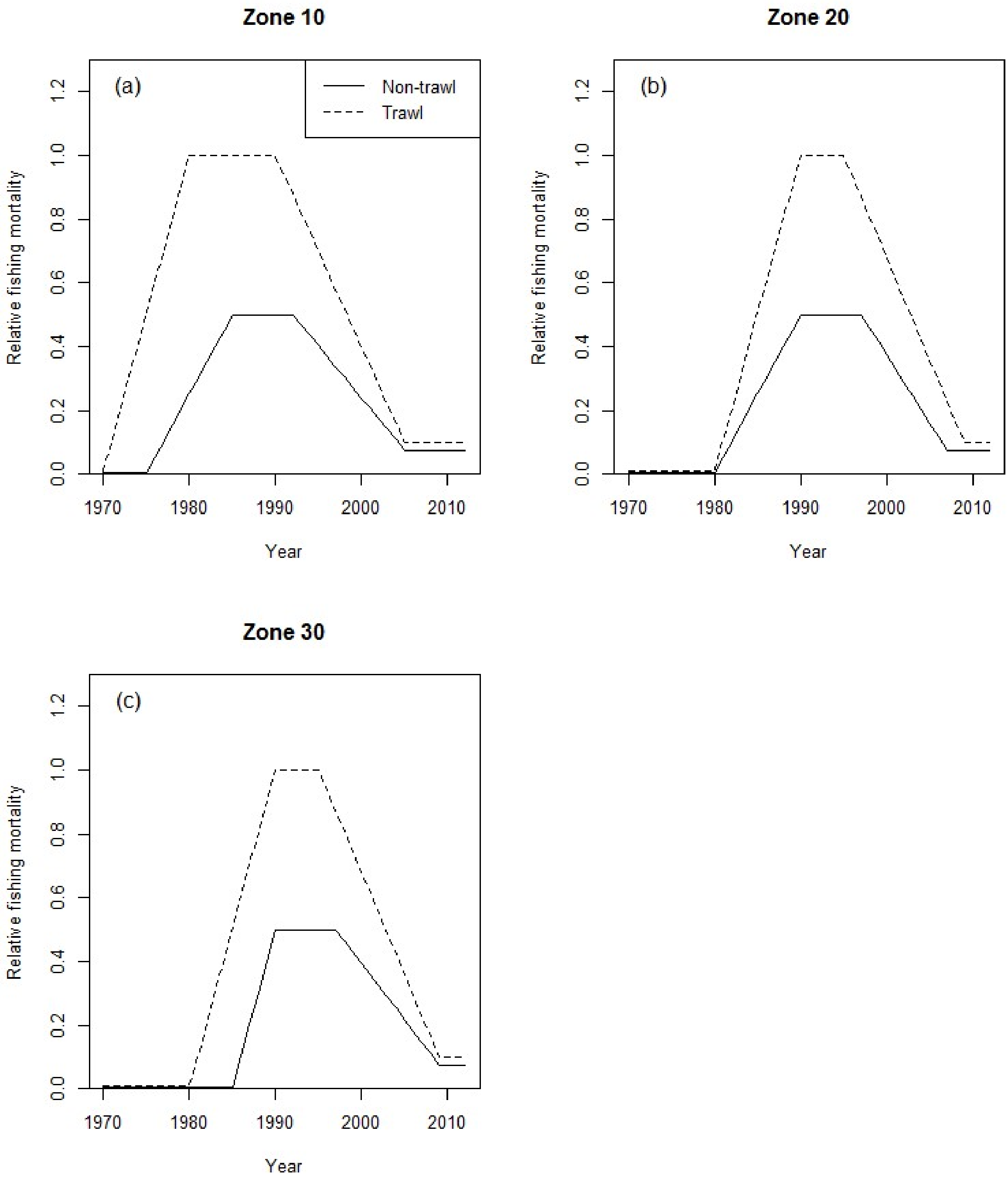


Figure 5.2: Relative fishing mortality rates by gear and zone in the absence of closures.

Allowance for closed areas and simulation scenarios

Two sets of simulations are undertaken to evaluate the impact of closed areas on the performance of five assessment configurations. The two sets differ in terms of whether survey data (indices of abundance, survey length-frequency data, and survey conditional age-at-length data) are available. The bulk of the analyses consider two scenarios regarding closures:

- Zone 10 is closed from 2000 onwards
- Zone 10 is closed from 2005 onwards

These two scenarios are considered because the status of the closed area will be nearer to its unfished level when it is closed earlier. When a closed area is implemented in one zone, the exploitation rates in the other zones are increased so that overall, and to ensure comparability among simulation scenarios, the ratio of the total (over zone) spawning stock biomass in 2012 to the total unfished spawning biomass (B_0) is the same (0.4) under deterministic projections (i.e. no spatial variation in settlement and no variation in recruitment about the stock-recruitment relationship). This is to avoid the impacts of closed areas on estimation performance being confounded with changes in overall population biomass. The total amount of length and conditional age-at-length data is the same among scenarios, i.e. closing one zone will lead to more length and conditional age-at-length data for the other zones. This specification is made to avoid the effects of closures being confounded with the total amount of data available for assessment purposes.

Each of the scenarios, as well as the baseline scenario in which there are no closed areas, are conducted with and without survey data. Surveys are assumed to be conducted from 2000 (i.e. at the start of one of the sets of closures). Surveys are assumed to occur every 2nd year. The selectivity of the survey gear is assumed to be the same as that of the trawl fishery (but the assessments do not known this). The survey CV for each zone is assumed to be 0.1 and the survey length-frequency data are assumed to be a multinomial sample with an effective sample size of 100, while the survey conditional age-at-length data are assumed to be a multinomial sample of size 500.

All of the scenarios are conducted when growth either does or does not vary spatially. Figure 5.3 shows the growth curves by zone when growth varies spatially. The growth curve when growth is spatially-invariant is set equal to that for zone 20 in Figure 5.3. This choice was made because the biomass in zone 20 is larger than that in the other zones.

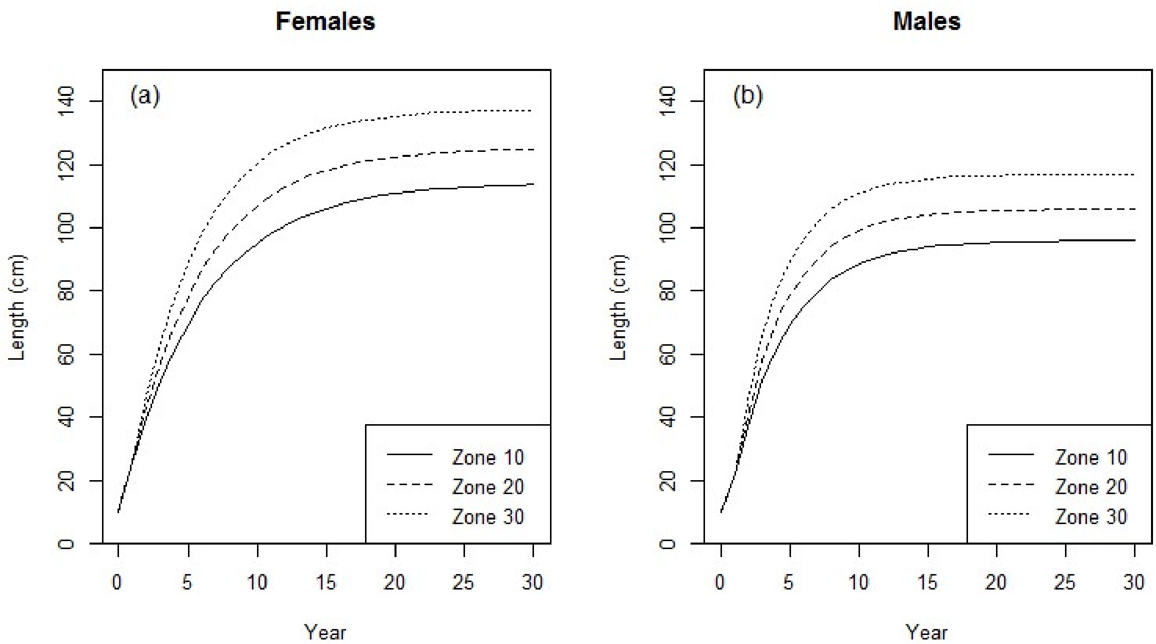


Figure 5.3: Von Bertalanffy growth curves by zone.

Sensitivity is explored to making the closures occur in zone 20 rather than zone 10. Closing zone 20 rather than zone 10 is likely to exacerbate the impacts of the closures given the biomass in zone 20 is larger than that in zone 10 (Figure 5.4). Zone 20 was closed in 2000 for comparability with the scenario when zone 10 was closed in 2000.

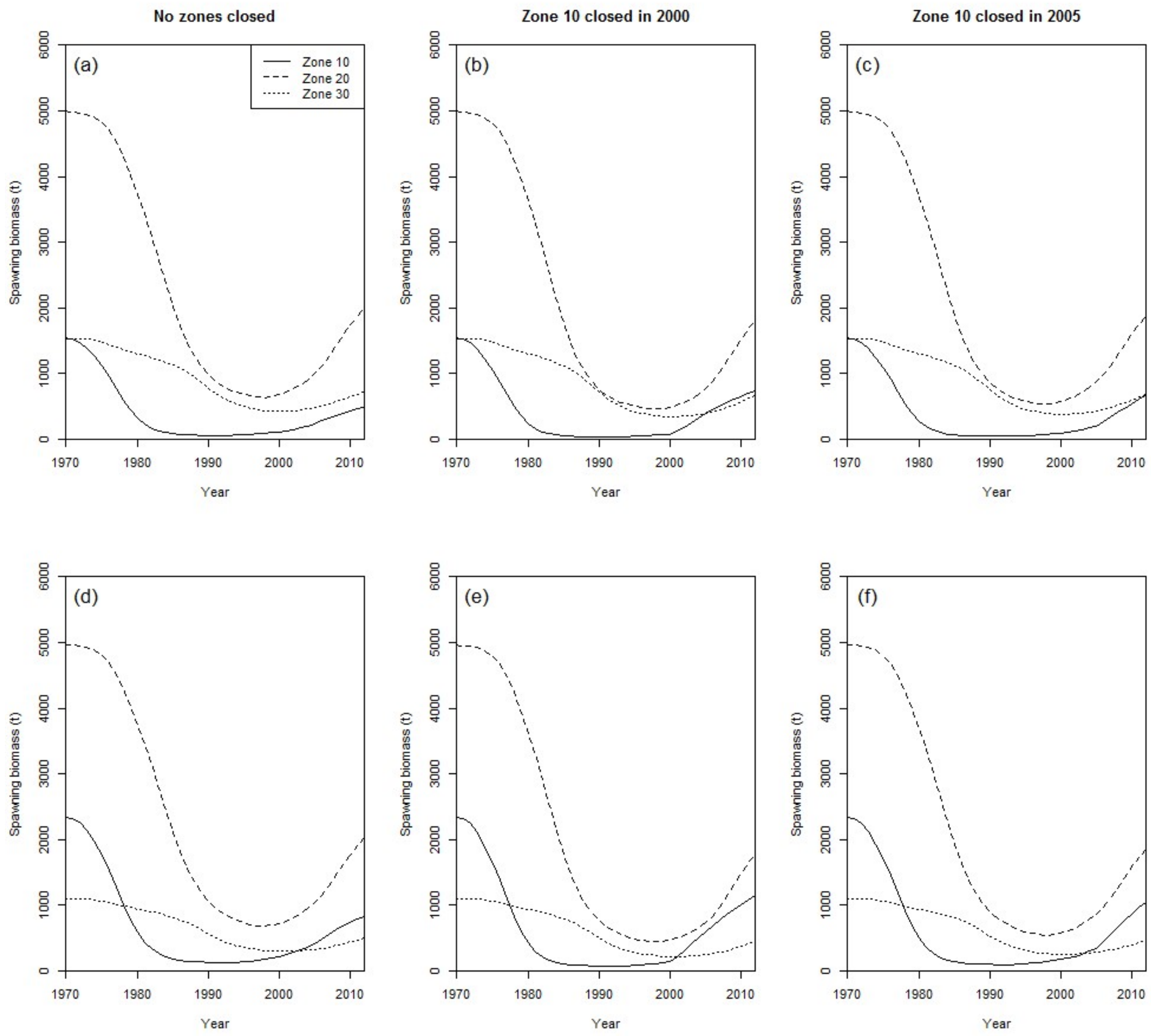


Figure 5.4: Deterministic time-trajectories of female spawning biomass. Results are shown in the upper panels when growth varies spatially and in the lower panel when growth is the same for all zones.

Sensitivity is explored to having survey data for a longer period of time (from 1990 rather than 2000). This involved initiating a survey in zone 10 in 1990 prior to its closure in 2000. Sensitivity to ignoring the CPUE indices and using only survey data after the closed areas are implemented is also examined. The last sensitivity analysis was conducted because the trend in CPUE (which is based on catch and effort for the areas open to fishing) will be biased relative to total biomass across open and closed areas; ignoring these CPUE data will lead to the assessment results being based primarily on the trends in survey data.

The simulations are based on 200 replicates for each scenario. Results are reported for all 200 replicates for each stock assessment configuration. There was evidence that some of the stock assessment configurations failed to converge (i.e. did not result in a positive definite Hessian matrix) for some of the replicates. However, the results were robust to basing analyses on only the replicates that appear to have converged (i.e. that did result in a positive definite Hessian matrix).

5.3.2 Stock assessment configurations

Punt et al. (2015) considered seven stock assessment configurations. This paper considers four of these configurations, two of those used most commonly in reality (NSWA and FAA; definitions below) and two that closely mimic the way the data are generated (SSTVR and FULL). This paper also considers an assessment configuration not considered by Punt et al. (2015), but on which a recent assessment of canary rockfish *Sebastes pinniger* was based (Thorson and Wetzel, 2015) (SSTVRSEL). The SSTVRSEL configuration (defined below) would be expected to outperform the other two spatially-structured assessment configurations when growth is the same for all zones because it matches the operating model and has fewer parameters than the SSTVR and FULL configurations. However, this configuration may perform poorly when growth differs spatially.

The five assessment configurations estimate unfished recruitment, natural mortality (assumed to be the same for males and females), growth by sex (five parameters per sex: the parameters that govern von Bertalanffy growth and the CVs of length-at-age for ages 1 and 20), length-specific selectivity parameters (logistic for the trawl fleet, double-normal for the non-trawl fleet, i.e. based on the correct selectivity patterns), catchability for the CPUE indices, and recruitment for simulated years 1963-2013 [1]. The estimation methods are assumed to know the correct form of the stock-recruitment relationship, the true value of steepness and the true value for the extent of variation about the stock-recruitment relationship. This is to ensure that the focus of the study is on effects of closed areas, given the well-known difficulty in estimating steepness (Conn et al., 2012; Lee et al., 2012). Recruitment estimation includes an initial bias-ramp (Methot and Taylor, 2011) to avoid bias when estimating the deviations about the stock-recruitment relationship given length and CPUE data are only available for the more recent years. The spatial configurations (FULL, SSTVR, and SSTVRSEL) can estimate the proportion of the unfished biomass in each zone, zone-specific deviations in the recruitment, and spatial variation in growth.

Data weighting is a key component of any stock assessment, and the results of an assessment can be sensitive to approaches to data weighting (McAllister and Ianelli, 1997; Francis, 2011, 2014). As in Punt et al. (2015), the assumed extent of inter-annual variation in catchability is set to 0.1, while the effective sample sizes for the length and conditional age-at-length data are set to 20% of the actual sample sizes.

Naïve spatially-weighted aggregated (NSWA)

This configuration does not recognize that there are spatial differences in population structure and abundance. It involves conducting a spatially-aggregated assessment, and combines the data spatially:

- The catch data are summed over zones.
- The catch rate data are aggregated across areas, defining the catch-rate for year y as total catch for year y divided by the total effort for year *y*.
- The annual trawl and non-trawl catch length-frequency data by fleet are pooled over zones, weighting the data for each zone by the annual catch by the zone.
- The annual age-length keys (i.e. the conditional age-at-length data) are summed over zones (without catch weighting). This reflects how data have been aggregated in actual assessments for Pink Ling (Whitten and Punt, 2014).

Fleets-as-areas (FAA)

In common with the NSWA configuration, FAA (Punt et al., 2014; Waterhouse et al., 2014) is based on a population dynamics model that assumes that the areas being assessed contain a single homogenous stock, with recruitment estimated as annual deviations about a stock-recruitment relationship. However, unlike the NSWA configuration, FAA assumes that each zone (10, 20 and 30) contains one trawl and one non-trawl fleet,

[1] The period for which recruitment is estimated is longer than the period for which catches are available to allow the initial age-structure to differ from that corresponding to an unfished population.

with the selectivity patterns by fleet estimated separately for each zone. The catch-rate, length-frequency and conditional age-at-length data for each fleet are kept separate by zone in this method rather than being aggregated spatially.

Spatially-structured with time-varying spatial recruitment and spatial growth (FULL)

This configuration models the animals in each zone separately within a single assessment model. Animals of age 1+ are assumed to be sedentary, with the annual proportions of the total number of age 0 animals from the stock-recruitment relationship allocated to each area treated as estimable parameters. It allows the parameters determining growth (L_∞, k, length-at-age 0 and the CVs for animals of ages 0 and 20) to differ spatially, and also allows the proportion of the total recruitment settling to each area to vary by year.

Spatially-structured with time-varying spatial recruitment (SSTVR)

This approach is the same as FULL, except growth is assumed to be the same spatially.

Spatially-structured with time-varying spatial recruitment and constant selectivity (SSTVRSEL)

This approach is the same as SSTRV, except that selectivity is assumed to be the same for each gear in each zone (i.e. selectivity is only estimated for the two fleets across all three zones).

5.3.3 Performance metrics

The assessment method provides many outputs (Methot and Wetzel, 2013). However, the focus for this paper are the estimates of initial total (female) spawning biomass[2] (B_0), final (2012) total spawning biomass (BCUR) and relative spawning biomass (B_{CUR}/B_0). Within Australia and the U.S., biomass in absolute terms directly impacts estimates of sustainable catch limits, while relative spawning biomass is used to assess stock status relative to biological reference points. The results of the simulations are summarized by relative error distributions as well as by the median over simulations of the absolute relative errors (MARE). The relative error for a given quantity is the estimated value of the quantity less its true value, divided by its true value, and multiplied by 100, i.e. a positive value indicates an overestimate of the quantity and vice versa.

In principle, the performance of assessments could be improved by applying model selection methods. However, it is not possible to compare all of the configurations. Specifically, NSWA, FAA and the spatial configurations are not comparable as they use different data. However, it is possible to use model selection methods to compare the FULL, SSTVR and SSTVRSEL configurations. Consequently, AIC is used to select a 'best' model from the FULL, SSTVR and SSTVRSEL configurations in each replicate, which enables a comparison of which of the three is optimal most often.

5.4 Results

5.4.1 No closed areas

Figure 5.5 shows time-trajectories of relative errors for the scenarios in which there are no closed areas and there are no survey data. Results are shown for the five assessment configurations and when growth in the operating model is the same for all zones and when growth differs spatially. As expected, the results for the NSWA, FAA, SSTVR and FULL configurations are essentially identical to those in Punt et al. (2015) when growth varies spatially. The results in Figure 5.5 are based on all 200 simulations even though the Hessian

[2] Total spawning biomass is spawning biomass summed over the three zones.

matrices for some of the simulations were not positive definite (Supplementary Table S5.1). Results based on only those simulations for which the Hessian matrix was positive definite are essentially identical to those in Figure 5.5.

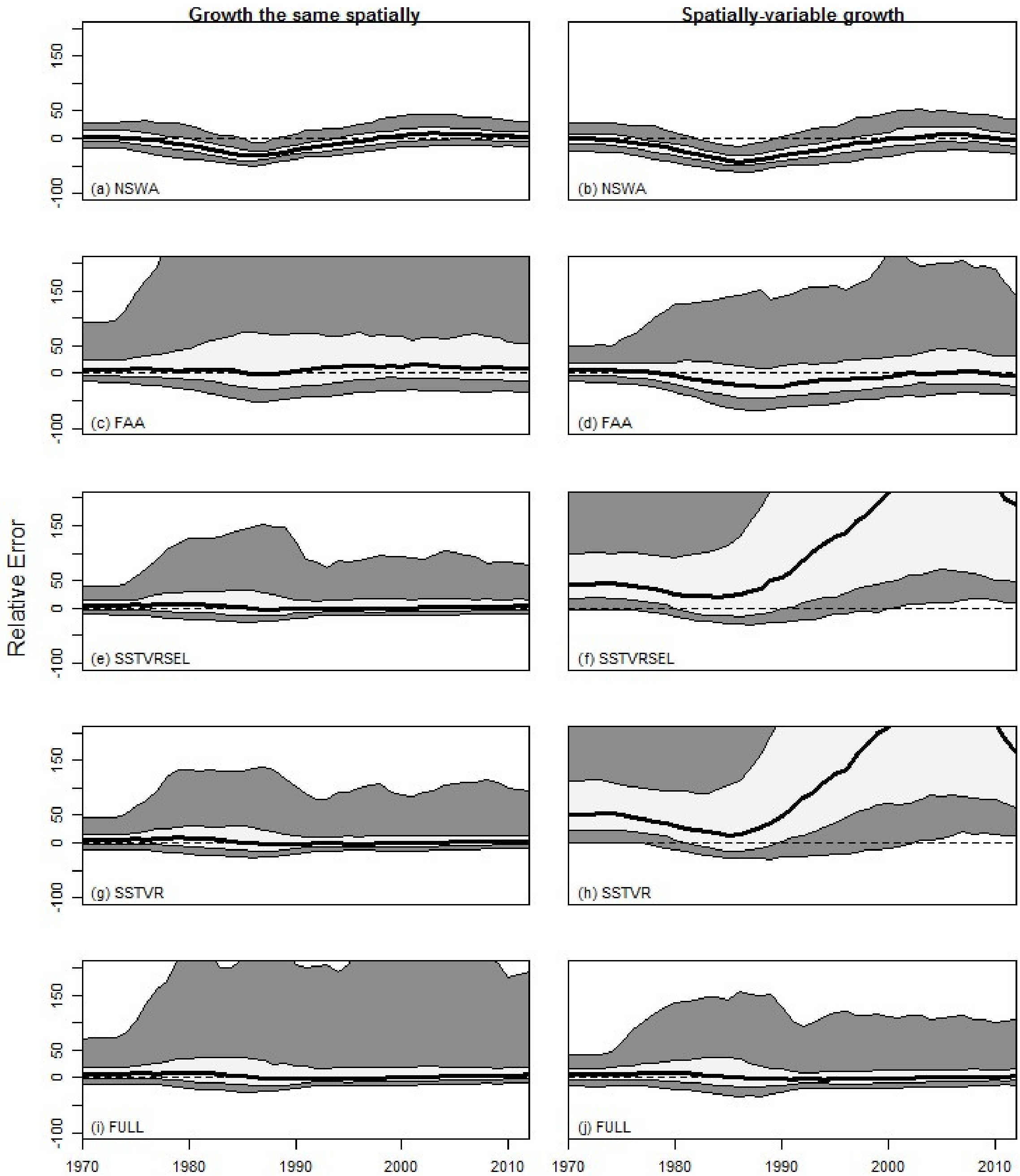


Figure 5.5: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the baseline scenario (no closed areas) and there are no survey data. Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels).

The estimates of total spawning biomass for the FAA, SSTVRSEL, SSTVR and FULL configurations are close to the true values in median terms when growth does not vary spatially in the operating model (Figure 5.5c, e, g, i). In contrast, the estimates of total spawning biomass from the NSWA configuration differ from the true values from 1980 to 1998 even when growth does not vary spatially (Figure 5.5a). This can be attributed to combining population (and hence catch rate) trends from zones that differ spatially. However, the estimates of total spawning biomass from the NSWA configuration are the most precise, likely due to this configuration estimating the fewest parameters and have the fewest years with missing data, irrespective of whether growth varies spatially or not. When growth does not vary spatially, the precision of the estimates of total spawning biomass from the spatially-structured assessment configurations is best (lowest level of between-simulation variation) for the SSTVRSEL configuration and precision becomes poorer as the number of parameters are increased (i.e. SSTVR and FULL).

The ability to estimate total spawning biomass, summarized in terms of MAREs for initial, and final total spawning biomass and for the ratio of final to initial total spawning biomass, is best for SSTVRSEL (initial total spawning biomass) and SSTVR (final total spawning biomass and the ratio of final to initial total spawning biomass) when growth is the same spatially (Table 5.1a, left columns). However, apart from FAA, there is little difference in MARE among the configurations when growth is the same spatially.

Table 5.1: Median (over simulations) absolute relative errors for three quantities of management interest for the scenarios in which there are no closed areas. The assessment configurations that achieve the lowest median absolute relative errors by performance measure are indicated in bold underline.

Assessment configuration	Growth is the same spatially			Spatially-variable growth		
	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass
(a) No survey data						
NSWA	9.10	8.55	12.17	**<u>9.66</u>**	12.38	17.07
FAA	11.91	26.02	23.72	10.75	26.22	23.70
SSTVRSEL	**<u>8.44</u>**	7.94	10.03	43.50	187.90	102.72
SSTVR	9.10	**<u>7.81</u>**	**<u>10.00</u>**	51.99	163.91	97.45
FULL	9.80	7.83	11.47	10.33	**<u>8.66</u>**	**<u>11.59</u>**
(b) With survey data						
NSWA	9.94	14.01	13.95	**<u>9.07</u>**	15.45	17.62
FAA	11.39	24.11	21.96	11.07	21.44	21.19
SSTVRSEL	**<u>8.87</u>**	7.08	10.18	77.41	339.76	153.99
SSTVR	9.12	**<u>6.69</u>**	**<u>9.77</u>**	94.78	379.24	153.90
FULL	9.72	7.55	10.84	10.49	**<u>8.65</u>**	**<u>11.46</u>**

The quantitative ranking of the assessment configurations differs markedly when growth varies spatially in the operating model than when growth does not spatially vary (Table 5.1; Figure 5.5). The two spatially-structured assessment configurations that assume that growth is the same for all zones (SSTVR and SSTVRSEL) are markedly biased and imprecise when grows varies spatially (Figure 5.5f, h). In contrast, the bias for the NSWA and FULL configurations are essentially the same irrespective of whether growth varies spatially in the operating model or not (Figure 5.5a, b, i, j). The estimates from the FAA configuration are negatively biased when growth varies spatially (Figure 5.5d). When expressed in terms of MARE, the NSWA (initial total spawning biomass) and the FULL assessment configurations (final total spawning biomass and final to initial total spawning biomass) perform best when growth varies spatially (Table 5.1a). However, the FULL assessment performs considerably better (i.e., less bias) than the NSWA configuration when spatial growth occurs, and is the most consistent model across underlying growth assumptions in the operating model.

The availability of survey data every 2nd year from 2000 does not change the relative ranking of the assessment configurations, irrespective of spatial variation in growth (Table 5.1b). However, the MAREs for the SSTVR and SSTVRSEL configurations are even larger when growth varies spatially and there are survey data, while the MAREs for the other assessment configurations are, with a few exceptions, lower, as might be expected given the availability of additional data.

5.4.2 Zone 10 is closed

Figure 5.6 shows relative error distributions for the scenarios in which zone 10 is closed to fishing in 2000 and there are no survey data (the results are qualitatively the same, but less extreme, for the scenarios in which zone 10 is closed to fishing in 2005; see Supplementary Figure S5.1 and Supplementary Figure S5.2 and Supplementary Table S5.2). Considering the results for the case in which growth is the same for all zones (Figure 5.6a, c, e, g, i), application of the spatially-structured assessment configurations leads to unbiased results (Figure 5.6e, g, i), but as expected from previous papers that have evaluated the consequences of closed areas on the performance of stock assessment methods (e.g., Punt and Methot, 2004; McGilliard et al., 2015), the estimates from the NSWA and FAA configurations are biased (Figure 5.6a, c). For this variant of the operating model, the bias for the FAA configuration is positive for much of the assessment period while the bias for the NSWA configuration increases over time for most of the assessment period. Therefore, in contrast to the operating model with no closed areas (Figure 5.5, Table 5.1), performance of the NSWA and FAA configurations are markedly poorer than the spatially-structured configurations in terms of bias (Figure 5.6a, c, e, g, i, Table 5.2a, left columns). The MAREs for the spatially-structured configurations are higher for the scenarios with closed areas than when there are no closed areas (compare the MAREs in Table 5.1a with those in Table 5.2a), even though the amount of length and age data is the same. This is particularly the case for SSTVSEL and SSTVR.

Allowing for spatially-varying growth in the operating model leads to larger relative errors in final total spawning biomass in the later years and a clear preference for the FULL configuration (Figure 5.6b, d, f, h, j; Table 5.2a, right columns). The FAA configuration generally performs better than the NSWA configuration in terms of MARE when zone 10 is closed to fishing (Figure 5.6a, b, c, d; Table 5.2a).

The availability of survey data from 2000 has little impact on the performance of the assessment configurations when there is no spatial variation in growth, except that the precision of the estimates is improved, particularly for SSTVRSEL, SSTRVR and FULL (Figure 5.7a, c, e, g, i; Table 5.2b). All assessment configurations, except FULL, provide biased estimates (Figure 5.7). The availability of survey data from 1990, i.e. before closures were first implemented, generally leads to lower MAREs compared to lack of such data, but this is not always the case, e.g. initial total spawning biomass from FAA when growth is spatially-varying (Table 5.2a vs Table 5.2).

Table 5.2: Median (over simulations) absolute relative errors for three quantities of management interest for the scenario in which a closed area is established in zone 10 in 2000. The assessment configurations that achieve the lowest median absolute relative errors by performance measure are indicated in bold underline.

Assessment configuration	Growth is the same spatially			Spatially-variable growth		
	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass
(a) No survey data						
NSWA	13.41	46.19	32.52	14.87	87.63	64.72
FAA	12.53	23.63	19.55	18.13	48.19	37.76
SSTVRSEL	**<u>8.88</u>**	9.86	**<u>12.01</u>**	418.39	863.32	76.30
SSTVR	9.57	**<u>8.73</u>**	13.43	478.67	940.82	72.82
FULL	11.71	10.86	14.19	**<u>11.24</u>**	**<u>10.81</u>**	**<u>12.41</u>**
(b) With survey data from 2000						
NSWA	16.28	44.89	24.47	17.61	72.51	50.71
FAA	10.96	24.57	20.48	12.32	22.36	20.25
SSTVRSEL	**<u>8.87</u>**	12.40	**<u>11.39</u>**	133.35	424.35	119.86
SSTVR	13.13	11.92	10.52	197.85	578.29	108.85
FULL	12.19	**<u>11.30</u>**	10.79	**<u>11.83</u>**	**<u>9.46</u>**	**<u>11.20</u>**
(c) With survey data from 1990						
NSWA	15.38	46.34	27.13	12.95	64.64	47.05
FAA	9.34	19.48	17.37	19.94	62.57	43.45
SSTVRSEL	**<u>9.19</u>**	**<u>8.48</u>**	10.77	132.32	618.47	191.87
SSTVR	11.12	7.59	11.07	254.50	1163.65	210.87
FULL	12.45	8.86	**<u>10.29</u>**	**<u>11.22</u>**	**<u>7.65</u>**	**<u>10.19</u>**

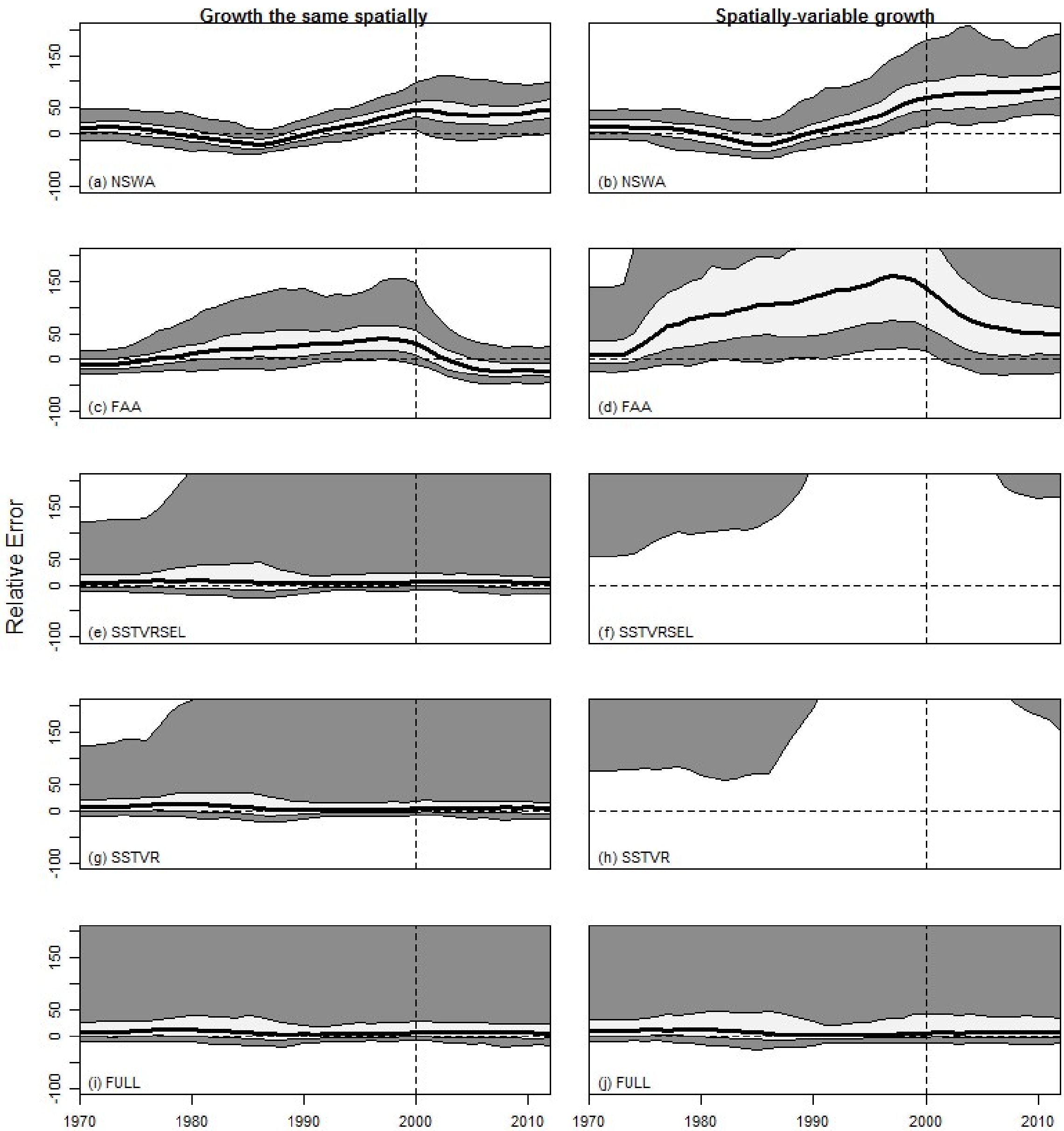


Figure 5.6: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the scenario in which zone 10 is closed to fishing in 2000 (vertical dashed lines) and there are no survey data. Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels).

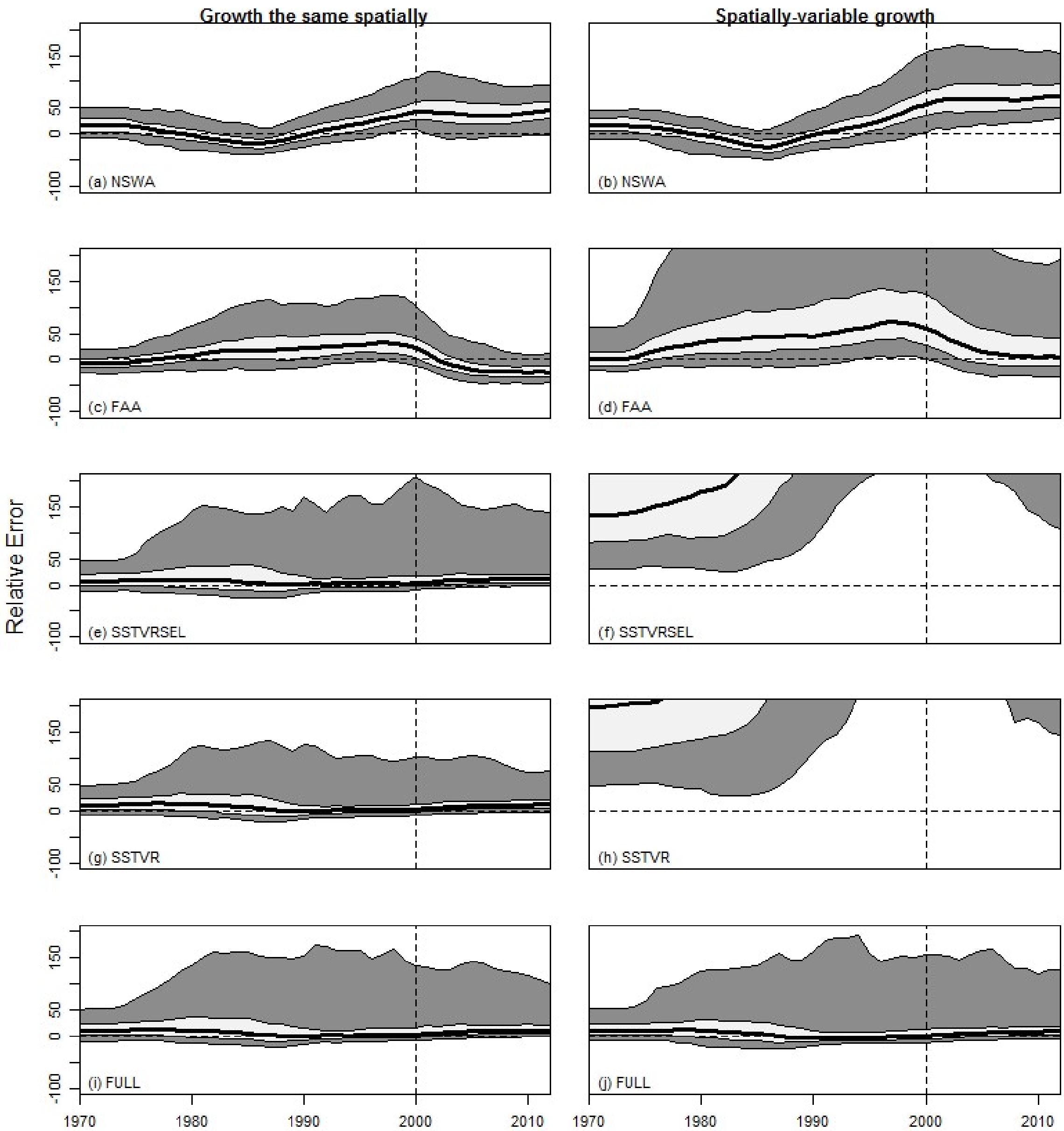


Figure 5.7: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the scenario in which zone 10 is closed to fishing in 2000 (vertical dashed lines). Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels). Survey data are available from 2000.

5.4.3 Sensitivity analyses

Zone 20 is closed

Zone 20 was assumed in the operating model to have a substantially larger biomass in an unfished state than zone 10 (Figure 5.4). The qualitative results for four of the five assessment configurations (FAA, SSTVRSEL, SSTVR, and FULL) are the same as for when zone 10 is closed (Figure 5.8). However, the MAREs are generally larger when zone 20 is closed (Table 5.2a, b and 3a, b). In contrast to these four configurations, the trend in relative error over recent years differs for the NSWA configuration between when zones 10 and 20 are closed (Figure 5.6a, b; Figure 5.8a, b). The relative error for NSWA increases when zone 10 is closed and decreases when zone 20 is closed. The magnitude of the trend in relative error is greatest when there is spatial variation in growth, irrespective of which zone is closed. The declining relative errors over time mean that the MAREs for the NSWA configuration for final total spawning biomass and the ratio of final to initial total spawning biomass are lower when zone 20 is closed than when zone 10 is closed (Table 5.2a and Table 5.3a).

Table 5.3: Median (over simulations) absolute relative errors for three quantities of management interest for the scenario in which a closed area is established in zone 20 in 2000. The assessment configurations that achieve the lowest median absolute relative errors by performance measure are indicated in bold underline.

Assessment configuration	Growth is the same spatially			Spatially-variable growth		
	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass
(a) No survey data						
NSWA	22.97	39.58	27.81	34.11	48.01	27.58
FAA	18.60	40.07	39.98	13.71	41.67	40.34
SSTVRSEL	**<u>12.11</u>**	**<u>19.34</u>**	**<u>16.57</u>**	49.94	98.80	40.03
SSTVR	12.18	19.84	17.86	56.78	95.52	39.35
FULL	13.97	20.35	20.06	**<u>13.23</u>**	**<u>23.01</u>**	**<u>20.89</u>**
(b) With survey data from 2000						
NSWA	17.97	39.34	32.29	31.11	47.50	28.57
FAA	15.83	38.20	38.49	**<u>13.56</u>**	33.49	33.72
SSTVRSEL	**<u>12.24</u>**	17.10	14.59	67.90	231.14	130.14
SSTVR	13.62	**<u>16.88</u>**	**<u>13.67</u>**	75.97	210.25	96.29
FULL	12.99	17.48	14.58	14.57	**<u>20.42</u>**	**<u>17.80</u>**

Ignoring the catch-rate data when applying the NSWA configuration

Ignoring the catch-rate data after 2000 did not lead to marked reductions in bias. Instead any reductions in bias were more than offset by higher variation, in particular, with no composition data after 2000, the ability to estimate fishery selectivity was poorer (results not shown).

Can estimation performance be improved by using model selection methods?

Table 5.4Table 5.4 shows MAREs for the selected configurations when AIC is used to select a best model from the spatial assessment configurations. The MAREs for the assessment configurations selected using AIC are lower than for any individual configuration, i.e. the results in Table 5.4 highlight the value of using model selection criteria to select whether to estimate spatial variation in growth and selectivity. In general, SSTVRSEL was selected for the scenarios in which growth did not vary spatially while FULL was selected when growth varied spatially, i.e. Stock Synthesis is able to correctly detect spatial variation in growth, at least at the level indicated in Figure 5.3.

Table 5.4: Median (over simulations) absolute relative errors for three quantities of management interest when the optimum spatial model for each replicate is selected using AIC. The scenarios on which these analyses are based did not involve survey data.

Operating model scenario	Growth is the same spatially			Spatially-variable growth		
	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass
No closed area	8.74	7.81	10.24	10.33	8.66	11.59
Closed area in Zone 10	9.33	9.86	12.12	12.24	10.81	12.41
Closed area in Zone 20	11.91	18.90	17.20	13.23	22.94	17.80

Estimates of spawning biomass by zone

The aim of this paper was to determine how well alternative assessment configurations are able to estimate total spawning biomass. However, the spatial assessment configurations estimate spawning biomass by zone. The ability to estimate zone-specific spawning biomass for a subset of the scenarios is summarized in the Supplementary Material. The extent of error depends on zone (least for zone 30 and largest for zone 10), irrespective of whether there are closed areas and whether growth varies spatially or not. Estimates of current and final spawning biomass are essentially unbiased for zones 10 and 20 and negatively biased for zone 30 for all assessment configurations when there are no closed areas and growth is the same spatially. As expected, spatially-varying growth leads to very poor performance for SSTVRSEL and SSTVR, with the estimates of spawning biomass for zone 10 being more positively biased and those for zone 30 more negatively biased. The patterns evident in Supplementary Figure S5.3 are also evident when zone 10 is closed (Supplementary Figure S5.4), but the magnitude of positive bias for zone 10 is larger than when there were no closed areas.

5.5 Discussion

The selection of an assessment configuration to apply in the face of spatial variation in exploitation rate, including closed areas, has been explored by Punt and Methot (2004), Garrison et al. (2011) and McGilliard et al. (2015). Punt and Methot (2004) recommended conducting assessments for each zone while Garrisson et al. (2011) and McGilliard et al. (2015) advocated the use of spatially-structured stock assessment models even if it is not possible to estimate movement rates among zones adequately. Punt et al. (2015) drew a similar conclusion that it is desirable to use spatially-structured assessment methods in the face of spatial variation in exploitation rates and biological parameters, but highlighted the consequences of selecting a mis-specified spatial model. Use of such methods in actual stock assessments (rather than in research on stock assessments) is still uncommon, but the number of such assessments is increasing steadily over time. For example, the most recent assessment of canary rockfish off the US west coast was based on a spatially-structured assessment, although the results from the spatially-structured version of the assessment were similar to those of a spatially-aggregated (fleets-as-areas) assessment (Thorson and Wetzel, 2015).

5.5.1 Selection of a best configuration

This study has highlighted the impact of uncertainty in growth rates on the performance of stock assessment methods. The 'best' assessment configuration in the face of spatial variation in growth was FULL whereas SSTVR and SSTVRSEL outperformed FULL when growth was the same spatially, a result that is consistent with those reported by Punt et al. (2015). This effect was present whether there were closed areas or not. This raises the question of whether it is ‘safer’ to estimate more parameters even when they are not supported by the data. In general, the study showed that estimating spatially-varying growth did not lead to much poorer estimation performance when growth did not vary spatially, but led to markedly improved estimation performance when this was the case. In contrast, ignoring spatial growth could lead to large biases when growth actually varied spatially.

It should be noted that the best performing (least biased) configurations (SSTVRSEL when growth does not vary spatially and FULL when growth does vary spatially) matched the operating model exactly so the performances of these best-performing configurations are likely over-estimated compared to the other configurations.

A common approach to dealing with spatial variation in biological parameters and exploitation rate is to divide the area into small regions and conduct assessments for each region, and this approach was evaluated using simulation by Punt and Methot (2004) and McGilliard et al. (2015) when the area being assessed contains a closed area. This approach is equivalent to the FULL approach, except that the FULL configuration can estimate common parameter values when the values of some of the parameters are shared across regions. For example, SSTVRSEL is equivalent to conducting assessments by zone, except that selectivity and growth are assumed to be the same for all three zones. An advantage of assessing multiple zones within one assessment is that it becomes possible to “borrow strength”, e.g. the trend in recruitment for the closed zones is informed by the data for the zones for which data are available.

The qualitative trends in relative error for four of the assessment configurations (FAA, SSTVR, SSTVRSEL, and FULL) are the same irrespective of which zone is closed. However, the trends in relative error for the NSWA configuration over recent years differed between whether zones 10 or 20 were closed (Figure 5.6 and Figure 5.8). The magnitudes of the trends in relative error also differed depending on whether growth was the same spatially or varied spatially. Thus, there was an interaction between spatial variation in growth, the size of the area closed, and trend in exploitation rate by area. A declining trend over time in relative error over recent years (e.g. as seen for the NSWA configuration for zone 20, Figure 5.8a, b) is expected from previous research (e.g. Punt and Methot, 2004; McGilliard et al., 2015) because catch-rate trends for the final years of the assessment period for the area open to fishing do not reflect the trend in the overall biomass (declining even though the total population size is increasing). In contrast, the rate of increase in biomass for zone 10 is less than for zones 20 and 30, so having a closed area in zone 10 means that the rate of increase is overestimated – the size of the effect of trends in catch rate not mimicking those of the population size are less when zone 10 is closed because zone 10 is smaller than zones 20 and 30.

This paper explored the performance of five assessment configurations, three of which were also considered by Punt et al. (2015). The performance of the other two assessment configurations (RSWA and BZSA) would have behaved similarly to NSWA and FAA respectively. Other assessment configurations could have been considered, including assessment configurations that extend FULL by allowing for movement among zones. However, such a configuration would not likely have performed better than the FULL configuration because for this study the FULL configuration matches the operating model used to generate the data.

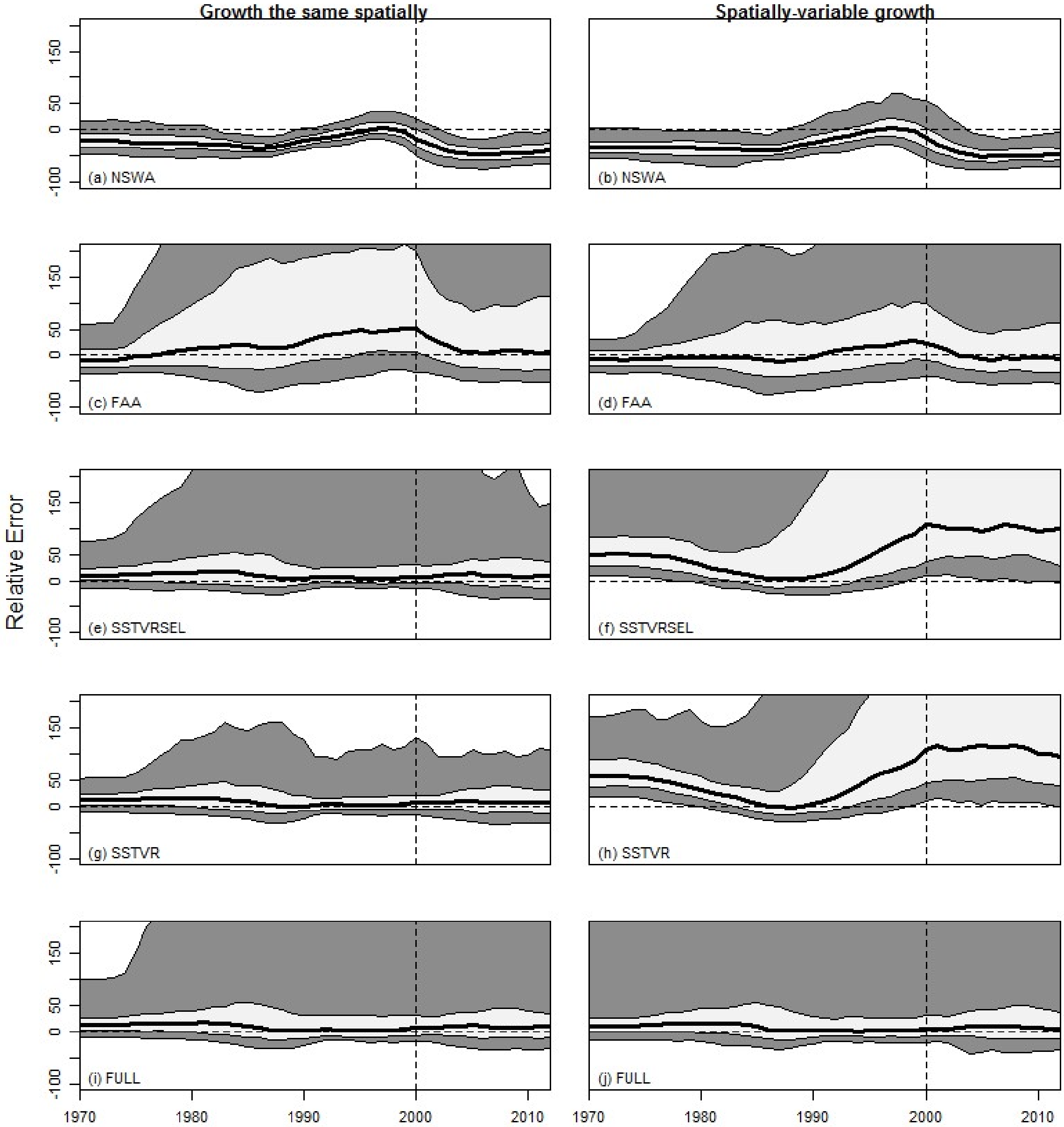


Figure 5.8: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the scenario in which zone 20 is closed to fishing in 2000 (vertical dashed lines) and there no survey data. Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels).

5.5.2 Value of survey data

This study explores the value of survey data in the face of closures. In general, the effects of model mis-specification dominate those of additional data, although, as expected, the MAREs for the assessment configuration that is correctly specified are lower when relatively precise survey data are available, particularly when the survey data started before the closed areas were implemented. However, the magnitude of improved performance associated with survey data is fairly limited even though the assumed survey CV is very small (0.1). Punt et al. (2015) note that a key reason for the inability to estimate biomass is the lack of catch-rate and compositional data for the early years of the fishery and this problem is not overcome by collecting survey data for recent years. Moreover, survey data for the actual situation considered in this paper only started in 2008.

5.5.3 Caveats

The caveats associated with the analyses of this paper match those of (Punt et al. 2015), namely that some of the values for the parameters of the operating model are assumed to be known exactly. Specifically, the steepness of the stock–recruitment relationship and σ_R are assumed to be known exactly. In addition, the data are generated under the assumption that fish are homogenously distributed throughout each zone. This is unlikely to be correct in general and for Pink Ling in particular. This assumption should provide spatially-structured assessment configurations with an advantage over spatially-aggregated assessment configurations (Punt et al., 2015). Another key assumption is that the boundaries between the zones are correctly located with respect to differences in exploitation rate. The estimation performance of all of the configurations, especially those that are spatially-structured are likely to have been poorer had the boundaries not been correctly delineated. Finally, the analyses of this paper pertain to the case when there is no movement between zones, except at the larval level. This case applies fairly generally, including several rockfishes off the US west coast, but is simplified compared to situations for species such as tunas that exhibit considerable adult movement. It is unclear whether the general conclusions of this paper pertain to the case where there is substantial movement of animals of ages 1 and older (see, for example, Ying *et al*., 2011).

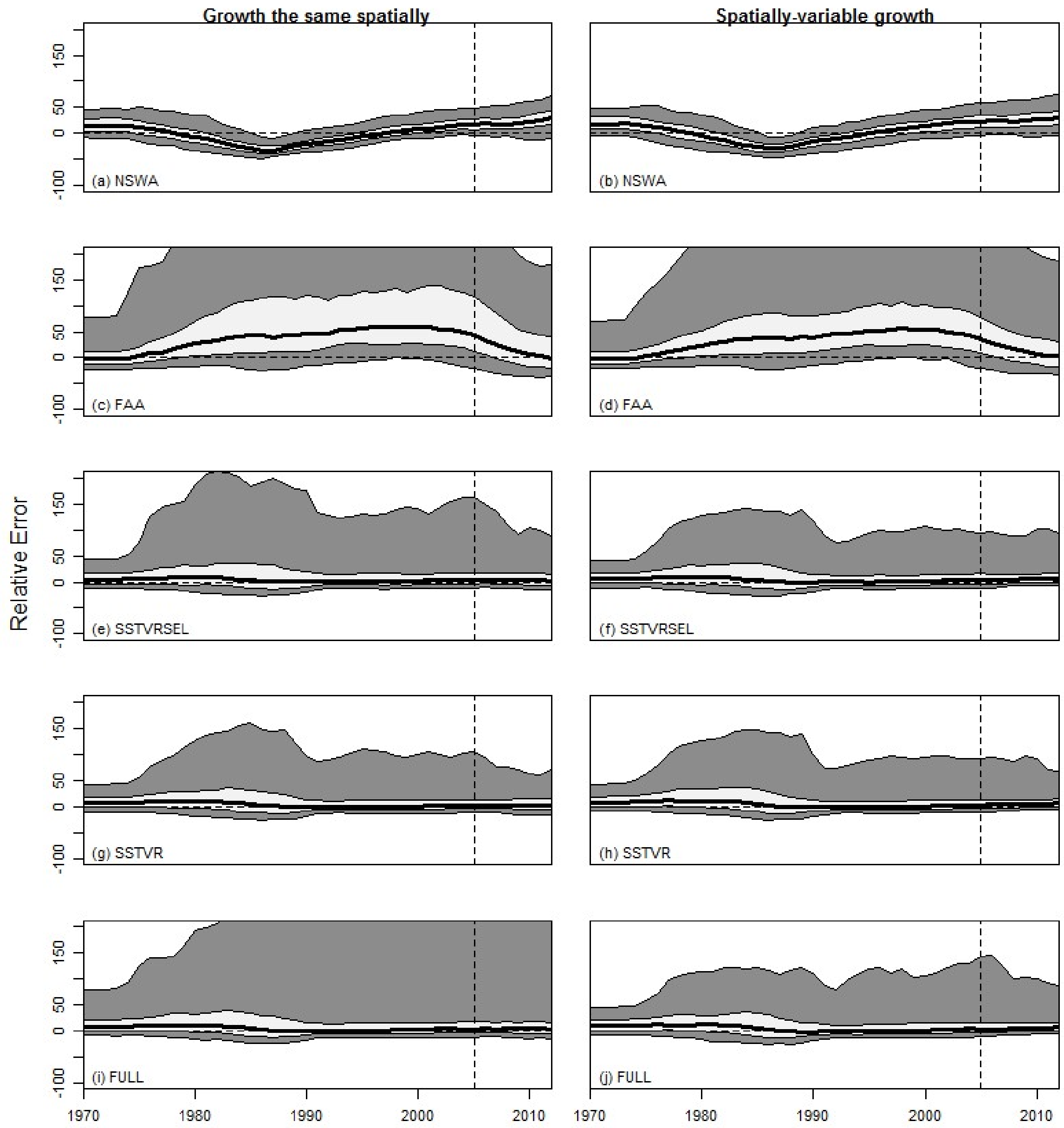
5.5.4 Management implications

The results of this paper (and other papers that have explored the implications of closed areas) highlight that closed areas increase uncertainty (and bias when the assessment is mis-specified). However, it is not clear whether basing management advice on the outcomes of such assessments will lead to an inability to achieve management goals with respect to sustainability. This question could be addressed using Management Strategy Evaluation, MSE (Punt et al., in press).

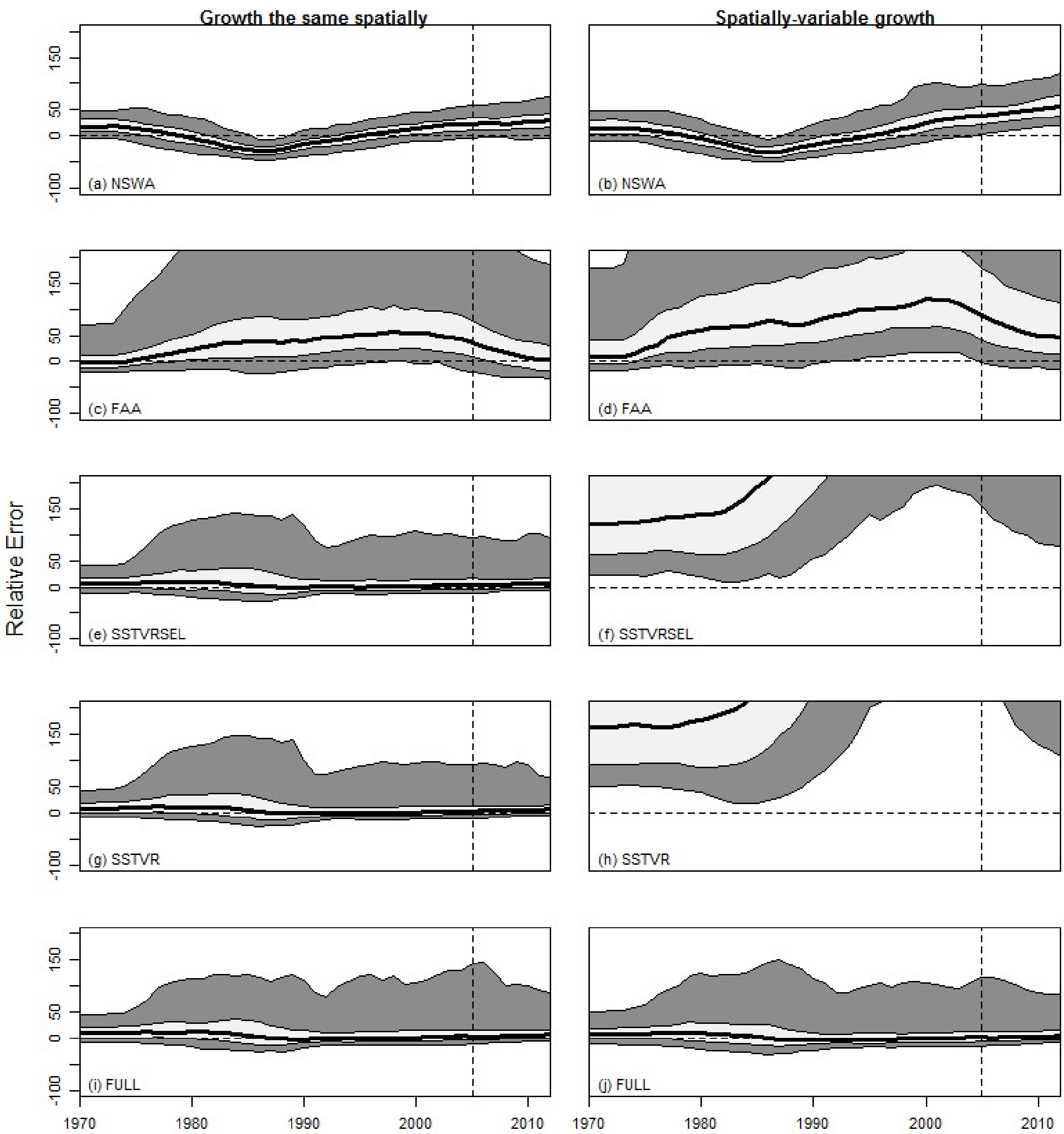
The spatial assessment configurations estimate biomass by zone. In principle, the results from such configurations could form the basis for spatial management (e.g. assessing stock status by zone and perhaps setting catch limits spatially based on spatial stock status). A variety of ways of calculating catch limits could be used and evaluated using MSE. For example, the Revised Management Procedure of the International Whaling Commission (IWC, 2012) includes options to spread catches spatially. How well these types of options perform could be explored using an MSE based on the operating model of this paper.

5.5.5 Conclusion

The study extends previous evaluations of the performance of assessment methods in the face of spatial variation in exploitation rate by highlighting that there are interactions between spatial trends in growth rates and those in exploitation rate as well as with the sizes of the areas over which growth and exploitation rate vary.

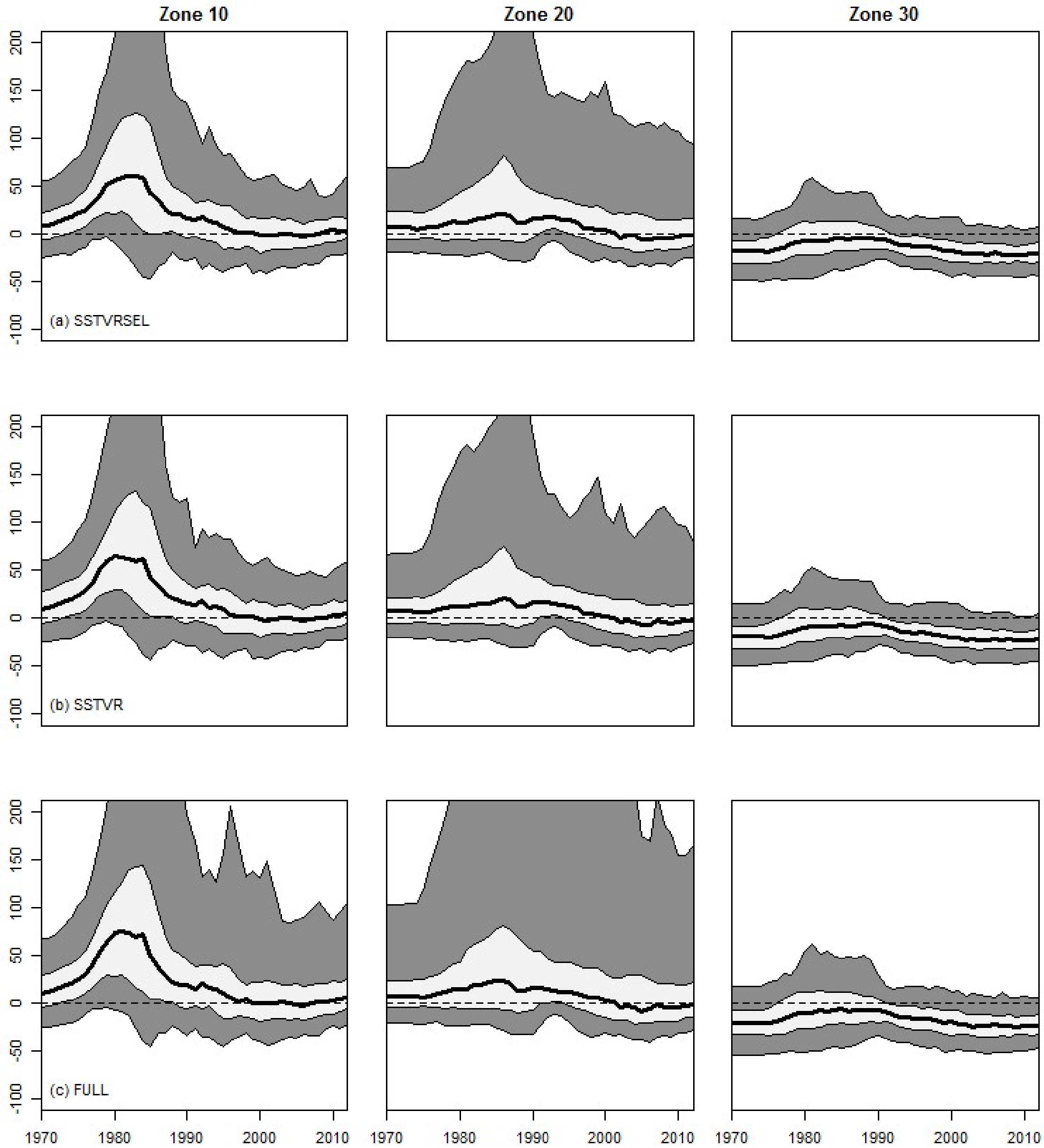


Supplementary Figure S5.1: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the scenario in which zone 10 is closed to fishing in 2005 (vertical dashed lines). Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels).

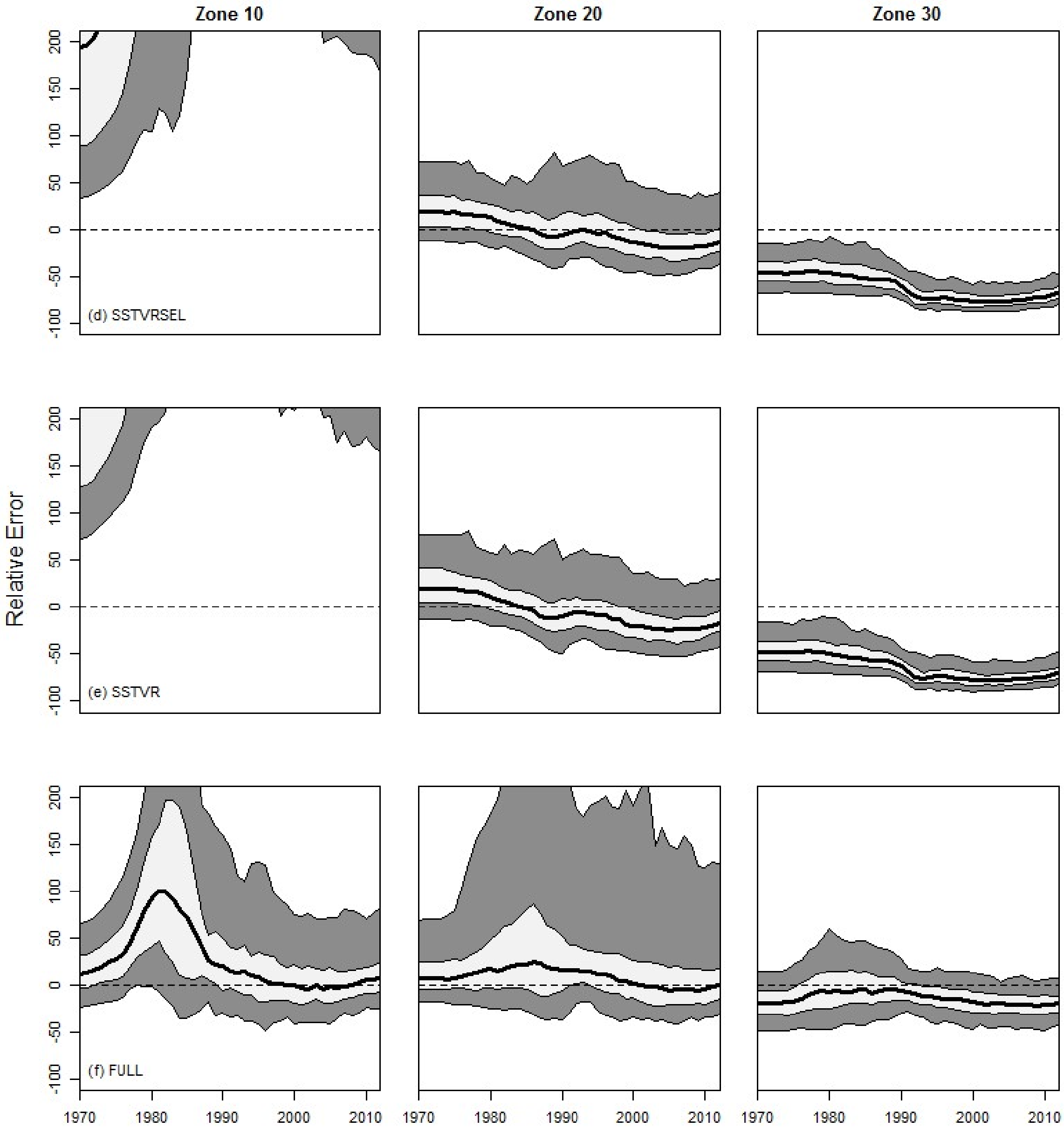


Supplementary Figure S5.2: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of total spawning biomass for five assessment configurations for the scenario in which zone 10 is closed to fishing in 2005 (vertical dashed lines). Results are shown when growth is the same in all zones (left columns) and when growth varies spatially (right panels). Survey data are available from 2005.

(i) Growth is the same in all zones.

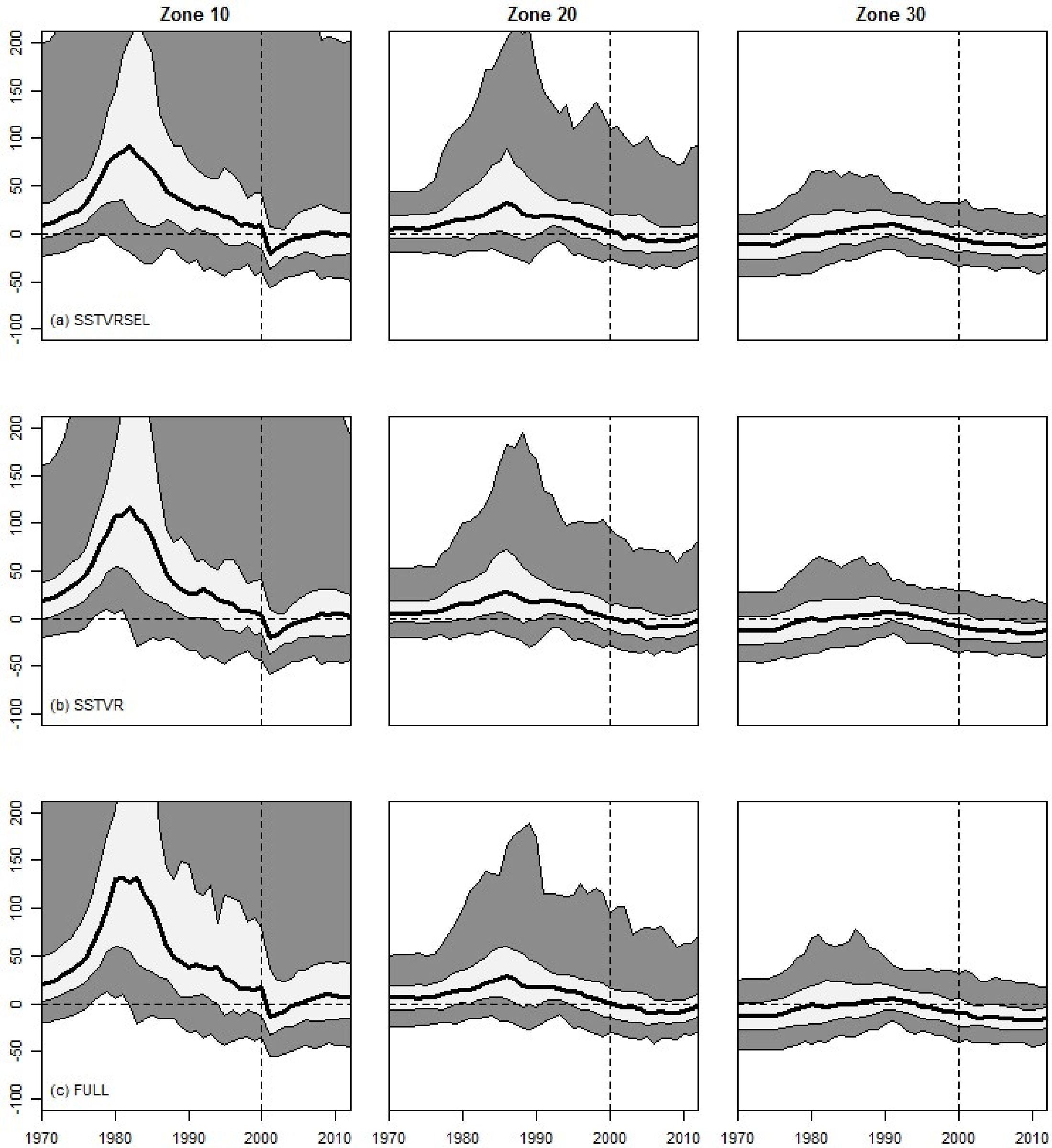


(ii) Growth varies spatially.

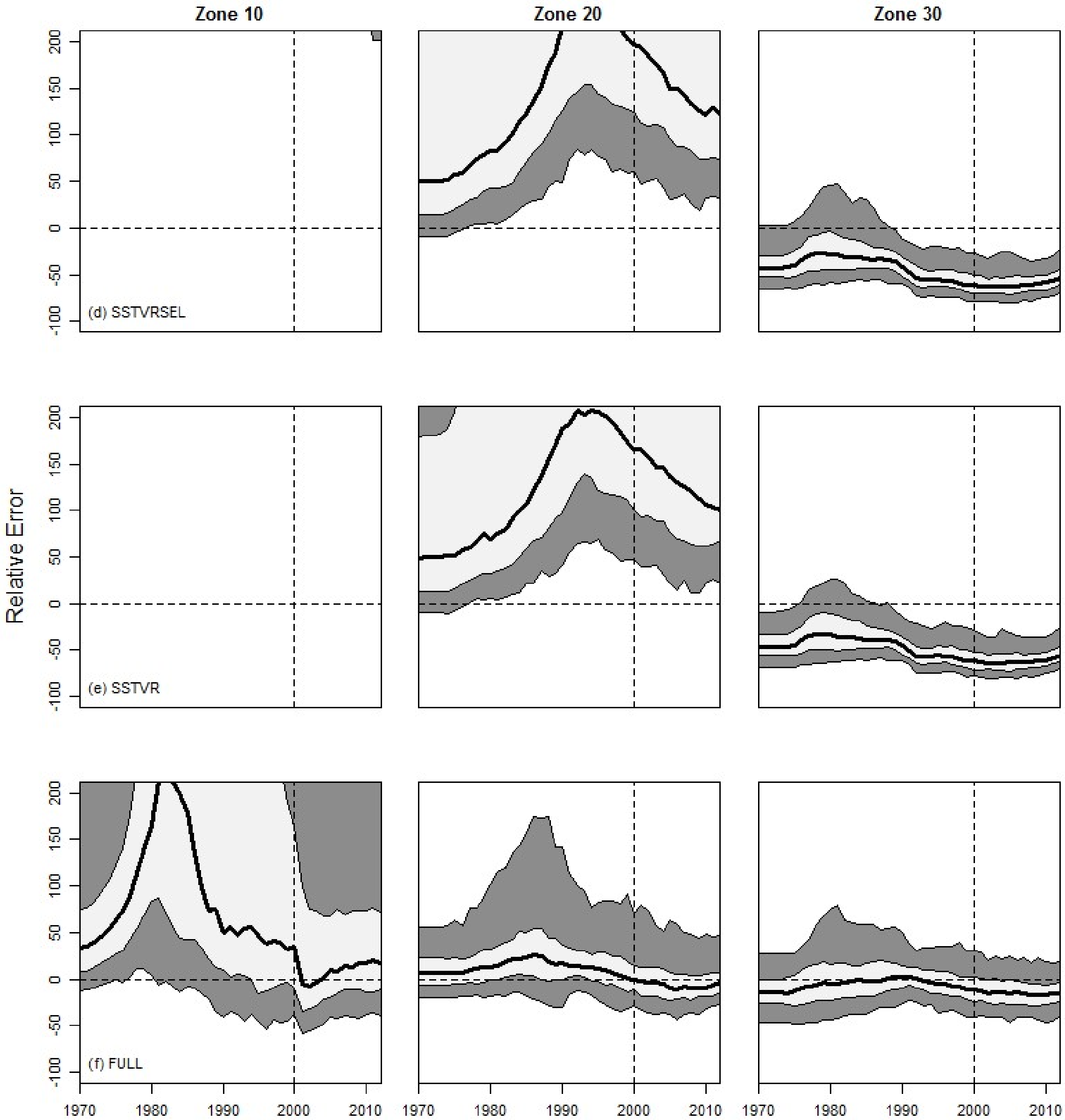


Supplementary Figure S5.3: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of spawning biomass by zone for the three spatial assessment configurations for the scenario in which there are no closed areas.

(i) Growth is the same in all zones.



(ii) Growth varies spatially.



Supplementary Figure S5.4: Relative error distributions (median relative errors, with 50% and 90% intervals) for the time-trajectory of spawning biomass by zone for the three spatial assessment configurations for the scenario in which zone 10 is closed to fishing in 2000.

Supplementary Table S5.1: Percentage of simulations for which the Hessian matrix is positive definite.

Assessment configuration	No MPA		Zone 10 closed (from 2000)		Zone 10 closed (from 2005)		Zone 20 closed (from 2000)	
	No	Yes	No	Yes	No	Yes	No	Yes
Spatially-variable growth								
NSWA	91	91	98	93	96	90	80	80
FAA	78	76	92	91	95	90	60	66
SSTVRSEL	93	91	95	99	98	93	90	92
SSTVR	78	77	97	83	86	77	71	58
FULL	70	80	71	69	81	66	65	66
Growth is the same spatially								
NSWA	95	86	94	91	92	88	69	69
FAA	83	76	76	78	93	85	81	69
SSTVRSEL	83	88	88	90	88	83	79	84
SSTVR	71	73	74	77	78	69	63	63
FULL	77	73	77	75	77	68	67	69

Supplementary Table S5.2: Median (over simulations) absolute relative errors for three quantities of management interest for the scenario in which a closed area is established in zone 10 in 2005. The assessment configurations that achieve the lowest median absolute relative errors by performance measure are indicated in bold underline.

Assessment configuration	Growth is the same spatially			Spatially-variable growth		
	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass	Initial spawning biomass	Final spawning biomass	Final to initial spawning biomass
(a) No survey data						
NSWA	14.29	28.95	18.84	17.88	30.61	15.76
FAA	11.95	26.17	21.94	10.98	19.92	18.65
SSTVRSEL	**<u>8.92</u>**	**<u>9.39</u>**	11.47	<u>8.74</u>	7.80	**<u>10.19</u>**
SSTVR	9.03	9.65	**<u>11.38</u>**	10.45	7.38	10.11
FULL	9.96	9.54	11.84	11.01	**<u>7.60</u>**	11.32
(b) With survey data from 2000						
NSWA	17.88	30.61	15.76	15.50	56.77	33.90
FAA	10.98	19.92	18.65	14.45	46.26	36.18
SSTVRSEL	**<u>8.74</u>**	7.80	**<u>10.19</u>**	119.74	450.38	149.25
SSTVR	10.45	**<u>7.38</u>**	10.11	161.99	582.47	145.48
FULL	11.01	7.60	11.32	**<u>9.62</u>**	**<u>6.80</u>**	**<u>10.86</u>**

6 The effect of marine closures on a feedback control management strategy used in a spatially-aggregated stock assessment: A case study based on Pink Ling in Australia

Punt, A.E.[1,2], Haddon, M.[2], Little, L.R.[2], Tuck, G.N.[2]

[1]*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, USA*

[2]*CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, Tasmania, Australia*

This has been published as André E. Punt, Malcolm Haddon, L. Richard Little, Geoffrey N. Tuck 2016. The effect of marine closures on a feedback control management strategy used in a spatially aggregated stock assessment: a case study based on pink ling in Australia, Canadian Journal of Fisheries and Aquatic Sciences 74, 1960- 1973.

6.1 Abstract

Simulation is used to explore the effect of spatial heterogeneity, including spatial closures, on the ability of feedback-control management strategies to achieve goals relating to conservation and utilization of fishery resources. The operating model underlying the projections is based on pink ling, *Genypterus blacodes,* off southern Australia assumes that animals are sedentary following settlement. The management strategies are able to move the resource towards the target level in the absence of spatial closures even though assessment results are biased. The probability of reducing the stock below its limit reference point is higher when growth rates vary spatially, but the effect is small. The probability of the stock being above its target reference point is lower when one of the smaller spatial strata is closed. However, performance is markedly different when a large fraction of the area is closed, with the stock substantially larger than the target level at the end of the projection period.

Keywords: age-structured stock assessment methods, closed areas, harvest control rules, simulation, spatial structure

6.2 Introduction

Spatial closures have always been a standard component of the toolbox used to achieve single-species goals in marine fisheries. Closures off the US west coast are used to reduce fishing mortality on stocks that have been declared to be overfished and are in need of rebuilding (Field et al. 2006); and closures in Australia's Northern Prawn Fishery have been introduced to reduce fishing mortality on juvenile prawns (AFMA 2015). Closures can also be implemented for biodiversity conservation (e.g., Anon 2015), reduction of impacts on threatened and endangered species (e.g., Greenstreet et al. 2006; Daley et al. 2015; AFMA 2012), and to protect habitat (e.g., Field et al., 2006).

Although the number of spatial closures has increased over time, so has their diversity. Some spatial closures are entirely no-take, while others are restricted to gear types such as bottom trawling restrictions on sea mounts (with pelagic gear allowed); still others, like the closures to protect overfished rockfishes *Sebastes* species off the US west coast, are seasonal. Spatial closures are often capable of achieving single-species and ecosystem goals (Field et al. 2006), but may come at the cost of increasing spatial heterogeneity in population abundance, and size- and age-structure compared to unfished conditions (e.g., Lester et al. 2009; McCook et al. 2010; Aburto-Oropeza et al. 2011).

It is common to ignore closures and associated spatial heterogeneity when conducting stock assessments (Punt et al. 2015), particularly when there is limited movement of animals post-recruitment. Several studies (e.g., Punt and Methot 2004; Garrison et al. 2011; McGillaird et al. 2015; Punt et al. 2016) have explored the impact of no-take spatial closures on the ability of stock assessments to provide accurate and precise estimates of management quantities. These studies have generally shown that ignoring spatial structure leads to bias, and this bias is exacerbated by closures. They also show that spatially-structured stock assessments reduce bias in estimates of biomass, at the cost of lower precision (Punt et al. 2015), but may lead to biased estimates of movement parameters (McGillaird et al. 2015). Ideally, spatially-structured stock assessments should be based on tagging data to inform estimates of movement, but such data are available for very few stocks (exceptions in Australia include the assessment of gummy shark *Mustelus antarcticus*, Pribac et al. 2005, and toothfish *Dissostichus eleginoides,* Day et al. 2015). Consequently, it is common to either ignore spatial structure and instead apply the fleets-as-areas approach, which accounts approximately for spatial structure (Punt et al. 2014; Waterhouse et al. 2014) or use spatial models with no tagging data (McGillaird, et al. 2015; Szuwalski and Punt 2015).

Although the estimates from assessments that ignore spatial structure, or use approximate methods for addressing spatial structure such as the fleets-as-areas approach, are biased (and imprecise), the use of such estimates in harvest control rules may not necessarily lead to an inability to achieve management goals. This is because the feedback nature of harvest control rules mean that errors may be corrected over time. This paper starts to address the question of whether closing areas negatively impacts the ability to achieve management goals by conducting one of the first management strategy evaluations (MSE; Smith 1993, 1994; Punt et al. In press) in which stock assessments that ignore spatial-structure are used to manage a system in which there is a large no-take closed area.

The simulations are based on pink ling, *Genypterus blacodes*, off southern Australia, a component of Australia's Southern and Eastern Scalefish and Shark Fishery (SESSF). The SESSF covers the region from southern Queensland, around Tasmania, to Cape Leeuwin in Western Australia, and is one of Australia's largest and most valuable Commonwealth (federal)-managed fisheries in Australia. Pink Ling is appropriate as the basis for the study because it is assessed as two separate stocks, separated east and west at 147˚E, and is a species that is believed not to move much after recruitment. Since 2005, four seasonal closures (from approximately September to November) have been in place to protect the spawning stock and reduce fishing mortality at Maria Island, Seiner's Horseshoe and Everard Horseshoe in the east, and the Ling Hole in the west (Figure 6.1). Closures within these areas have been both voluntary and legislated (SEMAC 2012). These closures are relatively small in area, but are considered to be located over some the most productive and previously favoured fishing grounds.

The MSE allows for three spatial zones (nominally zones 10, 20, and 30 of the SESSF; Figure 6.1), one of which (either zone 10 or zone 20) can be assumed to be closed. The stock structure hypothesis for this paper is that there is a single biological stock across the three spatial zones, i.e. the recruitment for each zone is determined by the total spawning biomass across all three zones. The fish populations in these three zones are assumed to be mixed through the settlement of age-0 animals, with animals of age-1 and older not moving between zones. Two fleets (essentially trawl and non-trawl) are assumed to operate in each zone, growth can be assumed to differ among zones, and recruitment is assumed to be stochastic, with spatial variation in the proportion of the total recruitment that settles to each zone, as well as temporal variation in total recruitment. Management for Pink Ling is based on an Individual Transferable Quota system in which the Total Allowable Catch (TAC) is set based on a Recommended Biological Catch (RBC). A tier system of harvest control rules is available to set the RBCs for all SESSF species (Dichmont et al. in press). Pink Ling is a 'tier 1' stock and its RBC is based on applying a harvest control rule to the outcomes from stock assessments (separately east and west of 147°E) implemented using Stock Synthesis (Methot and Wetzel 2013) and CASAL (Bull et al. 2005). TACs are usually set below RBCs to account for discarding and state catches (Smith et al. 2008), but for the purposes of this paper, the TAC is set equal to the RBC.

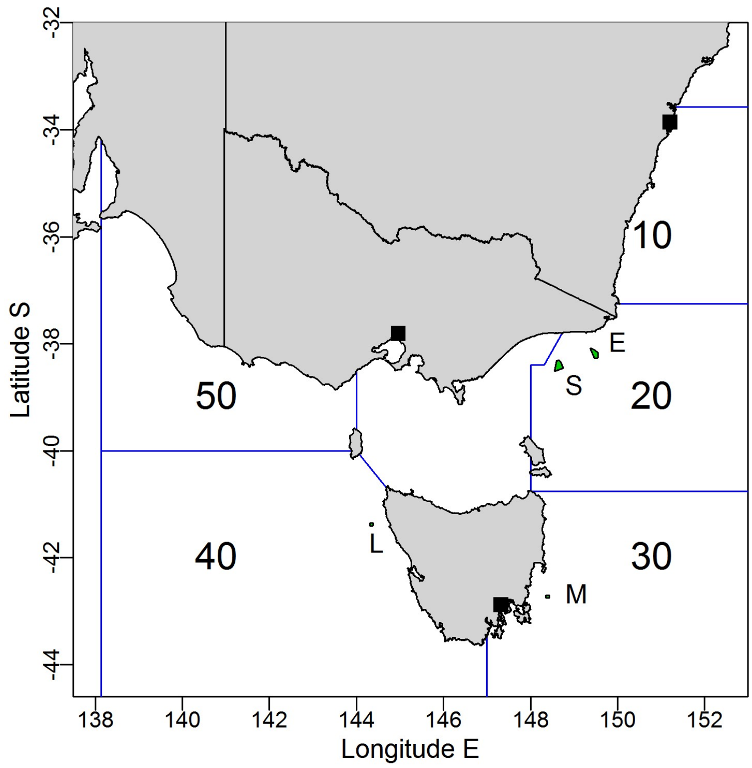


Figure 6.1: Schematic map of SESSF reporting blocks 10 – 50, with the fine blue lines representing block boundaries. The locations of Sydney, Melbourne, and Hobart are indicated by black squares from top to bottom. The east stock of pink link is found in zones 10, 20 and 30; the line between zones 30 and 40 is at 1470E. The real world closures include M, Maria Island; S, Seiners Horseshoe; E, Everard Horseshoe, and L, the Ling Hole.

The analyses of this paper compare the performance of the management system (spatially-aggregated stock assessments and linked harvest control rules) for Pink Ling when there is heterogeneity in population structure caused by spatial variation in fishing mortality and growth (but no spatial closures) and when the degree of spatial heterogeneity is increased by the introduction of spatial closures. The analyses, which extend earlier explorations of the impact of spatial structure on the estimation performance of stock assessments (Punt et al. 2015, 2016), also explore the sensitivity of the performance metrics to the availability of survey data from the closures, as well as the area that is closed.

The primary aim of the study is to examine whether biases in assessment outcomes caused by spatial heterogeneity in population structure will lead to an inability to achieve management goals.

6.3 Methods and Materials

The simulation evaluation is based on specifying a model of the population dynamics and projecting it forward using catch limits based on a management strategy. This ‘operating’ model is assumed to represent the truth for the projections, and is used to generate pseudo data sets. The projections of the operating model assume that the TACs set by the management strategy are taken exactly, i.e. no account is taken of mis-reporting, including catches being taken illegally from the closed area; a satellite vessel monitoring system is used to enforce spatial boundaries within the SESSF.

6.3.1 The operating model

The operating model includes spatial variation in growth and in the proportion of the total recruitment that settles by zone. The historical period covered by the operating model is a 43-year period (nominally ‘1970 ‘to ‘2012’), and the projections involve a 20-year period during which catch limits are updated every third year based on applying a harvest control rule to the outputs from a stock assessment. The choice of conducting assessments every third year mimics the 2012 move in the SESSF to less frequent stock assessments, and reduces the computational demands of the analyses performed here.

The three zones represented in the operating model are assumed to receive different proportions of the total recruitment in an unfished state (0.28, 0.49, and 0.23 respectively for zones 10, 20 and 30, which reflect roughly the relative amount of habitat for Pink Ling off southeastern Australia; Whitten and Punt 2014), with the extent of variation in spatial distribution, σ_ϕ, set to 0.7. Given a Beverton Holt stock-recruitment relationship, the recruitment (at age-0) to zone z at the start of year y, R_y^z, is given by:

$$R_y^z = \frac{e^{\phi^z+\eta_y^z}}{\sum\limits_z e^{\phi^z+\eta_y^z}} \frac{4\,h\,R_0\,\tilde{S}_y/\tilde{S}_0}{(1-h)+(5h-1)\tilde{S}_y/\tilde{S}_0} e^{\varepsilon_y-\sigma_R^2/2} \quad ; \quad \varepsilon_y \sim N(0;\sigma_R^2) \quad ; \quad \eta_y^z \sim N(0;\sigma_\phi^2) \tag{1}$$

where h is the “steepness” of the stock-recruitment relationship (Francis 1992), R_0 is the unfished equilibrium recruitment, $\tilde{S}_y$ is total (over zones) spawning biomass, $\tilde{S}_0$ is the unfished total spawning biomass, ϕ^z defines the expected proportion of the total recruitment that settles to zone z, σ_ϕ determines the variation about the expected proportion recruiting by zone across years, and σ_R is the standard deviation among recruitment deviations in log space. Spawning biomass is defined as:

$$\tilde{S}_y = \sum_z \sum_a O_a^z \, N_{y,a}^{\mathrm{fem},z} \tag{2}$$

where O_a^z is the product of maturity-at-age and weight-at-age and $N_{y,a}^{fem,z}$ is the number of females in zone z, of age a in year y (see Methot and Wetzel [2013] for details of how O_a^z is calculated).

The value for h is set to 0.75 and that for σ_R to 0.7 (Whitten and Punt 2014). The stock assessment method on which management advice is based (implemented in Stock Synthesis) is assumed to know the correct form of the stock-recruitment relationship, the true value of steepness and the true value for the extent of variation about the stock-recruitment relationship. This is to ensure that the focus of the study is on effects of closed areas, given the well-known difficulty in estimating steepness (Conn et al. 2012; Lee et al. 2012). Punt et al. (2015) outline the selectivity patterns by gear (assumed to be the same among zones). Supplementary Figure S6.1 shows the spatial variation in historical fishing mortality in the absence of closed areas. The fishery is assumed to have started in zone 10 and then moved progressively southward over time – reflecting the fact that the fisheries off southeast Australia started in the mainland ports (and within zone 10). As in Punt et al. (2015, 2016), the maximum level of fishing mortality is assumed to be the same spatially, while the fully-selected fishing mortality for the non-trawl fleet is assumed to be half that for the trawl fleet. For consistency with the actual assessment for Pink Ling (Whitten and Punt 2014), selectivity for the trawl fleet is assumed to be a monotonic logistic function of length while that for the non-trawl fleet is modelled using a unimodal double-normal selectivity function.

As in Punt et al. (2015, 2016), catch-rate data are only assumed to be available from 1986 (Haddon 2014), while collection of age- and length-composition data is assumed to start in 1975. Between 1975 and 1985, length data are assumed to be available for 5% of the combinations of years, gears and zones, with years selected at random, and gear-zone combinations for the years with samples for data on length selected in proportion to the size of the catch in weight (reflecting that catches are sampled by port samplers and onboard vessels). From 1986 to 1997, length data are assumed to be available for 20% of year-gear-zone combinations, and 70% thereafter. The observed catch length-composition data for a year-gear-zone combination is a Dirichlet sample from the true catch length-composition, with an effective sample size of 100. This is lower than the actual sample sizes for pink ling, but the length-composition data for Pink Ling are known to be over-dispersed, and the available length-composition data are considerably down-weighted to reflect this when assessments are conducted (Whitten and Punt 2014; Cordue 2015). The length-composition data are assumed to be unsexed, as is the case in reality for Pink Ling and most fish stocks. The age-length keys for the historical period are assumed to be obtained from a subset of the length-frequencies (56%), with a sample size of 500 for each sex. Age-length keys are assumed to be available for all year-gear-zone combinations in the future. Given a year-gear-zone, the age data are assumed to be a simple random sample from the catches by age and length, for simply and to mimic the intent of the sampling program. The age-estimates are not subject to age-reading error for simplicity.

Allowance for closed areas and simulation scenarios

The scenarios (Table 6.1) explore the impact of factors that could impact estimation and management performance:

- Whether there are closures, and if there are closures, are they in zone 10 or zone 20.
- Whether there are spatial differences in growth (see Supplementary Figure S6.2 for how growth varies spatially; the growth curve when growth is spatially-invariant is set equal to that for zone 20 - this choice was made because the biomass in zone 20 is larger than that in the other zones).
- Whether survey data are available for assessment purposes.

The closures are assumed to be implemented in 2000. When a closed area is implemented in zone 10 or zone 20, to ensure comparability among analyses the exploitation rates in the zones open to fishing are increased so that the ratio of the total (over zones) spawning stock biomass at the start of the projections in 2012 to the total unfished spawning biomass (B_0) is the same (0.4) under deterministic projections (i.e., no spatial variation in settlement and no variation in recruitment about the stock-recruitment relationship). This is to avoid the impacts of closed areas on performance being confounded with changes in overall population biomass. This assumption means, for example, that zones 20 and 30 will be more depleted when the management strategy is first applied if zone 10 is closed.

The total amount of historical length and conditional age-at-length data is the same among scenarios, i.e. closing one zone will lead to more length and conditional age-at-length data for the other zones. This specification is made to avoid the effects of closures being confounded with the total amount of data available for assessment purposes; although in reality one effect of closures is likely to be that less fisheries data is collected.

Table 6.1: The scenarios considered in the analyses of this paper

Scenario	Closures	Spatial variation in growth	Survey data
A	No	No	No
B	No	No	Yes
C	Zone 10	No	No
D	Zone 10	No	Yes
E	No	Yes	No
F	No	Yes	Yes
G	Zone 10	Yes	No
H	Zone 10	Yes	Yes
I	Zone 20	No	No
J	Zone 20	Yes	No

Surveys are assumed to be conducted from 2000 (i.e. at the start of the closures). Surveys are assumed to occur every 2nd year. The selectivity of the survey gear is assumed to be the same as that of the trawl fishery. The survey CV for each zone is assumed to be 0.1 and the survey length-frequency data are assumed to be a multinomial sample with an effective sample size 100, while the survey conditional age-at-length data are assumed to be a multinomial sample of size 500. Thus, the simulations are based on a highly informative survey.

The projections are replicated 100 times for each scenario. This is adequate to detect differences among operating models and management strategies that are consequential at a management level.

6.3.2 The management strategy

The management strategy consists of stock assessment and harvest strategy components. These components are specified to mimic (to the extent possible) how assessments are conducted, and how management advice is provided for Pink Ling off southeastern Australia. Although assessments are assumed to be conducted every three years, RBCs are set annually, and between assessments are based on RBC projections from the previous assessment.

The stock assessment

Two assessment methods are examined.

Naïve spatially-weighted aggregated assessment method (NSWA).

The primary stock assessment considered in this work mimics how management advice is currently provided for pink ling. It does not recognize that there are spatial differences in population structure and abundance, involves conducting a spatially-aggregated assessment, and combines the data available for assessment purposes spatially:

- The catch data are summed over zones.
- The catch rate data are aggregated across zones, defining the catch-rate for year *y* as total catch for year *y* divided by the total effort for year *y*.
- The annual trawl and non-trawl catch length-frequency data by fleet are pooled over zones, weighting the data for each zone by the annual catch by the zone.

- The annual age-length keys (i.e., the conditional age-at-length data) are summed over zones (without catch weighting). This reflects how data have been aggregated in actual assessments for Pink Ling (Whitten and Punt 2014).

The assessment estimates unfished recruitment, natural mortality (assumed to be the same for males and females), growth by sex (five parameters per sex: the parameters that govern von Bertalanffy growth and the CVs of length-at-age for ages of 1 and 20), length-specific selectivity parameters (logistic for the trawl fleet, double-normal for the non-trawl fleet, i.e. based on the correct selectivity patterns), catchability for the CPUE indices, and recruitment for simulated years 1963 onwards. Recruitment estimation includes an initial bias-ramp (Methot and Taylor 2011) to avoid bias when estimating the deviations about the stock-recruitment relationship given length and CPUE data are only available for the more recent years. The difference between the last year with data and the point at which the bias-ramp declines is always four years. Thus, for the first year of projection (2013) the bias-ramp begins to decline in 2009 and this point is incremented by three years each time a stock assessment is conducted

Fleets-as-areas assessment method (FAA)

This assessment method assumes a single homogenous population and that zones 10, 20 and 30 each contain a separate trawl and non-trawl fleet. Separate selectivity patterns are estimated for the three trawl fleets and for the three non-trawl fleets. Thus, this assessment method makes the same population structure assumption as the spatially-aggregated NSWA method, but the catch-rate, length-frequency and conditional age-at-length data for each fleet are kept separate by zone in this method rather than being aggregated spatially. This approach to stock assessment has been applied widely in Australia (e.g., Whitten and Punt 2014; Blue Grenadier *Macruronus novaezelandiae*, Tuck 2014) and off the US west coast (e.g., Petrale Sole *Eopsetta jordani*, Haltuch et al. 2013; Canary Rockfish Sebastes pinniger, Stewart 2009).

The harvest control rule

Catch limits are set using the B_{20}:B_{35}:B_{48} harvest control rule (Figure 6.2). This harvest control rule, which is the default for data-rich SESSF stocks (Dichmont el al. in press; Smith et al. 2008), sets the catch limit to zero if the stock is assessed to be below the limit reference point of 20% of the unfished spawning biomass, B_{20} (in reality, targeted fishing would cease and only bycatch allowed, but in the projections the fishery is closed). If the stock is assessed to be above B_{35} the fishing mortality rate used to determine the catch limit is set to F_{48}, the fishing mortality rate that is estimated to correspond to a depletion of spawning biomass to 48% of its unfished level (B_{48}), with this fishing mortality declining linearly between B_{35} and B_{20}. B_{48} is a proxy for B_{MEY}, the biomass corresponding to Maximum Economic Yield, which is the target biomass for Australian Commonwealth fisheries. The Australian Harvest Strategy Policy (DAFF 2007) allows for the use of proxies for B_{MEY} ($1.2 \times B_{MSY}$), where the proxy for B_{MSY} is taken to be 40% of the unfished spawning biomass, i.e. $0.4B_0$ (Rayns 2007).

The outcomes from the harvest control rule are subject to meta-rules (Stobutzki et al. 2001; Dichmont et al. in press). Specifically, catch limits are not permitted to change by more than 50% from one year to the next. Note that because catch limits are a set of 3-year blocks, the change in catch limit from one time-block to the next can substantially exceed 50%.

The allocation of the catch limit to zone and gear is based on the relative catch by gear and zone in the last year of the assessment period, i.e. the split of the total RBC for all future years equals that for last year for which catch data are available (2012).

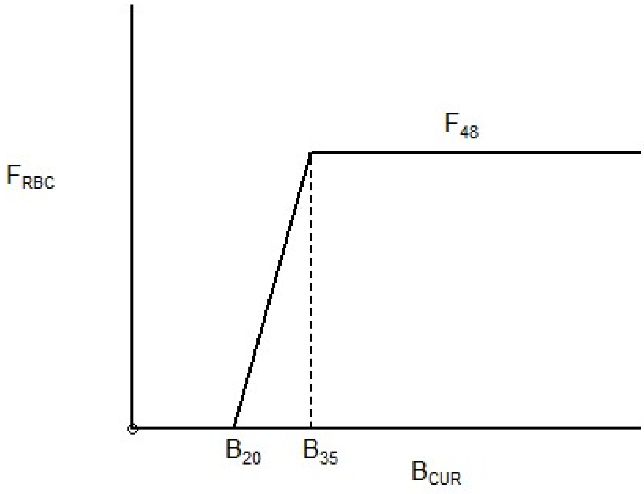


Figure 6.2: The B20: B35: B48 harvest control rule.

6.3.3 Performance metrics

Two types of performance metrics are reported – those pertaining to the ability of the stock assessment method to estimate spawning biomass adequately and those related to achieving management goals.

The assessment method provides many outputs (Methot and Wetzel 2013). However, the focus for this paper, and following Punt et al. (2015, 2016), are the estimates of initial total (female) spawning biomass[3] (B_0), final (2012) total female spawning biomass (B_{CUR}) and relative spawning biomass (B_{CUR}/B_0). The results of the projections are summarized by relative error distributions, as well as by the median over replicates of the absolute relative errors (MARE). The relative error for a given quantity is the estimated value of the quantity less its true value divided by its true value and multiplied by 100, i.e. a positive value indicates an overestimate of the quantity and vice versa.

The management-related performance metrics are those used when the management strategies for the SESSF were developed (e.g., Wayte and Klaer 2010; Little et al. 2011; Fay et al. 2011; Klaer et al. 2012):

- the annual probability over the entire projection period that the stock is below the limit reference point of 20% of the unfished spawning biomass;
- the probability that the stock is above the target reference point of 48% of the unfished spawning biomass at the end of the projection period;
- the average catch over the projection period; and
- the annual average variation in catch, i.e.:

$$AAV = \sum_{y=2013}^{2042} \left| C_y - C_{y-1} \right| / \sum_{y=2013}^{2042} C_y$$

3 Total spawning biomass is spawning biomass summed over the three zones.

The Harvest Strategy Policy for fisheries managed by the Australian Commonwealth (federal) government (DAFF 2007) stipulates that the probability of being below 20% of the unfished spawning biomass should not exceed 10% and this is treated as a 'performance standard' when evaluating the results of the projections.

6.4 Results

6.4.1 Results for a single replicate

Prior to exploring the results for multiple replicates, it is illustrative to examine the results for single replicates. Figure 6.3 shows results for projections when growth does not vary spatially, when there are no survey data, for cases when there are no spatial closures (Figure 6.3a-d), and when zone 10 is closed in 2000 (Figure 6.3e-h) (scenarios A and C in Table 6.1). Each assessment estimate (dashed lines Figure 6.3a,e) is based on the same data as the previous assessment, except that an additional three years of data are included in each subsequent assessment. The estimates of spawning biomass from the assessments (dashed lines in Figure 6.3a,e) are consequently correlated and often systematically different from the true values for several groups of years. For the two projections considered in Figure 6.3, there is an early period when the assessment overestimates recent spawning biomass followed by a period when spawning biomass is under-estimated (Figure 6.3a,e). The estimates of spawning biomass for the years before the first application of the management strategy also differ quite substantially from the true values for both projections, i.e. irrespective of whether there is a closed area. For the projection with a closed area, the estimates of biomass for the early years (1970 to ~1982) are always over-estimated (Figure 6.3e).

The time-trajectories of spawning biomass by zone (Figure 6.3b,f) differ prior to implementation of the closed area as well as thereafter owing to different (random) allocations of total recruitment spatially as well as differences in catches spatially (Figure 6.3c,d,g,h). The spawning biomass for zone 10 increases to above B0 when there are closures (Figure 6.3f). This is because the trend in zone 10 differs markedly from the trends for the other two zones when there is a closed area (Figure 6.3f). The catches in the zones open to fishing increase after 2012 as the stock is assessed to be above the target biomass of 0.48B0 (green horizontal line in the upper left panels of Figure 6.3a,e).

6.4.2 Performance of the management strategy in the absence of closed areas

The upper two rows of panels in Figure 6.4 shows distributions for the time-trajectories of total spawning biomass relative to the unfished level and catch when management advice is based on the spatially-aggregated NSWA assessment, and there are no closed areas. Spawning biomass is relatively close to the target reference point of $0.48B_0$ throughout the projection period (Figure 6.4a,c,e,g; median spawning biomass in 2042 relative to B_0 between 0.43 and 0.49 [Table 6.2]), with the results not differing substantially depending on whether survey data are available in addition to catch rate data nor whether there is spatial variation in growth. The probability that the total spawning biomass exceeds $0.48B_0$ varies between 0.3 (scenario B) and 0.55 (scenario F) (Table 6.2). In all cases, catches increase in the first year of the projection period as the stock is assessed to be above the target reference point, but subsequently decline to levels consistent with a spawning biomass of $0.48B_0$. The between-replicate variation in spawning biomass is largely independent of scenario (Figure 6.4a,c,e,g). In contrast, between-replicate variation in catch is higher when growth varies spatially (Figure 6.4b,d,f,h). Average catches are higher when growth varies spatially but among-year variation in catch is also higher (Table 6.2).

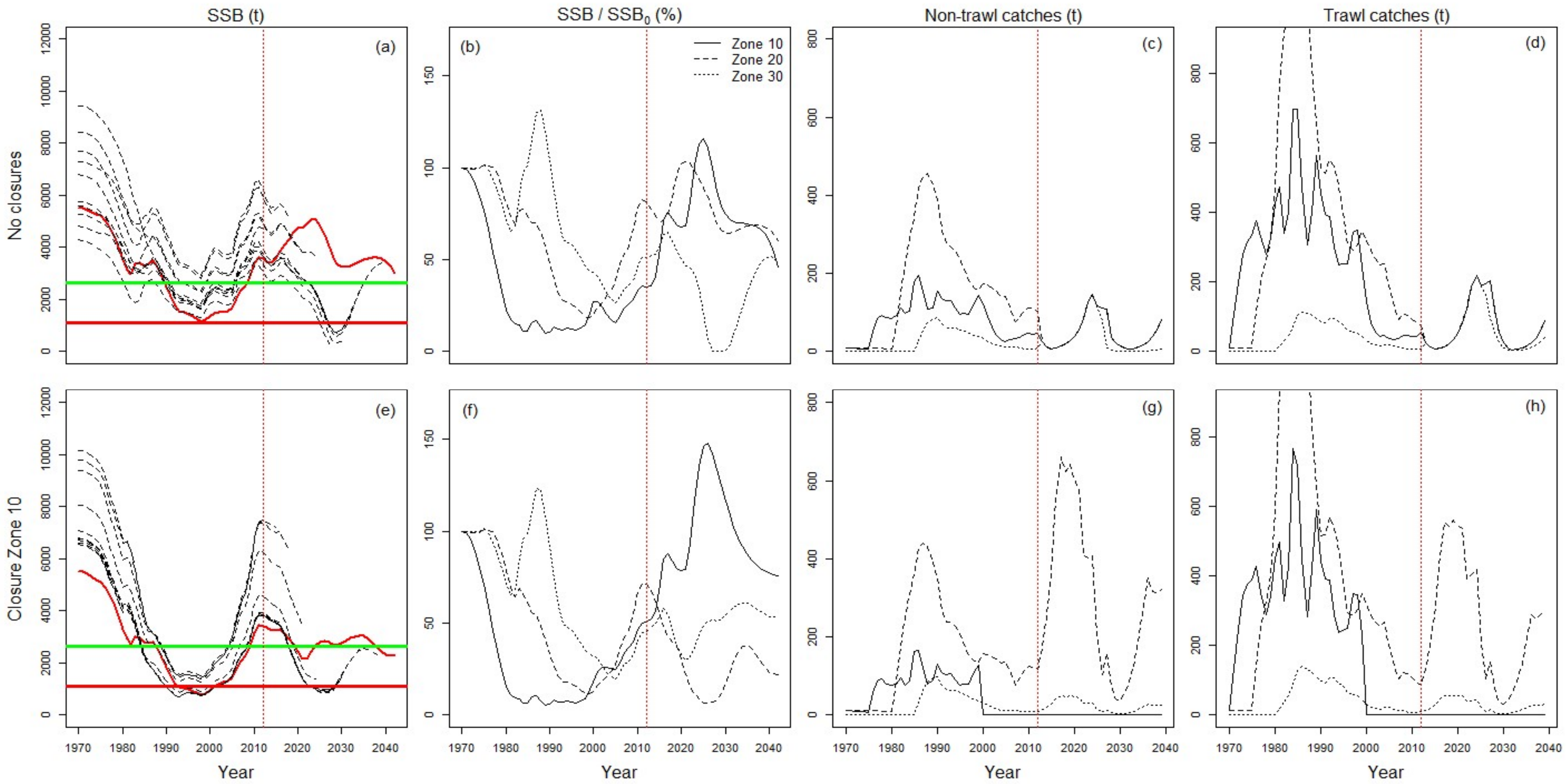


Figure 6.3: (a,e) true time-trajectory of total spawning biomass (thick solid line) and the estimates of the time-trajectories of spawning biomass each time an assessment is conducted (thin dashed lines), (b,f) true time-trajectory of spawning biomass by zone, (c,g) time-trajectory of non-trawl catches by zone, and (d,h) time-trajectory of trawl catches by zone. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines in panels (a,e) are respectively the limit and target reference points. Results are shown in (a-d) when there are no spatial closures and in (e-h) when there is a closure in zone 10 starting in 2000. The assessments are based on the spatially-aggregated NSWA method.

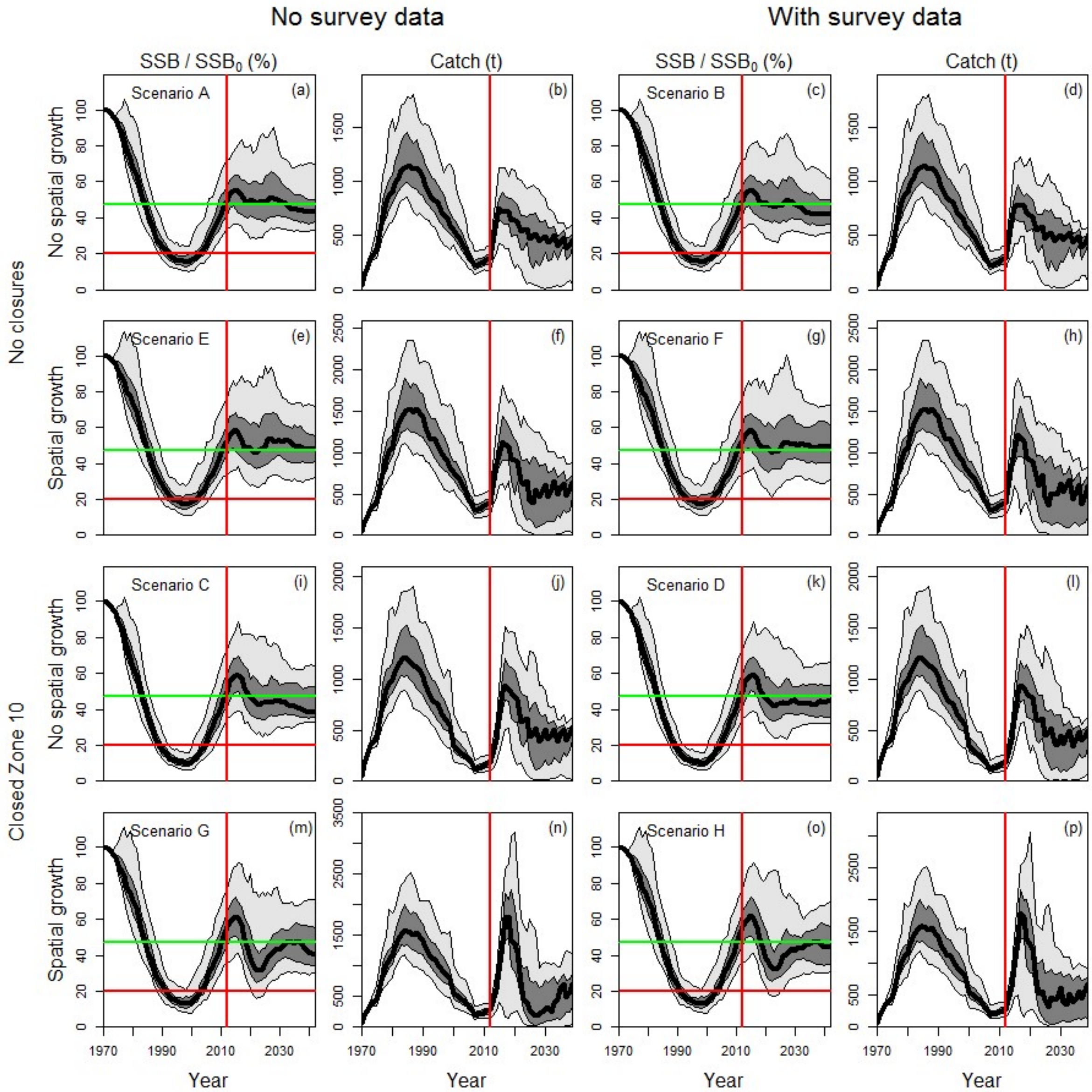


Figure 6.4: Time-trajectories of total spawning biomass relative to unfished spawning biomass (columns 1 and 3) and catch (columns 2 and 4). Results are shown when there are no closed areas (scenarios A, B, E, and F) and when zone 10 is closed (scenarios C, D, G and H). There are survey data for scenarios B, D, F and H. Growth does not vary spatially for scenarios A, B, C and D, but does vary for scenarios E, F, G and H. The solid line is the distribution median, the dark shading covers the central 50% of the distributions and the light shading 90% of the distributions. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines are respectively the limit and target reference points. Management advice for the analyses in this figure are based on the spatially-aggregated NSWA approach.

Table 6.2: Performance measures (probability of being below the limit reference point, probability of being above the target reference point, final relative spawning biomass, average catch, and AAV) for the two assessment methods and the various scenarios.
Table 6.2a: Zone 10 closed.

Scenario											
	P <			S_{Fin} / S_0		Average catch				AAV	
	LRP	P > TRP	Low 5	Median	Up 5	Low 5	Median	Up 5	Low 5	Median	Up 5
Spatially-aggregated NSWA assessment											
A*	0.001	0.37	32.7	43.3	68.9	307	525	727	0.040	0.277	0.670
B*	0.002	0.30	31.8	41.9	61.7	312	524	780	0.043	0.275	1.177
E*	0.002	0.51	32.2	48.1	71.5	386	660	973	0.037	0.352	1.100
F*	0.003	0.55	31.2	49.4	72.3	393	646	979	0.013	0.362	0.934
C	0.002	0.36	32.8	38.9	63.4	283	538	764	0.014	0.180	1.248
D	0.003	0.43	32.5	44.5	66.8	279	535	776	0.012	0.213	0.884
G	0.016	0.39	30.4	40.7	71.3	409	755	1036	0.017	0.339	0.899
H	0.018	0.44	30.8	45.2	69.6	415	739	1081	0.016	0.126	0.840
Spatially-aggregated fleets-as-areas assessment											
A*	0.002	0.32	29.2	41.4	71.1	161	509	744	0.009	0.093	0.805
E*	0.006	0.37	28.0	44.3	67.6	394	713	1073	0.015	0.169	1.032
C	0.000	0.54	36.2	49.1	74.7	160	421	603	0.013	0.120	0.792
F	0.019	0.48	30.6	47.6	66.8	423	735	1101	0.006	0.307	1.350

*No closures

Table 6.2b: Zone 20 closed.

Scenario				$S^{\%}_{Fin} / S^{\%}_{0}$		Average catch				AAV	
	P < LRP	P > TRP	Low 5	Median	Up 5	Low 5	Median	Up 5	Low 5	Median	Up 5
Spatially-aggregated NSWA assessment											
I	0.000	1.000	54.0	63.6	77.5	106	256	397	0.001	0.363	1.166
J	0.000	0.980	50.3	66.9	77.8	137	292	485	0.000	0.179	1.374

There is no apparent benefit of survey data on management performance (contrast the results of scenario B with those for scenario A, and the results for scenario F with those for scenario E; Figure 6.4). This is because even though surveys provide an index of abundance that is linearly proportional to biomass for each zone, the assessment model is mis-specified through it not accounting for spatial structure, and this means that the results remain biased.

6.4.3 Performance of the management strategy when there are closed areas

Spawning biomass equilibrates close to the target level when zone 10 is closed in 2000 and management is based on the spatially-aggregated NSWA assessment (Figure 6.4i,k,m,o). The total spawning biomass initially rebuilds with correspondingly higher catches, to a level greater than the target reference point; given the higher catches the biomass then declines to below the target biomass. This pattern is most evident when growth varies spatially (Figure 6.4m,o), with the probability that the total spawning biomass is below the limit reference point being higher than when growth does not vary spatially (roughly 2% of all years for scenarios G and H compared to less than 1% for scenarios A-E; Table 6.2). Average catches are again higher when growth varies spatially, but unlike the case when there were no spatial closures, catch variation (AAV) is highest when there are no survey data (Table 6.2). Consequently, as was the case when there were no spatial closures, the availability of survey data only leads to a limited impact on management performance.

6.4.4 Assessments based on the fleets-as-area approach

The time-trajectories of spawning biomass relative to the unfished level and catch are similar for the two assessment methods (NSWA and FAA) when there are no spatial closures (Figure 6.4 vs Figure 6.5). However, the results differ between the two assessment methods when zone 10 is closed. Specifically, spawning biomass is consistently above the target reference point (but does eventually reach $B_{48\%}$) when growth does not vary spatially and zone 10 is closed if the assessment is based on the fleets-as-areas approach (Figure 6.5e). The pattern of results when zone 10 is closed and growth varies spatially is close to that when assessments are based on the NSWA assessment method. Inter-annual catch variation (AAV) is lower when management advice is based on the FAA assessment method (Table 6.2).

6.4.5 Closing zone 20 instead of zone 10

Qualitatively, the time-trajectories of spawning biomass and catch are markedly different, both historically and into the future, when zone 20 rather than zone 10 is closed in 2000 (Figure 6.4 and Figure 6.6). The difference in the historical period is most evident for the catches between 2000 and 2012, which are much lower when zone 20 is closed than when zone 10 is closed. This occurs because the biomass in zone 20 is much larger than in zone 10 and the assessment is negatively biased when zone 20 is closed (see below) so future TACs are consequently lower (contrast Figure 6.4 and Figure 6.6; Table 6.2a and Table 6.2b). One implication of this is that the total spawning biomass stabilizes at a biomass that is higher than the target level of $B_{48\%}$ (Figure 6.6;Table 6.2b).

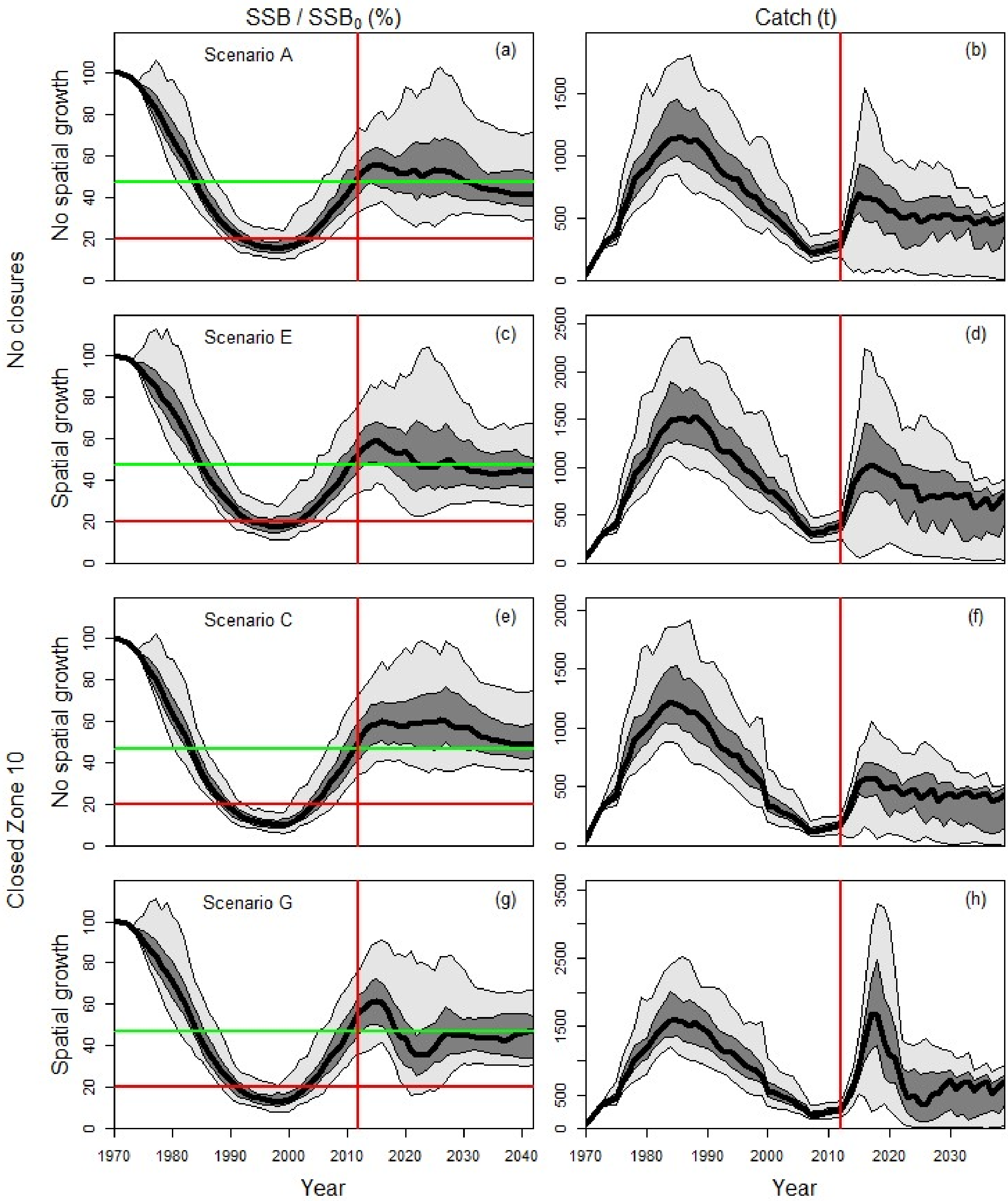


Figure 6.5: Time-trajectories of total spawning biomass relative to unfished spawning biomass (left columns) and catch (right columns) when management advice is based on the FAA method and there are no survey data. Results are shown in the upper two rows when there are no spatial closures (scenarios A and E) and when zone 10 is closed in 2000 (scenarios C and G). Growth does not vary spatially for scenarios A and C, but does vary for scenarios E and G. The solid line is the distribution median, the dark shading covers the central 50% of the distributions and the light shading 90% of the distributions. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines are respectively the limit and target reference points.

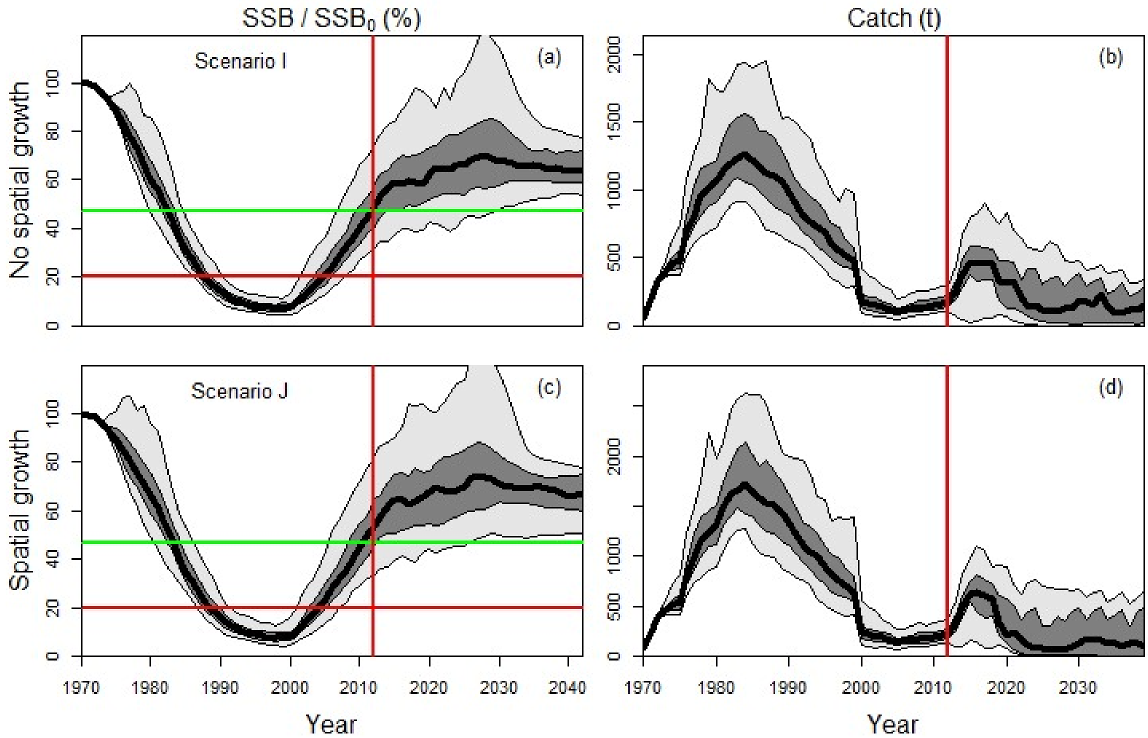


Figure 6.6: Time-trajectories of total spawning biomass relative to unfished spawning biomass (left columns) and catch (right columns) when zone 20 is closed in 2000 and the assessment is based on the spatially-aggregated NSWA method and there are no survey data. Results are shown in the upper row when growth does not vary spatially (scenario I) and the lower row when growth varies spatially (scenario J). The solid line is the distribution median, the dark shading covers the central 50% of the distributions and the light shading 90% of the distributions. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines are respectively the limit and target reference points.

6.4.6 Results by zone

The time-trajectories of spawning biomass differ among zones even when there no closures (Scenarios A, E), primarily because of the way catches have been taken historically (greater depletion between 1980 and 2000 for zone 10 than zone 20 and particularly zone 30) (Figure 6.7). Zone 10 is depleted to below the limit reference point when there no closures, given the assumption that future catches are taken spatially in proportion the spatial distribution of historical catches (Figure 6.7a,d). All zones are close to the target level when the management strategy is first applied when there are no closures (Figure 6.7a-f). When zone 10 is closed in 2000 the biomass starts the projection well above the target reference point (green line, Figure 6.7g,j) and recovers to close to the unfished level when it is closed (Figure 6.7g,j). Zone 20 however, starts at the target biomass and is depleted to close to the limit reference point, in median terms when this is the case (Supplementary Figure S6.3). This is because the exploitation rate for zone 10 is higher than that for zone 30. The depletion of zone 20 is more extreme when growth varies spatially. Thus, the spatially-aggregated results differ from those by zone.

6.4.7 Error estimating biomass

The estimates of total spawning biomass from the spatially-aggregated NSWA assessment method are biased (the median error differs from zone) prior to the first application of the management strategy irrespective of whether or not there are closures, with the bias larger when growth varies spatially (Figure 6.8 a,c,e,g,i,k). These biases are due to spatial heterogeneity in the abundance index as well as the length- and conditional age-at-length data that cannot be addressed using a spatially-aggregated assessment model such as NSWA (Punt et al. 2016). The extent of bias changes between the 3rd and 5th assessments (those conducted in 2018 and 2024), although the effects are more marked when zone 10 is closed (Figure 6.8a,c,e,g,i,k). The assessment leads to negatively biased estimates of biomass towards the end of the projection period irrespective of whether zone 10 is closed or not. Thus, the additional 30 years of data does not improve the ability to estimate trends in spawning biomass. There is also considerable among-replicate variation in errors in estimating spawning biomass (Figure 6.8, columns 2 and 4). The relative errors of the estimates of total spawning biomass exhibit the same general patterns irrespective of whether zone 10 or zone 20 is closed, but the estimates are negatively biased when zone 20 is closed (and for scenario J show extreme variation among replicates, in particular the possibility exists of occasional highly positive biased estimates of total spawning biomass).

The estimates from the FAA method are less biased compared to the NSWA method (and the pattern in bias is quite different for scenario C than for the other scenarios) (Figure 6.8m-p). However, and consistent with Punt et al (2016), the among-replicate variation in errors estimating spawning biomass is larger for the FAA method (e.g. Figure 6.8b vs Figure 6.8n).

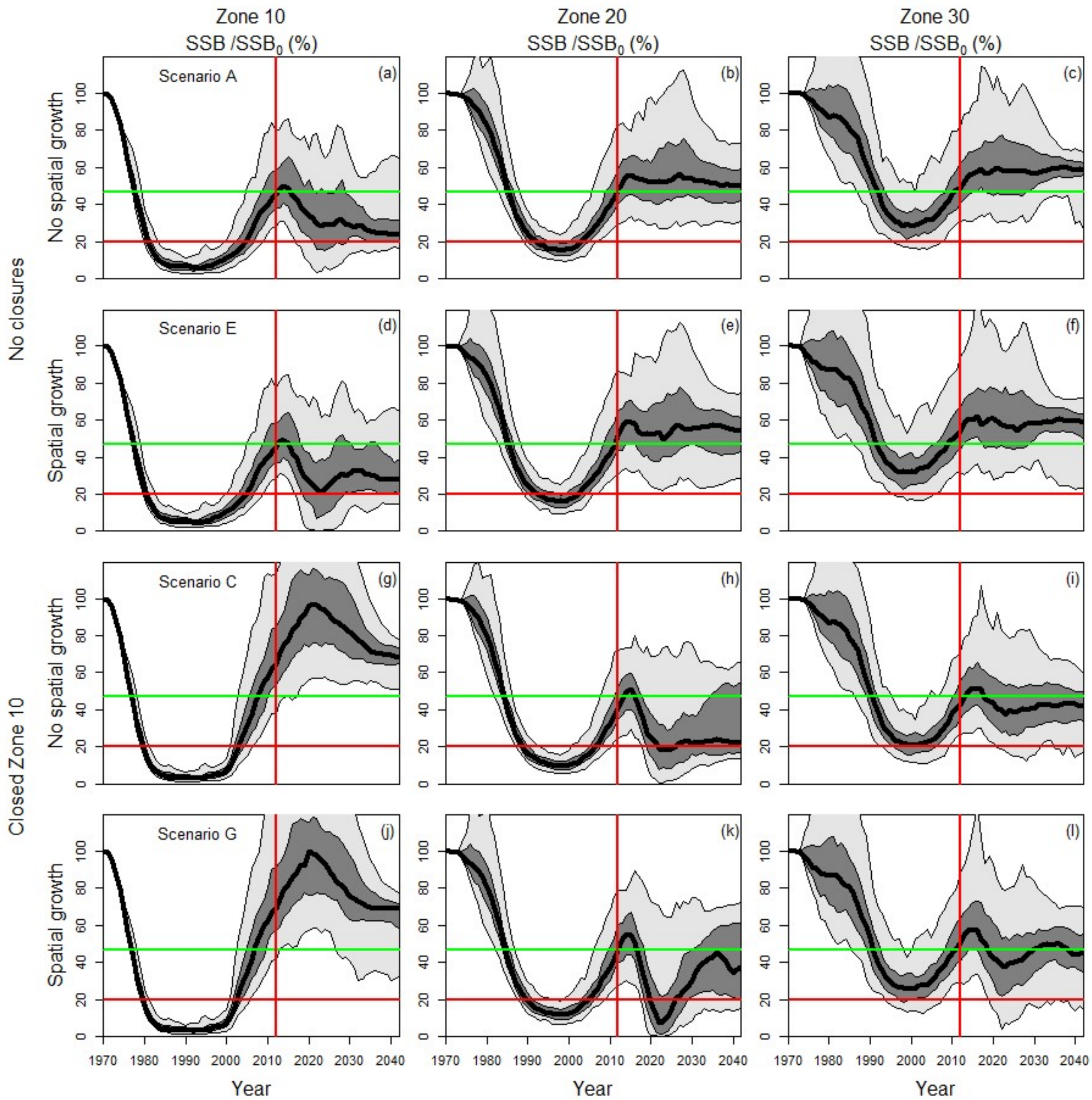


Figure 6.7: Time-trajectories of spawning biomass relative to unfished spawning biomass by zone (10, 20 and 30, left, center, and right columns). Results are shown in the upper two rows when there are no spatial closures (scenarios A and E) and when zone 10 is closed in 2000 (scenarios C and G). Growth does not vary spatially for scenarios A and C, but does vary for scenarios E and G. The solid line is the distribution median, the dark shading covers the central 50% of the distributions and the light shading 90% of the distributions. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines are respectively the limit and target reference points. The results in this figure are based on the spatially-aggregated NSWA method with no survey data.

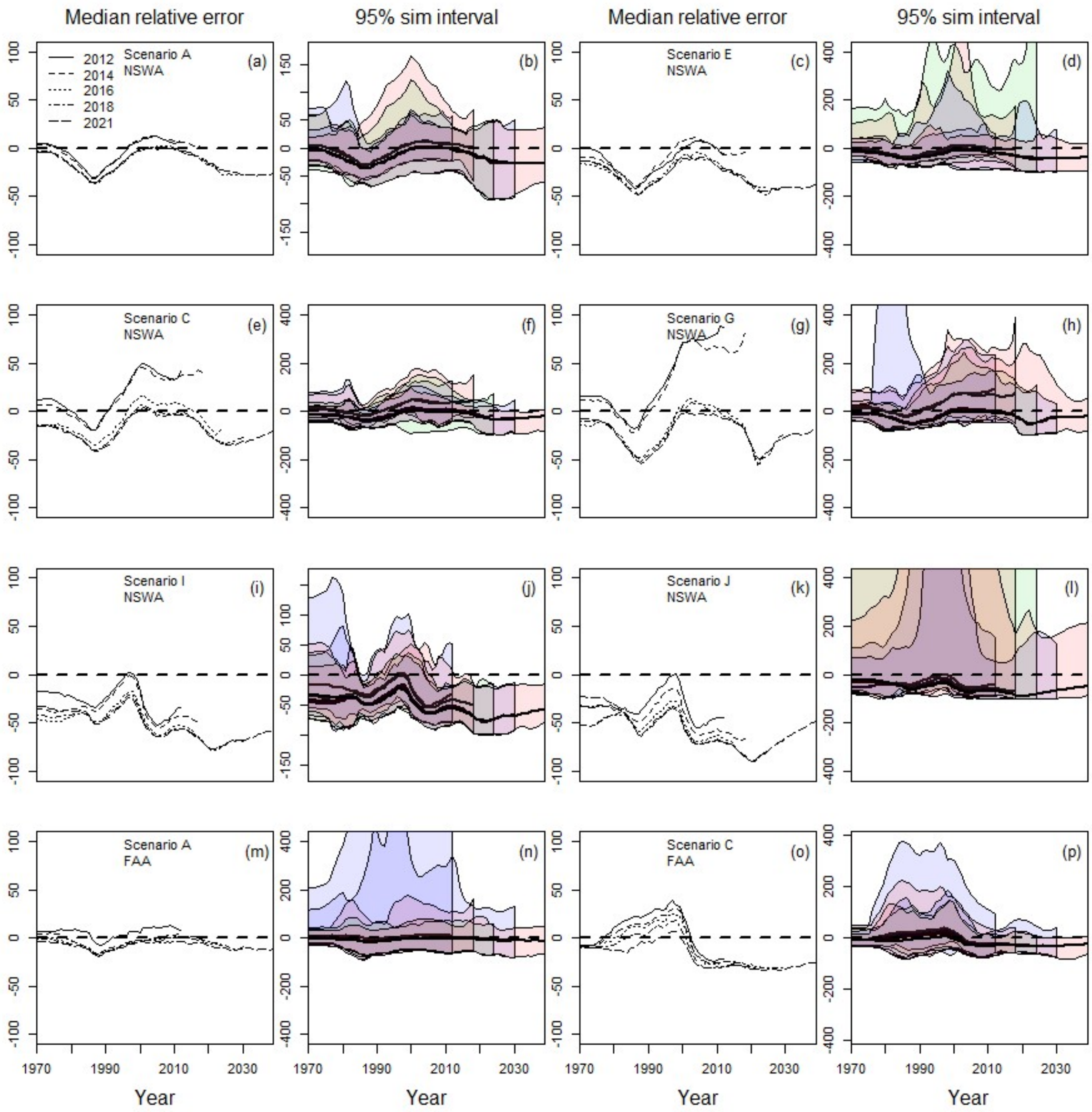


Figure 6.8: Median relative errors for first 3rd, 5th, 7th and 10th assessments and the 95% simulation intervals for the relative errors for these assessments. Results are shown for management based on the NSWA method in (a, b) for scenario A, in (c, d) for scenario E, in (e, f) for scenario C, in (g, h) for scenario G, in (i, j) for scenario I, and in (k, l) for scenario J and for the FAA method in (m, n) for scenario A, and in (o, p) for scenario C.

6.5 Discussion

Several studies have explored the behaviour of stock assessment methods in the face of spatial heterogeneity in population structure, including when a portion of habitat is closed (e.g. Punt 2003; Punt and Methot 2005; Garrison et al. 2011; Guan et al. 2013; McGilliard et al. 2015). It is well known that attempting to manage two (or more) stocks as one can lead to an inability to achieve conservation and utilization objectives (e.g. Hilborn 1985; Fu and Fanning 2004). However, management strategy evaluation has not previously evaluated the ability of management strategies to achieve management goals in the face of spatial heterogeneity and spatial closures when there is a single stock (i.e. simulating the assessment method rather than setting a fishing mortality rate). Punt and Hobday (2009) explored the performance of a management strategy for rock lobsters that was based on a size-structured population dynamics model when the region to be managed contained multiple stocks, but they did not examine the consequences of spatial closures, while Fay et al. (2011) explored the performance of an empirical management strategy, again in the absence of spatial closures.

The results of this paper confirm the results of earlier studies that spatial heterogeneity in abundance and age structure will lead to bias for assessments based on spatially-aggregated population dynamics models, irrespective of whether the NSWA or FAA methods are applied. They further confirm that the extent of bias in estimates of total spawning biomass is exacerbated in the presence of spatial closures. Although the simulations involved closures that were in operation for more than 50 years, there was no evidence that biases that arise due to spatial closures (or even spatial heterogeneity in population structure) reduce over time, even when high precision survey data are available for assessment purposes.

However, biased estimates of total spawning biomass do not necessarily lead to a complete inability to achieve management goals. While the time-trajectories of catches and biomass are substantially more variable than would be expected had a single stock been managed, the stock tends (in median terms) to be fairly close to the target level at the end of the projection period, with a low probability of being depleted below the limit reference point, at least in the absence of closures and for closed areas that are up ~25% of the stock area[4]. Larger closed areas lead to lower catches and stock sizes in excess of the target level. The extent to which the stock is above the target level depends on the size of the closed area and the target level. In the context of scenarios I and J, the target is 48% of the unfished level and close to 50% of the stock biomass (in an unfished state) is contained in the closed area. Consequently, it is not unexpected that the stock is well above the target biomass.

There was a notable impact of spatial variation in growth rates on the performance of the management strategy, owing to impacts of this variation on the bias of the assessment methods. The bias arises because the estimated growth curve in the assessment does not match those used when the data (particularly the length-frequency data) are generated.

The management strategies considered in this paper are based on one harvest control rule and two assessment methods. The particular harvest control rule examined, along with the two assessment methods, are those used in actuality for Pink Ling off southern Australia where recent management advice has been based on the NSWA method (Whitten and Punt 2014; Cordue 2015). However, there are many other types of harvest control rules that could be used to determine catch limits (see the review of harvest control rule formulations in Deroba and Bence 2008), while the input for applying harvest control rules could be obtained from many types of assessment method. In principle, some of the biases associated with spatial heterogeneity could be removed by using a spatially-explicit stock assessment method, but the extent to

[4] The approximate size of zone 10

which this would be the case depends on how accurately the spatial model underlying the assessment method matches reality (Punt et al., 2016). An additional complication associated with management strategies based on spatially-explicit models is that they tend to be very imprecise (Punt et al. 2015, 2016), which would lead potentially to high between-year variation in assessment outcomes and catches.

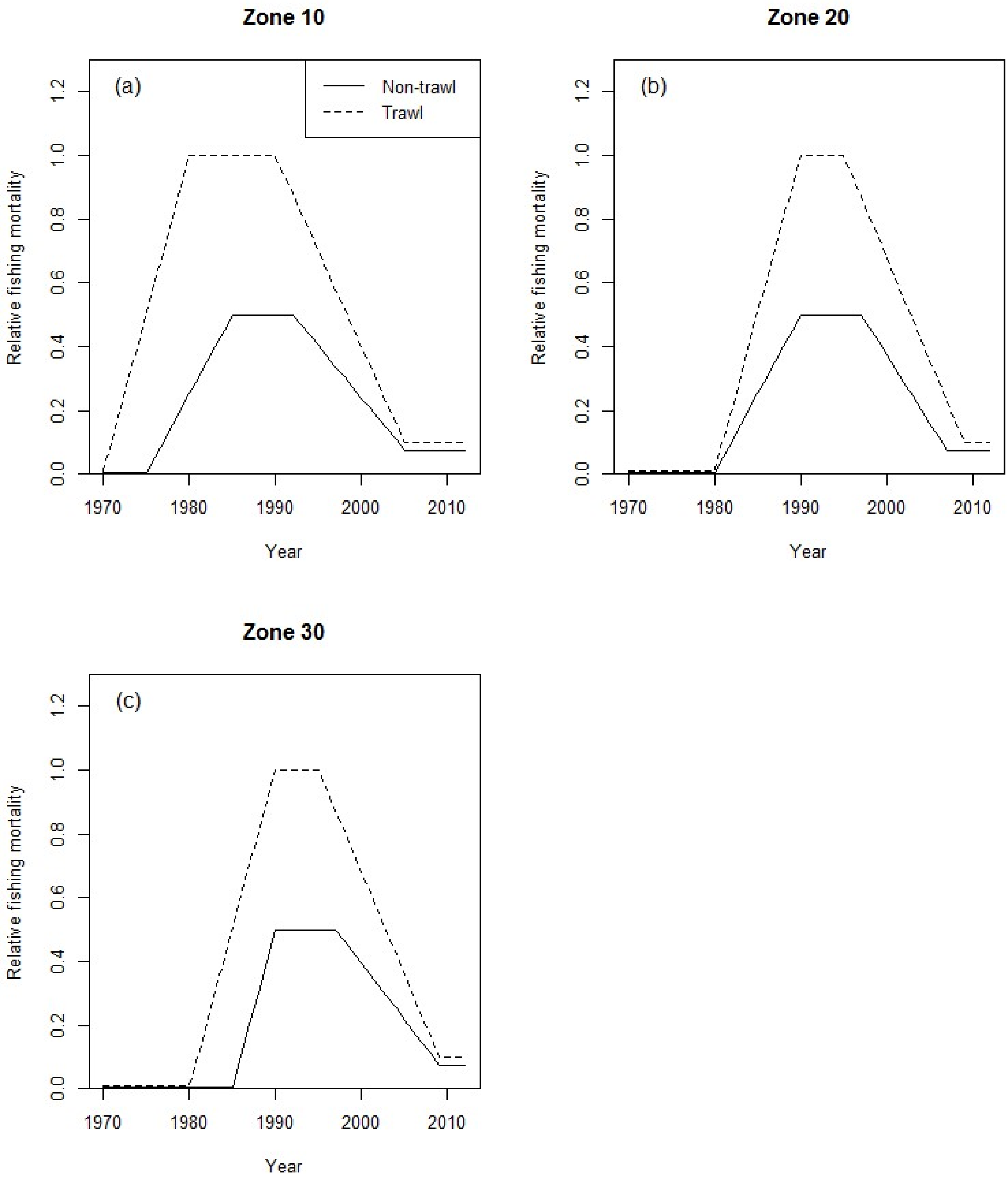
Several jurisdictions, including Australia, have adopted empirical rather than model-based management strategies (e.g., Plagányi et al. 2007; Smith et al. 2008). Empirical management strategies tend to lead to more variable catch limits than model-based approaches (Butterworth and Punt 1999), but may be less biased as they do not attempt to interpret monitoring data in the context of an incorrect population dynamics model. Whether empirical management strategies outperform model-based management strategies in the face of spatial heterogeneity should form the basis for future work. This may be the case when there is spatial heterogeneity in abundance trends as well as in trends in age and length data because empirical methods do not attempt analyse all available data using a model that would be mis-specified.

The operating model on which this study was based was simple compared to reality by including only three areas and not allowing adult movement. The assumption of no movement should have increased spatial heterogeneity, as movement tends to reduce the effects of the spatial variation in fishing mortality. Other areas where future work should focus include (a) changing the number of areas and how assessment regions match population structure, (b) conducting projections where catch limits are not taken in proportion to the spatial distribution of the catch in 2012 (fishers are to some extent place-based and hence likely to continue to use the same gear and fish in the same zone, but this can and will change over time), and (c) considering performance metrics related to catch-rates as well as to catches (which will, all things being equal, be lower in the presence of a large closed area).

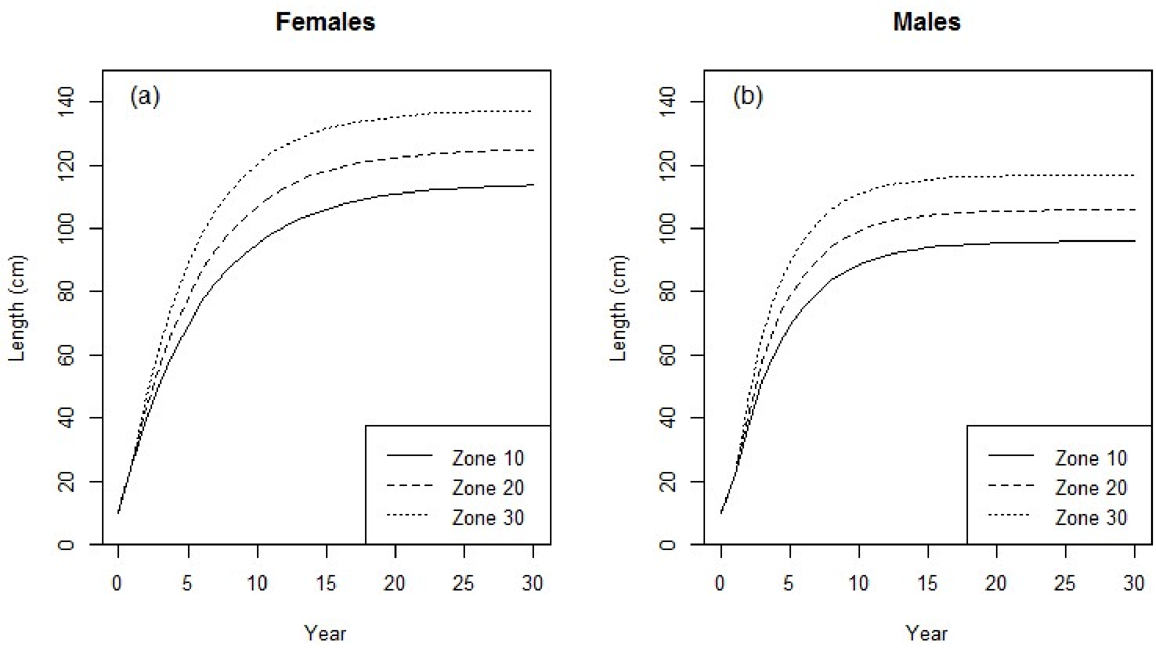
Although the results of this study are case-specific (in common with all simulation studies), several general conclusions can be drawn:

- Assessment outcomes will be biased in the face of unmodelled spatial heterogeneity, and the biases are unlikely to disappear given additional data, and are exacerbated by spatial closures.
- Biased assessments due to unmodelled spatial heterogeneity do not necessarily lead to an inability to achieve management goals.
- The ability to achieve management goals is affected by closed areas, with the effects larger for large closed areas.
- Survey data from within the closed area do not necessarily improve management outcomes.

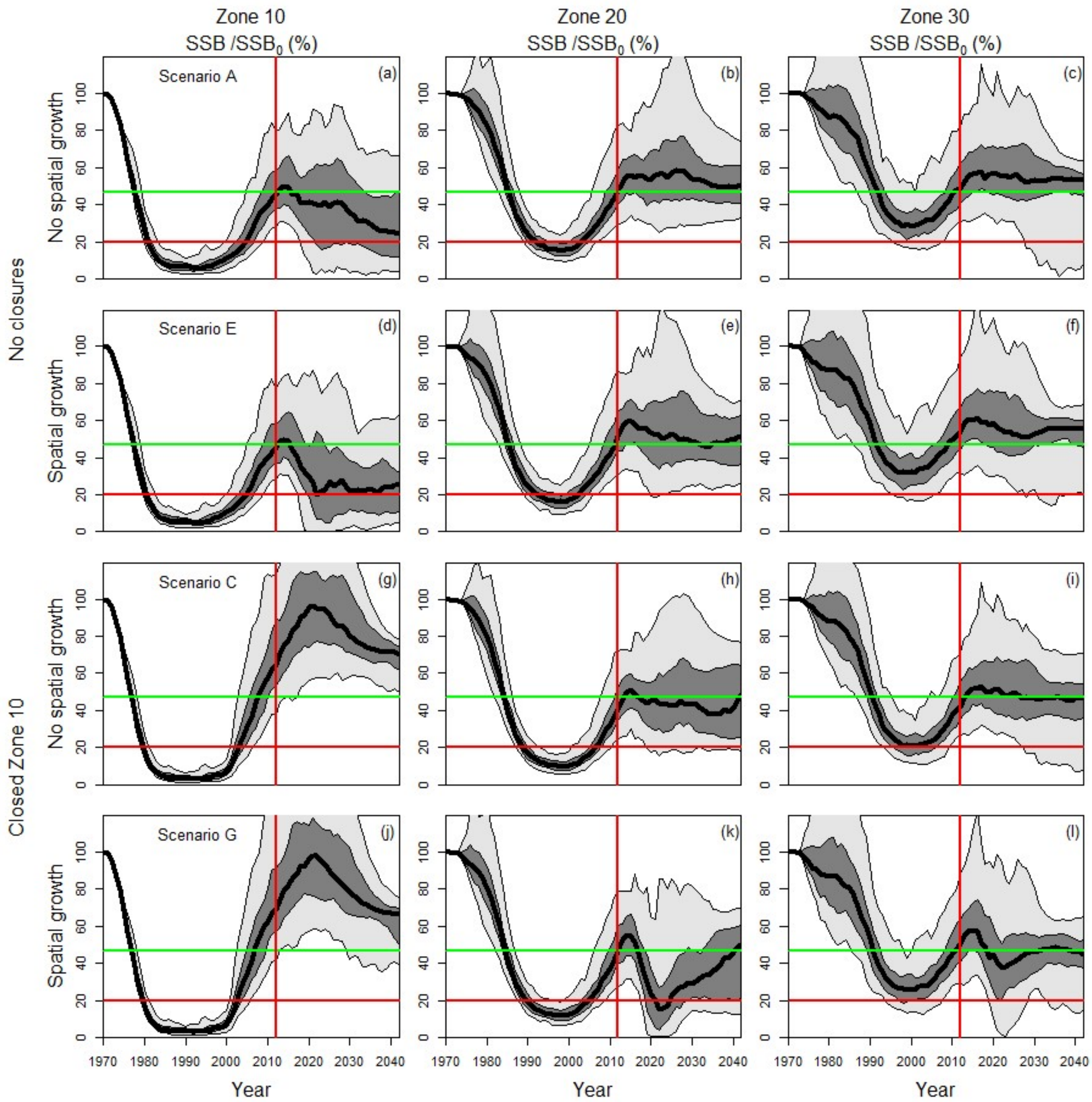
Spatial heterogeneity and closed areas have always been a reality for fisheries management (Field et al. 2006). However, the results of studies such as the present one highlight that effective management may require sampling programs and assessment frameworks designed to support management strategies tailored to there being closed areas. Such programs could involve tagging programs as well as sampling of age structure in closed areas. Frameworks such as those outlined in this paper could be used to evaluate the extent to which alternative management strategies can outperform current approaches.



Supplementary Figure S6.1: Fishing mortality rates by gear and zone in the absence of closures.



Supplementary Figure S6.2: Growth rates by zone.



Supplementary Figure S6.3: Time-trajectories of spawning biomass relative to unfished spawning biomass by zone (10, 20 and 30, left, center, and right columns). Results are shown when growth does not vary spatially (scenarios A and C) and when growth varies spatially (scenarios E and G). The solid line is the distribution median, the dark shading covers the central 50% of the distributions and the light shading 90% of the distributions. The vertical red line denotes when the management strategy is first applied, and the red and green horizontal lines are respectively the limit and target reference points. The results in this figure are based on the FAA assessment method.

7 An Empirical Examination of the Effect of Marine Closures upon CPUE Standardization

Haddon, M.[1]

[1]*CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, Tasmania, Australia*

7.1 Abstract

The work in this Appendix constituted an empirical evaluation of the effects of closures upon catch-rate standardizations through using actual standardizations from the Southern and Eastern Scalefish and Shark Fishery and applying three different treatments relating to the inclusion or exclusion of data from closures and then comparing the outcomes of the analyses. The treatments were 1) to ignore the advent of closures (possibly valid if the closure was small or little catch was taken from them), 2) treat the closure as a factor in the standardization where all data from the closure area is treated as one level and data from outside the clsoure as a different level of a single categorical factor (possibly valid if the closure has not been present for too many years), and 3) exclude all data ever taken from within the closure region. While very many Commonwealth Marine Protected Areas have been declared only those in the south-east are currently active so the analyses were restricted to the SESSF. The trawl fisheries off of eastern Australia whose catch and effort data were standardized were those for Flathead (*Neoplatycephalus richardsoni*), Pink Ling (*Genypterus blacodes*), and John Dory (*Zeus faber*). In all the trawl fisheries considered the outcomes from the three treatments (as measured by the trend of standardized CPUE through the years) barely differed from each other. This was not surprising as in most cases the closures present only influenced a very small proportion of the catching area and an equivalently small proportion of the catch. In addition to the trawl fisheries the auto-line fishery for Blue-Eye Trevalla fishery (*Hyperoglyphe antarctica*) was included and results from this fishery differed from the trawl fisheries because some of the recent closures, such as the Flinders Research Zone and around the St Helens Hill, together accounted for just over 20% of all Blue-Eye catches in Zones 20 and 30. Even in that case the differences between treatments were minor and mostly occurring up to from 1997 – 2007, but even those were within the bounds of uncertainty of each of the standardizations.

The lack of influence or effect of the current marine protected areas within the south-east on the trawl fisheries should not be surprising. Before the Commonwealth closures were first introduced a separate research project was initiated tasked with producing alternative closure definitions that attempted to minimize the effect of those closures upon commercial fisheries, which was relatively successful. This is reflected in the minor amounts of catches excluded from the trawl fisheries by those closures, which in turn flows on to the lack of any significant effects upon the standardizations. This could not be an explanation for the lack of effect in the Blue-Eye auto-line fishery because the closures with most influence on that species were introduced later than the Commonwealth closures. Instead, it appears that the Industry vessels and their skippers are capable of rapidly adapting to the advent of even effectively large closures so that any potential effects they might have are masked by the vessels altering their fishing behaviour and developing alternative fishing grounds.

In conclusion, the current south-east closures have only had a minimal effect upon the standardized CPUE trends through a combination of the closures being designed to have minimal effects on fisheries or the fishers themselves adapting to the closure of some of their favoured fishing grounds by finding alternative localities. Despite closures having a minimal effect when standardizing catch and effort data from a region that contains closures the optimum strategy for any standardization is to exclude all data taken within the closed area from subsequent analysis. If this restricts the amount of data too much for a usable standardization then the next best approach would be to treat the inside and outside of a closure as a factor in a standardization (akin to declaring two, or more, new areas within the data being standardized. By definition the assumption would be that there would be no data in the closures after they were installed.

7.2 Introduction

The introduction of marine closures has proceeded in the South-East of Australia and beyond, leading to an array of closures of varying size, intent, and difficulty of avoidance by fishing vessels. Some have a simple shape while others are more complex and thus present a greater challenge for avoidance by commercial fishers (for example the closures to the east of Flinders Island have a very complex boundary; Figure 7.1). The Commonwealth closures in the South-East were introduced in 2007; those areas initially proposed were modified after analysis and consultation with Fishing Industry members (Buxton *et al*., 2006) so as to avoid areas historically important to the SESSF although some of the South-East closures have been modified since then and other fishery closures have been introduced subsequently. The details (dates and boundaries) of closures and their changes will need to be carefully re-constructed for CPUE standardizations consistently to take closures into account.

Many of the south-east closures overlap to a degree with areas where commercial fishing occurs so they can be expected to interfere with fishing operations. The objective of this Appendix is to explore the potential effects upon commercial catch-per-unit-effort and their standardization. It is common practice to use commercial CPUE as an index of relative abundance within the stock assessments of many of the species assessed in the Southern and Eastern Scalefish and Shark Fishery (SESSF). Where a closure is large and overlaps extensively with the distribution of a fishery for a particular species (e.g. Silver Trevally and the Batemans Bay Marine Protected Area) then the displacement of the fishery into different areas may be expected to have an effect upon the commercial CPUE. Such effects certainly occurred when the initial 700m deepwater closure was imposed as it effectively shut down both the eastern and western deepwater shark fisheries as well as the Orange Roughy fisheries, which was the primary intended effect (Haddon, 2014). Catches of deepwater sharks recovered somewhat after the first modification to the deepwater closure in 2009; with the re-opening of the eastern Orange Roughy fishery other changes to the deepwater closure have occurred more recently.

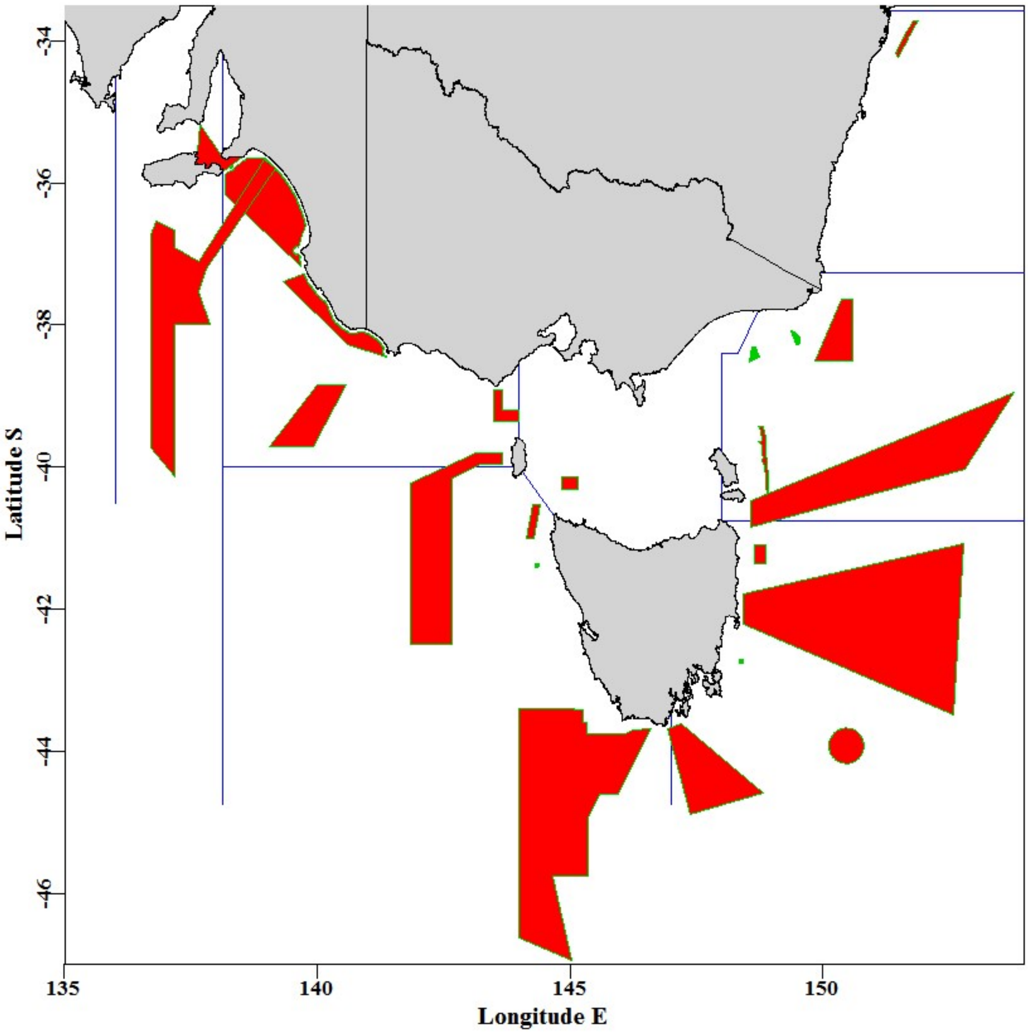


Figure 7.1: A schematic map of an array of the closures installed in the south-east SESSF region. The red closures are permanent while the four much smaller green areas are seasonal and not present in all years. In all cases trawling is banned within the closed areas, although other fishing methods might be acceptable. Missing from this plot are the very extensive deepwater closures where all waters > 700m depth were closed (although this was revised after two years to open a few subsets of that area). Also missing is the extensive trawl closure in Bass Strait.

With respect to the practicalities of CPUE standardization, the introduction of a closure is a singular event so there can be no simple comparison of the CPUE before the closure with that after because there is no period of overlap where some vessels are closed out and some remain in. Any such on/off global event within a fishery, such as the introduction of a closure, is a problem for CPUE standardization as there is no valid way to calibrate CPUE for a species before and after the closure introduction. There are three obvious options available for subsequent CPUE standardization following introduction of a closure (Table 7.1); other options may exist but available data does not make alternatives currently viable.

Table 7.1: Three possible options for how to conduct CPUE standardization analyses following the introduction of closure(s) in a given fishery.

Option	Action
Ignore	Ignore the complication of the new closures and proceed as before; this might be valid if the closure only overlaps with the fishery in a very minor manner.
Factor	Include the closure as a new factor in the standardization; this might be valid for a few years but eventually, due to no ongoing new data coming from inside the closure, may become misleading once too many years have elapsed.
Exclude	Once a closure is in place then remove any data within the closure from consideration right back through the history of the fishery; This may be valid but could potentially mislead a formal stock assessment model that attempted to use the time-series.

7.2.1 Objectives

1. Determine the differences that can arise when conducting CPUE standardizations based around the three options investigated given closures by standardizing CPUE for several SESSF species using the different data selection approaches.
2. Determine whether the practical examples enable recommendations to be made about which approach to use under what circumstances.

7.3 Methods

7.3.1 Data Used

Catch and effort data from the SESSF up to the end of 2015 were used in the analyses. Given the complexity of the implementation of the many different closures (Figure 7.1) only trawl and auto-line fisheries were examined as it is clear that no trawling is allowed in the Commonwealth closures introduced in the SESSF in 2007 (and some modified in 2009 and beyond). There are extra uncertainties associated with the effect of closures on auto-line CPUE, with some auto-line vessels avoiding some closures (e.g. St Helen's Hill) even when it remains open to the auto-line method. Nevertheless, the effect of closures can be investigated for auto-line using the recently developed time-series of catch-per-hook CPUE (Haddon, 2016a).

7.3.2 The Standardizations Considered

Three standardizations were conducted for each species differing either in terms of the data selected for the standardization (select all data – "ignore the closures", or only select that data that occurred outside of the imposed closure – "exclude the closures") or the structure of the model used (whether closures were included as a factor within the standardization; Table 7.1).

7.3.3 The Species and Closures Examined

The effect of ten closures (Table 7.2; Figure 7.2) in the eastern SESSF zones (10, 20, and 30), on CPUE was examined by conducting CPUE standardizations using standard methods (see Haddon, 2014; Sporcic and

Haddon, 2016). The data selection criteria in each case and the optimum models were as used previously for each species considered (Sporcic and Haddon, 2016). These selections did not always include all three zones.

Table 7.2: Some characteristics of the 23 closures illustrated in Figure 7.1 sorted in order of minimum Longitude. The temporary, occasional, and seasonal closures are small and green in Figure 7.1.

MinLong	MaxLong	MinLat	MaxLat	Name
146.938	148.828	-44.885	-43.607	Huon
148.350	148.433	-42.750	-42.700	MariaIsland
148.429	152.800	-43.483	-41.067	Freycinet
148.545	148.752	-38.530	-38.306	Seiners
148.587	153.783	-40.850	-38.950	Flinders
148.650	148.875	-41.350	-41.083	StHelens
148.718	148.933	-40.377	-39.433	FlindersRZ
149.367	149.550	-38.267	-38.083	Everard
149.850	150.600	-38.500	-37.633	EastGippland
151.437	151.892	-34.228	-33.724	Endeavour

The trawl fisheries for Flathead in zones 10 – 20, Flathead in zone 30, Pink Ling in zones 10 – 30, and John Dory in zone 10 and 20 were examined. The auto-line fishery for Blue-Eye in zones 20 and 30 was also examined. The catch and effort data were extracted from the log-book database for each species and then used in the five standardizations and the outcomes compared.

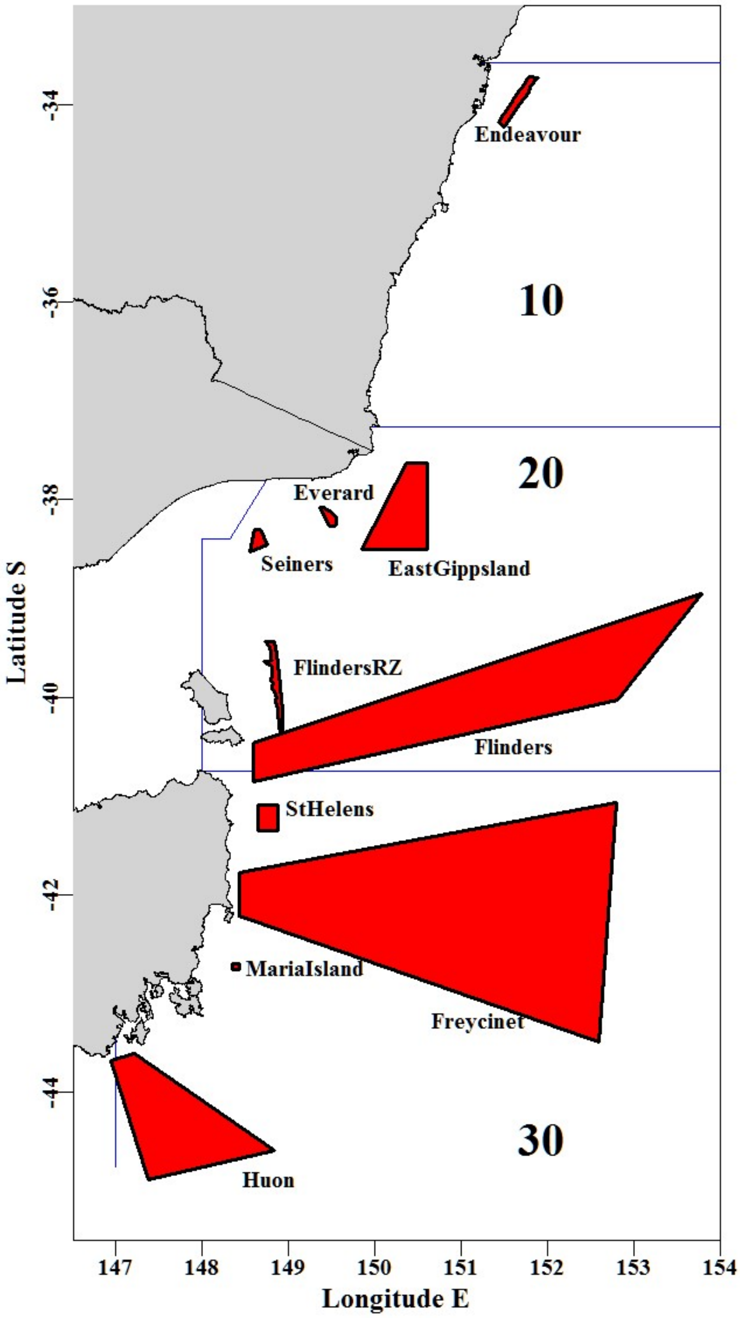


Figure 7.2: A schematic map of the SESSF zones used and the closures within those zones that were explored for their effects upon CPUE standardization.

7.3.4 The National Network of Reserves

The South-East closures were the only ones used in this empirical analysis of the potential effects of closures on commercial CPUE. This was because the reserves in the South-East were the only ones activated at the time of writing. In fact, the national network of reserves is very extensive (Figure 7.3). However, apart from the reserves in the South-East, established in 2007, all other reserves are currently under review. The text in the middle of Figure 7.3 reads:

“*IMPORTANT INFORMATION FOR MARINE USERS*

Transitional arrangements apply to the South-west, North-west, North and Temperate East Networks and the Coral Sea reserve. These arrangements involve NO CHANGES ON THE WATER for marine users. Note, there are no changes to management arrangements in the marine reserves that existed prior to the establishment of the new reserves, that is, the same restrictions on activities will continue to apply even where those reserves have been incorporated into new reserves. More information is available at www.environment.gov.au/marinereserves “.

What this means is that while the closures have been defined they are not currently active. Exceptions are found in the Heard and McDonald and Macquarie Island Toothfish fisheries and some other areas (e.g. in the Northern Prawn fishery), nevertheless, analyses on the effects of the Commonwealth closures are currently restricted to being made in the South-East.

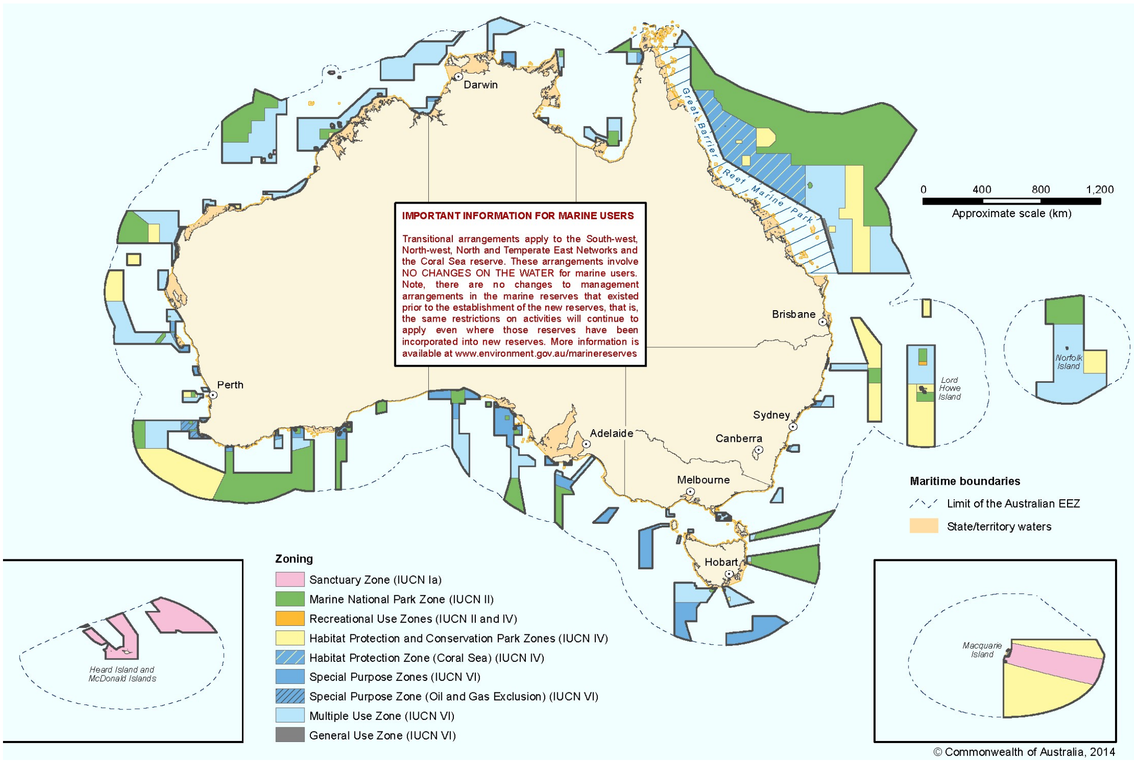


Figure 7.3: A schematic map of the network of Commonwealth Marine Reserves taken on 09/12/2016 from http://www.environment.gov.au/system/files/pages/2ed9e96f-d06b-460b-81de-8cd11f2ea66f/files/national-map_0.pdf

7.4 Results

7.4.1 Flathead (*Neoplatycephalus richardsoni*) Zones 10 – 20

The Flathead fishery in zone 30 is of a different character to that in zones 10 and 20 so zones 10+20 and 30 are standardized separately. The catches by zone are variable through time (Table 7.3; Figure 7.3). The catch by method is also variable, with most of the rest of the total catch being taken using Danish Seine (Table 7.4; see Sporcic and Haddon, 2016).

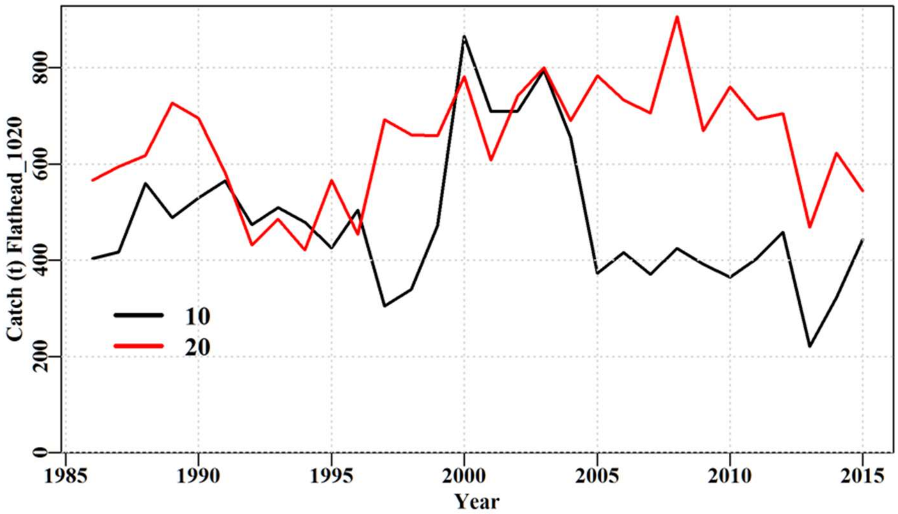


Figure 7.4: Reported catches of flathead taken by demersal trawl from 1986 – 2015 in SESSF zones 10 and 20 (see Table 7.3).

Table 7.3: Reported catches of flathead taken by demersal trawl from 1986 – 2015 in SESSF zones 10 and 20 (see Figure 7.3).

Year	10	20	Year	10	20
1986	403.218	565.578	2001	708.336	608.087
1987	417.102	594.312	2002	710.263	741.638
1988	559.980	616.996	2003	795.247	800.549
1989	488.102	726.587	2004	654.343	689.967
1990	529.143	695.352	2005	372.877	783.117
1991	565.381	581.834	2006	415.900	733.009
1992	473.601	431.413	2007	370.755	705.708
1993	508.967	485.208	2008	424.911	905.909
1994	478.752	421.547	2009	392.244	668.469
1995	425.257	565.637	2010	364.544	759.808
1996	504.020	453.345	2011	403.752	692.747
1997	305.240	691.438	2012	458.407	704.087
1998	339.052	660.639	2013	221.140	468.321
1999	471.779	657.891	2014	322.842	623.086
2000	865.298	780.980	2015	443.268	544.406

Table 7.4: Catch by trawl code within Zones 10 and 20, with TDO (Trawl Demersal Otter) being a relatively new code used in eLogs.

Year	TW	Year	TW	Year	TDO	TW
1986	968.796	1996	957.365	2006		1148.909
1987	1011.414	1997	996.678	2007		1076.463
1988	1176.976	1998	999.691	2008		1330.820
1989	1214.689	1999	1129.670	2009		1060.713
1990	1224.495	2000	1646.278	2010		1124.352
1991	1147.215	2001	1316.423	2011		1096.500
1992	905.014	2002	1451.900	2012		1162.494
1993	994.175	2003	1595.795	2013	63.633	625.828
1994	900.299	2004	1344.310	2014	481.593	464.335
1995	990.894	2005	1155.994	2015	466.161	521.513

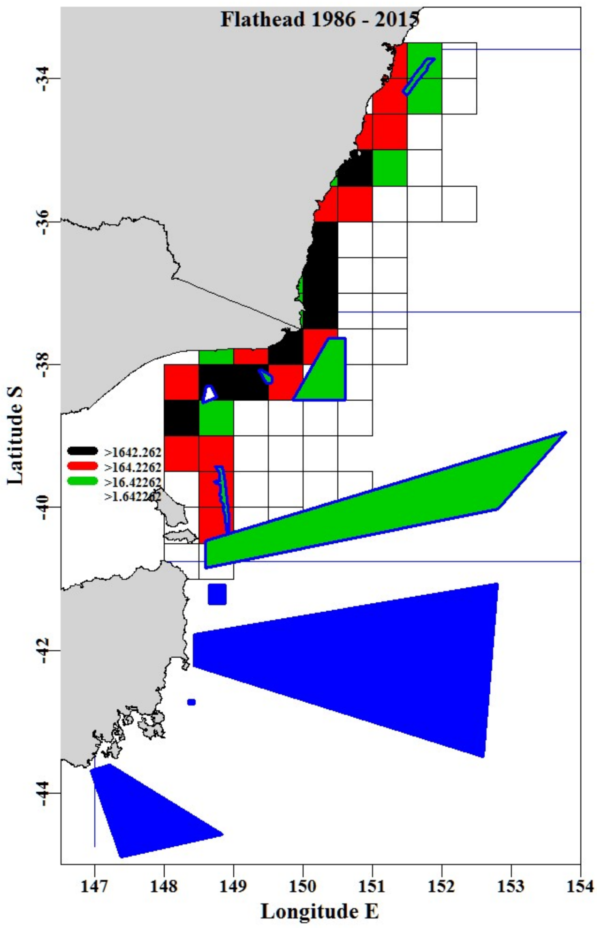


Figure 7.5: Schematic map with all reported records of trawl caught Flathead in zones 10 and 20 from 1986 – 2015. Black-filled closures had more than 1,642 tonnes, Red-filled more than 164 tonnes, green-filled more than 16 tonnes, white-filled closures had >1.6 tonnes, and blue-filled closures had no records (see Table 7.5).

The total catch within all closures across the years 1986 – 2015 amounted to less than three quarters of 1% (any catches after about 2007 reflect the inaccuracy of the recorded GPS data rather than any real catches that have been taken inside the closures). Given there were only 1,568 records within the closures and 271,003 records outside, it is not surprising that the effect of the closures on the standardization for flathead in zones 10 and 20 was effectively trivial (Table 7.5; Figure 7.5 and Figure 7.6).

Table 7.5: The total number of records and catches in the open areas and closures for Flathead in zones 10 – 20 from 1986 – 2015.

Closure Name	Records	Catch (t)	%Catch
Open	271003	33501.460	99.26
Endeavour	289	26.095	0.077
FlindersRZ	332	46.275	0.137
EastGippland	223	27.353	0.081
Flinders	261	46.745	0.138
Seiners	405	95.146	0.282
Everard	58	8.305	0.025
Total	272571	33751.383	100
Total Out	271003	33501.464	99.26
Total In	1568	249.919	0.74

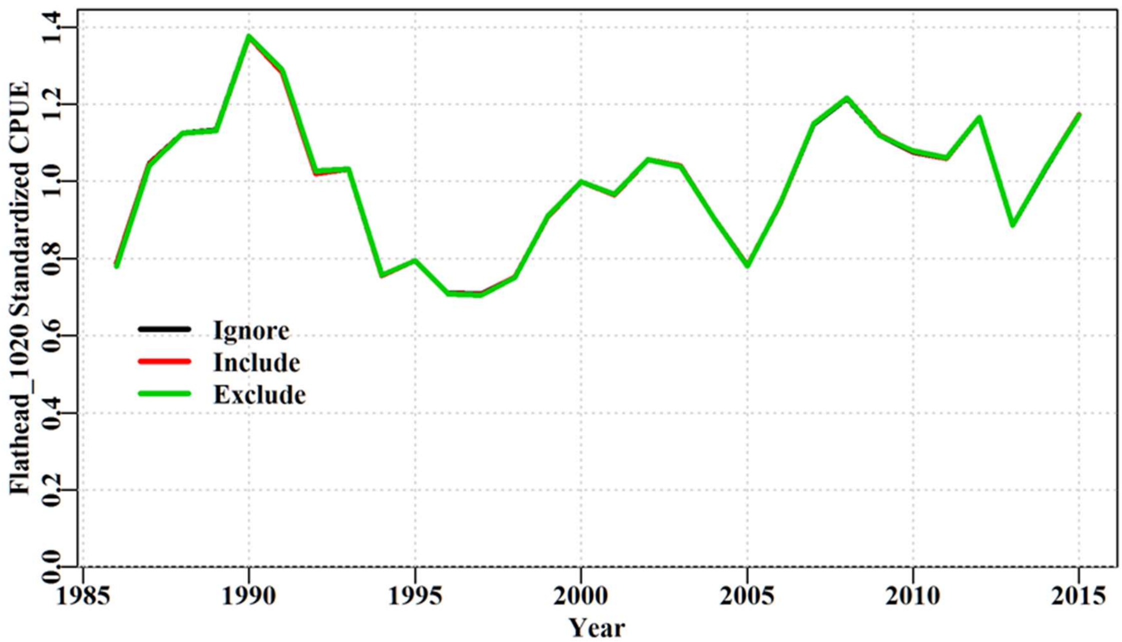


Figure 7.6: The outcome of the three alternative analyses on the standardization of Flathead in zones 10 – 20 taken by trawl. Only extremely small effects are just visible in years such as 1991, 1992, 1997, and 2010; otherwise the standardizations lie almost precisely on top of each other.

The optimum standardizations for the ‘ignore closures’, ‘include as a factor’, and ‘remove closure’ data analytical options all have very similar values differing only at the third of fourth decimal place (Table 7.6).

Table 7.6: The optimum standardizations for the ignore closures, include as a factor, and remove closure data analytical options. The simple geometric mean (LnCE ~ Year) is included for comparison. Each series can be multiplied by the arithmetic average of the yearly geometric mean (43.02kg/hr) to put them all on the same scale. Currently each series has a mean of 1.0.

Year	Geometric	Ignore	Include	Exclude	Year	Geometric	Ignore	Include	Exclude
1986	0.6879	0.7877	0.7865	0.7795	2001	0.9023	0.9655	0.9656	0.9662
1987	0.8300	1.0463	1.0444	1.0412	2002	0.9657	1.0556	1.0564	1.0563
1988	0.9662	1.1251	1.1248	1.1246	2003	0.9514	1.0394	1.0404	1.0388
1989	0.9690	1.1337	1.1320	1.1313	2004	0.8319	0.9038	0.9043	0.9046
1990	1.2524	1.3771	1.3756	1.3759	2005	0.7494	0.7814	0.7817	0.7802
1991	1.1835	1.2851	1.2821	1.2890	2006	0.9427	0.9421	0.9427	0.9443
1992	1.0163	1.0215	1.0211	1.0270	2007	1.3285	1.1485	1.1493	1.1487
1993	0.9924	1.0317	1.0316	1.0326	2008	1.3391	1.2151	1.2166	1.2171
1994	0.7612	0.7564	0.7564	0.7570	2009	1.2720	1.1181	1.1193	1.1188
1995	0.7633	0.7945	0.7941	0.7947	2010	1.2473	1.0767	1.0779	1.0786
1996	0.6972	0.7093	0.7087	0.7084	2011	1.2006	1.0592	1.0601	1.0614
1997	0.7125	0.7080	0.7071	0.7057	2012	1.2909	1.1652	1.1663	1.1662
1998	0.7510	0.7531	0.7527	0.7515	2013	0.9933	0.8862	0.8870	0.8863
1999	0.8642	0.9077	0.9075	0.9091	2014	1.1921	1.0355	1.0366	1.0364
2000	1.0021	0.9992	0.9984	0.9984	2015	1.3436	1.1716	1.1729	1.1702

7.4.2 Flathead (*Neoplatycephalus richardsoni*) Zone 30

Catches of Flathead by trawl in SESSF Zone 30 only rose above 100 t per year after 1997. They reached a peak in 2004 after which they dropped sharply, but have been rising again almost every year since 2010 (Figure 7.6; Table 7.7).

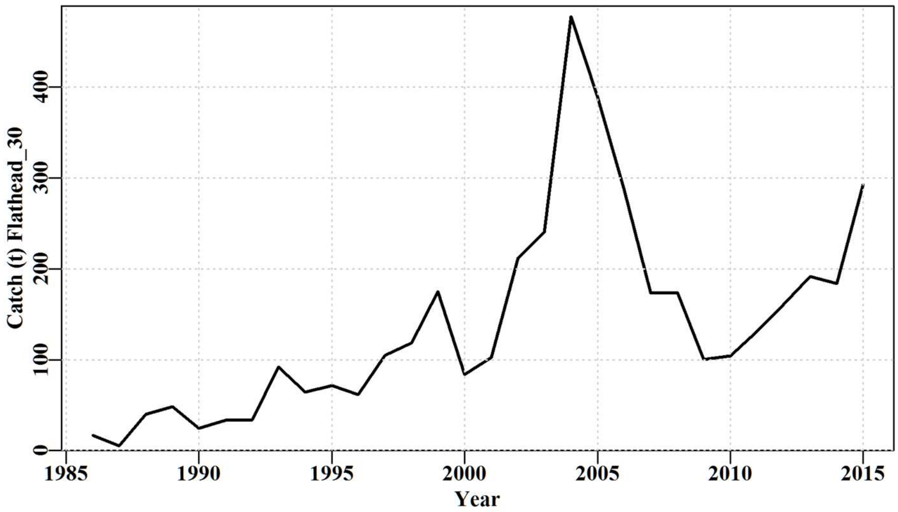


Figure 7.7: Reported catches of Flathead taken by demersal trawl from 1986 – 2015 in SESSF zones 30 (see Table 7.7).

Table 7.7: Reported catches of Flathead taken by demersal trawl from 1986 – 2015 in SESSF zones 30 (see Figure 7.6).

Year	TW	Year	TDO	TW	Total
1986	16.754	2001		102.749	
1987	5.155	2002		212.158	
1988	40.256	2003		240.110	
1989	48.473	2004		477.416	
1990	24.619	2005		388.325	
1991	33.413	2006		287.968	
1992	33.897	2007		173.155	
1993	92.079	2008		173.739	
1994	64.487	2009		100.225	
1995	71.349	2010		104.186	
1996	61.425	2011		131.274	
1997	104.875	2012		160.746	
1998	118.552	2013	9.072	182.273	191.345
1999	175.052	2014	69.279	114.408	183.687
2000	83.664	2015	153.849	139.04	292.889

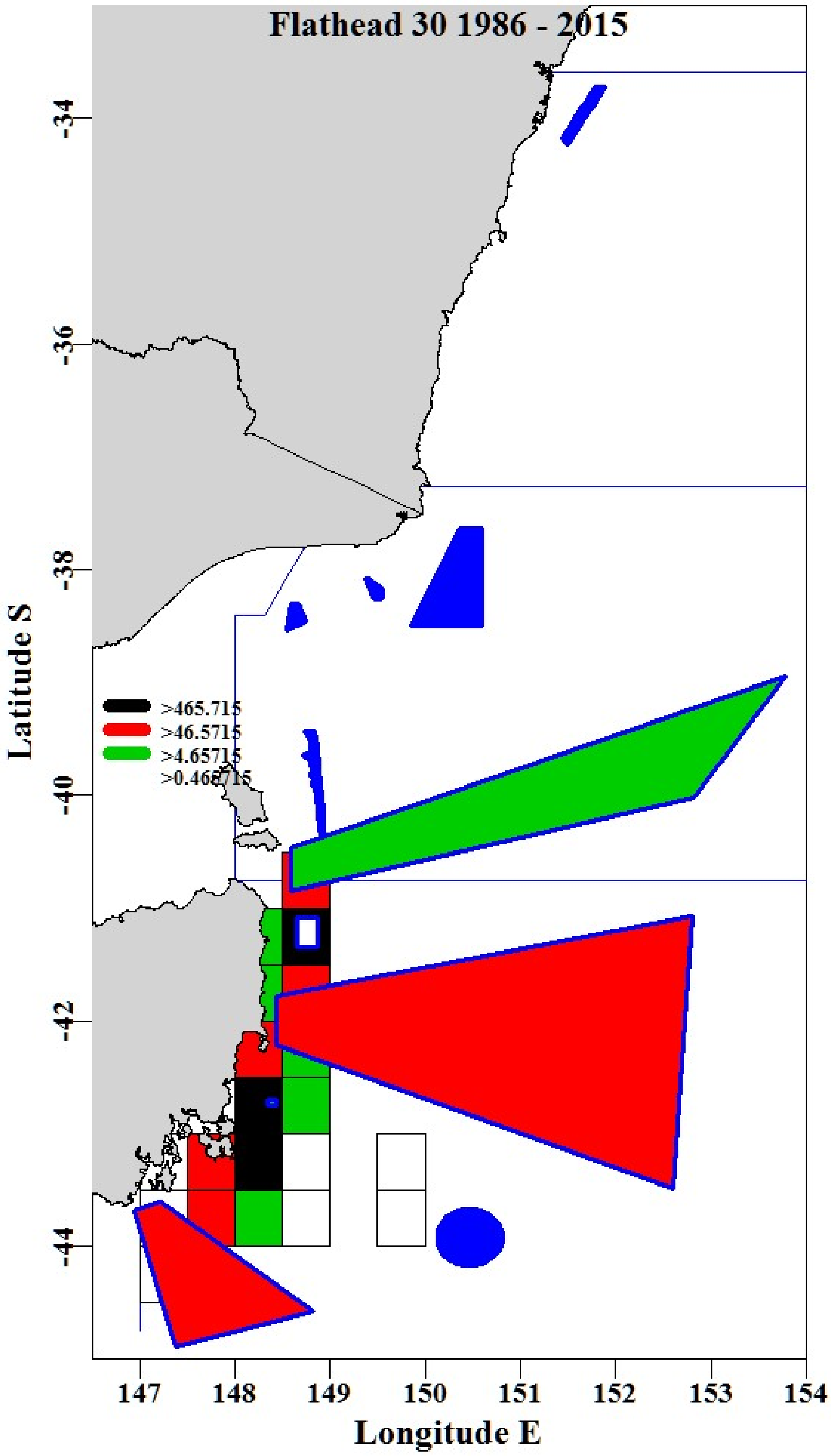


Figure 7.8: Schematic map with all reported records of trawl caught Flathead from zone 30, from 1986 – 2015. Black-filled closures had more than 465 tonnes, red-filled closures more than 46 tonnes, green-filled closures had >4.6 tonnes, white-filled closures had more than 0.46 tonnes, and blue-filled closures had no records (see Table 7.8).

The total catch within all closures across the years 1986 – 2015 amounted to about 4.6% and similarly for the number of records at about 5.2% (any catches after about 2007 reflect the inaccuracy of the recorded GPS data rather than any real catches that have been taken inside the closures). Despite this higher proportion of catches in closures in zone 30 than in zones 10 and 20, their effect on the standardization remained effectively trivial. There are very minor deviations in many places along the times series (Table 7.8; Figure 7.8 and Figure 7.9). Note that the CPUE trajectory for zone 30 is rather different to that for zones 10 and 20, confirming that their separate treatment avoids some complex spatial interactions with year.

Table 7.8: The total number of records and catches in the open areas and closures for Flathead in zone 30 from 1986 – 2015.

Name	Records	Catch	pCatch
Open	21838	4000.664	95.4
StHelens	77	6.221	0.1
Huon	560	82.762	2.0
Freycinet	516	96.227	2.3
Flinders	45	7.812	0.2
MariaIsland	6	0.335	0.0
Total	23042	4194.021	100.0
Total Out	21838	4000.664	95.39
Total In	1204	193.357	4.61

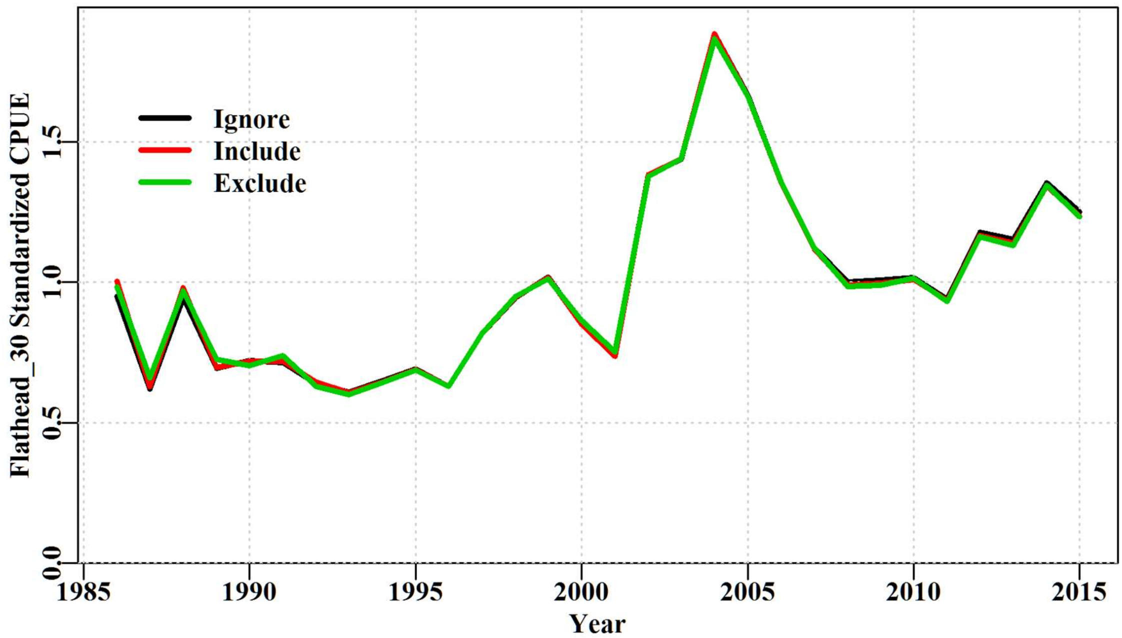


Figure 7.9: The outcome of the three alternative analyses on the standardization of Flathead in zone 30 taken by trawl. Only very small effects are visible in a number of years across the time series.

The optimal standardizations for the Ignore, the Include, and the Exclude options for Flathead in Zone 30 mostly differ at the second decimal place although there are still years that differ at the third and even fourth decimal place (Table 7.9). Visually they remain very similar (Figure 7.9).

Table 7.9: The optimum standardizations for the ignore closures, include as a factor, and remove closure data analytical options. The simple geometric mean (LnCE ~ Year) is included for comparison. Each series can be multiplied by the arithmetic average of the yearly geometric mean (49.994kg/hr) to put them all on the same scale. Currently each series has a mean of 1.0.

Year	Geometric	Ignore	Include	Exclude	Year	Geometric	Ignore	Include	Exclude
1986	0.8452	0.9491	1.0029	0.9832	2001	0.6655	0.7411	0.7370	0.7535
1987	0.4152	0.6198	0.6290	0.6616	2002	1.1107	1.3840	1.3832	1.3773
1988	0.7692	0.9453	0.9815	0.9703	2003	1.1081	1.4364	1.4391	1.4395
1989	0.7557	0.6935	0.6961	0.7261	2004	1.7333	1.8854	1.8848	1.8671
1990	0.7507	0.7211	0.7196	0.7025	2005	1.6040	1.6647	1.6601	1.6609
1991	0.5859	0.7154	0.7192	0.7392	2006	1.3837	1.3593	1.3586	1.3620
1992	0.8290	0.6389	0.6457	0.6305	2007	1.2187	1.1231	1.1197	1.1250
1993	0.6310	0.6095	0.6084	0.6015	2008	1.0813	1.0002	0.9890	0.9862
1994	0.6837	0.6493	0.6463	0.6423	2009	1.0735	1.0080	0.9971	0.9901
1995	0.7338	0.6922	0.6900	0.6882	2010	1.0453	1.0175	1.0097	1.0135
1996	0.5814	0.6303	0.6304	0.6304	2011	1.0746	0.9416	0.9353	0.9320
1997	0.7638	0.8179	0.8168	0.8173	2012	1.2964	1.1783	1.1656	1.1614
1998	1.0650	0.9458	0.9467	0.9504	2013	1.2012	1.1522	1.1401	1.1313
1999	1.1473	1.0199	1.0178	1.0126	2014	1.4727	1.3544	1.3428	1.3450
2000	0.9390	0.8539	0.8509	0.8646	2015	1.4350	1.2521	1.2365	1.2344

7.4.3 Pink Ling (*Genypterus blacodes*) Zones 10 – 30

The fishery for Pink Ling has a more complex history than the fishery for Flathead because it can also be targeted by auto-line vessels. The rapid drop in catch from zone 10 by trawl in 2000 and 2001 has been explained by a transfer/sale of Pink Ling quota from the trawl fleet to the auto-line fleet (Table 7.10; Figure 7.10). There was also a large drop in both the geometric and standardized CPUE over the same period (Table 7.12; Figure 7.12). This led to the stock assessments from 2011 onwards to consider the eastern Pink Ling stock to be just below the Limit Reference Point (Punt and Taylor, 2012).

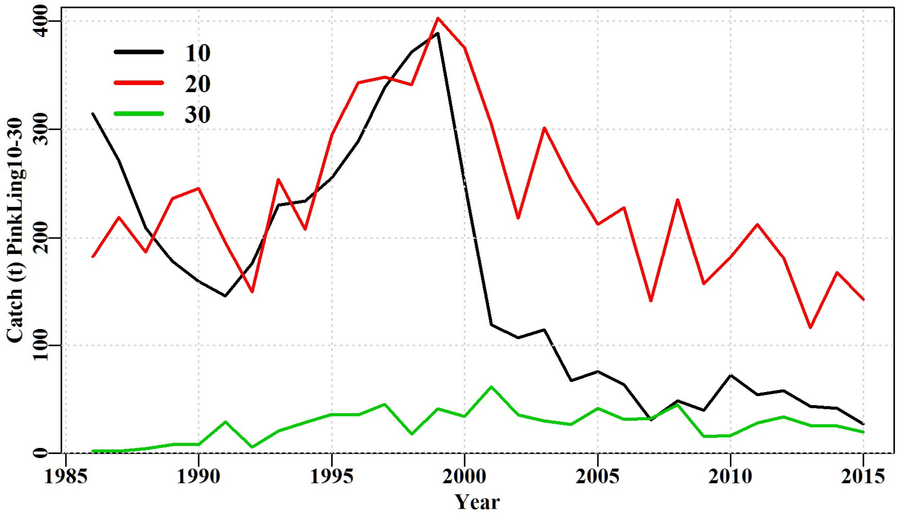


Figure 7.10: Reported catches of Pink Ling taken by demersal trawl from 1986 – 2015 in SESSF zones 10 – 30 (see Table 7.10).

Table 7.10: Reported catches of Pink Ling taken by demersal trawl from 1986 – 2015 in SESSF zones 10 – 30 (see Figure 7.9).

Year	10	20	30	Year	10	20	30
1986	314.477	181.925	1.896	2001	119.013	304.944	61.674
1987	271.077	218.992	2.245	2002	106.938	218.057	35.598
1988	209.126	186.573	4.378	2003	114.393	301.477	29.893
1989	178.035	236.054	7.988	2004	67.395	252.875	26.968
1990	159.261	245.436	8.385	2005	75.758	212.448	41.743
1991	145.704	195.643	28.950	2006	63.499	228.071	31.531
1992	176.153	149.347	5.806	2007	31.023	141.086	32.198
1993	230.026	253.845	20.603	2008	48.896	235.294	44.846
1994	234.002	207.749	28.514	2009	39.817	156.773	15.772
1995	255.548	294.988	36.150	2010	72.535	182.205	16.392
1996	288.862	342.913	35.808	2011	54.275	212.576	28.045
1997	338.960	348.020	45.674	2012	58.242	181.406	33.675
1998	371.650	341.020	17.788	2013	43.485	116.173	25.748
1999	388.577	402.701	41.377	2014	41.826	167.527	25.464
2000	250.831	375.397	34.098	2015	27.215	142.546	19.687

The effect of closures on the Pink Ling trawl fishery was to exclude areas where previously up to about 4.25% of all catches were taken. Some closures in the Horseshoe region of eastern Bass Strait included catches > 100t (Table 7.11; Figure 7.11). However, these were distributed through time and the effects on the CPUE trend were very minor (Figure 7.12).

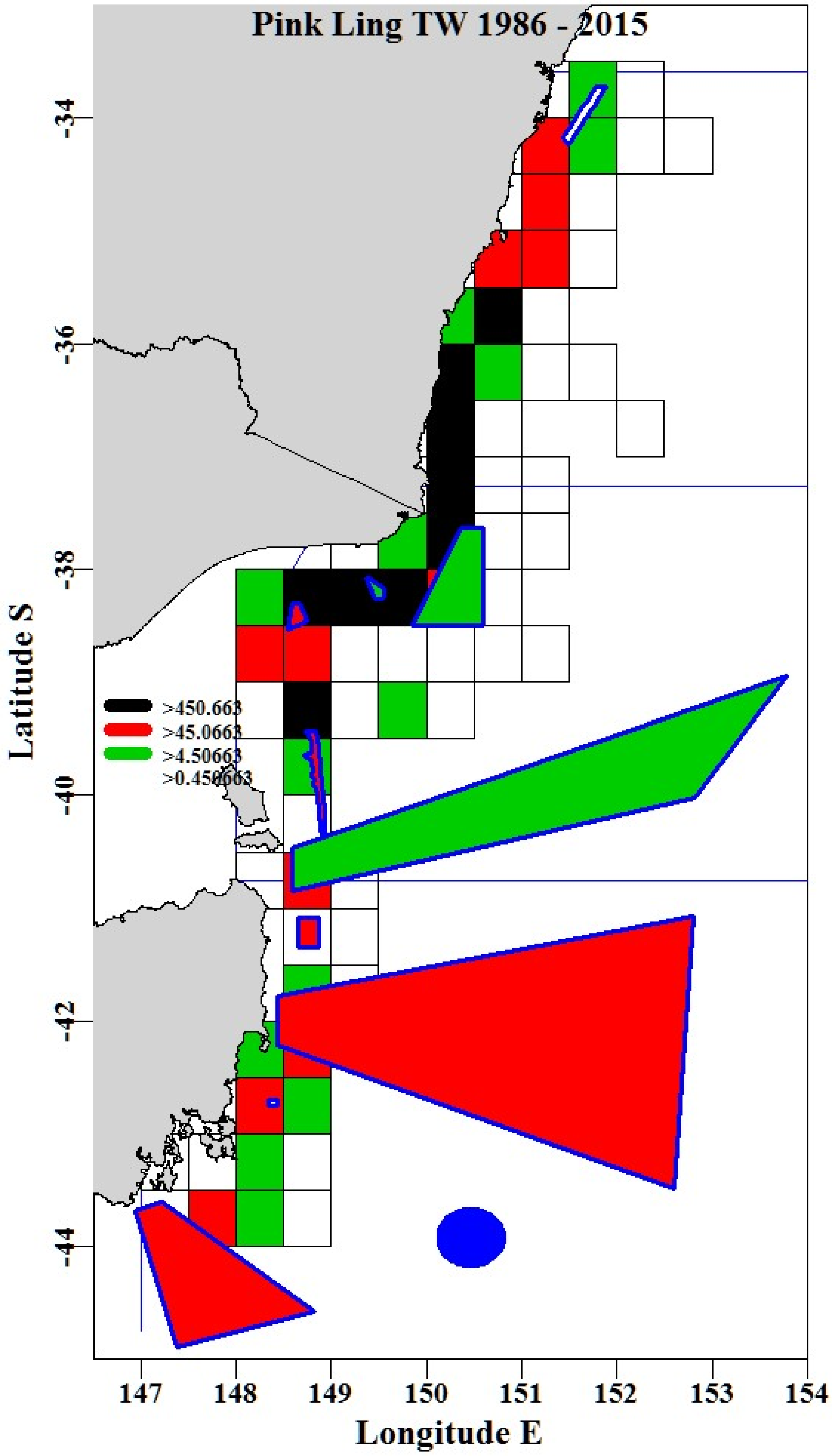


Figure 7.11: Schematic map with all reported records of trawl caught Pink Ling in zones 10 – 30 from 1986 – 2015 in depths 250 – 600m. Red-filled closures had more than 45 tonnes, green-filled closures had >4.5 tonnes, and white-filled closures had > 0.46 tonnes, and blue-filled closures had no records (see Table 7.11).

Table 7.11: The total number of records and catches in the open areas and closures for Pink Ling taken by trawl in depths of 250 – 600m in zines 10 – 30 from 1986 – 2015.

Closure	Records	Catch	% Total Catch
Open	96034	12060.157	95.72
Endeavour	941	17.526	0.14
StHelens	28	1.599	0.01
FlindersRZ	479	35.067	0.28
Huon	81	5.120	0.04
Freycinet	454	48.062	0.38
EastGippland	286	55.082	0.44
Flinders	719	73.472	0.58
MariaIsland	56	3.993	0.03
Seiners	805	155.589	1.24
Everard	774	143.883	1.14
Total	100657	12599.550	100.00
Total Out	96034	12060.157	95.72
Total In	4623	539.393	4.28

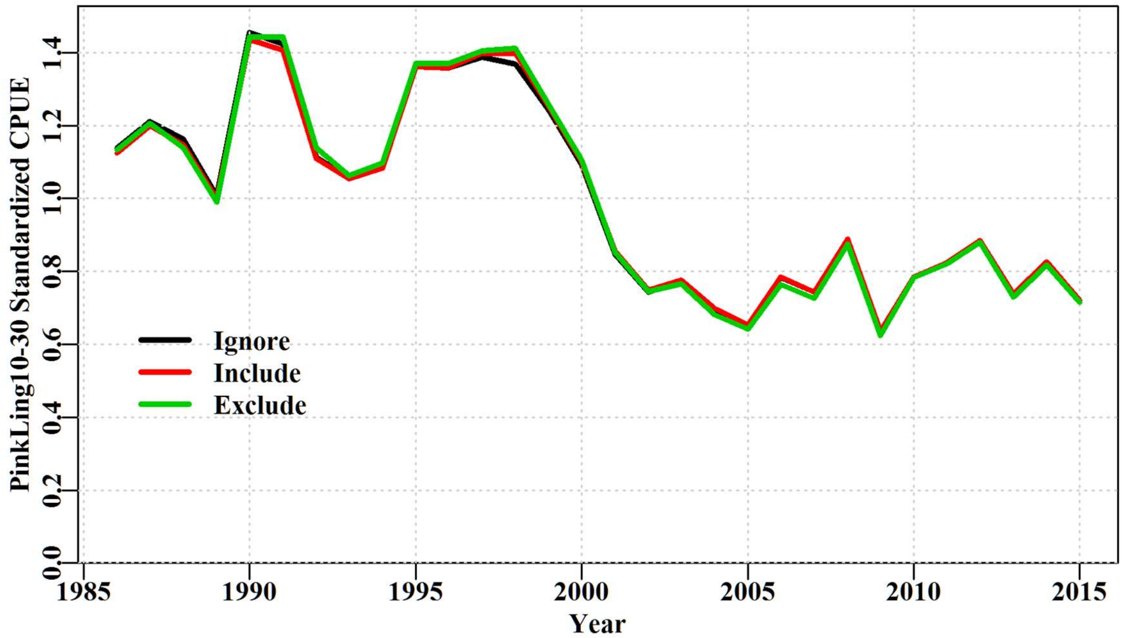


Figure 7.12: The outcome of the three alternative analyses on the standardizations of Pink Ling in zones 10 – 30, in depths 250 – 600m, taken by trawl. Only minor effects are visible in a number of years across the time series.

Table 7.12: The optimum standardizations for the ignore closures, include as a factor, and remove closure data analytical options for Pink Ling in zones 10 – 30. The simple geometric mean (LnCE ~ Year) is included for comparison. Each series can be multiplied by the arithmetic average of the yearly geometric mean (44.57kg/hr) to put them all on the same scale. Currently each series has a mean of 1.0.

Year	Geometric	Ignore	Include	Exclude	Year	Geometric	Ignore	Include	Exclude
1986	0.9263	1.1390	1.1244	1.1345	2001	0.8542	0.8479	0.8549	0.8527
1987	0.8709	1.2111	1.2002	1.2065	2002	0.7106	0.7438	0.7486	0.7445
1988	0.9085	1.1630	1.1477	1.1396	2003	0.8198	0.7751	0.7762	0.7666
1989	0.8590	1.0072	0.9987	0.9901	2004	0.7531	0.6966	0.6990	0.6806
1990	1.2027	1.4556	1.4366	1.4432	2005	0.7324	0.6510	0.6540	0.6409
1991	1.1796	1.4246	1.4068	1.4425	2006	0.9560	0.7835	0.7850	0.7641
1992	1.1243	1.1132	1.1104	1.1382	2007	0.9195	0.7431	0.7442	0.7256
1993	1.1348	1.0581	1.0549	1.0626	2008	1.1279	0.8878	0.8877	0.8755
1994	1.0544	1.0856	1.0834	1.0979	2009	0.8205	0.6337	0.6364	0.6244
1995	1.1573	1.3626	1.3618	1.3701	2010	0.9284	0.7848	0.7849	0.7824
1996	1.2401	1.3575	1.3587	1.3696	2011	1.0508	0.8250	0.8240	0.8210
1997	1.2527	1.3868	1.3978	1.4042	2012	1.0922	0.8842	0.8845	0.8803
1998	1.1665	1.3694	1.3980	1.4118	2013	0.9526	0.7358	0.7387	0.7302
1999	1.1312	1.2437	1.2539	1.2607	2014	1.0982	0.8229	0.8263	0.8173
2000	1.0051	1.0917	1.1017	1.1067	2015	0.9701	0.7156	0.7209	0.7154

7.4.4 John Dory (*Zeus faber*) Zones 10 – 20

Most catches of John Dory are considered as a desirable byproduct because targeting is generally not thought to be common. Catches in zone 10 have declined to levels more similar to those taken in zone 20, which has been more stable through time (Table 7.13; Figure 7.13). Consistent with the notion that this species cannot be targeted very successfully the CPUE appears to closely reflect the catches (Figure 7.14 and Figure 7.15).

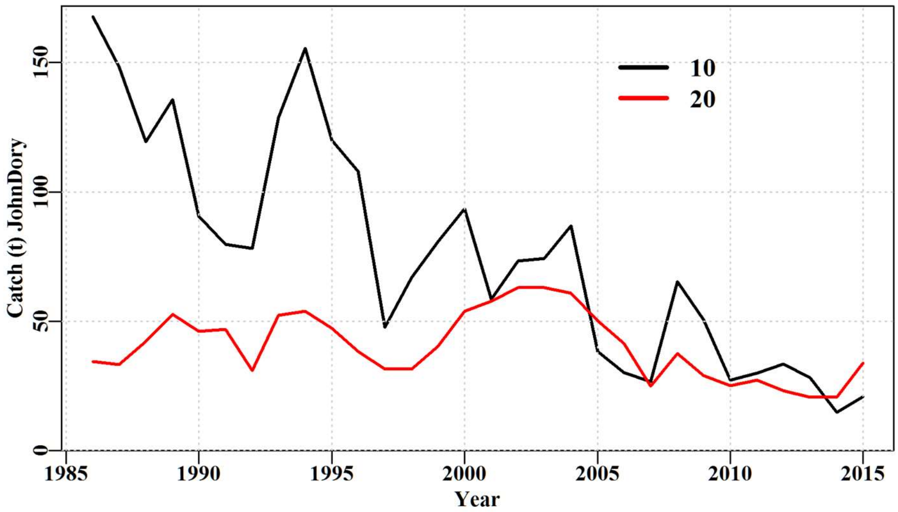


Figure 7.13: Reported catches of John Dory taken by demersal trawl from 1986 – 2015 in SESSF zones 10 and 20 (see Table 7.13).

Table 7.13: Reported catches of John Dory taken by demersal trawl from 1986 – 2015 in SESSF zones 10 and 20 (see Figure 7.12).

Year	10	20	Year	10	20
1986	167.704	34.519	2001	58.597	57.716
1987	148.226	33.355	2002	73.398	63.012
1988	119.396	42.232	2003	74.315	63.006
1989	135.747	52.721	2004	86.895	60.801
1990	90.548	46.226	2005	38.335	50.305
1991	79.755	46.941	2006	30.197	41.429
1992	78.095	31.021	2007	26.789	24.896
1993	128.763	52.304	2008	65.402	37.590
1994	155.507	53.878	2009	50.673	29.073
1995	119.969	47.317	2010	27.289	25.159
1996	107.976	38.369	2011	30.044	27.356
1997	47.710	31.483	2012	33.437	23.142
1998	66.843	31.644	2013	28.309	20.604
1999	80.703	40.291	2014	14.775	20.647
2000	93.444	53.901	2015	20.965	33.801

The catches and relative catch levels that have been taken within closures through 1986 – 2015 are effectively trivial (Figure 7.14; Table 7.14). With the total percentage of catches previously taken in closures being only about 0.16% it is not surprising that the closures appear visually to have had no effect on the standardizations. Most of the differences between the trends are at the fourth decimal place (Table 7.15).

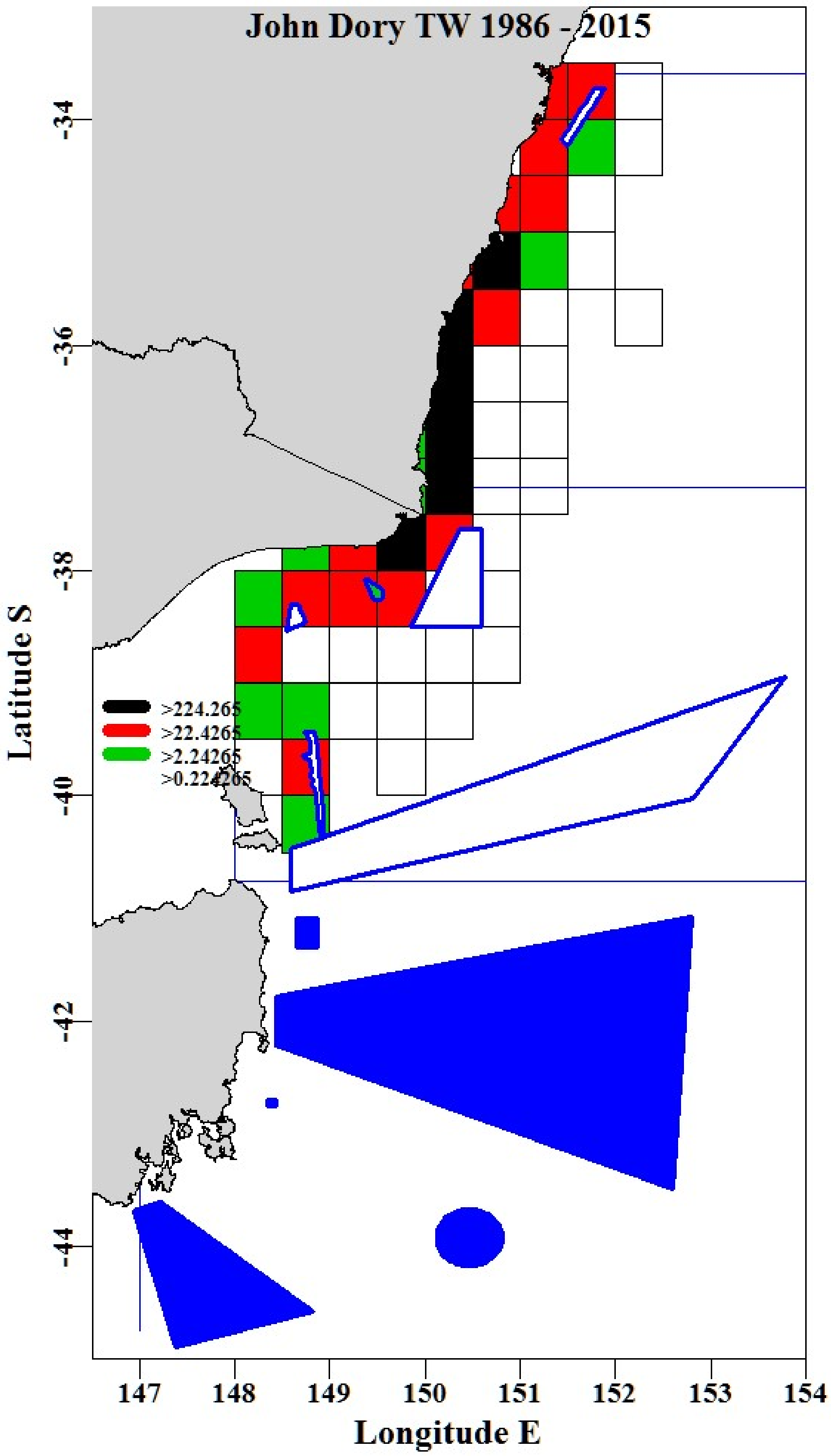


Figure 7.14: Schematic map with all reported records of trawl of John Dory in zones 10 and 20 from 1986 – 2015. The green-filled closure (Everard) had >2.2 tonnes, and white-filled closures had < 2.2 tonnes, and blue-filled closures had no records (see Table 7.14).

Table 7.14: The total number of records and catches of John Dory taken by trawl in the open areas and closures in zones 10 – 20 from 1986 – 2015.

Name	Records	Catch	pCatch
Open	141812	3488.839	99.84
Endeavour	2	0.110	0.00
FlindersRZ	34	0.755	0.02
EastGippland	103	2.576	0.07
Flinders	56	0.202	0.01
Seiners	72	1.691	0.05
Everard	21	0.370	0.01
Total	142100	3494.543	100.00
Total Out	141812	3488.839	99.84
Total In	288	5.704	0.16

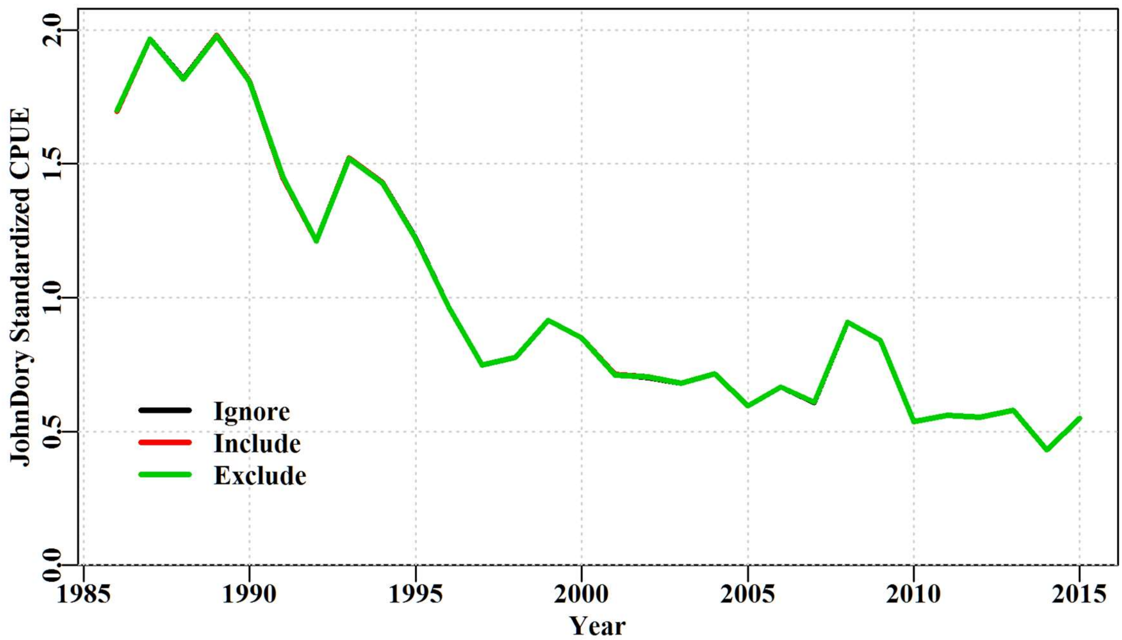


Figure 7.15: The outcome of the three alternative analyses on the standardization of John Dory in zones 10 – 20 taken by trawl. There are no visible effects on the different standardizations which lie almost precisely on top of each other.

Table 7.15: The optimum standardizations for the ignore closures, include as a factor, and remove closure data analytical options. The simple geometric mean (LnCE ~ Year) is included for comparison. Each series can be multiplied by the arithmetic average of the yearly geometric mean (8.33kg/hr) to put them all on the same scale. Currently each series has a mean of 1.0.

Year	Geometric	Ignore	Include	Exclude	Year	Geometric	Ignore	Include	Exclude
1986	1.6472	1.6962	1.6968	1.6997	2001	0.6448	0.7126	0.7124	0.7121
1987	1.8236	1.9657	1.9655	1.9672	2002	0.6743	0.7025	0.7052	0.7049
1988	1.7957	1.8173	1.8169	1.8170	2003	0.6752	0.6796	0.6800	0.6800
1989	2.0406	1.9811	1.9814	1.9782	2004	0.7323	0.7165	0.7164	0.7166
1990	1.8719	1.8092	1.8089	1.8056	2005	0.5732	0.5953	0.5963	0.5952
1991	1.5406	1.4460	1.4474	1.4507	2006	0.6094	0.6674	0.6664	0.6666
1992	1.2304	1.2124	1.2121	1.2114	2007	0.6000	0.6075	0.6089	0.6098
1993	1.5211	1.5223	1.5218	1.5209	2008	0.9210	0.9097	0.9088	0.9094
1994	1.4456	1.4300	1.4298	1.4290	2009	0.8976	0.8405	0.8396	0.8403
1995	1.2728	1.2202	1.2195	1.2193	2010	0.5668	0.5365	0.5360	0.5363
1996	0.9696	0.9659	0.9657	0.9657	2011	0.5880	0.5603	0.5599	0.5599
1997	0.7238	0.7490	0.7486	0.7485	2012	0.6033	0.5542	0.5539	0.5541
1998	0.7786	0.7768	0.7768	0.7764	2013	0.6138	0.5802	0.5799	0.5801
1999	0.8442	0.9155	0.9154	0.9153	2014	0.4446	0.4320	0.4317	0.4319
2000	0.7743	0.8493	0.8497	0.8498	2015	0.5757	0.5484	0.5484	0.5479

7.4.5 Blue-Eye (*Hyperoglyphe Antarctica*) Zones 20 & 30

Blue-Eye are now assessed using a Tier 4 assessment, an empirical harvest strategy combined with a catch-per-hook analysis that combines the CPUE from Drop-Line and auto-line (Haddon, 2016a, 2016b).

Catches in Zone 20 have declined through time, which potentially reflects the introduction of the relatively influential closures near Flinders Island as the decline started in 2007 such that now the catches from zone 20 are minor (Table 7.16). Approximately 30% of all catches from 2002 – 2015 were taken from areas now inside closures (Table 7.17; Figure 7.17). The effect of such a large proportion of catches is apparent between the standardized CPUE trajectories from the three data selection treatments (Figure 7.18)._The changes prior to 2007 are more marked than those after but all of them remain relatively minor and barely affect the overall trend.

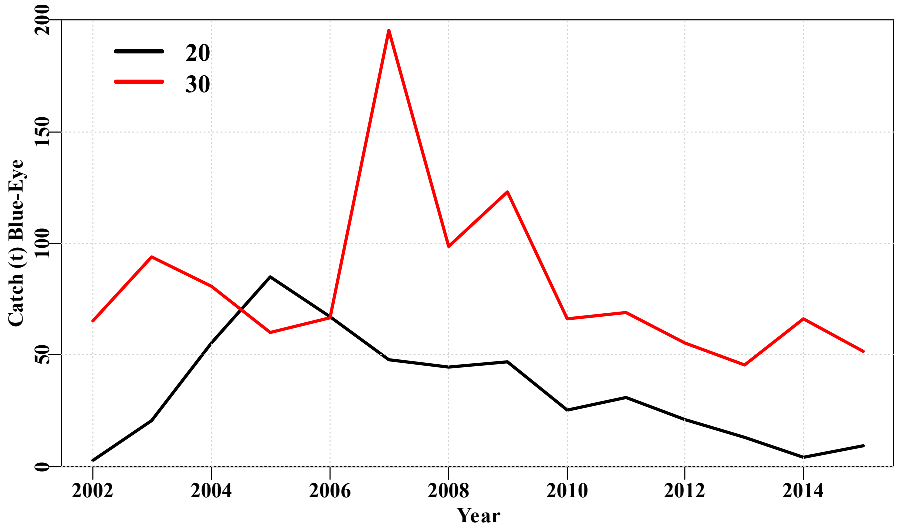


Figure 7.16: The catch of Blue-Eye by zone taken by auto-line. At least some of the decline in catch in Zone 20 is due to the advent of the closures.

Table 7.16: The catch by zone and the catches reported inside various closures and outside (Figure 7.16).

Year	Z20	Z30	Total	Outside	StHelens	FlindersRZ	Huon	Flinders	Other
2002	2.640	65.100	67.740	43.326	13.165	2.500	6.610	0.080	2.059
2003	20.574	93.768	114.342	46.457	36.039	8.075	16.465	2.780	4.526
2004	55.245	80.581	135.826	58.794	35.208	12.115	9.899	12.002	7.807
2005	84.748	59.833	144.581	68.804	19.127	34.170	2.400	14.478	5.602
2006	67.075	66.585	133.660	91.483	0.220	35.348	1.980	1.106	3.524
2007	48.001	195.263	243.264	204.795	10.980	17.623	0.400	4.525	4.942
2008	44.439	98.763	143.202	110.034	2.492	22.785	3.641	0.770	3.481
2009	47.014	122.952	169.966	127.608	4.092	29.058	3.657	3.003	2.548
2010	25.422	66.128	91.550	69.620	6.210	13.073	0.322	1.058	1.268
2011	30.835	68.834	99.669	78.586	0.411	17.478	1.362	0.440	1.391
2012	21.176	55.333	76.509	58.285	0.210	10.003	2.159	0.385	5.467
2013	13.151	45.406	58.557	49.083	0.151	3.890	3.745	0.144	1.544
2014	3.867	66.351	70.218	68.014			1.787		0.417
2015	9.031	51.790	60.821	52.734	0.631		4.814	0.106	2.536

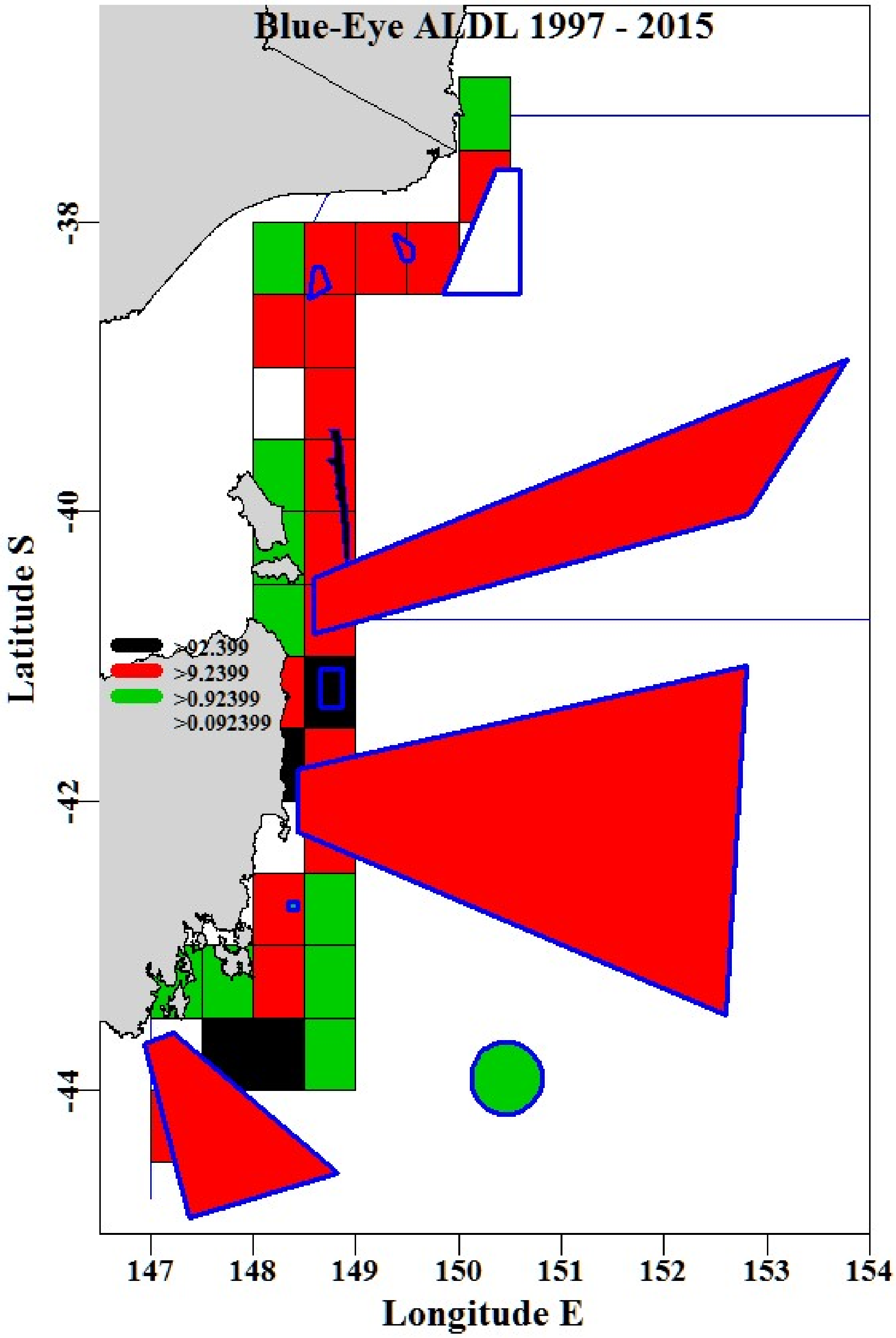


Figure 7.17: Schematic map with all reported records of Blue-Eye caught either by Drop-Line or auto-line in zones 20 and 30 from 1997 – 2015. Black-filled closures had more than 92 tonnes (St Helens and Freycinet), red-filled closures had >9.2 tonnes, green-filled closures had > 0.92 tonnes, and white-filled closures had < 0.09 tonnes (see Table 7.17).

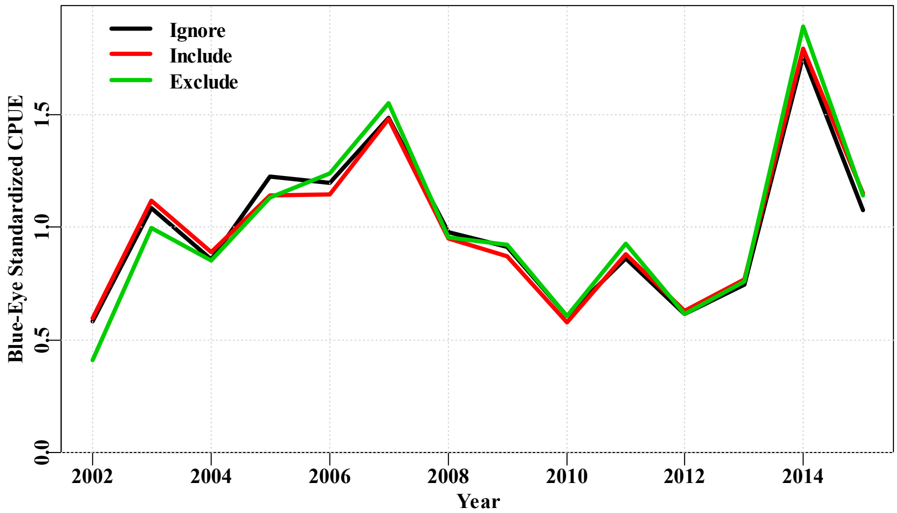


Figure 7.18: The outcome of the three alternative analyses on the standardization of Blue-Eye in zones 20 – 30 taken by auto-line. There are minor differences apparent between the 'ignore' and 'include' options, but there are larger effects with the 'exclude' option.

Table 7.17: The number of records, catch and percent of total catch reported by auto-line both outside of closures and inside particular closures.

Name	Records	Catch	%Catch
Outside	2549	1127.624	70.04
StHelens	182	128.935	8.01
FlindersRZ	368	206.116	12.80
Huon	68	59.241	3.68
Freycinet	87	15.571	0.97
EastGippland	1	0.03	0.00
Flinders	165	40.876	2.54
MariaIsland	22	3.624	0.23
Seiners	84	12.556	0.78
Everard	101	15.329	0.95

7.5 Discussion

Currently the only closures in the Commonwealth network of closures that are active (meaning they can exclude fishing) are those in the South-East, first established in 2007. However, efforts were made to minimize their effects upon the commercial fisheries when they were first proposed (Buxton *et al*, 2006). The objectives of that project (Buxton *et al*, 2006) that reviewed the closures proposed by the then Department of Environment and Heritage (DEH, approximately equivalent to the Department of the Environment and Energy in 2016) in an attempt to minimize the effects upon commercial fisheries were:

1. To quantify the commercial fisheries catch for key species within the proposed MPAs for the South-east region
2. To quantify the commercial fisheries economic value associated with the catch within the proposed MPAs for the South-east region
3. To quantify the socio-economic impact of the proposed MPAs on the commercial fishing industry
4. To outline in terms of 1, 2 & 3, alternative approaches that minimize impacts on the fishing Industry without compromising the biodiversity objectives of DEH.

Thus, it is not surprising that the effects of those closures on the CPUE of commercial fishing is relatively minor. In the case of Blue-Eye taken by the auto-line fishing method the closure that had the most effect was the relatively new Flinders Research Zone, which was closed specifically to reduce the effects of fishing on deepwater dogfish within the enclosed area (Figure 7.17). However, even this targeted closure, which was over one of the primary fishing locations for Blue-Eye (more than 15% of all reported catches in Zones 20 and 30 were taken inside the Flinders RZ closure) had only relatively minor effects on CPUE. Mostly, the effect on the trend of CPUE was minor. Overall, while there may be an immediate negative effect on CPUE following the introduction of a closure, after a short period industry appear to adapt and it becomes difficult to detect the effects of closures on CPUE standardizations.

The fact that the effects of so many closures on fishery catch rates is invariably relatively minor suggests that the fishing industry are capable of adapting to a changing management regime by displacing the effort they would have expressed in what are now closed areas in a manner that works to maintain their catch rates. Targeted closures that cover prime fishing areas can have an effect, but industry appears capable of minimizing even such targeted closures. Presumably their operations become somewhat less efficient, at least until or if they manage to find alternative fishing grounds that provide them with similar fishing opportunities. As far as is known, however, there is no data available to determine whether fishers travel further to fish than they did before closures were introduced. So no specific statements regarding the economic effects can be made.

It needs to be noted, however, that these analyses were conducted mainly on closures that had been modified to minimize their effects on fishing. If in the rest of the Commonwealth Marine Reserve network there are other closures that have not taken into account their potential effect on fishing then it may still be possible that if a large proportion of a fishery becomes closed a negative effect on CPUE could be observed. In this study of Flathead, Pink Ling, John Dory and Blue-Eye in the South-East little or no effects on CPUE were detectable.

An exception to these conclusion is the 700m deepwater closure where the principle fishing area for species such as Orange Roughy (*Hoplostethus atlanticus*) and the eastern and western deepwater sharks (basket TAC species) have been closed. For species where most of the fishable preferred habitat is closed the meaningfulness of any CPUE standardization becomes questionable. Generally the number of available records is greatly reduced and there are repeated reports of fishers altering their fishing behaviour near and around the deepwater closure. Such changes in fishing behaviour would imply that the catches taken by a given amount of effort after imposition of the closure are not necessarily comparable to catches tkane by the same amount of effort prior to the closure. Once more marine closures come into effect and exclude fishing more examples of different degrees of overlap with active fishing areas will become available and further empirical studies of the effects of such closures can be made.

8 Simulation study on the effect of CPUE resource standardization with and without marine closures

Miriana Sporcic[1]

[1]*CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, Tasmania, Australia*

8.1 Executive summary

This study examined the effect of marine closures on standardized catch-per-unit-effort (CPUE) to determine how well CPUE indexes true abundance according to various scenarios regarding resource (fish) movement dynamics and fisher behaviours. This involved simulating CPUE data using an agent based model (Resource-Fisher Integrated Model – RESFIM; Sporcic 2007; Sporcic and Smith 2009) across selected resource movements based on a generic platycephalid (i.e., tiger flathead) frequently occurring in the Southern and Eastern Scalefish and Shark Fishery (SESSF) and selected fisher behaviours that differ in terms of the degree of knowledge of resource dynamics. Generated RESFIM CPUE data were then standardized using generalized linear models (GLIMs), and resultant CPUE indices compared with true abundance to examine the effect that marine closures have on the CPUE-abundance relationship. In addition, the estimated bias of the proportionality parameter (which links CPUE to abundance), the degree of improvement of CPUE-abundance linearity following standardisation, and relative errors of annual indices (temporal bias) were also examined in the context of marine closures.

Overall, linearity between CPUE and abundance is desired for CPUE to adequately index abundance. Linear CPUE-abundance relationships were achieved in approximately 30% of resource/fisher scenarios following standardizations and mostly when resource movement was non-random. Standardisations also led to improvements towards linearity across other scenarios but which resulted in non-linearity (i.e. mostly hyperstable or hyperdepleted) indicating biased CPUE indices. Nominal CPUE-abundance relationships were non-linear (i.e. biased), and therefore not an adequate abundance index.

Key findings and recommendations are based on the results of this study and are provided below.

Key findings

- Significant improvements (in terms of linearity) over unstandardized indices for most resource/fisher scenarios support the use of standardized CPUE estimates as proxies for abundance.
- Statistical standardizations improved CPUE-abundance linearity, and the degree of improvement depends on the resource/fisher scenario. Significant improvements occurred when resource movement was non-random and when areas were not closed to fishing.
- Standardized CPUE were least effective (i.e. provided minimal improvements towards linearity) at indexing abundance under random resource movement (i.e. fish move randomly in space), irrespective of fisher behaviours with or without marine closures.

- Different factors associated with resource movement and fisher behaviour affected CPUE-abundance linearity, and linearity significantly improved following GLIM standardizations, with improvements mostly greater under information sharing and perfect fisher behaviour scenarios, with or without marine closures.
- Including year (Y), vessel (V), month (M), grid-location (G), G x M terms in standardizations led to the greatest improvements with or without closed areas across the resource/fisher scenarios.
- Including the Year term only in standardizations led to most biased standardized indices.
- The presence of non-linear CPUE-abundance relationships (i.e., β_S <1; β_S >1) with marine closures may bias stock assessments, as it deviates from CPUE-abundance linearity generally assumed in assessments. However, non-linearity is most likely when a resource is randomly dispersed either with or without marine closures.
- The bias in standardized CPUE-abundance relationships (β_S) were greater across each of the resource/fisher scenarios with than without closures.
- Relative error (which gives an indication of bias) was greater without marine closures than with marine closures based on either nominal or standardized CPUE across selected resource/fisher scenarios.
- Relative errors of nominal CPUE was mostly greater when resource movement was non-random (i.e. under habitat attraction) and without marine closures.
- Overall, relative errors of standardized CPUE reduced compared to nominal CPUE with or without marine closures across most resource/fisher scenarios.
- No trends in estimated bias of annual standardized CPUE were obtained in this study with or without marine closures.
- Relative errors of standardized CPUE were generally similar and mostly positive with or without marine closures.
- The utility of standardized CPUE to index abundance (as indicated by the relative error) was better under habitat attraction compared to random resource movement across fisher behaviours with our without marine closures.
- The utility of standardized CPUE to index abundance (as indicated by CPUE-abundance linearity (β_S)) was better across all resource/fisher scenarios, with or without marine closures, except when resource movement was random and fisher behaviour was non-random and under marine closures.
- Non-linear statistical standardized CPUE-abundance relationships were obtained across resource/fisher scenarios.

Recommendations

- Statistical standardization analyses should be employed to improve CPUE-abundance relationships and reduce temporal biases in standardized indices either with or without marine closures.
- At least year (Y), vessel (V), month within year (M), grid-location (G) and interaction term grid-location x month (G × M) should be employed in standardizations to obtain greatest improvements towards linearity and least biased estimates.
- Exploratory analyses should be undertaken to determine the appropriate statistical distribution of the response. Distributions to consider are: Normal, log-normal, Poisson, negative binomial if there when there are no zeros catches. Other statistical distributions (e.g. Tweedie distributions) should be considered when there are zero catches.

- GLIM standardizations should be performed in all resource/fisher scenarios, under no marine closures. The terms to use are: Y, V, G, M and G×M.
- Standardizations should be performed when resource movement is non-random under marine closures using terms Y, V, G, M and G×M.
- When resource movement is random and fisher behaviour is non-random (under marine closures), fitting the Year term should be used.
- Standardizations could be employed (fitting all main effects and G × M) when both resource movement and fisher are random, under marine closures, although the degree of improvement to CPUE-abundance linearity may be minimal.
- Statistical power relationships that account for non-linearity could be incorporated in stock assessments for data-rich fisheries (i.e. for fisheries that are subject to Tier 1 stock assessments in the SESSF).

8.2 Introduction

Marine spatial closures implemented in Australian waters have aimed to (i) generate a comprehensive, adequate and representative system of marine protected areas, (ii) contribute to the long term ecological viability of marine and estuarine systems, (iii) maintain ecological processes and systems and (iv) protect Australia's biological diversity at all levels (Commonwealth of Australia, 2003). They may also may help reduce fishing pressure on stocks, allow juvenile fish to grow, protect spawning/breeding areas and protect marine habitats or particular resident species (e.g. Knuckey et al., 2009).

If marine closures aim to reduce fishing pressure in particular areas and are sufficiently large, they may affect fishery-dependent catch-per-unit-effort (CPUE) data. Also, such closures could affect how the population status of a species is perceived if CPUE is influential in single-species stock assessments. Commercial fisheries CPUE indices are used to index population abundance for many Australian fisheries. As such, it is important to determine to what extent marine closures affect CPUE indices in terms of their ability to proxy species abundance. CPUE may change following the introduction of closures if the areas closed were important fishing locations for a particular fish species, or if fishers alter where and when they fish in response to closures excluding them from preferred and/or 'high yield' locations.

Most fisheries stock assessment models rely on long term indices of absolute or relative abundance. With notable exceptions (e.g. north-east and north-west Atlantic groundfishes; e.g. Azarovitz, 1981; NOAA, 2013), indices based on fishery independent surveys are either seldom available or available for only a small number of years. Instead, commercial CPUE indices are often used to index abundance, although this usually involves the assumption of a linear relationship between CPUE and abundance. However, hyperstability, where CPUE remains approximately the same as biomass declines, is known to occur for some species, and other non-linear relationships would also invalidate the basis of the use of CPUE upon which many current analyses depend on. Alternative recommendations exist, (i.e., observation error models to estimate non-linearity; Harley et al., 2001), but are rarely implemented. Moreover, commercial CPUE indices are typically used to index abundance for data-poor fisheries and when no formal stock assessment models are used to assess stock status. These indices are generally computed from annual means (nominal CPUE) or from CPUE standardization analyses that aim to remove factors unrelated to changes in abundance (e.g. vessel, fishing time, fishing location).

Standardization of CPUE data, aimed to provide less biased estimates of relative annual abundance indices, traditionally employs well known statistical techniques such as general linear models (GLMs), generalized linear models (GLIMs) and generalized linear mixed models (GLMMs) (see review by Maunder and Punt, 2004). The procedure involves fitting variables to remove or reduce the effects of factors that may otherwise distort changes in abundance. Factors often incorporated in standardizations include season, year, area, gear type, vessel characteristics (e.g. efficiency, power) and/or harvest depth. Despite the vast array of CPUE standardization- methods in fisheries literature, few studies have been conducted to examine factors that may improve CPUE-abundance linearity, and to ascertain the extent to which fisher and fish movement behaviours influence the interpretation of CPUE indices.

Although developments on theoretical simulation modelling have helped to evaluate the use of CPUE as a proxy of abundance, most have been limited to fisher behaviours involving perfect information (Goodyear, 2003, 2006), individual and information sharing (Gaertner and Dreyfus-Leon, 2004), while others have excluded any type of fisher behaviour (McDonald et al., 2001; Campbell, 2004; Thorson et al., 2012; Oro et al., 2015a). In addition, few studies have incorporated a range of fisher behaviours and resource movements to address this problem (Sporcic, 2007; Sporcic and Smith, 2009).

In this study, an agent-based simulation model is employed to determine the effect of marine closures on standardized CPUE, in the context of different fisher behaviour types and resource (fish) movement dynamics. Resource movement types are based on a generic platycephalid (i.e. tiger flathead) frequently occurring in the Southern and Eastern Scalefish and Shark Fishery (SESSF). Generated CPUE data are standardized using GLIMs, and these CPUE indices then related to true abundance to examine the effect that marine closures have on the CPUE-abundance relationship. Standardizations are performed across a range of resource movements, in an attempt to ascertain to what extent the resultant CPUE indices represent trends in resource abundance for selected fisher behaviour scenarios with and without marine closures. Results are discussed in terms of estimated bias of the proportionality parameter, relative errors of annual indices (temporal bias) and the degree of improvement of CPUE-abundance linearity following standardisation.

8.3 Methods

8.3.1 Simulation model – RESFIM

An integrated resource/fisher operating model (RESFIM – **RES**ource-**F**isher **I**ntegrated **M**odel) has been developed to examine how adequately CPUE indexes resource abundance (Sporcic, 2007; Sporcic and Smith, 2009). This agent-based model links resource movement and fisher behaviour, an approach that is lacking in most current individual-based models employed to examine CPUE-abundance relationships. RESFIM incorporates the main characteristics of a fishery process, such as spatial resource movement and recruitment (year 5), but excludes age or length data as CPUE indices are seldom stratified by these factors, at least within Australia. For each resource movement scenario, namely (a) random and (b) habitat attraction, RESFIM simulates daily resource dynamics for 10 years and behaviour of 10 fishers harvesting the resource. A Bayesian belief network (BBN) system is used to determine fisher's harvest locations. The decision on where to harvest is based on various factors, including information sharing and environmental conditions. The general model structure and links between main model components is shown in Figure 8.1. The model adopts a two dimensional (x-y) spatial coordinate structure comprising 2 × 15 grid system.

Resource movement scenarios

a) Random: This scenario assumes that a resource is randomly distributed in space, and does not exhibit any spatial density dependence. Movement into and out of a location is generated from a uniform U[0,1] distribution, which determines movement probabilities between locations.

b) Habitat attraction: This scenario assumes that resource movement to preferred locations is independent of resource density or biomass, and that it varies within and between seasons over a restricted number of nearest neighbour locations, hence allowing for seasonal migrations to occur. Biomass redistribution is equally likely at each time step and movement is based on a normally distributed random variable, with a mean μs centred at a preferred location. The dispersion parameter (standard deviation, σs) controls the variability of biomass across locations and hence the extent of clustering of an aggregation. Two forcing parameters, μ_s and σ_s, therefore describe the distinct seasonal cycles which persist inter-annually (Sporcic, 2007).

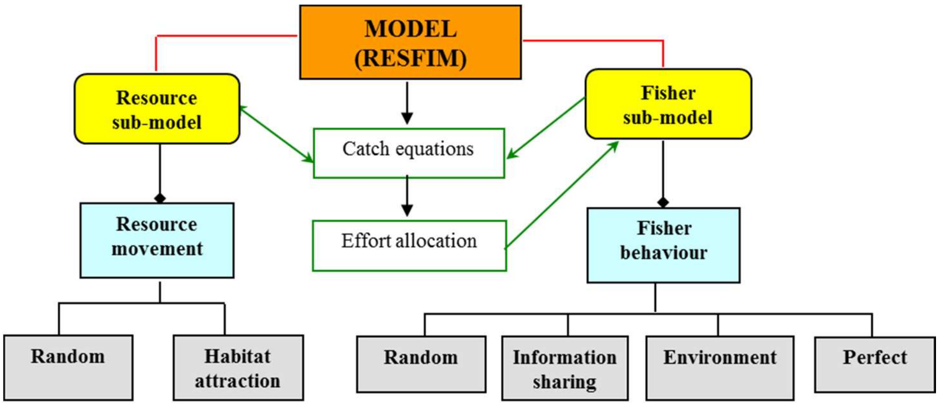


Figure 8.1: Schematic of general RESFIM model structure used in this study. Closures occur in the Resource sub-model.

Fisher behaviour scenarios

a) Random: The choices on where to harvest are made randomly and repeatedly over successive days, and fishers harvest independently of each other.

b) Information sharing: Information exchange between fishers was simulated using probability density functions (pdfs), estimated as the probability of a fisher moving from one location to another. A fisher's pdf is updated daily using historical CPUE and information shared by other fishers, using a link matrix which describes the degree of sharing (Sporcic, 2007). This matrix does not describe the causal mechanism for information exchange between fishers, but implicitly uses information gained from other fishers. The harvesting location of an individual fisher is chosen based on the fisher's updated pdf. The same amount of information sharing takes place for all ten fishers during each fishing trip.

c) Environmental conditions: Both CPUE and current environmental conditions (SST) determine where a fisher will harvest. SST was simulated based on an autoregressive integrated moving average model (STARIMA). Optimized parameters, based on a polynomial model that uses the previous year's CPUE and environmental conditions, were updated at the start of each year to account for changes in relative resource depletion. The model was used daily to determine harvest location corresponding to the fisher's highest predicted CPUE. If the daily SST at each location does not correspond to the maximum predicted CPUE, then a location corresponding to where the predicted CPUE is closest to the maximum is chosen for harvesting.

d) Perfect information: This scenario assumes that fishers have perfect information of the resource being exploited. Fishers move to a harvest location with the greatest available biomass. No updating is required under this hypothesis, since perfect knowledge of the spatial biomass is assumed.

Marine closures: Approximately 33% of the area simulated was closed to fishing. These contiguous locations corresponded to grids 12 through to 22. Also, marine closures were introduced at the commencement of the fishery.

8.4 Influence of marine closures on CPUE standardizations

The effect of marine closures on CPUE indices was investigated by simulating how adequately standardized indices reflect true abundance under two resource and four fisher behaviour scenarios with and without closures. Firstly, RESFIM was employed to generate daily CPUE data of a platycephalid similar to the Tiger Flathead (*Neoplatycephalus richardsoni*) occurring in the SESSF in the context of (i) no-marine closures (i.e. all areas open) and (ii) the introduction of marine closures (where simulated CPUE data are based on only the open portion of the stock) (e.g. see Table 8.1). In each instance, the generated series were standardized to examine the effect(s) of introducing marine closures using statistical standardization methods commonly employed in fisheries studies (including the SESSF (Sporcic, 2016), i.e., generalized linear models (GLMs; e.g. McCullagh and Nelder, 1989). How well resulting standardized CPUE indices reflect true abundance was then investigated by examining CPUE-abundance relationships and associated bias as well as relative error measures (Equations 2, 3). Results were examined in the context of the interactions between resource movement dynamics and fisher behaviour, as well as the role closures have on these two dynamics.

8.4.1 Simulations performed

Simulations were performed using RESFIM for several scenarios to assess how adequately standardized indices reflect true abundance in the context of marine closures, resource movement dynamics and fisher behaviour. Four fisher behaviour scenarios were simulated: random, information sharing, environmental conditions and perfect, in conjunction with two resource movement scenarios: random and habitat attraction. Simulations were performed across these eight resource/fisher scenarios with and without marine closures. Marine closures are introduced at the beginning of each simulation period across the eight resource-fisher scenario combinations. Simulated CPUE data were standardized to assess the adequacy of the relative indices to reflect true abundance.

Table 8.1: No marine closures. Estimated mean proportionality parameter using nominal-CPUE (θ_N), standardized-CPUE (θ_S) and corresponding coefficient of variation (CV) for selected resource/fisher scenarios and no marine closure (MC). Mean statistical bias (relative to θ_S) and percentage change between θ_N and θ_S are also provided. Median relative error (MedRE), median absolute relative error (MedARE) are also estimated based on nominal [N] and standardized [S] CPUE. Note: an overall mean of the median estimates across the simulation runs were estimated.

Resource	Fisher	Mean θ_N (CV)	Mean of $MedRE_N$	Mean of $MedARE_N$	Mean θ_S (CV)	Mean bias θ_S	Change mean θ_S from θ_N (%)	Mean of $MedRE_S$	Mean of $MedARE_S$
Random	Random	0.854 (0.10)	0.031	0.073	0.860 (0.10)	-0.140	0.51	0.030	0.074
	Share	0.755 (0.15)	0.064	0.097	0.872 (0.15)	-0.129	15.49	0.040	0.096
	Environment	1.019 (0.14)	0.035	0.049	1.075 (0.14)	0.075	5.46	0.031	0.046
	Perfect	0.566 (0.23)	0.004	0.058	0.805 (0.10)	-0.195	42.26	0.024	0.071
Habitat attraction	Random	1.105 (0.15)	0.018	0.087	1.086 (0.15)	0.086	-1.70	0.013	0.085
	Share	0.385 (0.16)	0.196	0.288	0.718 (0.12)	-0.282	87.02	0.056	0.186
	Environment	1.211 (0.15)	0.013	0.113	1.084 (0.15)	0.084	-10.24	0.014	0.091
	Perfect	0.573 (0.15)	0.170	0.225	0.876 (0.08)	-0.124	52.96	0.070	0.112

8.4.2 Statistical standardizations

GLIM-derived standardized CPUE indices were obtained from RESFIM-generated CPUE data on vessel catches (kg/day) spanning 10 years (McCullagh and Nelder, 1989; Chambers and Hastie, 1992). All GLIM analyses were performed using R® statistical software.

Exploratory data analyses were initially employed to determine the most appropriate distribution of the response to be used in GLIMs by fitting log-normal, negative binomial, Poisson and gamma distributions to CPUE for selected resource/fisher scenarios (and simulation runs). The mean-variance relationship of the response variable (CPUE) was compared to known theoretical relationships of the fitted distribution. Since the negative binomial distribution was found to be the most appropriate (using model diagnostics and minimum AIC), all subsequent CPUE standardization analyses were performed using this distribution and a log-link function. Hence, estimated standard errors and corresponding confidence intervals should better account for the observed variability in the response.

In addition, terms fitted in each model were also assessed to determine the best fitting model and whether the additional covariates significantly added to the overall model fit. The model chosen was the one which corresponded to the best fit according to several criteria (corrected Akaike Information Criterion (AICc; Burnham and Anderson, 2002) and model diagnostics.

Variables examined for their contribution to model fit in each GLIM analysis were overall mean (ψ_1), year (Y), month within year (M), grid-location (G), vessel (V), interaction terms and an error term (ϵ) based on a negative binominal distribution (Equation 1). Vessel consists of vessel number, fisher and skipper. Sea surface temperature (SST) was also employed, but was omitted from subsequent analyses since it accounted for less seasonal variation in CPUE than M.

$$CPUE_{v,y,c,t} = \psi_1 + Y + V + G + M + G \times M + V \times G + V \times M + \varepsilon \quad (1)$$

Homogeneity of variance was tested, and variables transformed if variance in-homogeneity was statistically significant ($P<0.05$; Zar, 1984). Models consisting of five fixed effects were employed for each resource/fisher scenario (and runs). Sub-models were then tested for significance using the likelihood ratio test at the 5% significance level based on the AIC statistic (Akaike, 1974; Burnham and Anderson, 1998, 2002). Variables were tested for significance in each sub-model using Type 3 sums of squares, which accounts for other fitted terms. The presence of over-dispersion was tested using the deviance statistic and model degrees of freedom.

8.4.3 Performance criteria

Performance measures were calculated using annual standardized-CPUE indices relative to the overall mean of the series. Other metrics were also evaluated, but they were insensitive to the overall results and therefore are not reported further.

Error measures employed

Relative errors used across each scenario comprised the median relative error (MedRE, Equation 3) and the median absolute relative error (MedARE, Equation 4) as defined below. These metrics provide information on the bias of the standardized index and true abundance, and determined whether (i) there was an overall positive or negative bias (MedRE) or whether there was a trend in annual standardized CPUE (MedARE).

Relative error (RE) at year *y* is defined as

$RE_y = (\text{Std-CPUE}_y - B_y)/B_y$, where B_y refers to biomass at year *y* and Std-$CPUE_y$ refers to standardized CPUE in year *y*. (2)

Median relative error is defined as:

$$MedRE = \text{median}\ (RE_{y=1} \ldots RE_{max\ y}) \quad (3)$$

Median absolute relative error is defined as:

$$MedARE = \text{median}\ (ARE_{y=1} \ldots ARE_{max\ y}), \text{ where } ARE_y = |RE_y|. \quad (4)$$

These estimated errors (Equation 3, 4) were averaged across each of the five simulation runs per scenario. Relative errors were also estimated for nominal CPUE by replacing Std-CPUE with nominal CPUE.

Power relationship

Standardized [S] annual CPUE indices were computed. Using a non-linear statistical estimation technique (NLIN in R), both annual standardized CPUE and abundance indices were used to estimate the proportionality (shape) parameter (β_S) for each resource/fisher scenario using the equation:

$$CPUE_y = \alpha B_y^{\beta_S} e^{\varepsilon_y} \quad (5)$$

where α is the scaling parameter, *B* the stock biomass during year *y* and ε_y an error term. The CPUE-abundance relationship was deemed to be hyperstable if $\beta_S<1$, proportional if $\beta_S=1$ or hyperdepleted if $\beta_S>1$. Approximate 95% confidence intervals were used to determine if the estimated shape parameter differed significantly from 1 ($P<0.05$). The mean of each estimated shape parameter was obtained across five simulation runs for each scenario, as was the percentage of times the true value of 1 was either within or outside the estimated 95% confidence intervals. A significant result was obtained if the true value was outside the confidence intervals on at least 95% of runs (i.e. across five simulation runs). The corresponding mean coefficient of variation of β_S was also estimated across simulation runs for each scenario, as was the mean statistical bias of β_S which was then tested if it differed significantly from zero ($P<0.05$). A significant negative mean bias indicates a hyperstable CPUE-abundance relationship ($\beta_S<1$), whereas a positive mean bias indicates a CPUE-abundance hyperdepleted relationship ($\beta_S>1$). If the estimated mean bias is zero, then the CPUE-abundance relationship is deemed to be linear. A hyperstable relationship occurs when CPUE remains high as abundance drops, while a hyperdepleted relationship occurs when CPUE drops faster than abundance declines.

All CPUE-abundance relationships obtained in this study assumed a zero x-y intercept (Equation 5). Since the principal objective was to estimate the exponent (β_S) of this curve, alternative non-linear curves incorporating non-zero x-y intercepts were not employed (*c.f.* Richards and Schnute, 1986). The percent change between nominal [N] β_N and standardized [S] β_S were also estimated. The coefficient of variation (CV) of mean β_S were also estimated, and the mean bias of β_S across the five simulations per resource/fisher scenario.

8.5 Results

8.5.1 Effect of spatial closures on CPUE standardization

No marine closures

Random resource movement: Mean β_N was 0.566 – 1.020 across all fisher behaviours in the absence of marine closures (Table 8.1, Table 8.3; Figure 8.2). The greatest hyperstable relationship occurred when fishers had perfect information, followed by fishers sharing information. Median relative errors and median absolute errors were <0.1 ($MedRE_N$, $MedARE_N$; Table 8.1; Figure 8.4A, Figure 8.5A). Moderate hyperstable and linear relationships were obtained across the four fisher behaviours, following CPUE standardizations (β_S 0.805-1.086) with improvements towards linearity of up to 42% (Table 8.1, Table 8.3). Corresponding median relative errors were ≤0.03 ($MedRE_S$ 0.02-0.03; Figure 8.4B) and ≤0.1 ($MedARE_S$ 0.05-0.10; Figure 8.5B).

Habitat attraction resource movement: Mean β_N was 0.385 – 1.211 across all fisher behaviours in the absence of marine closures (Table 8.1, Table 8.3). The greatest hyperstable relationship occurred when fishers shared information, followed by perfect information. Both median relative errors and median absolute errors were <0.29 ($MedRE_N$ 0.02-0.20; $MedARE_N$ 0.09-0.29; Table 8.1; Figure 8.4A, Figure 8.5A). Moderate hyperstable and linear relationships were obtained across the four fisher behaviours, following CPUE standardizations (β_S 0.805-1.086) with improvements towards linearity of up to 53% (Table 8.1, Table 8.3). Corresponding median relative errors were ≤0.07 ($MedRE_S$ 0.01-0.07; Figure 8.4B) and ≤0.11 ($MedARE_S$ 0.08-0.11; Figure 8.5B).

Summary

Hyperstable and weakly hyperdepleted relationships were obtained based on nominal and standardized-CPUE proportionality relationships across the resource/fisher scenarios (Table 8.1, Table 8.3). Temporal bias in estimated standardized abundance ($MedRE_S$ and $MedARE_S$) were similar compared to the bias in nominal CPUE for each resource/fisher scenario (Table 8.1; Figure 8.4, Figure 8.5).

Summary

Hyperstable and weakly hyperdepleted relationships were obtained based on nominal and standardized-CPUE proportionality relationships across the resource/fisher scenarios (Table 8.1, Table 8.3). Temporal bias in estimated standardized abundance ($MedRE_S$ and $MedARE_S$) were similar compared to the bias in nominal CPUE for each resource/fisher scenario (Table 8.1; Figure 8.4, Figure 8.5).

Table 8.2. Marine closures. Estimated mean proportionality parameter using nominal-CPUE (β_N), standardized-CPUE (β_S) and corresponding coefficient of variation (CV) for selected resource/fisher scenarios and marine closure (MC) scenario. Mean statistical bias (relative to β_S) and percentage change between β_N and β_S are also provided. Median relative error (MedRE), median absolute relative error (MedARE) are also estimated based on nominal [N] and standardized [S] CPUE. Note: an overall mean of the median estimates across the simulation runs were estimated.

Resource	Fisher	Mean β_N (CV)	Mean of $MedRE_N$	Mean of $MedARE_N$	Mean β_S (CV)	Mean bias β_S	Change mean β_S from β_N (%)	Mean of $MedRE_S$	Mean of $MedARE_S$
Random	Random	0.888 (0.26)	0.005	0.072	0.881 (0.26)	-0.119	-0.61	0.009	0.069
	Share	1.084 (0.13)	0.029	0.090	1.284 (0.14)	0.284	18.47	0.008	0.101
	Environment	1.390 (0.14)	-0.016	0.076	1.426 (0.14)	0.426	2.68	-0.018	0.078
	Perfect	0.837 (0.18)	0.076	0.134	0.863 (0.18)	-0.137	3.20	0.070	0.139
Habitat attraction	Random	1.481 (0.29)	-0.016	0.110	1.343 (0.26)	0.343	-9.05	-0.005	0.091
	Share	0.531 (0.14)	0.113	0.157	1.004 (0.11)	0.004	89.66	-0.011	0.068
	Environment	1.099 (0.17)	0.010	0.064	0.967 (0.15)	-0.033	-12.10	0.008	0.035
	Perfect	0.523 (0.23)	0.121	0.173	0.969 (0.10)	-0.031	88.12	0.034	0.066

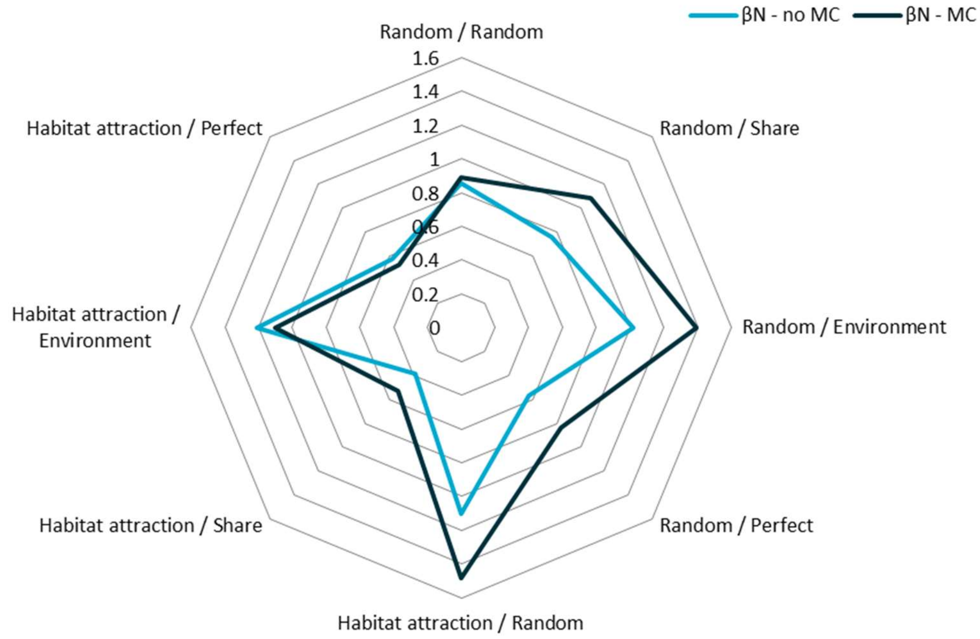


Figure 8.2: Radar representation of the mean proportionality parameter (β_N) based on nominal CPUE for each resource/fisher scenario with and without marine closures (MC).

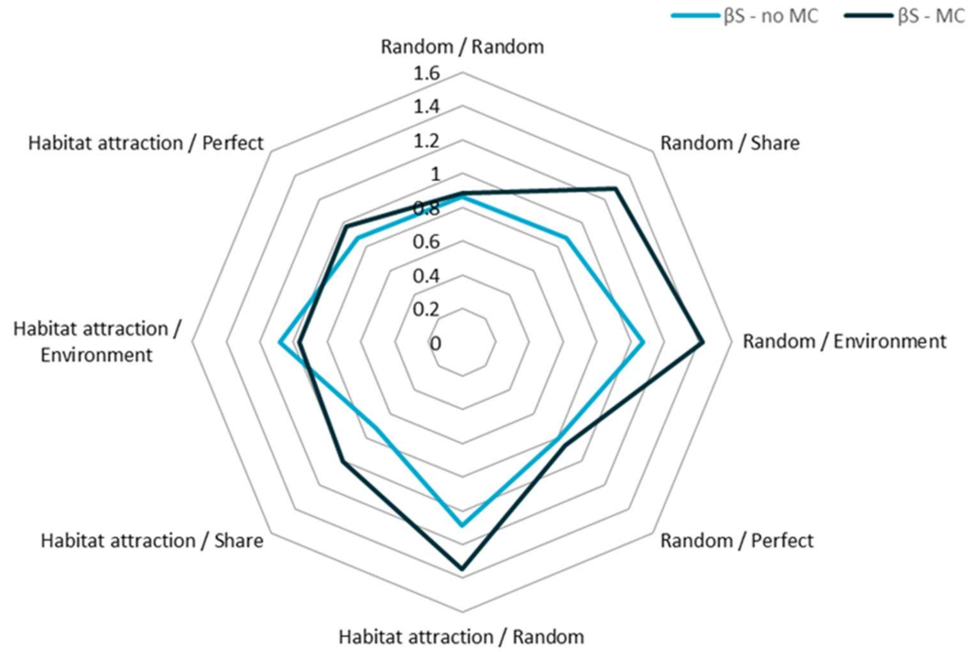


Figure 8.3: Radar representation of the mean proportionality parameter (β_S) based on standardized CPUE for each resource/fisher scenario with and without marine closures (MC).

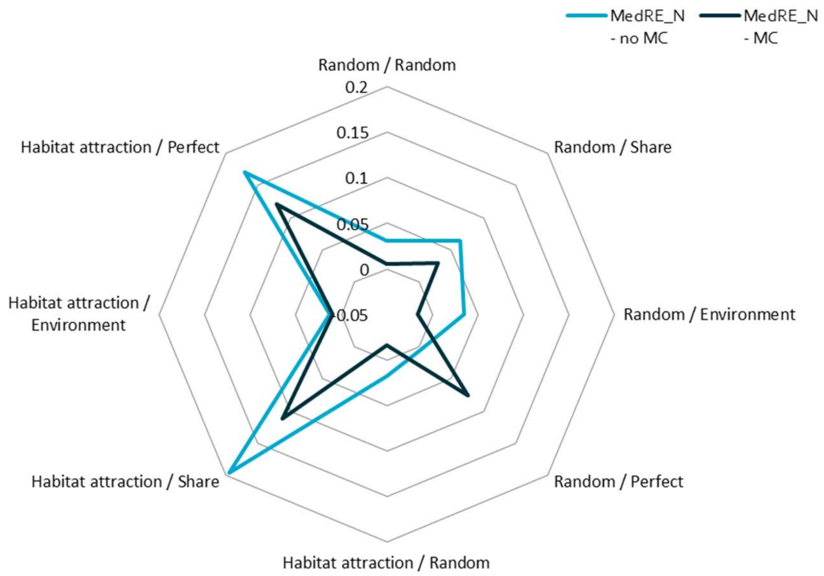
Table 8.3. Comparison of CPUE-abundance relationships across the resource/fisher scenarios and with or without marine closures. Proportionality relationship is based on nominal [N] and standardized [S] CPUE. Marine closure (MC), no marine closure (no MC). Shading: rose (hyperstable; β<1); green (linear; β=1); blue (hyperdepleted; β>1).

Resource/fisher scenario		Protected area	Proportionality relationship (β)					
Resource movement	**Fisher behaviour**		$\beta_N<1$	$\beta_N=1$	$\beta_N>1$	$\beta_S<1$	$\beta_S=1$	$\beta_S>1$
Random	Random	No MC	0.854			0.860		
	Share		0.755			0.872		
	Environment			1.019				1.075
	Perfect		0.566			0.805		
Habitat attraction	Random	No MC		1.105			1.086	
	Share		0.385			0.718		
	Environment				1.211		1.084	
	Perfect		0.573			0.876		
Random	Random	MC	0.888			0.881		
	Share			1.084				1.284
	Environment				1.390			1.426
	Perfect		0.837			0.863		
Habitat attraction	Random	MC			1.481			1.343
	Share		0.531				1.004	
	Environment			1.099			0.967	
	Perfect		0.523				0.969	

Marine closures

Random resource movement: Mean β_N was 0.837– 1.390 across all fisher behaviours in the presence of marine closures (Table 8.2; Figure 8.2). The greatest hyperstable relationship occurred when fishers had perfect information. Moderate hyperdepleted relationships also occurred when fishers used environmental conditions, which also resulted in an overall negative bias (-0.016) in nominal abundance indices ($MedRE_N$ -0.016-0.03, Table 8.2; Figure 8.4A). Also, overall absolute bias in nominal CPUE indices were ≤0.134 ($MedARE_N$ 0.072-0.134, Table 8.2; Figure 8.5A) across the fisher behaviour scenarios.

Moderate hyperstable and hyperdepleted relationships were obtained across the four fisher behaviours following CPUE standardizations (β_S 0.863-1.426) with small improvements towards linearity (Table 8.2). Estimated bias in standardized abundance indices were ≤0.07 ($MedRE_S$ -0.018-0.07; Figure 8.4B) or ≤0.139 ($MedARE_S$ 0.078-0.139; Figure 8.5B). Positive overall bias ($MedRE_S$) occurred for all fisher behaviour scenarios except when fishers used environmental conditions to determine where to fish.



A

B

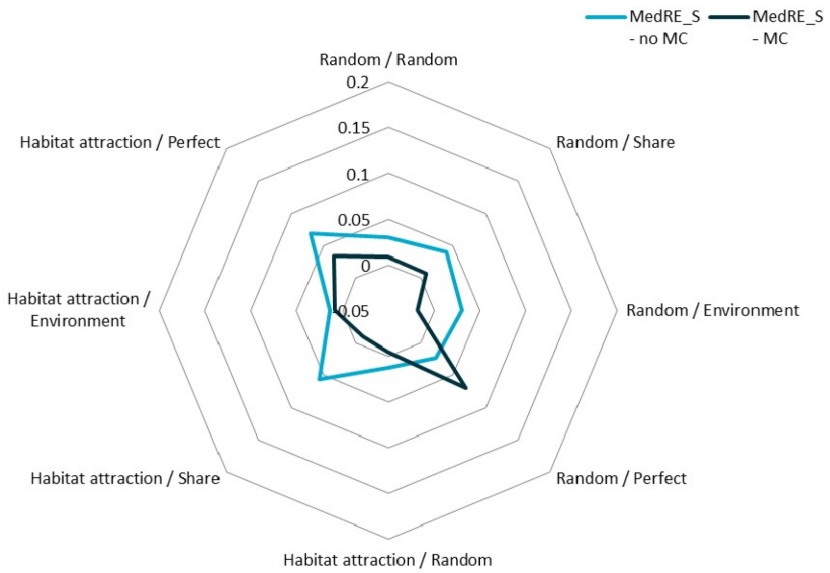
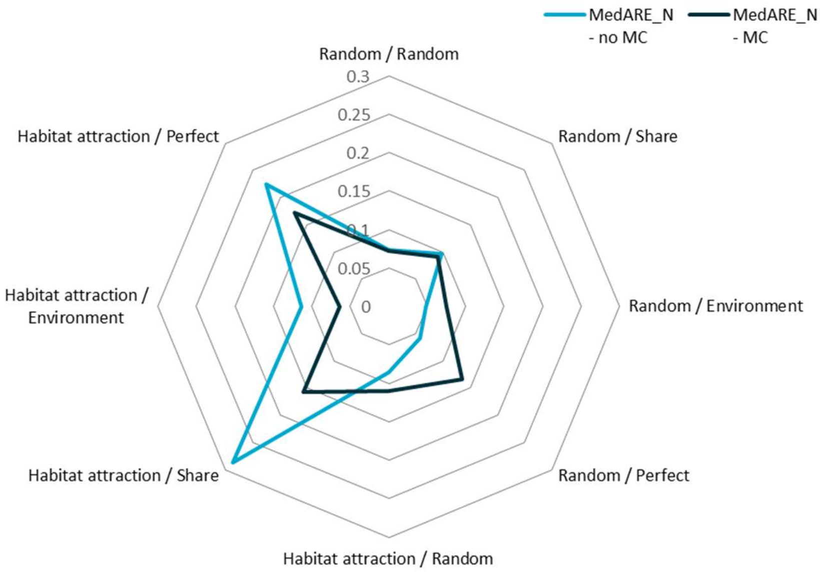


Figure 8.4 A: Radar representation of the median relative error (MedRE$_N$) based on nominal CPUE for each resource/fisher scenario with and without marine closures (MC). B: Radar representation of the median relative error (MedRE$_S$) based on standardized CPUE for each resource/fisher scenario with and without marine closures (MC).

A



B

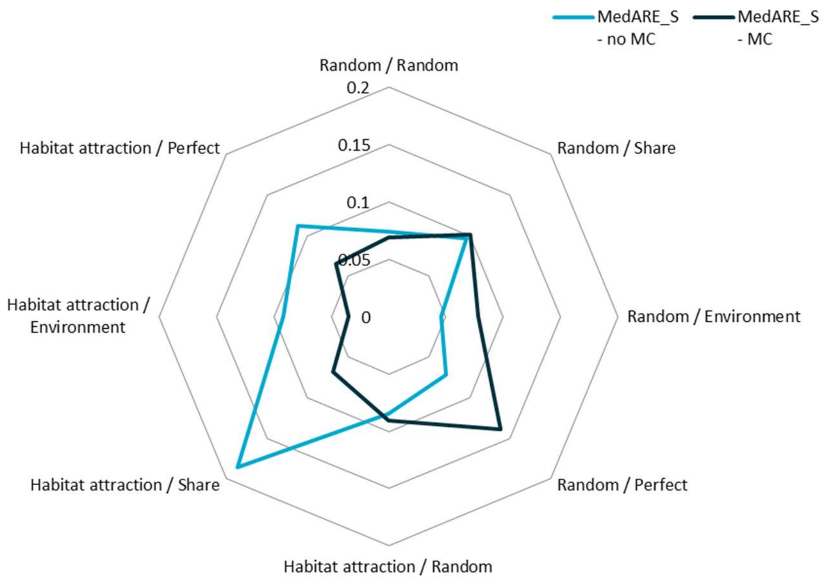


Figure 8.5 A: Radar representation of the median relative error ($MedARE_N$) based on nominal CPUE for each resource/fisher scenario with and without marine closures (MC). B: Radar representation of the median relative error ($MedARE_S$) based on standardized CPUE for each resource/fisher scenario with and without marine closures (MC).

Habitat attraction resource movement: Mean β_N was 0.523 – 1.481 across all fisher behaviours in the presence of marine closures (Table 8.2; Figure 8.2). The greatest hyperstable relationship occurred when fishers had perfect information, closely followed by sharing information. By contrast, the greatest hyperdepleted relationship occurred when fishers fished randomly (β_N 1.481). Both median relative errors and median absolute errors were <0.171 ($MedRE_N$ -0.016-0.121; $MedARE_N$ 0.06-0.171; Table 8.2, Figure 8.4A, Figure 8.5A).

Moderate hyperstable and linear relationships were obtained across the four fisher behaviours following CPUE standardizations (β_S 0.967-1.343) with improvements towards linearity of up to 90%, where fishers shared information (Table 8.2). Corresponding median relative errors were ≤0.034 ($MedRE_S$ -0.011-0.034; Figure 8.4B) and ≤0.091 ($MedARE_S$ 0.035-0.091, Figure 8.5B).

Summary

Very little improvement towards linearity occurred when nominal CPUE-abundance relationships (β_N) were hyperstable under random resource movement. By contrast, such relationships considerably improved under habitat attraction resource movement.

Estimated temporal bias in standardized indices ($MedARE_S$) was similar to bias in nominal CPUE indices for the same resource/fisher scenario in the presence of marine closures (Table 8.2; Figure 8.5). Overall, positive biases were apparent in standardized indices ($MedRE_S$, $MedARE_S$; Figure 8.4B, Figure 8.5B).

Overall summary – with and without marine closures

Random resource movement: All CPUE-abundance relationships (β_N) were closer to linearity with marine closures than without marine closures (e.g. β_N 0.888 vs 0.854 – random; β_N 1.084 vs 0.755 - information sharing; β_N 0.837 vs 0.566 – perfect information; Table 8.1, Table 8.2; Figure 8.2). Standardizations improved CPUE-abundance relationships across all fisher behaviours except when fishers used environmental conditions to determine where to fish (Figure 8.8, Figure 8.9). There was a negative bias of β_S (corresponding to hyperstable relationship) when fishers randomly fished or had perfect information, with or without marine closures (Table 8.1, Table 8.2). By contrast, approximate linear relationships occurred when fishers randomly fished or used environmental conditions under no marine closures (β_S 1.086 and 1.075; Figure 8.2) and hyper-depleted relationships when fishers shared information or used environmental conditions in the presence of marine closures (β_S 1.284 and 1.426; Table 8.2; Figure 8.3).

Relative errors ($MedRE_S$) of standardized CPUE were mostly less under closures compared to no closures, but insignificantly (Table 8.1, Table 8.2; Figure 8.4B). By contrast, $MedARE_S$ were greater under closures compared to no closures (Table 8.1, Table 8.2; Figure 8.5B). Also, relative errors (i.e. temporal biases) of nominal CPUE were mostly smaller under no marine closures compared to those under marine closures and under the same fisher behaviour scenarios (*cf*. Figure 8.4A, Figure 8.5A).

Habitat attraction resource movement: CPUE-abundance relationships (β_N) were closer to linearity with a closure than without marine closures (e.g.; β_N 0.531 vs 0.385 - information sharing; β_N 1.099 vs 1.211 – environment), except for random (β_N 1.481 vs 1.105) and perfect (β_N 0.523 vs 0.573) fisher behaviours (Table 8.1; Figure 8.2). Standardized CPUE-abundance relationship mostly improved this relationship across all fisher behaviours with or without closures (Table 8.2; Figure 8.3).

Approximate linear relationships were estimated when fishers shared information, used environmental conditions or had perfect information under marine closures (β_S 1.004, 0.967, 0.969; Table 8.2; Figure 8.9). By contrast, moderate hyperdepleted relationship was estimated for random fisher behaviour under the same marine closures (β_S 1.343; Table 8.2).

Approximate linear relationships were obtained when fishers randomly fished or used environmental conditions under no marine closures (β_S 1.086 and 1.084; Table 8.1). By contrast, moderately hyperstable relationships were estimated under information sharing and perfect fisher behaviours under no marine closures (β_S 0.718 and 0.876; Table 8.1).

Relative errors ($MedRE_S$) of standardized CPUE were generally similar and mostly positive with or without marine closures (Table 8.1, Table 8.2; Figure 8.4B, Figure 8.5B). Also, these relative errors (i.e. temporal biases) of nominal CPUE were less than corresponding relative errors under the same fisher behaviour scenarios (*cf*. Figure 8.4A, Figure 8.5A).

Relative errors ($MedRE_S$, $MedARE_S$) of standardized CPUE were mostly less than corresponding errors under closures compared to no closures, but insignificantly (Table 8.1, Table 8.2; Figure 8.4B, Figure 8.5B).

Overall summary of Results (across resource/fisher scenarios):

Generally, there were fewer hyperstable relationships (β_S) across the resource/fisher scenarios with than without closures (25% hyperstable - MC vs 63% hyperstable – no MC; Table 8.3). By contrast, there were more linear or hyperdepleted CPUE-abundance relationships (β_S) across the resource/fisher scenarios with than without closures (38% linear - MC vs 25% linear – no MC; 38% hyperdepleted - MC vs 13% hyperdepleted – no MC; Table 8.3). All these relationships occurred under habitat attraction resource movement. If a proportion of the resource is closed to fishing, this will lead to fishers actively avoiding areas that could be fished. Standardization analyses reduced the bias in β_S.

Linear or hyperdepleted CPUE-abundance relationships (β_N or β_S) were obtained under the random/environment scenario, without closures. After closures, these relationships were more hyperdepleted (β_S >1), which suggests that fishers were unable to harvest the stock that would otherwise be available to fishing. Also, the use of historical environmental conditions did not lead to higher catches. Therefore, the degree of improvement following standardization analyses is reduced under hyperdepletion, when fishers fish randomly or fish are more dispersed under marine closures (Table 8.3).

Hyperstable relationships resulted when fishers shared or had perfect information under the habitat resource movement scenario with no closures (β_S <1, Table 8.3). By contrast, these relationships were linear under the same resource/fisher scenario with closures (β_S =1, Table 8.3). This suggests that closing some of the stock available to fishing leads to a less pessimistic view of resource abundance.

Overall, nominal CPUE-abundance relationships across all resource/fisher scenarios with and without closures were mostly hyperstable (Table 8.1 - Table 8.3) and so were poor as an indicator of abundance relative to standardized CPUE. Also, the bias in standardized CPUE-abundance relationships (β_S) were greater across each of the resource/fisher scenarios with than without closures.

No trends in estimated bias of annual standardized CPUE were obtained in this study with or without marine closures (Figure 8.6, Figure 8.7).

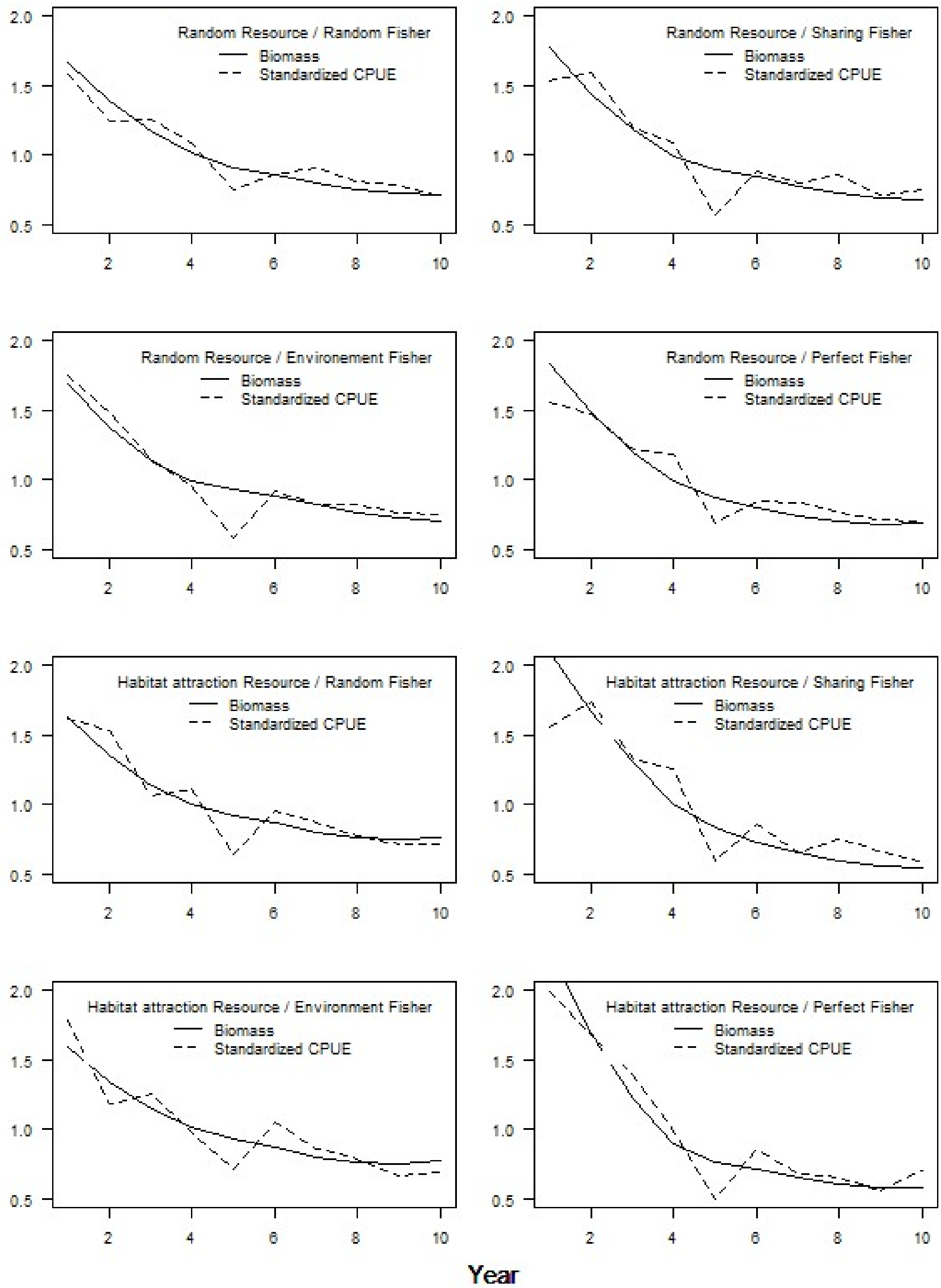


Figure 8.6: No marine closures (one simulation). Top (row 1, 2): Relative annual biomass and standardized-CPUE for Random resource movement across four fisher behaviours (Random, Sharing, Environment and Perfect). Bottom (row 3, 4): Relative annual biomass and standardized-CPUE for Habitat attraction resource movement across four fisher behaviours (Random, Sharing, Environment and Perfect).

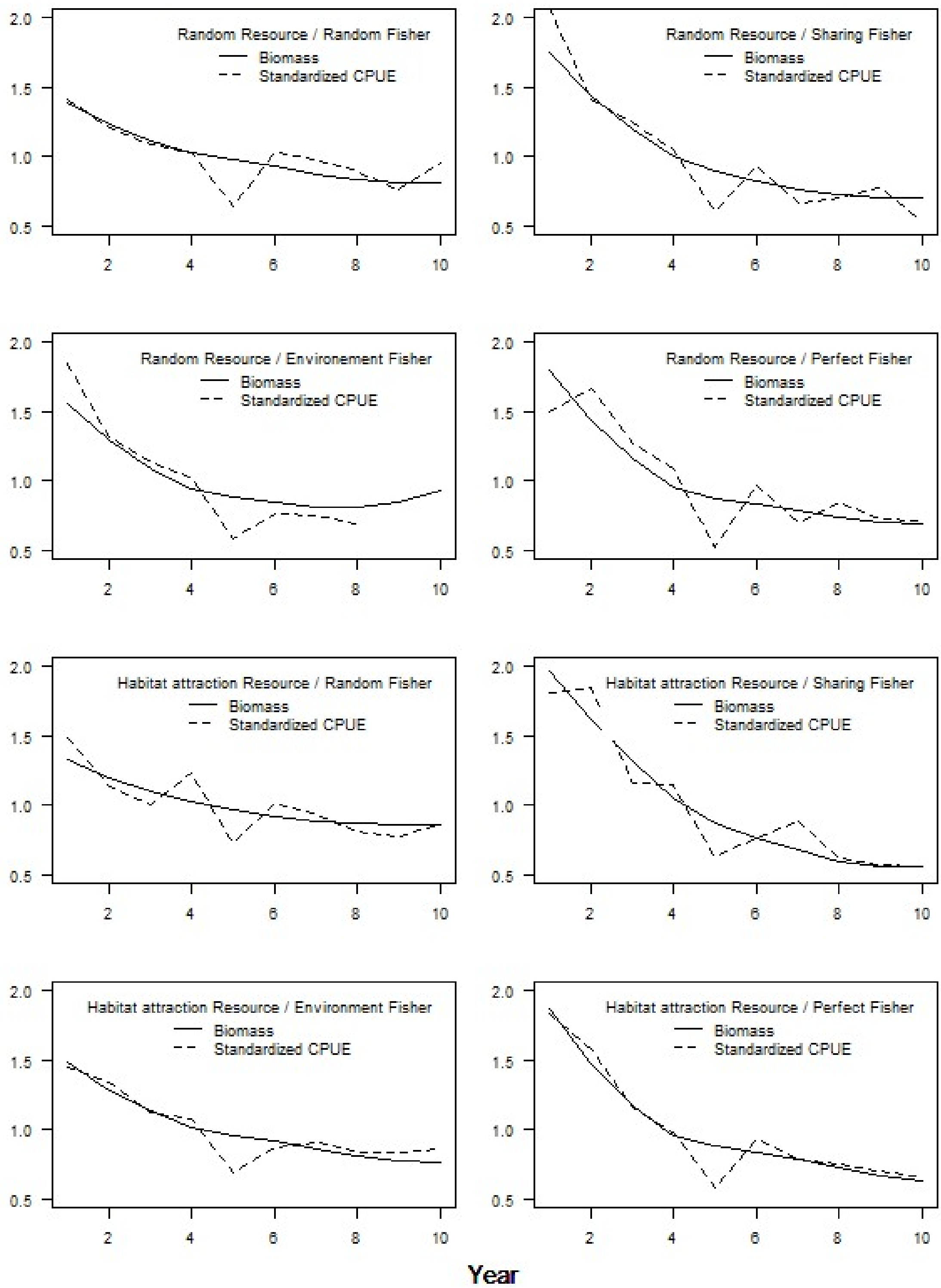


Figure 8.7: Marine Closures (one simulation). Top (row 1, 2): Relative annual biomass and standardized-CPUE for Random resource movement across four fisher behaviours (Random, Sharing, Environment and Perfect). Bottom (row 3, 4): Relative annual biomass and standardized-CPUE for Habitat attraction resource movement across four fisher behaviours (Random, Sharing, Environment and Perfect).

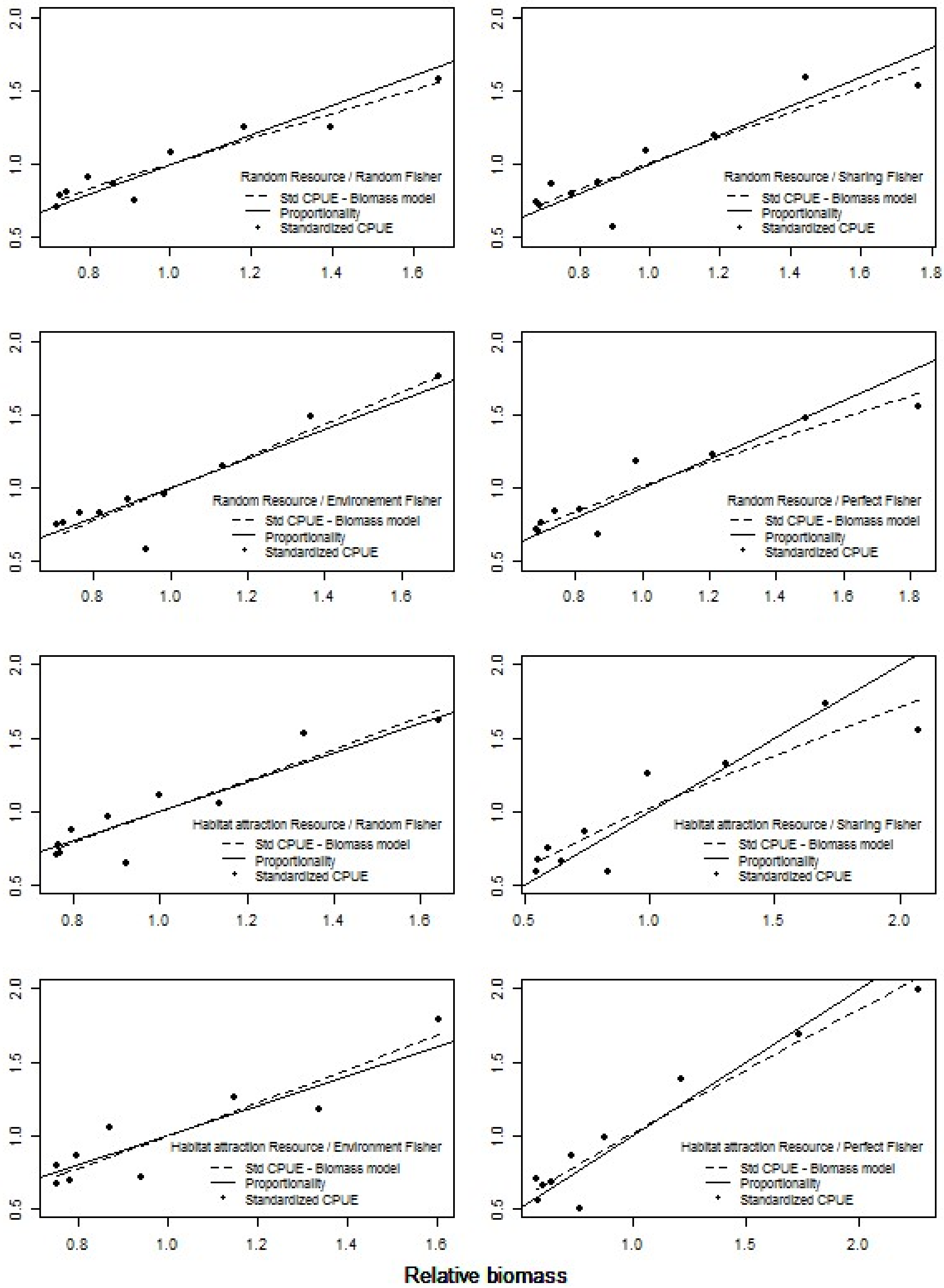


Figure 8.8: No marine closures (one simulation). CPUE-abundance relationships across resource/fisher scenarios. Refer to Table 8.3 for estimated proportionality parameter (β_S) across each resource/fisher scenario.

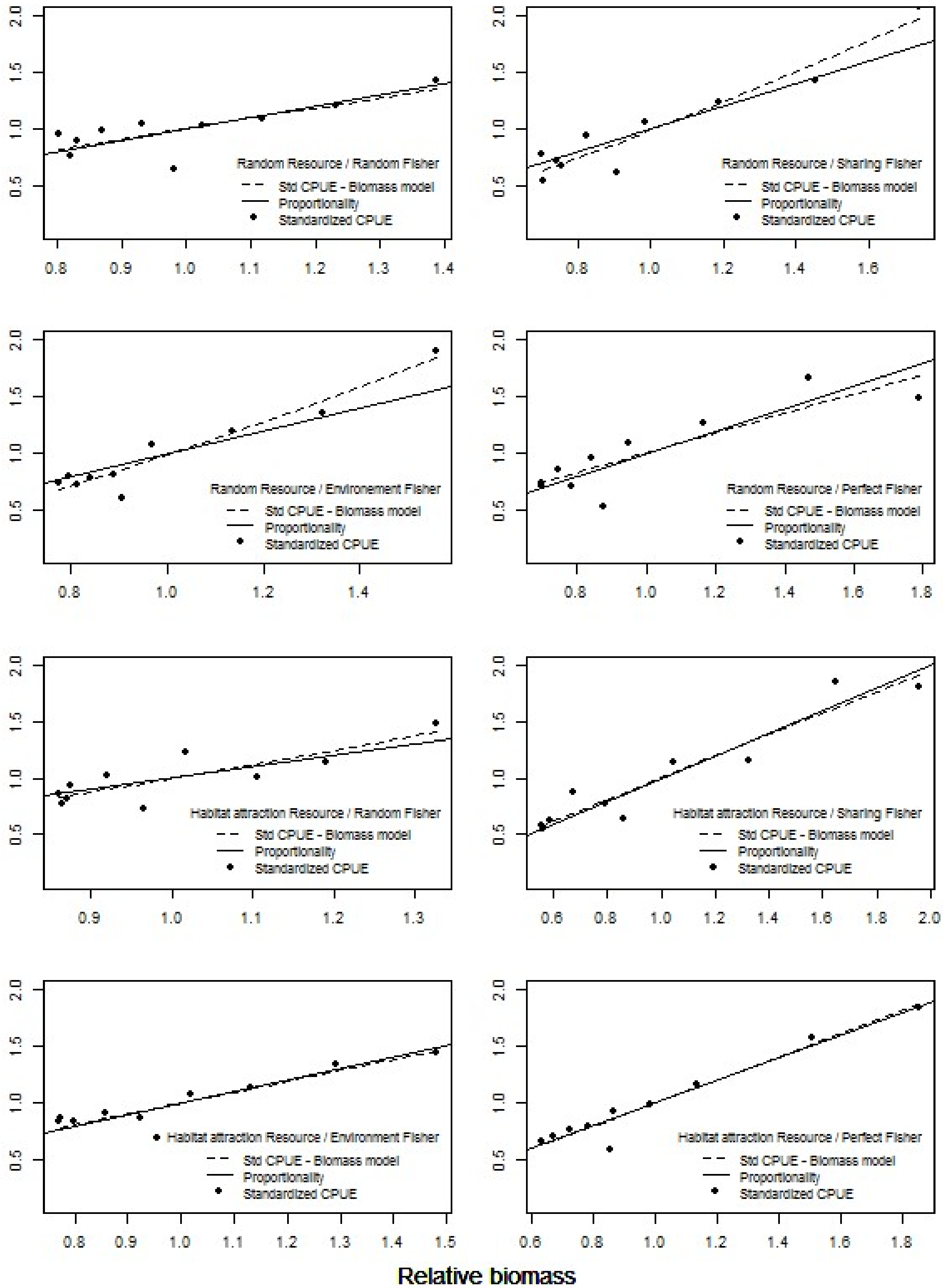


Figure 8.9: Marine closures (one simulation). CPUE-abundance relationships across resource/fisher scenarios. Refer to Table 8.3 for estimated proportionality parameter (β_S) across each resource/fisher scenario.

8.6 Discussion

The main aim of this study was to determine how well standardized CPUE indices represent resource abundance for selected resource/fisher scenarios, with and without marine closures, by examining CPUE-abundance relationships and relative errors based on data generated using an individual based simulation model (RESFIM; Sporcic, 2007). Results support the use of CPUE indices to index resource abundance following the application of GLIMs as a tool for standardizing CPUE. Whether CPUE adequately reflects abundance largely depends on the interaction between selected resource movement and fisher behaviour types and the implementation of closed areas to fishing.

This discussion focuses on CPUE-abundance relationships, relative errors, initial exploratory analyses undertaken and chosen standardization techniques, observations and variables used and diagnostics tests. Finally, a detailed overview is provided on which techniques and terms should be employed to improve linearity of CPUE-abundance relationships, as well as limitations of the statistical model. Comparisons made through this discussion with similar work are greatly limited by the fact that there are only three available studies known to the author that consider (i) CPUE-abundance using GLIM-based indices without marine closures (Hanchet et al., 2005) and (ii) standardized CPUE and relative errors with and without marine closures (Ono et al., 2015a,b).

8.6.1 Estimation of CPUE-abundance relationship and relative errors

CPUE-abundance comparison – with and without marine closures

Strict linearity between CPUE and abundance is desired for CPUE to adequately index abundance. Linear standardized CPUE-abundance relationships were achieved in approximately 30% of resource-fisher scenarios and mostly when resource movement was non-random. All other scenarios were non-linear., i.e. mostly hyperstable or hyperdepleted, which indicate biased CPUE indices.

Generally, there were fewer hyperstable relationships (β_S <1) across the resource/fisher scenarios with closures than without closures. By contrast, there were more linear or hyperdepleted CPUE-abundance relationships (β_S>1) across the resource/fisher scenarios with than without closures. This can be explained by the fact that if a proportion of the area is closed to fishing, fishers effectively change their behaviour and avoid areas which otherwise would be have been fished if these areas were not protected. However, the type of relationship (hyperstable or hyperdepleted) and degree of bias in β_S is dependent on the interaction between the resource, fisher dynamics and closed areas.

There were more (approximate) linear standardized CPUE-abundance relationships (β_S) under habitat attraction resource movement compared with random resource movement when part of the resource was closed to fishing. The presence of mostly linear or hyperdepleted relationships (β_S) under habitat attraction resource movement and spatial closures is directly related to the underlying resource movement dynamics and fisher behaviour. Standardizations improved CPUE-abundance linearity relationships in all scenarios under this resource movement, with greatest improvements over nominal CPUE when fishers shared or had perfect information with or without marine closures. This suggests that standardizations can effectively index abundance using covariates commonly available for standardization analyses under the habitat attraction resource movement scenario. By contrast, the effectiveness of standardizations to proxy abundance under random resource movement is less pronounced as CPUE-abundance relationships were mostly hyperstable or hyperdepleted across fisher behaviours.

A greater number of hyperdepleted standardized CPUE-abundance relationships occurred when fishers shared information or used environmental cues in the presence of marine closures compared to no marine closures. This suggests that standardizations are least effective in indexing abundance under random resource movement and with marine closures. Therefore, the utility of standardized indices to proxy abundance should be treated with caution if fisher behaviour is not random or perfect under random resource movement.

Summary

The existence of hyperstability or hyperdepletion (with or without marine closures) may bias stock assessments, as CPUE would deviate from indexing abundance in a linear fashion. Also, the use of CPUE as an index of abundance under hyperstable or hyperdepleted relationships will lead to either overly optimistic or pessimistic views of population levels respectively. As such, statistical power relationships that account for non-linearity could be incorporated in stock assessments for data-rich fisheries (i.e. for fisheries that are subject to Tier 1 stock assessments in the SESSF).

Where standardized-CPUE is used to proxy abundance (i.e. for data-poor fisheries), standardizations should be performed in all resource/fisher scenarios, under no marine closures, as small to large improvements to CPUE-abundance linearity were obtained.

In addition, standardizations should be performed when resource movement is non-random under marine closures. However, when resource movement is random and fisher behaviour is non-random (under marine closures), the utility of standardized CPUE as an abundance index was minimal as proportionality relationships led to greater hyperdepletion, given that fishers were less effective at fishing 'high yield' locations. In such instances, fitting the Year term led to better CPUE-abundance relationships. Also, standardizations could be employed (fitting all main effects and G×M term) when both resource movement and fisher were random, under marine closures, although the degree of improvement was minimal.

Relative errors

Overall, the relative error in this study was larger under no marine closures compared to marine closures using either nominal or standardized CPUE, across the selected resource/fisher scenarios. Part of this finding (i.e. larger relative errors) is also supported by another simulation study based on nominal CPUE without marine closures (Ono et al., 2015b).

Also, this error (temporal bias) was mostly positive ($MedRE_N$; $MedRE_S$) with or without marine closures, suggesting that estimated annual nominal or standardised CPUE abundance indices over-estimated the true underlying abundance. Temporal bias of nominal CPUE was mostly larger when resource movement was not random (i.e. under habitat attraction) and without marine closures. This temporal bias was further reduced using standardized CPUE ($MedRE_S$; $MedARE_S$) with or without marine closures. However, further work is needed to assess whether this can also be concluded under different resource and fisher behaviour dynamics.

No trends in estimated bias of annual standardized CPUE were obtained in this study (with or without marine closures), in contrast to estimated bias trends obtained in Ono et al. (2015a). Possible differences between these studies may be attributed to different resource/fisher dynamics, the number of years simulated (10 years this study *vs* 30 years by Ono et al., 2015a) and the proportion of area closed to fishing (33% this study). Also, while not considered in this study, adopting different fisher behaviours to include the use of acoustics and/or groups of fishers using different criteria to determine where to fish may also influence how well standardized-CPUE can index abundance. These behaviours have been previously examined in a similar study (Sporcic, 2007; Sporcic and Smith, 2009).

8.6.2 Statistical considerations

Selection of response distribution

A series of exploratory analyses performed to determine the appropriate distribution of the response, mean-variance CPUE plots (grouped by year) suggested that either the gamma or negative binomial distributions were appropriate, i.e. variance proportional to the square of the mean response. By contrast, standardized indices assuming a log-normal or Poisson distribution differed significantly from those based on the former two distributions using the same model. Fitted models using either the gamma or negative binomial distributions yielded similar point estimates, including standardized CPUE indices and CPUE-abundance relationships. However, for each fitted model in each scenario, the minimum AIC statistic resulted from assuming a negative binomial distribution of the response, suggesting that this be used for all subsequent GLIM analyses in this study. The decision to use the negative binomial distribution mirrors the fact that this distribution is accepted in statistical and fisheries literature (e.g. McCullagh and Nelder, 1989; Punt et al., 2000; Maunder and Punt, 2004; Ortiz et al., 2012).

Terms and interactions that enhanced linearity

Overall results showed that year (Y), month (M), grid-location (G), vessel (V) and G x M interaction terms were important terms in GLIM standardizations across each resource/fisher scenarios. These findings are consistent with terms most commonly used in various Australian fisheries (Klaer, 2004; Sporcic, 2016), but differ from those of Hanchet et al., (2005) who reported V as the most important term for New Zealand's southern blue whiting fishery. Although, it is noted that this result could be obtained in this study if there was greater variability in catch rates among simulated fishers.

Diagnostic tests

The AIC statistic (Akaike, 1974) was chosen to determine the best model fit instead of the deviance statistic. Unlike the deviance statistic, the AIC uses a penalty term that adds the number of fitted parameters (+2*k*), thereby penalizing larger models. In this study, standardized CPUE indices resulting from best fitting models using this statistic did not necessarily lead to β_S closest to 1. Instead, greater non-linear relationships occurred mostly after using either minimum AIC across scenarios.

In many instances, fitting too few or too many terms when using either statistic did not lead to β_S closest to 1, and particularly in the case of multiple interactions which resulted in increased hyperstability and/or hyperdepletion relative to smaller nested models (exploratory analyses), which was also found in Sporcic (2007). In addition, the inclusion of higher order interactions did not significantly improve CPUE-abundance linearity, suggesting that such terms should be fitted only when there is reason to believe they may influence CPUE, and/or if there are very few or no missing observations across term levels.

8.6.3 Recommendations on CPUE standardization analyses

The GLIM-based CPUE standardizations performed in this study using RESFIM data have demonstrated that key decisions on specific factors should be considered for such analyses. Factors include data availability, choice of technique and/or response distribution, variable and model selection, diagnostic tests and residual plots, all of which will ultimately influence CPUE-abundance linearity and therefore interpretations of biomass trends.

Initially, it is recommended that initial exploratory analyses be undertaken to identify the likely response distribution(s) using mean-variance CPUE plots. If more than one distribution is identified, e.g. gamma or negative binomial, then the smallest AIC statistic could be used to select an appropriate distribution from various fitted models. Mis-specifying the distribution, e.g. by wrongly assuming the mean-variance relationship, may have a substantial effect on estimated CPUE indices, relative errors and ultimately on CPUE-abundance linearity.

Fitted terms which account for small, but significant changes to CPUE should be included when they are believed to influence CPUE, and are well estimated. However, caution is required when fitting higher order terms, particularly if they are not well represented by the data as they may lead to increasing hyperstability or hyperdepletion (Sporcic, 2007). The same applies when using minimum AIC to determine the best fitted model, which may not result in linearity if there are too many or too few fitted terms, particularly if they correspond to higher order interactions.

In summary, different terms should be fitted depending on specific aspects of resource dynamics and fisher behaviour. Also, several competing models should be tested in the absence of true abundance data, and standardized indices examined for each model to decide which CPUE indices to use as proxies of abundance. In the end, the chosen model should be closely linked to knowledge of the fishery and diagnostic tests, and may not necessarily correspond to the best fitted model based on minimum AIC.

8.6.4 Limitations and future work

This study did not compare the performance of different statistical methods (e.g. generalized additive models; delta models; zero inflation models; generalized linear mixed models) used to standardize CPUE, in the context of marine closures. Instead, it assessed different distributions within the generalized linear modelling framework to determine the appropriate error distribution for use across each simulated scenario. While the comparison of different statistical techniques was previously investigated in the context of marine closures (Ono et al., 2015a), it was not examined in the context of fisher behaviour, as in this study. The existence of marine closures leads to changes in fisher behaviour, as areas that were once fished are no longer available to fishing and therefore new locations are subsequently fished. This in turn leads to spatial heterogeneity in the analyses of CPUE data. Further work is therefore required to assess the effect of different statistical standardization techniques in the presence of marine closures across different resource/fisher behaviours.

Also, this study introduced marine closures at the commencement of the fishery. Increasingly, closures have been introduced in fisheries years after its inception (e.g. Australia's SESSF; Sporcic, 2016), which correspond to periods of pre-closure and post-closures in the CPUE time series. While such pre-closure and post-closure periods has been examined in CPUE analyses in the study by Ono et al. (2015a), it was not examined in the context of fisher behaviour, as in this study. Further work is therefore required to assess whether there are systematic and/or increasing biases in standardized CPUE indices in each of the pre- and post-closure periods in the context of resource/fisher scenarios.

Marine closures have been established in core areas of Australian fisheries and used to manage fisheries e.g. deep-water closures in the SESSF. In the absence of long-term abundance indices derived from fishery independent surveys or other methods that estimate population status (e.g. integrated stock assessments, genetic and acoustic techniques), time series of standardized CPUE may be the only available indicator of stock status for many quota and non-quota SESSF fish species. How well these perform is dependent on the interaction of resource/fisher behaviours and on the information that can be measured and subsequently used in standardization analyses. More so, marine closures can influence estimated biases in standardized CPUE and also the proportionality between CPUE and abundance, and should be accounted for when these indices are used in stock assessments or used as a proxy for population abundance.

8.6.5 Appendix summary

This study showed that different factors associated with resource movement and fisher behaviour affected CPUE-abundance linearity, and that linearity significantly improved following GLIM standardizations. Overall, improvements towards linearity were greatest under information sharing and perfect fisher behaviour scenarios with or without marine closures. This confirmed that unstandardized CPUE indices were severely biased when fisher behaviour is non-random, with the degree of bias largely depending on the ability of fishers to successfully harvest a mobile resource. Results also suggested that the negative binomial distribution should be employed in the GLIM standardizations across the resource/fisher scenarios.

Results also indicated that Y, V, G, M and G×M were important terms in GLIM standardizations, and that the latter should be considered in such analyses when a resource aggregates and moves seasonally between different locations.

Interactions such as Y×G, Y×M or Y×V should also be considered if it is likely that annual CPUE changes differ across locations, seasons or vessels, respectively. Otherwise, incorporating higher order terms could lead to small improvements towards linearity or, in some cases, to greater degrees of non-linearity, particularly if higher order terms are not well represented by the data.

If standardized-CPUE is used to proxy abundance, then GLIM standardizations should be performed in all resource/fisher scenarios, under no marine closures. The terms to use are: Y, V, G, M and G×M.

In addition, standardizations should be performed when resource movement is non-random under marine closures using terms Y, V, G, M and G×M. However, when resource movement is random and fisher behaviour is non-random (under marine closures), fitting the Year (Y) term should be used. Also, standardizations could be employed (fitting all main effects and the G × M term) when both resource movement and fisher were random, under marine closures, although the degree of CPUE-abundance linearity improvement was minimal.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

FOR FURTHER INFORMATION

CSIRO Oceans & Atmosphere
Geoff Tuck
t +61 03 6232 5106
e Geoff.tuck@csiro.au
w www.csiro.au

AT CSIRO, WE DO THE
EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy
every year. As the largest patent holder
in the nation, our vast wealth of intellectual property has led to more
than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE.
WE INNOVATE.