Fisheries and Oceans Canada
Science

Pêches et Océans Canada
Sciences

C S A S

Canadian Science Advisory Secretariat

S C C S

Secrétariat canadien de consultation scientifique

Research Document 2009/043

Document de recherche 2009/043

Development of precautionary management strategies for the British Columbia sablefish (*Anoplopoma fimbria*) fishery

Élaboration de stratégies de gestion préventives pour la pêche de la morue charbonnière (*Anoplopoma fimbria*) en Colombie-Britannique.

S.P. Cox[1]
A.R. Kronlund[2]
M.R. Wyeth[2]

[1] School of Resource and Environmental Management
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6

[2] Fisheries and Oceans Canada
Science Branch, Pacific Region
Pacific Biological Station
3190 Hammond Bay Road
Nanaimo, BC V9T 6N7

ISSN 1499-3848 (Printed / Imprimé)
ISSN 1919-5044 (Online / En ligne)

Canada

Table of Contents

LIST OF TABLES II
LIST OF FIGURES III
ABSTRACT VI
RÉSUMÉ VII
1 INTRODUCTION 1
2 DEVELOPING MANAGEMENT PROCEDURES FOR B.C. SABLEFISH 3
  2.1 OVERVIEW 3
  2.2 FISHERY OBJECTIVES 5
  2.3 DATA CHOICES 6
  2.4 ASSESSMENT METHODS 8
  2.5 HARVEST CONTROL RULES 12
3 SCENARIOS FOR B.C. SABLEFISH 17
  3.1 OPERATING MODEL 17
  3.2 KEY UNCERTAINTIES – STOCK PRODUCTIVITY 18
  3.3 KEY UNCERTAINTIES – PRESENT STOCK STATUS 19
  3.4 PERFECT INFORMATION MANAGEMENT PROCEDURES 19
4 PERFORMANCE MEASURES 20
5 RESULTS 21
  5.1 GENERAL TRENDS IN FISHERY AND CONSERVATION PERFORMANCE 21
  5.2 TRADE-OFF RELATIONSHIPS BETWEEN YIELD AND CONSERVATION 23
  5.3 INDIVIDUAL PERFORMANCE METRICS FOR SELECTED MANAGEMENT PROCEDURES 26
6 DISCUSSION 29
  6.1 A STRATEGY FOR CHOOSING A MANAGEMENT PROCEDURE 31
  6.2 LIMITATIONS AND FUTURE WORK 33
7 CONCLUSION 35
ACKNOWLEDGEMENTS 37
LITERATURE CITED 38
APPENDIX A REQUEST FOR SCIENCE INFORMATION AND/OR ADVICE 67
APPENDIX B DATA 69
APPENDIX C SABLEFISH OPERATING MODEL FOR POPULATION AND FISHERY DYNAMICS 83
APPENDIX D PRODUCTION MODELS FOR SABLEFISH STOCK ASSESSMENT 91
APPENDIX E CATCH-AGE MODELING APPROACH TO SABLEFISH STOCK ASSESSMENTS 101
APPENDIX F TRADE-OFF RELATIONSHIPS AND PERFORMANCE STATISTICS 116

List of Tables

Table 1 Summary of management procedures indicating the data, assessment method, and harvest control rule type and parameters. 45

Table 2 Distinguishing features of operating model scenarios S1-S4. Parameters represent the steepness of the Beverton-Holt stock-recruitment function (H) and trap fishery hyperstability parameter ($Q_{2,TRAP}$). The initial depletion of the stock as of 2006 (D_{INIT}) depends on the operating model parameters. 46

Table 3 Definitions of performance statistics used for sablefish management strategy evaluation. Symbols T_1 and T_2 define the period over which performance measures are computed. 46

Table 4 Summary of selected management procedures. 47

Table 5 Summary of performance statistics and ranks by management procedure for Scenario 1. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario. 48

Table 6 Summary of performance statistics and ranks by management procedure for Scenario 2. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario. 49

Table 7 Summary of performance statistics and ranks by management procedure for Scenario 3. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario. 50

Table 8 Summary of performance statistics and ranks by management procedure for Scenario 4. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario. 51

List of Figures

Figure 1 Schematic representation of the management strategy evaluation process including the simulation feedback loop between the state of the stock and the iterative application of management procedures. 52

Figure 2 Annual sablefish landings (a) and stock abundance indices include (b) nominal fishery CPUE (kg/trap), (c) standardized survey mean CPUE (kg/trap), (d) tagging biomass index, and (e) Japanese LL CPUE index. 53

Figure 3 Conceptual diagram of a harvest strategy consistent with the precautionary approach (from DFO 2006). 54

Figure 4 Observed annual quotas for B.C. sablefish (1998-2006; open circles) and predicted (solid circles) values based on equation (1) with $\lambda_1 = 0.75$ and $\lambda_2 = 299$ 55

Figure 5 Example trade-offs between median average catch and median average depletion. This figure demonstrates how these plots are constructed and therefore does not represent actual results. Panel (a) compares catch-age procedures CA1 - CA3 with lines to indicate the difference in average catch (vertical) and depletion (horizontal) between CA1 and CA3; panel (b) compares three data-based (DB1 - DB3) and three catch-age model-based procedures; and panel (c) shows the same three catch-age procedures for two hypothetical stock scenarios S1 (red) and S2 (black). Note the colored and labeled indicators along the top of the plot showing the initial stock depletion (e.g. 2006) for each scenario. 56

Figure 6 Results for a single replicate of scenario S3 when spawning biomass shows a strongly increasing trend under specific (a) data-based, (b) production model-based, and (c) CA model-based management procedures. Biomass and catch units in metric tons. All procedures use VHR decision rules with the first five years constrained to a maximum 15% change in catch. The vertical dotted line indicates the start of the simulation period and the horizontal dotted line indicates spawning biomass in the first year of procedure implementation. 57

Figure 7 Results for a single replicate of scenario S1 when spawning biomass shows a stable future trend under specific (a) data-based, (b) production model-based, and (c) CA model-based management procedures. Biomass and catch units in metric tons. All procedures use VHR decision rules with the first five years constrained to a maximum 15% change in catch. The vertical dotted line indicates the start of the simulation period and the horizontal dotted line indicates spawning biomass in the first year of procedure implementation.. 58

Figure 8 Trajectories of catch under scenario S1 using data-based procedures with (a) $\lambda_2 = 150$ and (b) $\lambda_2 = 210$; production model with (c) $U^{REF} = 0.06$ and (d) $U^{REF} = 0.10$; and catch-age model with (e) $U^{REF} = 0.06$ and (f) $U^{REF} = 0.10$. Vertical dashed lines indicate year 2006. Trajectories are summarized by the median (thick black line), three individual simulation replicates (thin black lines), and 5th to 95th percentiles (shaded area). 59

FIGURE 9 TRAJECTORIES OF CATCH UNDER SCENARIO S4 USING DATA-BASED PROCEDURES WITH (A) $\lambda_2 = 150$ AND (B) $\lambda_2 = 210$; PRODUCTION MODEL WITH (C) $U^{REF} = 0.06$ AND (D) $U^{REF} = 0.10$; AND CATCH-AGE MODEL WITH (E) $U^{REF} = 0.06$ AND (F) $U^{REF} = 0.10$. VERTICAL DASHED LINES INDICATE YEAR 2006. TRAJECTORIES ARE SUMMARIZED BY THE MEDIAN (THICK BLACK LINE), THREE INDIVIDUAL SIMULATION REPLICATES (THIN BLACK LINES), AND 5TH TO 95TH PERCENTILES (SHADED AREA). 60

FIGURE 10 TRADE-OFF RELATIONSHIP BETWEEN MEDIAN AVERAGE CATCH (T) AND MEDIAN AVERAGE SPAWNING BIOMASS DEPLETION FOR SCENARIOS S1-S4 MEASURED OVER (A) 1-5, (B) 6-10, (C) 11-20 AND (D) 21-40 YEAR TIME INTERVALS. SYMBOLS ARE ORDERED FROM LEFT TO RIGHT CORRESPONDING TO DATA-BASED WITH $\lambda_2 = \{240, 210, 180, 150, 120\}$ (♦, ▲, ●, ■, ▼), PRODUCTION MODEL-BASED WITH $U^{REF} = \{0.10, 0.08, 0.06\}$ (☒, ×, +), AND CA MODEL-BASED WITH $U^{REF} = \{0.10, 0.08, 0.06, 0.04\}$ (△, ○, □, ▽) PROCEDURES. PERFECT-INFORMATION PROCEDURES WITH $U_{MSY} = \{0.04, 0.06, 0.08, 0.10\}$ ARE INDICATED BY ASTERISKS. INVERTED TRIANGLES AND LABELS ALONG THE UPPER *X*-AXIS INDICATE INITIAL DEPLETION VALUES FOR EACH SCENARIO. 61

FIGURE 11 SUMMARY OF CATCH VARIABILITY PERFORMANCE OVER PROJECTION YEARS 11-20 (PANEL A) AND 21-40 (PANEL B) FOR THREE SELECTED CLASSES OF MANAGEMENT PROCEDURES. RESULTS ARE SHOWN FOR SCENARIOS S1 (BLACK), S2 (RED), S3 (GREEN) AND S4 (BLUE). RESULTS FOR THE QUASI-PERFECT INFORMATION PROCEDURES ARE SHOWN AS “TRUE”. SYMBOLS INDICATE THE MEDIAN VALUE AND THE VERTICAL BARS INDICATE THE 5^{TH} AND 95^{TH} PERCENTILES OVER 50 REPLICATES. 62

FIGURE 12 AVERAGE ANNUAL CATCH OVER PROJECTION YEARS 11-20 (PANEL A) AND 21-40 (PANEL B) FOR THREE SELECTED CLASSES OF MANAGEMENT PROCEDURES UNDER SCENARIOS S1 (BLACK), S2 (RED), S3 (GREEN) AND S4 (BLUE). PERFECT-INFORMATION PROCEDURES ARE SHOWN AS “TRUE”. SYMBOLS INDICATE THE MEDIAN VALUE AND THE VERTICAL BARS INDICATE THE 5^{TH} AND 95^{TH} PERCENTILES OVER 50 REPLICATES. 63

FIGURE 13 RETROSPECTIVE PATTERNS IN ESTIMATED SPAWNING BIOMASS FOR DECLINING/STABLE (LEFT COLUMN), INCREASING (CENTER COLUMN), AND INCREASING/DECLINING (RIGHT COLUMN) TRUE SPAWNING BIOMASS TRAJECTORIES (SOLID BLACK LINES). SOLID GRAY LINES REPRESENT SUCCESSIVE SPAWNING BIOMASS TRAJECTORIES ESTIMATED BY THE CATCH-AGE MODEL UNDER A VHR RULE WITH $U^{REF}=0.08$. 64

FIGURE 14 RETROSPECTIVE PATTERNS IN ESTIMATED SPAWNING BIOMASS FOR DECLINING/STABLE (LEFT COLUMN), INCREASING (CENTER COLUMN), AND INCREASING/DECLINING (RIGHT COLUMN) TRUE SPAWNING BIOMASS TRAJECTORIES (SOLID BLACK LINES). SOLID GRAY LINES REPRESENT SUCCESSIVE SPAWNING BIOMASS TRAJECTORIES ESTIMATED BY THE PRODUCTION MODEL UNDER A VHR RULE WITH $U^{REF}=0.08$. 65

FIGURE 15 AVERAGE SPAWNING BIOMASS DEPLETION OVER PROJECTION YEARS 11-20 (PANEL A) AND 21-40 (PANEL B) FOR THREE SELECTED CLASSES OF MANAGEMENT PROCEDURES UNDER SCENARIOS S1 (BLACK), S2 (RED), S3 (GREEN) AND S4 (BLUE). PERFECT-INFORMATION PROCEDURES ARE SHOWN AS “TRUE”. SYMBOLS INDICATE THE MEDIAN VALUE AND THE VERTICAL BARS INDICATE THE 5^{TH} AND 95^{TH} PERCENTILES OVER 50 REPLICATES. 66

Correct citation for this publication:

Cox, S.P., Kronlund, A.R., and Wyeth, M.R. 2010. Development of precautionary management strategies for the British Columbia sablefish (*Anoplopoma fimbria*) fishery. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/043. vi + 145 p.

Abstract

A fishery decision-making framework was recently advanced by Fisheries and Oceans Canada (DFO) that requires application of stock reference points, harvest rules and compliance with the Precautionary Approach. The framework is intended to assure resource sustainability and meet the requirements of various eco-certification programs. This paper describes management procedures for the sablefish (*Anoplopoma fimbria*) fishery in British Columbia that address each requirement of the DFO precautionary framework. We develop and compare the performance of relatively simple data-based fishery management procedures, which set annual catch limits by combining the preceding year's catch limits with the recent average of fishery-independent surveys, with model-based procedures that set annual catch limits using constant exploitation rate policies and estimates of stock biomass from production or catch-age models. The data-based and model-based procedures we examined employed either constant harvest rate (CHR) or variable harvest rate (VHR) decision rules for setting annual catch limits where the latter addresses a specific DFO precautionary requirement to adjust fishery exploitation rates in response to changes in stock status. All candidate management procedures were tested in stochastic simulations against four operating model scenarios that reflect uncertainties about productivity and current status of the B.C. sablefish stock. In general, VHR decision rules provided consistently better conservation outcomes compared to CHR rules, especially for low productivity scenarios. Data-based procedures provided similar trade-offs between catch and conservation as more elaborate model-based procedures and both types of procedures met inter-annual catch variability objectives. In terms of average annual catch, data-based rules outperformed management procedures based on aggregate production models, while procedures based on catch-age models performed better than data-based rules. Catch-age model procedures were able to track large increases in stock biomass and thus obtain larger average catches under these conditions. Production models consistently under-estimated biomass during periods of population growth and therefore under-exploited growing stocks. Future phases of this work will expand on candidate management procedures and the scenarios against which these procedures are evaluated.

Résumé

Le ministère des Pêches et Océans Canada (MPO) a récemment présenté un cadre décisionnel de pêche exigeant la mise en application de points de référence des stocks, de règles en matière de capture et de conformité à l'approche préventive. Le cadre vise à assurer la durabilité des ressources et à répondre aux exigences des divers programmes d'écocertification. Dans ce document, on décrit les procédures de gestion pour la pêche de la morue charbonnière (*Anoplopoma fimbria*) en Colombie-Britannique en ce qui concerne chaque exigence du cadre préventif du MPO. Nous expliquons et comparons le rendement de procédures de gestion de pêche relativement simples reposant sur les données, lesquelles fixent les limites de capture annuelles en combinant les limites de capture de l'année précédente avec la moyenne récente obtenue par des enquêtes indépendantes sur les pêches, et ce, selon des procédures reposant sur les modèles qui fixent les limites annuelles de capture à l'aide des politiques sur le taux d'exploitation constante et des estimations de biomasse du stock découlant de la production ou des modèles de capture selon l'âge. Les procédures reposant sur les données et sur les modèles que nous avons étudiées utilisaient soit les règles de décision du taux de capture constant (TCC) ou du taux de capture variable (TCV) pour fixer les limites de capture lorsqu'elles font l'objet d'une exigence préventive précise du MPO en vue d'ajuster le taux d'exploitation de pêche en fonction des changements pour l'état des stocks. Toutes les procédures de gestion à l'étude ont été analysées lors de simulations stochastiques avec quatre scénarios de modèle d'exploitation qui représentent les incertitudes quant à la productivité et à l'état actuel des stocks de morue charbonnière en C.-B. En général, les règles de décision TCV donnaient de façon constante de meilleurs résultats pour la conservation par rapport aux règles TCC, en particulier pour les scénarios de faible productivité. Les procédures reposant sur les données ont abouti à des résultats semblables entre la capture et la conservation, car les procédures plus complexes reposant sur les modèles et chacun des types de procédures répondaient aux objectifs de variabilité de capture interannuels. En ce qui concerne la moyenne annuelle de capture, les règles reposant sur les données ont donné un rendement supérieur aux procédures de gestion reposant sur les modèles de production globaux, tandis que les procédures reposant sur les modèles d'âge des prises ont donné un meilleur rendement que les règles reposant sur les données. Les procédures pour le modèle de l'âge des prises ont permis de faire le suivi des importantes augmentations de la biomasse du stock et d'obtenir ainsi des captures moyennes plus importantes dans ces conditions. Pour les modèles de production, on a eu de façon constante une sous-estimation de la biomasse pendant les périodes de croissance de la population et, par conséquent, une sous-exploitation de la croissance des stocks. Les étapes ultérieures de ce travail porteront sur les procédures de gestion à l'étude et sur les scénarios à partir desquels on a évalué ces procédures.

1 Introduction

A fishery decision-making framework incorporating the precautionary approach was recently advanced by the Fisheries and Aquaculture Management Branch of Fisheries and Oceans Canada (March 2007, http://www.dfo-mpo.gc.ca/fm-gp/peches-fisheries/fish-ren-peche/sff-cpd/overview-cadre-eng.htm). The framework is intended to assure resource sustainability and meet the requirements of various eco-certification programs by demonstrating the application of reference points, harvest rules and compliance with the precautionary approach. Canada's national policy on the precautionary (DFO 2006; FAO 1995) approach demands: (*i*) prior identification of undesirable outcomes and measures that avoid or correct them promptly, (*ii*) the formulation of decision rules that specify what actions will be taken when deviations from operational targets and constraints are detected, and (*iii*) that the management plan should be adopted only after it has been demonstrated to effectively avoid undesirable outcomes for both the resource and fishing communities. Central to the framework is the treatment of uncertainty and risk, with explicit requirements to communicate the risk of resource decline associated with management actions. Participation of fishery stakeholders in the development of decision rules is identified as a requirement for successful application of the decision making framework, although such participation is secondary to national and international commitments.

The evaluation of fishery management strategies via closed-loop feedback simulations (Walters 1986, Cooke 1999, de la Mare 1986, 1996, 1998) offers a potential vehicle for implementing a collaborative, precautionary approach to fisheries management that involves stakeholders in all aspects of fisheries management strategy development. The management strategy evaluation (MSE) methodology is defined by four components: (*i*) operational objectives, (*ii*) specific fishery monitoring data and stock assessment methods, (*iii*) harvest control rules that set catch limits based on estimates of stock status, and (*iv*) a prospective evaluation of the entire procedure using a set of performance statistics (de la Mare 1996). Operational objectives reflect national and international policy commitments as well as specific statements by stakeholders that identify their interests for the conduct of the fishery. Fishery monitoring, stock assessments, and harvest control rules represent the decision-making framework whereby scientific information is collected, processed, and used in setting fishery regulations (e.g., annual catch limits). The prospective evaluation involves simulation testing of management procedure performance against a fixed set of goals (ecological, economic, and sociological) over a range of possible scenarios for the stock and fishery.

Scenarios represent structural hypotheses about the fish stock and/or fishery dynamics that are not currently resolved by the available data or those that may never be resolved. Evaluation of procedures across scenarios and the incorporation of uncertainty into the simulations is the basis of robustness testing. Development and evaluation of management procedures therefore addresses all three demands of the precautionary approach to fisheries management as well as DFO's decision-making framework. In particular, the approach: (*i*) considers alternative stock assessment approaches for identifying undesirable outcomes; (*ii*) evaluates alternative forms of decision rules that specify how regulations will change (i.e., corrective actions) in response to estimates of stock status relative to operational targets; and (*iii*) demonstrates, via computer simulation, how effectively whole management procedures meet fishery management objectives while avoiding undesirable outcomes. Thus, management strategy evaluation changes the focus from the traditional "best assessment" to identifying a "best procedure" from a set of candidates that is robust to uncertainties about the real world. Here, "best" implies the procedure that most closely meets the desired objectives over a wide range of plausible scenarios about the resource and fishery processes.

Like traditional stock assessments, the MSE approach also integrates resource monitoring data into stock assessments and management advice; however, it differs from the traditional approach because it includes a simulation step to test whether application of stock assessment methods and decision rules provide outcomes that are consistent with fishery objectives. Typically, fishery objectives fall within the three broad categories of catch, inter-annual stability of catch, and acceptable risks of irreversible or economically undesirable stock depletion (Butterworth 2007). The relative success of candidate management procedures at achieving fishery objectives is judged by comparing a set of performance statistics obtained from simulating the consistent application of the procedures into the future using the data collected up to each point in time (de la Mare 1998, Punt et al. 2002a,b,c). Successful management procedures must, on average, achieve the desired objectives even if the stock assessment component of the procedure is in error. Here, management strategy evaluation is compatible with the precautionary approach because it demonstrates robustness in the face of uncertainty (FAO 1995). A key feature that distinguishes robust management procedures is deliberate negative feedback control that provides potential corrective actions in response to declines in stock size or erroneous perceptions about the current status of the resource.

Management strategy evaluation is not without limitations and pitfalls. For example, although specification of realistic stock and fishery dynamics models ("operating models") for simulation testing of candidate procedures is not particularly difficult, determining the relative credibility of these alternative operating models is often a critical limitation to management strategy adoption. On the other hand, most traditional stock assessments use models that also invoke strong assumptions about the true state of the stock without taking the step of determining whether the recommended harvests meet long-term catch and conservation objectives. Widespread acceptance and implementation of the management strategy evaluation approach has also been slow despite its attractiveness from a scientific and precautionary point-of-view (Smith et al. 1999; Butterworth 2007). Extensive work has been conducted in South Africa, Australia, and New Zealand where MSE has found application (Butterworth and Punt 1999; Smith et al. 1999; Bentley et al. 2005). Difficulties in adopting formal management procedures appear to derive from several causes, among them (*i*) a lack of stakeholder and management confidence in following a procedure that is derived in a more complicated way than a typical stock assessment, (*ii*) a lack of stakeholder ownership of the process, (*iii*) difficulty integrating results across a range of possible stock scenarios where the "plausibility" weighting of each scenario is usually unspecified (Butterworth and Punt 1999), and (*iv*) lack of a policy decision regarding the appropriate trade-offs among objectives (e.g., maximizing catch, reducing inter-annual variation in catches, minimizing the risk of serious stock depletion). The latter policy decision can be difficult to obtain if fishery objectives among stakeholders and fishery managers are not well articulated and understood by the participants in the evaluation process (Butterworth 2007).

This paper develops a management strategy evaluation approach for the sablefish (*Anoplopoma fimbria*) fishery in British Columbia. The MSE approach does not have a documented history of application to fisheries on the Pacific coast of Canada, although Logan et al. (2005) developed a suite of simulation scenarios for the Strait of Georgia lingcod (*Ophiodon elongatus*) stock designed to evaluate catch levels consistent with fishery conservation objectives using MSE principles. Our methodology and results represent the first stage of work that involved consultations with stakeholders, managers, and other stock assessment scientists in developing a management strategy evaluation approach. We develop and compare the performance of relatively simple data-based fishery management procedures, which are attractive to stakeholders, with model-based procedures that tend to be attractive to scientists. Data-based harvest rules set annual catch limits by combining the preceding year's catch limits with the recent average of fishery-independent surveys, thus eliminating the traditional stock assessment modeling component. In contrast, model-based procedures set annual catch limits, C_t, using the constant harvest rate policy $C_t = U^{ref}B_t$, where U^{ref} is a reference exploitation rate and B_t is an estimate of stock biomass from a production model or more sophisticated statistical catch-age model. The model-based procedures attempt to mimic more elaborate management systems that depend heavily on catch sampling and stock assessment modeling. Presumably, the more complex model-based procedures have a greater chance of utilizing the resource in an optimal way if they produce unbiased estimates of stock size. Model-based procedures may act to control inter-annual variability in catch limits. Constant harvest rate rules, though simple, are not compliant with national policy since they make no adjustment to the removal rate based on estimated stock status. Therefore, we develop variable harvest rate rules for both data-based and model-based procedures to comply with DFO (2006) by adjusting the exploitation rate or catch in pre-defined ways depending on the estimated stock status. Results of constant harvest rate rules are retained, however, to allow comparison with precautionary derivatives of these rules. Regardless of their form, all candidate management procedures are tested against four alternative operating model scenarios for sablefish that are distinguished by two levels each of stock productivity and the perception of current stock status. We show that, in general, data-based procedures that employ variable harvest rate decision rules provide reasonable catch and conservation performance compared to more elaborate model-based procedures. In fact, some data-based procedures actually performed better overall than production model procedures under these scenarios. The MSE approach can easily be extended to incorporate a wide range of stock scenarios (including species other than sablefish), assessment approaches, decision rules, and performance measures as required in the development of precautionary fisheries management strategies for Canadian fisheries.

2 Developing Management Procedures for B.C. Sablefish

2.1 Overview

Work on management strategy evaluation for sablefish was prompted by a review of sablefish stock assessment commissioned jointly by Fisheries and Oceans Canada (DFO) and the Canadian Sablefish Association in 2005. The review documented the rapid turnover in stock assessment methods for sablefish over the preceding 20 years, the absence of clearly articulated objectives for the fishery, and absence of a consistent procedure for setting quotas based on scientific advice (S.P. Cox and S. Martell, Independent review of the British Columbia sablefish (*Anoplopoma fimbria*) scientific research and assessment program, unpublished). Previous stock assessments of B.C. sablefish have followed a "traditional approach" (Butterworth 2007) of developing a best mathematical assessment of the resource that integrates available data and current understanding of the structural components of the population dynamics. Advice has been provided as specific total allowable catch (TAC) recommendations, or as decision tables that present the probability of specific future stock sizes given alternative fixed catch projections (e.g., Haist et al. 2005). The utility of this traditional scientific approach is predicated on the assumption that the stock assessment method used accurately reconstructs the true state of the stock, and correctly predicts the range of stock response to future harvesting. A shortcoming of this approach is the attempt to capture the risk of following a single management action, for example a constant catch policy, since the long-term consequences to the stock depend on subsequent actions in response to new data (Cooke 1999). Typically catch, biological and possibly resource survey data may become available each year and assessments are revised over time resulting in updating of the management advice. The approach also fails to directly illustrate the trade-offs among yield, stability in yield, and fish stock conservation.

We initiated a consultation process in which industry and fishery manager stakeholders participated in the development and evaluation of management procedures that are consistent with their own objectives. In the following sections we discuss considerations that guided our initial choices of data, assessment method, and harvest control rule components of candidate management procedures. Note that this presentation does not constitute the final evaluation for B.C. sablefish since we do not consider several scenarios that we believe are quite plausible. Where appropriate, we identify key themes from Canadian fisheries policy, the management strategy evaluation literature, and stakeholder consultation that influenced our choices.

We followed a step-wise scheme to evaluate management procedures for B.C. sablefish that is depicted in Figure 1 and stated as an algorithm below:

1. Identify a working set of objectives through consultation with fishery managers and representatives of the sablefish industry with respect to catch and inter-annual catch variability, and by adopting conventions in the scientific literature for conservation objectives;

2. Define a range of alternative management procedures by considering combinations of (i) data types and data collection frequency, (ii) assessment methods, and (iii) harvest control rules;

3. Specify an operating model to enable simulation of alternative plausible scenarios for the sablefish population, fishery dynamics, and data generation mechanisms. This step involves fitting the operating model to available data to determine model parameters consistent with the stock history and the structural assumptions of the scenario, a process termed conditioning;

4. Project stock and fishery status for each management procedure into the future under each alternative scenario. Each iteration of the projection involves the following steps:

    a) Generate the data available for stock assessment;
    b) Apply the stock assessment method to the data to estimate quantities required by the control rule;
    c) Apply the harvest control rule to generate a catch limit;
    d) Subtract the catch limit from the simulated sablefish population as represented by the operating model.

5. Calculate a set of quantitative performance statistics that can be used to compare outputs of candidate management procedure against the management objectives.

Step 4a involves application of the operating model that was identified in Step 3, which maintains the state of the population over time and also generates the data that will be collected in the future. The operating model is described in detail in Appendix C. Data generated by the operating model are generally the fishery and survey data that are currently being accumulated by sampling programs, but these data could include new types for which cost-benefit analyses are required. A key feature of the evaluation process is that the assessment method applied in step 4b is blind to the operating model; that is, the assessment is only provided with data such as survey indices of abundance and catch-at-age. This "closed-loop" simulation strategy for testing harvest management procedures is well documented in the literature (e.g., Walters 1986, de la Mare 1998, Cooke 1999, Punt and Smith 1999, Sainsbury 2002, Butterworth 2007).

Each management procedure component in 4a – c requires a particular set of choices. For example, the data step could involve only a survey index of abundance, the assessment step could involve a simple or complex modeling approach and the harvest control rule may make adjustments for risk and uncertainty. The choices made will affect fishery performance and therefore are the main focus of management strategy evaluation. Details of the suite of management procedures, along with the choices involved in each, are summarized in Table 1.

2.2 Fishery objectives

During the course of several management strategy evaluation workshops involving industry stakeholders, we recorded objectives for fishery performance as well as suggestions about data and methods that should be evaluated. Stakeholders expressed consensus objectives that inter-annual catch variability should be restricted to less than 15% to 20%. They also indicated a desire for increasing and maintaining trap fishery catch-per-unit effort (CPUE) to at least 14 to 16 kg/trap, which represents an approximate 50% increase from current annual averages of near 10 kg/trap as determined from fishery logbooks. Stakeholder objectives for sablefish catch rates in longline hook fisheries have not been determined to date and may be difficult to apply in practice because that fishery is increasingly subject to multi-species bycatch accountability constraints as a result of the recently implemented Groundfish Pilot Integration Proposal (Fisheries and Oceans Canada 2006, 2007; Koolman et al. 2007). No specific objectives regarding desired catch levels were stated by stakeholders except to attain the maximum possible catches subject to conservation considerations and the constraint on catch variability. Additional constraints on the level of catch may result from non-directed fisheries. For example, the multi-species groundfish trawl sector is allocated 8.75% of the commercial sablefish quota. Sablefish represents an important species for the trawl fleet, both for economic reasons and in terms of a potential avoidance species when prosecuting their fishery for other groundfish species, but quantitative limits on inter-annual catch variability have not been identified.

We have not yet obtained any specific or general target for sablefish conservation objectives, although the above CPUE objective does imply certain increases in stock size and we designed the variable harvest rate rules presented below to avoid historic low survey results obtained in 2001. Yield and conservation objectives cannot be simultaneously maximized; pursuit of the former necessarily implies a trade-off reduction in stock abundance. Neither the Canadian national policy nor the scientific literature is specific on the degree to which stock abundance can be reduced, but 20% of the pre-exploitation equilibrium biomass has been suggested as a lower limit in several studies (e.g., Beddington and Cooke 1983, Francis 1992, Punt 1995, 1997). However, other choices such as $0.25B_0$ have been put forward (Hall et al. 1988; Quinn et al. 1990). Hilborn (1997) was critical of the use of $0.2B_0$ because the level is arbitrary, suggesting that some stocks depleted to very low levels have recovered and may even be capable of producing high sustainable yields. Hilborn (1997) further suggested that the mechanisms that would lead to undesirable states of the resource should be incorporated into the operating model and their effects would be manifest in future catches. However, we agree with Butterworth and Punt (1999) who note that few studies have incorporated the possibility of depensation at low stock levels, which has the obvious consequence of seriously compromising recovery potential. Thus, we adopted $0.2B_0$ as a preliminary conservation reference point for this paper.

2.3 Data choices

In this document we emphasize the development of a management strategy evaluation; therefore, we provide only a brief overview of the available data and particular choices that were made for this development. Readers interested in the details of fishery independent sablefish survey and tagging programs may consult background documents beginning with Wyeth et al. (2007). Landings data are available since 1913 (Haist et al. 2005), but were limited to 1965 to 2006 for the purposes of this analysis (Figure 2a) because this period marks the major re-development of the directed sablefish fishery following a long period of low annual catches of about 1000 t following the higher demand for fish protein during World War I. At the present time, we considered three types of data that provide either a relative index or potentially a direct measurement of stock biomass (Figure 2 b-d). Catch rate-based data types such as fishery or survey catch-per-unit-effort (CPUE) are relative abundance indices; that is, they only provide information about changes in stock biomass at a given time relative to biomass at some other time. Relative abundance indices for sablefish include (i) commercial trap fishery nominal CPUE (1979 – 2006), (ii) the standardized trap survey CPUE (1990 – 2006), and (iii) Japanese longline fishery CPUE (1965 – 1980, McFarlane and Beamish 1983, Stocker and Saunders 1997). In addition, an index derived from a mark-recapture program (1992-2006) can potentially yield an absolute estimate of stock abundance provided several assumptions are entertained, the most controversial of which revolves around the assumed values of the tag reporting rates over time, which are unknown and cannot be estimated with existing data.

Nominal sablefish trap fishery CPUE was calculated using a ratio-of-means estimator as the sum of trap catches divided by the sum of trap effort for all records that have valid observations for both catch and effort. Nominal trap CPUE shows a trend that suggests relatively high stock abundance in the late 1980s and early 1990s, followed by a period of lower and slowly declining catch rates from the mid-1990s to historic lows experienced in 2001 and 2002 (Figure 2b). Recruitment from the 1999/2000 year classes noted by Haist et al. (2005), Hanselman et al. (2007), and Schirripa (2007) apparently fueled a modest increase in catch rates which increased sharply in 2003, but have declined since that time to a mean annual rate near the 2001 level. Standardization of commercial trap catch rates using generalized linear models (GLMs) is only possible for data beginning in 1990 because earlier data aggregated multiple fishing events in the source database which precludes general linear model analysis over the entire 1979 to 2006 period. In view of this restriction we elected to utilize the longer nominal CPUE time series for this analysis. We also did not attempt to derive relative abundance indices from other commercial fisheries in B.C. that capture sablefish because it is unlikely that any potential index would be reliable in the long-term. Aside from the usual biases of commercial fishery CPUE data, fishery catch rates in B.C. are strongly affected by the multi-species groundfish management regime. For example, the recently implemented Commercial Groundfish Pilot Integration Proposal (Fisheries and Oceans 2006, 2007) which allows access to quota species across traditional gear and license boundaries was developed to improve accounting for catch and also to allow retention of species that would be discarded under restrictions of existing license regulations. As a consequence, the proportion of sablefish landings by longline hook gear has increased since the introduction of the pilot program in 2006, perhaps because the directed sablefish longline hook fishery can expand into locations that were unprofitable in the past due to bycatch implications. The trap fishery can also potentially expand into areas that were formerly costly owing to bycatch of rockfishes (*Sebastes* spp).

The standardized sablefish survey (1990-2006) used for abundance indexing includes nine survey localities that were intentionally selected because they were fished by commercial vessels and were spatially dispersed about 60 nm apart. This spatial arrangement permitted all localities to be visited within a 30 day period given favourable weather. Because only one set is conducted within each specified depth interval at each survey locality, there is no replication of sets within each combination of depth and locality. The exact spatial position of each set is also at the discretion of the fishing master rather than being randomly selected. Typically, survey localities include high-relief bathymetric features such as gullies or canyons, which reflect the original intent to index sablefish abundance in core fishing areas that represent prime fishable habitat.

A second annual fishery-independent survey that follows a depth and area stratified random sampling design was initiated in 2003, initially for the purposes of distributing tags coast-wide at random locations over five area strata and three depth strata of the offshore habitat range of sablefish (i.e., 183 to 1372 m; Wyeth et al. 2006). Fishing practices were standardized at the outset of the survey in hopes of yielding a second survey abundance index with statistical properties superior to the existing standardized survey. The design differences, as well as increased sample size for the stratified random survey, mean that the two surveys may react differently in response to changes in actual stock abundance. Potential differences between these surveys may not become apparent until major changes (increases or decreases) in abundance occur in the sablefish stock. At this time we have not conducted assessments using the stratified random survey because the time-series is short and ages for sablefish caught during this survey are not complete. Management strategy development for sablefish in the near future will investigate the feasibility of adopting the stratified random survey index as the primary abundance measure since it should *a priori* have better statistical properties and stock coverage.

There are currently two sablefish tagging programs that could provide direct estimates of sablefish biomass under the relatively strict assumptions (in addition to a few others) that tag reporting rates are constant and equal to 1.0 for all years of recoveries. In an attempt to make these assumptions hold as nearly as possible, we currently calculate the tagging biomass estimate from the “traditional tag program” (Wyeth et al. 2006) by including only tag recoveries that were captured in the 12 months following release and were released at 60cm from traditional offshore survey localities. The selection of 60cm or greater fish is an attempt to minimize effects of over-reporting under-sized fish because they have tags.

Proportion-at-age data are available from commercial trap fishery samples (1979 – 2002) and trap survey programs (1990 – 2006), though samples sizes may be low or absent in some years (Appendix B). To date very few fish captured during the stratified random survey have been aged, although work is progressing. All age data used in this paper were obtained using the burnt otolith section method (MacLellan 1997).

2.4 Assessment methods

Fisheries stock assessments provide a scientific opinion on the status and productivity of fish stocks that is intended to inform managers and stakeholders of the short- and long-term consequences of alternative management strategies (Walters and Martell 2004). The processes and methods for assessing fish stocks have generated a large literature on the subject of fisheries stock assessment modeling (c.f., Hilborn and Walters 1992; Quinn and Deriso 1997). In general, the level of detail represented in fisheries models usually varies in direct proportion to the level of detail in the available data. Simple aggregated biomass dynamics models are typically preferred where only fishery catch and effort are available. Complex models based on detailed accounting of changes in stock composition (age-, size-, gender-) can be applied where auxiliary data such as age-proportions in the catch or surveys exist. Such approaches attempt to separate fishing effects from other processes such as natural mortality and recruitment and therefore offer the ability to detect and account for changes in fishery selectivity, growth and size at age, and other time-varying changes in the population and fishery dynamics that might otherwise be assumed constant. Complex models also allow for evaluation of fishery regulations based on characteristics of fish populations (e.g., minimum size limits or male-only). However, it is not always true that more complex models perform better than simpler ones under all circumstances (NRC 1998). For example, Butterworth and Punt (1999) as well as Cooke (1999) concluded on the basis of the International Whaling Commission (IWC) management strategy simulation experience that management procedures based on complex population models offered few advantages over simpler approaches. In fact, harvest levels of some fisheries have been managed quite successfully using so-called "data-based" methods that simply adjust catch limits in response to directly-observable statistics such as survey catch rates or abundance estimates (Hilborn et al. 2002; Walters and Martell 2004). Under most circumstances, however, data-based approaches alone do not provide any assurance that a management procedure will produce an "optimal" outcome in terms of yield or conservation. Model-based stock assessment approaches, on the other hand, can be used to address the fisheries optimization problem of maximizing expected long-term yield.

Stock assessment models are often the most contentious component of fishery management procedures. The growing complexity of stock assessment models appears to lead to frustration among fishery managers and stakeholders (Cotter et al. 2004), which potentially limits the use of scientific advice. Typically, discussion of assessment models tends to focus on the technical aspects of model fitting at the expense of how best to provide management advice. In sablefish management procedure workshops, two divergent suggestions were expressed by industry stakeholders that represent the dichotomy between distrust of complex models and their desire to use available data to optimize the economic yield from the fishery. Some sablefish industry stakeholders requested that we examine a process for setting catch limits that "… reflect the abundance of fish on the grounds…" perhaps by using only the most recent survey or fishery CPUE. Although most industry stakeholders were skeptical of stock assessment models, others suggested that B.C. sablefish assessments should include the available commercial fishery and survey age composition data, a significant amount of which has been collected with industry support through formal collaborative agreements. Various age-structured stock assessments that have been applied to B.C. sablefish in the past are reviewed in Appendix B.

We developed and evaluated two alternative classes of management procedures in which stock assessment complexity ranged from simple data-based methods to complex model-based methods that employ either aggregate biomass or catch-age models (Table 1). Details of the two model-based approaches are given in Appendices D (production model) and E (catch-age model). The data-based assessment method, which depends primarily on a fishery-independent survey, is described in conjunction with harvest control rules in Section 2.5. Model-based procedures differ from data-based ones because they require several explicit assumptions about the underlying fish population dynamics and observations. In some cases, the strong assumptions made in these models about recruitment, growth, and mortality processes can lead to systematic trends in assessment biases (Walters 2004). On the other hand, classical stock assessment modeling provides a consistent and formal set of methods for evaluating potential biases, for example, by performing retrospective analysis (Mohn 1999). As mentioned above, fisheries stock assessment models serve to provide estimates of stock status and productivity, which essentially means providing annual estimates of (*i*) the stock biomass available to the fishery and the level of biomass relative to some reference point and (*ii*) the "optimum" fraction of the available biomass to harvest. Often, the latter task is difficult in fisheries systems where large variability in recruitment (i.e., process errors) and monitoring data (i.e., observation errors) tend to mask underlying stock production relationships (Walters 1986; Schnute and Kronlund 2002). Preliminary management procedure simulations indicated that sablefish are no exception to this pattern, so we removed the option of estimating the optimum harvest rate because these estimates tended to be highly variable, especially in the first few years of simulated management procedures. Thus, choices for stock assessment models in this paper represent a range of stock biomass estimators, whereas choices for harvest rates are treated as tuning parameters of the harvest control rule.

Production model

Although the at-sea catch sampling and ageing program for B.C. sablefish has generated catch-age samples from 1988 – 2000 (with some missing years), the program is inactive at the present time. Therefore, it is possible that future management procedures for sablefish will involve only catch-by-gear and abundance index data (e.g., survey catch rates). Stock assessments based on remaining catch-effort or catch-survey data will likely involve aggregated biomass dynamics models. Therefore, we developed a Schaefer biomass dynamics model that contained some added flexibility to deal with highly variable recruitment in a stock such as sablefish (Appendix D). In particular, we formulated an errors-in-variables production model estimator (Schnute and Richards 1995; Punt 2003) to account for process and observation errors when attempting to estimate harvestable biomass. Data used by the production model included only the standardized trap survey catch rate and total catch aggregated over trap, longline, and trawl fisheries. The production model estimator uses a tuning parameter, ρ, which represents the proportion of the total random error that is assigned to the observations. Thus, ρ controls how much of the variability in survey catch rates is assigned to random observation errors and how much is assigned to random process errors or unaccounted for changes in the stock biomass. We initially evaluated $\rho = 0.5$ but found that variability in catches far exceeded stated fishery objectives for this indicator. As a compromise we evaluated $\rho = 0.8$, which smoothed the estimated biomass trajectories and catches generated by the harvest control rules.

The production model contains several assumptions that are clearly violated given the underlying operating models that we used to test robustness. First, the production model assumes a single spawning-exploitable stock whereas in the operating model spawning stock and exploitable stocks are treated separately because fish recruit to the fishery before they recruit to the spawning population. Taking an aggregate catch from a single exploitable stock is also different from the true fishery structure in the operating model where different gear types exploit different components of the population (i.e., each gear type has a different selectivity function). Finally, the mean relationship between stock biomass and recruitment in the production model is mis-specified because the operating model's production function is not a symmetrical, dome-shaped function of total biomass as expressed by the Schaefer form of production model.

Catch-age model

Management procedures with a "complex" assessment method were investigated through the use of a state-space, catch-age model that is a multi-gear version of the approach developed by Schnute and Richards (1995) (Appendix E). Catch-age stock assessment modeling options are appealing for several scientific reasons. First, age-composition changes over time may contain information about temporal trends in fishing mortality and recruitment. Indeed, this particular capability is among the main reasons why so many fisheries agencies attempt to utilize ageing data. Second, in contrast to aggregate biomass production models, observed changes in fishery selectivity as measured by the annual sablefish tagging program can be accounted for in assessments as either fixed parameters or priors. Changes in fishery (and possibly survey) selectivity can have profound influences on abundance estimates from age-structured models, especially when there are few data to distinguish between dome-shaped and asymptotic selectivity functions. An extensive, industry-funded tag-recovery program for B.C. sablefish allows for direct estimation of length-based selectivity from tagging, and therefore potentially large improvements in age-structured assessment estimates. Third, a catch-age assessment approach provides the ability to use shorter times-series (< 20 years) of fishery-independent data alone, which potentially reduces many of the biases associated with fishery-dependent abundance indices. Finally, over the past two decades, sablefish industry stakeholders have made substantial investments in sampling and aging of commercial and survey catch as well as an annual tag release-recovery program. This information has not been consistently used to date, so questions invariably arise as to whether this information can actually contribute to higher fishery value.

Like the production model, the catch-age model also accounts for both process and observations errors. However, in this case, the model has a better basis for separating these effects because the recruitment signal appears in both catch-age and survey data; in fact, the catch-age model fits two sets of catch-age data (trap fishery and trap survey) and so has a distinct advantage over the production model, which must identify patterns of stochastic production from survey CPUE and catch alone. The catch-age assessment model also has the advantage of being similar in structure to the operating model used in simulation testing. The main difference is that the operating model employs a stock-recruitment relationship, while the catch-age assessment model only estimates average recruitment along with annual recruitment deviations from this average. The structural similarity combined with known selectivity should provide this catch-age model with a distinct advantage over the production model and the data-based procedures.

A key component of catch-age models is the selectivity function, which describes the proportion of each age-class that is available to the fishery. Typically, the selectivity function must be estimated from catch-age data at the same time as recruitment and fishing mortality rates. Thus, year-to-year variability in fishery and/or survey selectivity can mask inter-annual changes in recruitment and fishing mortality. For sablefish, selectivity can be estimated independent of the catch-age model by analyzing tag recovery patterns in relation to fish length-at-release (Appendix E, Figure E-1). A general approach to selectivity estimation from tagging is presented Myers and Hoenig (1997) with an example applied to Pacific halibut given in Clark and Kaimmer (2006), so we do not present details here. Our point is that independent estimates of fishery selectivity with respect to sablefish length can be provided to the catch-age model annually as known input parameters, much the same as, for example, growth and natural mortality rates. This approach, which is realistic as long as tagging programs continue into the future, would make it possible to relax the restrictive assumption within most catch-age models that selectivity is constant over time. On the other hand, estimated selectivity from tagged fish alone could provide biased estimates of fishery selectivity on the whole stock. A few possible sources of bias include, non-random release and recovery patterns, size-dependent reporting of tag returns (e.g., larger fish reported less frequently due to higher market value), or size-dependent tagging mortality. We are currently in the process of evaluating the tagging data and selectivity modeling approach.

2.5 Harvest control rules

A fishery harvest control rule, also known as decision rule, catch control law, feedback control rule, etc., represents a consistent procedure used to decide upon a total allowable catch (TAC) given some quantity (or quantities) such as a biomass estimate from a stock assessment. In some cases, the quantities of interest are probability distributions of certain outcomes given a range of catch limit options, which is one approach that has been used to set sablefish catch limits in the past (Haist et al. 2005). In either case, the main point of the rule is to provide a pre-defined means of changing the total allowable catch in response to changes in the condition of the stock. Without a deliberate and repeatable harvest control rule, appropriate changes to catch limits may be delayed, causing over-fishing. Indeed, previous management procedure simulation experience has indicated that there is little benefit to conducting stock assessments and changing TACs frequently in the absence of a consistent procedure (Punt et al. 2002b).

Our choice of harvest control rules to examine was determined by both stakeholder input and recent policy direction undertaken by Fisheries and Oceans Canada. DFO established national policy for harvest strategies consistent with the FAO Precautionary Approach (DFO 2006, FAO 1995) by prescribing generic harvest control rules that recognize three zones of stock abundance (Figure 3):

1. ***Healthy zone*** where the removal rate should not exceed the maximum acceptable *Removal reference*;
2. ***Cautious zone*** where fisheries management actions should promote stock rebuilding towards the ***Healthy zone*.** The removal rate should not exceed the *Removal reference*;
3. ***Critical zone*** where fishery management actions must promote stock growth. Removals by all human sources must be kept to the lowest possible level.

The *Removal reference* is the maximum acceptable removal rate, defined as the proportion of the total exploitable stock size removed by humans. Compliance with the United Nations Fish Stock Agreement requires that the *Removal reference* should be less than or equal to the removal rate associated with maximum sustained yield. The *Limit reference point* is taken to be the stock level below which the risk of impaired productivity increases to the point of serious harm, but not to the point of incurring a high risk of extinction. This definition implies an extinction reference point that remains unspecified by DFO (2006). Stock levels below the *Limit reference point* are considered to lie in the ***Critical zone***. The *Upper stock reference point* is that level where the removal rate is reduced from the *Removal reference* rate; stock levels greater than the *Upper stock reference point* are considered to be in the ***Healthy zone*** and those levels between the *Limit reference point* and *Upper stock reference point* are considered to lie in the ***Cautious zone***. The essential feature of the harvest strategy is that the removal rate is reduced from a maximum when the stock status declines below desirable levels and may be reduced to zero when stock status is at highly undesirable levels.

All harvest control rules we evaluate involve at least one key parameter that represents the removal rate, which ultimately controls the long-term average yield and stock size under a management procedure. For example, applying an exploitation rate that exceeds the production rate of the stock will eventually lead to long-term declines in stock biomass, fishery CPUE, and catch. We designate harvest rules that use a single exploitation rate-type parameter, regardless of stock biomass, as Constant Harvest Rate (CHR) rules and those that adjust exploitation rates in response to changes in stock condition as Variable Harvest Rate (VHR) rules. The latter form is particularly relevant to Canada's national policy on precautionary fishery management strategies. Specifically, VHR rules represent our implementation of the harvest strategy shown in Figure 3. The following sections formulate CHR and VHR rules that take stock assessment outputs (e.g., biomass or catch rate indices) and compute total allowable catch limits. We present the CHR rules first so that the use of stock assessment information is clear before moving on to VHR extensions. We categorize the harvest control rules generally into two classes depending on whether they use only observed data (data-based rules) or outputs from a population dynamics model (model-based rules).

Constant harvest rate rules

<u>Data-based CHR rules</u> – As noted above, the data-based method is simple enough to be described in conjunction with the harvest rule. For this paper, we narrow the definition of a data-based procedure to include only those that make no assumptions about the biological dynamics of the fish stock and thereby provide a "model-free" way to set annual quotas. One such data-based procedure computes a catch limit based on survey CPUE according to the exponentially weighted moving average formula (Harvey 1989)

$$(1) \qquad C_{T+1} = \lambda_1 C_T + (1-\lambda_1)\lambda_2 I_T^* \quad ,$$

where C_{T+1} is the catch limit for year $T+1$, I_T^* is a statistic computed from a relative abundance survey of the stock, the smoothing constant $0 \leq \lambda_1 \leq 1$ is the proportion of next year's catch limit that derives from the current one, and $\lambda_2 > 0$ is a parameter that converts the abundance index to a catch limit. In this paper, the statistic I_T^* is a 3-year moving average of the relative abundance survey catch rate. We examined other averaging windows and various weighting schemes for past surveys, but this 3-year average usually performed best at reducing inter-annual variability without compromising conservation performance. The smoothing parameter λ_1 reduces short-term fluctuations in catch by reducing the rate at which quotas are adjusted in response to changes in stock abundance. In Equation (1) is similar in appearance to the "hold-steady" harvest policy described and evaluated by Hilborn et al. (2002) for northeast Pacific rockfish (*Sebastes spp.*). However, equation (1) acts as a constant exploitation rate policy in contrast to Hilborn et al.'s formula, which is a constant escapement policy.

It is evident from equation (1) that parameter λ_2 represents an average exploitation rate that is scaled by survey catchability, i.e., $\lambda_2 = \bar{U}/q$ (Cox and Kronlund 2008). Thus, λ_2 is a key policy parameter of the procedure because it will determine the long-term stock size and yield from the fishery. Initial values for the parameters of equation (1) were determined by fitting a multiple linear regression of annual catch limits C_t on C_{t-1} and I_t^* (Figure 4). The resulting values $\lambda_1 = 0.75$ and $\lambda_2 = 299$ were treated as an upper limit because in preliminary simulations this procedure (*i*) was always the worst performing procedure in terms of depletion under all scenarios, (*ii*) lead to 40-year stock declines and fishery failure under two scenarios, and (*iii*) performed worst in average annual catch and depletion in two low productivity scenarios. Therefore, we examined combinations of $\lambda_1 = \{0.20, 0.50, 0.80\}$ and $\lambda_2 = \{120, 150, 180, 210, 240\}$ to represent both rapid to slow feedback responses to surveys and low to high average fishing mortality (Table 1).

Data-based procedures have the advantage that they are easy to understand and compute, are convenient for developing many replicates of a simulation over a wide range of scenarios and rule tuning parameters, and may often provide acceptable interim performance pending accumulation of data required to support more complex procedures.

Model-based CHR rules – Management procedures based on production and catch-age models represent more elaborate methods for setting annual catch limits that formally take into account uncertainty in both the recruitment process and the observations. Model-based procedures we consider each involve a three-step calculation of the catch limit in which (*i*) a point estimate is computed, (*ii*) uncertainty in the point estimated is computed or approximated, and (*iii*) a risk adjustment is applied. The point estimate is given by

$$(2) \qquad C_{T+1} = U^{ref} \hat{B}_{T+1} \quad ,$$

where $\hat{B}_{T+1}$ is the harvestable stock biomass projected to be present at the beginning of year T+1 and U^{ref} is a reference harvest rate. Here, the stock assessment model used to estimate and project the stock biomass $\hat{B}_{T+1}$ is a either a production model or a catch-age model as described in earlier sections. For the catch-age model procedure, $\hat{B}_{T+1}$ is the projected trap exploitable biomass. Production models estimate a single, aggregate biomass, so $\hat{B}_{T+1}$ does not represent any particular component of the stock (e.g., spawning, exploitable, survey, etc.). Step (*ii*) involves calculating a normal approximation to the Bayes posterior distribution for the catch limit. An approximation is required because both model-based estimators are non-linear, which means that the exact posterior distribution cannot be calculated directly. More accurate but computationally intensive approximations to the posterior distribution such as those obtained from Markov Chain Monte Carlo procedures are not practical in the context of management strategy evaluation at this time. Once a reasonable posterior approximation is obtained, step (*iii*) sets the catch limit equal to a pre-defined percentile of the posterior distribution. This latter step is appealing because uncertainty is incorporated into the catch limit algorithm in a direct way. For example, this approach uses the entire distribution of catch limit estimates rather than a single point such as the mode, median, or mean. Readers familiar with the management strategy evaluation literature will recognize the similarity between the latter two steps in our model-based procedure and the International Whaling Commission’s Catch Limit Algorithm (Cooke 1999).

Implementing model-based procedures requires a reference harvest rate U^{ref} and percentile Q of the posterior distribution of the catch limit. We implemented equation (2) using a range of exploitation rates $U^{ref} = \{0.04, 0.06, 0.08, 0.10\}$ which that encompasses the range of target fishing mortality rates used in both U.S. and Canadian sablefish assessments

(Haist et al. 2005; Hanselman 2006). We examined $Q = \{0.4,\ 0.5\}$ with the first value representing a risk-averse choice and the second a risk-neutral choice.

Variable harvest rate rules

Both the data-based and model-based procedures can be rendered compliant with DFO's precautionary approach to fisheries management (Figure 3; DFO 2006) by introducing variable harvest rate decision rules that reduce levels of exploitation if a stock declines below certain threshold and limit reference points. For model-based harvest policies this means adjusting the harvest rate U^{ref} downward if the estimated stock size decreases below some fixed, reference level. Similarly, we could reduce parameter λ_2 of the data-based rules as the survey moving average statistic I_T^* approaches some lower limit values. The next sections use terminology from Figure 3 to develop variable harvest rate rules.

Data-based VHR rules – We modified the data-based procedure to accommodate variable harvest rate decision rules by specifying standardized survey catch rates $\{I_{low}, I_{high}\}$ that define the *Limit Reference Point* and *Upper Stock Reference* points as displayed in Figure 3. Here, we assume that the 3-year average of the standardized survey is a reliable indicator of *Stock Status*, or indeed the weaker assumption that management of the fishery should respond to a 3-year trend in the index. The data-based exploitation rate parameter λ_2 is the *Removal Rate* when the *Stock Status* is above the *Upper Stock Reference*. This rule is implemented by computing $\tilde{\lambda}_{2,T+1}$ as survey catch rates change, i.e.,

$$(3) \qquad \tilde{\lambda}_{2,T+1} = \begin{cases} 0 & I_T^* < I_{low} \\ \lambda_2 \left(\dfrac{I_T^* - I_{low}}{I_{high} - I_{low}} \right) & I_{low} \le I_T^* < I_{high} \\ \lambda_2 & I_T^* \ge I_{high} \end{cases} .$$

The adjusted parameter $\tilde{\lambda}_{2,T+1}$ is then used in the data-based harvest control rule (Equation 1) in place of λ_2. We tested Limit Reference Point values, $I_{low} = \{3, 4\}$, which correspond approximately to the lowest standardized survey catch rates observed in 2001 when DFO and industry indicated their concerns about stock status and incurred an in-season reduction in the quota in 2002 (CSAS 2002, Fisheries and Oceans Canada 2002). Upper Stock Reference point values $I_{high} = \{10,15\}$ were chosen to reflect possible lower bounds on a "healthy" stock. For example, mean survey CPUE in the 1990s (excluding the extremely high value in 1993) was approximately 15 kg/trap. Although these reference values, along with $\{\lambda_1, \lambda_2\}$, could be optimized in a full management strategy evaluation, our intent here is to introduce one possible choice that is consistent with Canada's national policy.

Model-based VHR rules – Model-based decision rules that use variable harvest rates must similarly include *Limit Reference* point and *Upper Stock Reference* point values, which we call $\left\{D_{low}, D_{high}\right\}$, respectively. The symbol *D* denotes stock depletion relative to some predefined condition, which we use as the indicator of *Stock Status*. For model-based procedures where reference points are set relative to B_0, depletion is typically estimated by stock assessment models as the current fraction of the unfished biomass, i.e., estimated values of depletion near zero indicate stock extinction and those near one indicate an unfished stock.

For model-based VHR rules, the exploitation rate is set at U^{ref} when estimated *Stock Status* is above the *Upper Stock Reference* point and to zero at *Stock Status* levels below the *Limit Reference*. For estimated depletion between these reference points, the harvest rate is reduced via the function (Hilborn et al. 2002)

$$(4) \qquad U_{T+1} = \begin{cases} 0 & \hat{D}_T < D_{low} \\ U^{ref}\left(\dfrac{D_{high}}{\hat{D}_T}\right)\left(\dfrac{\hat{D}_T - D_{low}}{D_{high} - D_{low}}\right) & D_{low} \le \hat{D}_T < D_{high} \\ U^{ref} & \hat{D}_T \ge D_{high} \end{cases} .$$

Note that this harvest control rule is similar in form to the U.S. $F_{40\%}$ with 40-10 adjustment, which has a non-linear change in U_{T+1} with changes in stock status rather than the rectilinear form of the Canadian policy shown in Figure 3.

Our production model procedure estimates the unfished equilibrium biomass, B_0 and therefore estimates of depletion $\hat{D}_t$ are of the projected stock size relative to unfished conditions in 1965. Initial reference levels ($D_{low} = 0.1$ and $D_{high} = 0.4$) for production model VHR rules were chosen to mimic the U.S. $F_{40\%}$ with 40-10 adjustment harvest rule.

The catch-age model does not estimate unfished biomass. Instead, this stock assessment approach estimates biomass beginning in 1992, which is the start of reliable age composition sampling for the survey. We use this 1992 biomass as the "predefined condition" for determining stock depletion catch-age model-based management procedures; that is, the catch-age model estimates stock depletion by dividing the current stock assessment estimate of biomass by the estimated biomass in 1992, i.e., $\hat{D}_t = \hat{B}_t / \hat{B}_{1992}$. An estimated depletion value of $\hat{D}_t = 1.0$ means that the estimated Stock Status in year *t* is at the 1992 level. This approach of using recent points on the biomass trajectory is often more robust than estimating depletion relative to an unfished biomass, which is often poorly estimated (Punt et al. 2002b). Our initial choice of $D_{low} = 0.25$ for the Limit Reference Point corresponds approximately to the 2001 biomass, which was an all-time low sablefish biomass level acknowledged by the fishing industry. The Upper Stock Reference $D_{high} = 1.0$ corresponds to the biomass in 1992. Note that depletion levels estimated by the catch-age modeling approach are only used within management procedures and they should not be confused with true depletion levels in the operating model, which we use to determine conservation performance.

Allocation among gear-types

Once a catch limit is determined, it is then allocated among trap, longline, and trawl fisheries in the same proportion as occurred in 2006. This may not be realistic in the future because new regulations designed to reduce bycatch wastage and to promote greater accountability may affect how the final catch is distributed over gear-types (Koolman et al. 2007). However, the choice is reasonable until potentially new patterns of catch distribution among gear sectors emerge from the revised management regime. We recognize that this issue is important because different gear-types exploit different age-groups of sablefish and thus have differential impacts on the stock. In particular, gear types that concentrate on younger fish can intercept recruitment and potentially affect the fishing success of gears that target older individuals (Sinclair 1993).

3 Scenarios for B.C. Sablefish

3.1 Operating model

Regardless of type, candidate management procedures must be tested for robustness against known or potential sources of uncertainty and risk (Cooke 1999). This key step in the evaluation process implements many aspects of the precautionary approach, risk assessment, and risk management. Obviously, performing robustness tests on real fishery systems would not be precautionary because the results would take a long time to accumulate and risks to the fishery and the stock could be great. Therefore, most management procedures are evaluated in computer simulations that attempt to mimic the decision-making process and fishery dynamics in a realistic way. Defining “realistic” is often a challenge because many aspects of fish biology and fisheries ecology are difficult to observe, let alone understand well enough to implement in computer models. Nevertheless, testing management procedures against some well-known biological and fishery scenarios is a critical step in developing precautionary management procedures.

In this paper, we use the terms operating model and scenario interchangeably to represent the “true” simulated sablefish population, fishery, and data generation mechanisms. Operating models are used to simulate population and fishery dynamics that may be realized when a given management procedure is implemented. A single management procedure may be tested against multiple operating models where each is designed to represent a particular uncertainty or combination of uncertainties. A robust procedure is one that produces outcomes that meet predefined management objectives under any scenario that is considered plausible. The suite of operating models for a given fishery is called the reference set. As work progresses on a particular management strategy evaluation, operating models may be removed from the reference set if their effects are determined to be unimportant or implausible, while other operating models may be added based on new concerns about the stock and fishery. In this sense, management strategy evaluation is a process for continuous scrutiny of the management system rather than a static end-product. It is important to note that the goal of management strategy evaluation is to produce a stable decision-making process, so addition and deletion of scenarios should not always alter short-term decisions.

Candidate management procedures for sablefish were tested against alternative configurations of the age-structured population dynamics operating model specified in Appendix C. All configurations of the operating model assume that the B.C. sablefish spawning biomass was at unfished, deterministic equilibrium prior to directed fisheries in the mid-1960s. The operating model further assumes that the B.C. population is closed to immigration and emigration. The four scenarios we chose to consider for this paper result from setting two uncertain factors at two levels each, namely (*i*) stock productivity represented by two assumptions about the value of the operating model's stock-recruitment steepness parameter, and (*ii*) current stock status, which results from two assumptions about the relationship between trap fishery CPUE and stock abundance. The operating model scenarios, denoted S1 to S4 (Table 2), were parameterized by fitting the operating model to existing standardized survey, trap fishery, tagging, and catch-age data with operating model parameters either fixed at certain values or estimated. Conditioning the operating models in this way allowed us to maintain consistency between the historic data and the simulated future data. It is important to note that the operating model used to generate scenarios for the stock contains a number of fixed parameters such as growth, maturity, and natural mortality rates that may require specific scenarios because at least some will influence the long-term outcomes of management procedure simulations.

Uncertainties regarding fish movement, bycatch, and the spatial structure of the sablefish stock(s) were also key considerations raised in joint science-stakeholder-management workshops; however, all of these issues are beyond the scope of this paper. The following two sections describe in detail the rationale for our choices of operating model configurations for the key uncertainties of productivity and present stock status.

3.2 Key uncertainties – Stock productivity

Uncertainty related to stock productivity arises for two reasons. First, the B.C. sablefish fishery has taken a relatively steady average catch since the 1970s while fishery and survey catch per unit effort has with few exceptions remained steady or declined. Such a "one-way trip" (Hilborn and Walters 1992) pattern does not allow us to easily distinguish between a high unfished biomass combined with low productivity and low unfished biomass combined with high productivity. Second, estimates of stock productivity depend on what we assume about the natural mortality rates of sablefish. Unfortunately, the natural mortality parameter is very difficult to estimate for most fish stocks (Schnute and Richards 1995).

Alternative productivity assumptions can be represented in the operating models by adjusting the steepness of the stock-recruitment relationship, h, which is defined as the fraction of the unfished recruitment level that occurs when the spawning stock biomass is reduced to 20% of the unfished level. A steepness value near $h = 1.0$ means that recruitment is about the same as unfished when the spawning stock is reduced to only 20% of its unfished level. In an analysis of more than 700 stock-recruitment data sets, Myers et al. (1999) found that recruitment steepness averaged $h = 0.69$ over a wide range of fish families. Sablefish, which were included in the study, had the lowest steepness value in the entire study at $h = 0.26$ (which we ignored). Our estimates of recruitment steepness based on fitting the operating model to the above datasets are either $h = 0.49$ (± 0.11) or $h = 0.56$ (± 0.16) depending on assumptions about how well trap fishery CPUE reflects stock biomass (see below). Therefore, we chose $h = \{0.45, 0.65\}$ to bracket these values in the operating model.

3.3 Key uncertainties – Present stock status

The second factor distinguishing operating models is the current status of B.C. sablefish relative to average unexploited conditions. Similar to many stocks around the world, sablefish biomass estimates for the first two decades of commercial fishing depend strongly on fishery catch-per-unit-effort (CPUE). Obviously, there are clear dangers involved in using CPUE as an unbiased index of stock abundance (Hilborn and Walters 1992), especially over long time periods like 1970 – 2000 during which rapid evolution in fishing technology occurred. On the other hand, ignoring CPUE completely leaves a very short time-series of fishery-independent information that provides unreliable, and usually optimistic, estimates of unfished conditions and current stock status. As a compromise, we fitted two versions of the operating model to fishery CPUE in combination with surveys and catch-at-age. In the first, we fixed the trap fishery CPUE hyperstability parameter $q_{2,trap} = 1.0$ in the operating model, which assumes direct proportionality between trap fishery CPUE and exploitable sablefish biomass over all biomass levels. Under this scenario, estimated 2006 spawning biomass is 29% and 31% of the deterministic unfished equilibrium for recruitment steepness values $h = 0.45$ and $h = 0.65$, respectively. In our second approach, we admitted the possibility that CPUE could remain high and stable (i.e., hyperstable; Hilborn and Walters 1992) over a wide range of sablefish biomass. We implemented this assumption by treating hyperstability as a free parameter in the operating model. The estimated hyperstability parameter $q_{2,\mathrm{trap}} = 0.37$ (± 0.12) suggested that this effect could be substantial. The estimated 2006 spawning stock biomass corresponding to this hyperstability assumption is 18% and 20% of the unfished level for recruitment steepness values of $h = 0.45$ and $h = 0.65$, respectively (Table 2). Both scenarios are important because industry stakeholders are skeptical about data from the early fishery due to the systematic biases associated with hyperstability as well as lack of consistency between model results and personal experience during the 1970s and 1980s (e.g., biomasses in operating model fits appear too high in the 1970s). All classes of data-based and model-based procedures we evaluated ignored data collected before 1990, and thus present opportunities to deal with these concerns.

3.4 Perfect information management procedures

The four scenarios we used to test candidate management procedures differ in their fundamental population dynamics parameters, which means that equilibrium relationships between spawning stock biomass and yield also differ. Therefore, we developed a "perfect information" procedure for each scenario to provide a benchmark against which any other procedures could be compared. The perfect information procedure results are particularly useful for judging the impacts of stock assessment model errors on performance. Catch limits for perfect information procedures were computed using $C^*_{T+1} = U_{MSY} B_T$, where U_{MSY} is the true optimal exploitation rate and B_T is the true trap exploitable biomass in year T. These were always applied using CHR rules. Values of U_{MSY} are only affected by stock productivity; therefore, the operating model values $U_{\mathrm{MSY}} = \{0.045, 0.079\}$ represent the optimal exploitation rates of scenarios (S1, S2) and (S3, S4), respectively. Uncertainty in present stock status is revealed in the operating model parameters as uncertainty in the unfished spawning biomass estimates, which differ among all four scenarios. Therefore, operating model values of MSY = {2913, 3059, 4293, 4529} correspond to maximum sustainable yields for scenarios S1 to S4, respectively (Appendix C, Table C-4).

4 Performance measures

Harvest policies are typically evaluated based on three main performance categories: catch, catch variability, and conservation. The time horizon over which performance statistics are computed is also important because trade-offs among the three main categories tend to change over time. Thus, each performance statistic described below was computed for 1 – 5, 6 – 10, 11 – 20, and 21 – 40 year periods. Specific calculations are given in Table 3. Catch performance for each simulation is summarized by the average annual catch and the lowest 5th percentile of catch, the latter serving as an indication of "guaranteed" catch. Catch variability is summarized by the annual average of absolute variation in catch (*AAV*; Punt 2002b), which measures the average relative deviation in catch from year to year regardless of the direction of change (hence "absolute"). Conservation performance is measured by (*i*) average spawning stock biomass depletion ($\overline{D}$) and (*ii*) the proportion of years (P_{cons}) in which the spawning stock biomass remains above 20% of the unfished equilibrium level as defined by each scenario at any time during a replicate. Trap fishery CPUE, which stakeholders view as a measure of stock health in addition to a measure of relative profitability, is computed as a performance measure, but is not discussed in detail here. Performance statistics are summarized across 50 simulation replicates using medians of the above statistics, which were chosen to reduce the effects of extreme values.

Examination of the trade-off relationships among fishery performance indicators is critical to the evaluation process. However, presentation of trade-off relationships between, for example, median average catch and median average spawning biomass depletion are difficult to depict for the large number of scenario and management procedure combinations we examined. Therefore, we develop graphical presentations for hypothetical outcomes in a step-wise fashion here to aid interpretation of results reported below. Figure 5 is constructed using three panels to illustrate increasing complexity of comparisons among procedures and scenarios. Figure 5a shows the difference in trade-off performance between catch-age model procedures CA1 and CA3, where CA3 achieves a higher depletion value by 0.2 along the conservation *x*-axis, but gives up 2000 t of average annual catch along the yield *y*-axis. The relative performance of three data-based procedures (DB1-DB3, open symbols, dotted line) and three catch-age procedures (CA1-CA3, filled symbols, solid line) is shown in Figure 5b. The vertical dotted line helps to indicate that hypothetical procedure CA2 achieves 1000 t greater average annual catches than procedure DB2 for the same level of stock depletion. This difference in yield is less pronounced when comparing procedures CA3 and DB3, where DB3 achieves about the same depletion value as CA3 at the expense of 300 t reduction in average annual catches.

Figure 5c represents the relative performance of the same three catch-age procedures (CA1-CA3) but under two different hypothetical stock scenarios, S1 and S2. The initial depletion levels at the start of the simulation for S1 (black) and S2 (red) are indicated by the inverted triangle symbols along the upper *x*-axis of the panel. The trade-off lines between median average catch and median average depletion are similarly colour-coded to correspond to S1 and S2. For S1, management procedures CA2 and CA3 maintain or grow the stock relative to the initial depletion. Procedure CA1 declines the stock by 0.1 depletion units but achieves the highest average annual catches. Scenario S2 represents a less productive stock that starts the simulation at a lower initial depletion. Here procedures CA1 and CA2 both decline the stock from the initial depletion level. Only CA3 achieves stock growth, but this requires a sizeable reduction in average annual catches relative to the other candidate procedures.

5 Results

Although we evaluate a limited combination of management procedures and scenarios, the shear volume of simulation outputs can overwhelm the limits of interpretation when presented en masse. To facilitate comparisons among data-based, production model, and catch-age model management procedures, we present results for three "selected" management procedure classes, where selection was based on the following criteria:

(i) Management procedures must comply with Canada's national policy mandating precautionary variable harvest rate control rules (DFO 2006);
(ii) Candidate procedures must not generally result in further stock declines over 40 years under the worst-case scenario (S1).

Although the first criterion precludes consideration of constant harvest rate decision rules, many of these would be eliminated under criterion (*ii*) anyway. Table 4 provides the details of the candidate management procedure specifications within these three broad classes. Appendix F provides the full set of trade-off comparisons for all procedures considered in this document and full tables of performance statistics for the selected procedures (tabular results for all procedures evaluated are available upon request).

5.1 General trends in fishery and conservation performance

The transition between historical management outcomes and those simulated from application of management procedures was smooth for data-based procedures and quite abrupt for model-based procedures. For example, all catch-age model procedures gave large immediate reductions in catch from approximately 4,500 t in 2006 (actual outcome) to 1,200 – 2,000 t in 2007 (simulated outcome). These changes from existing catch levels arise because the stock assessment models both estimate low biomass from the existing 1990 - 2006 (1992 - 2006 for the catch-age model-based procedures) data regardless of the initial biomass for the scenarios (discussed further below). Such changes in catch limits are outside the 15 – 20% inter-annual catch variability objectives set by industry stakeholders despite potentially better performance of these procedures in the long-term. Therefore, we imposed a maximum 15% annual change constraint on catch limits derived from data-based, model-based, and perfect information procedures over the first 5 years of the simulations only. This tactic increases short-term conservation risks because catches are prevented from decreasing appropriately in the event of a rapid stock decline. Performance statistics for the first five years of the projections must therefore be interpreted with the understanding that range of short-term (5-year) outputs under each procedure is (usually) truncated by the constraint. For example, the catch-age model procedures would have reduced catches below that reported for the simulations during the first five years, occasionally to the point of reducing catches to zero under some VHR rules. Restricting the inter-annual change in catch causes lower depletion over the long term and longer rebuilding times from low initial depletion. Furthermore, under the maximum 15% annual change constraint, average catches are generally higher over the short-term than would have been achieved by strict application of the procedures. Low catches are maintained for a sustained period following the decline particularly for the low productivity scenarios. The consequent trade-off along the conservation axis that results from maintaining catch higher than specified under each procedure is, however, captured by performance statistics such as average depletion.

Most management procedures employing VHR rules lead to long-term increases in both the sablefish stock and catch, although the particular paths taken vary among management procedures. Figure 6 shows example trends in stock biomass and catch for one individual simulation replicate drawn from scenario S3, which represents a productive stock that is depleted to 20% of unfished biomass when the management procedures are first implemented in 2007. Under these conditions, data-based ($\lambda_2 = 180$; Figure 6a), production model ($U^{ref} = 0.08$; Figure 6b), and catch-age model ($U^{ref} = 0.08$; Figure 6c) procedures all provide sustained growth of the stock as indicated by the increasing biomass trajectories in the left panels. Note, however, that in all cases long-term increases in catch follow an initial period of declining catches during the first five years of management procedure implementation. The data-based and catch-age model procedures provide the greatest average catch, while the production model procedure maintains the stock at the highest depletion level because it consistently underestimates stock size for this particular replicate. When these same procedures are applied to scenario S1, where the stock is less productive and the initial depletion is low (Figure 7), the median stock level remains steady under the data-based procedure and declines slightly for both model-based procedures. Furthermore, in all cases the median stock level is never above the conservation reference point of $0.2B_0$. Both model-based approaches tend to over-estimate the spawning biomass in this simulation (Figure 7 b,c) as reflected by the retrospective biomass estimate traces. Note the large uncertainty in unfished biomass estimates for the production model (Figure 7b) and relatively precise biomass estimates for the catch-age model for 1992 (although with slight positive bias). The catch-age model over-estimates spawning biomass by up to one-third while the stock is declining between years 40 and 50. At about year 64 of the simulation, both model-based approaches make a relative sharp TAC adjustment downward, which reflects both procedures invoking VHR rules in response to the lowest biomass estimates of the simulation. The data-based procedure (Figure 7a) shows a much smaller decline in catch relative to the model-based approaches in these years.

Catch performance among management procedures also differed in expectation and range over simulation replicates. Figure 8 presents three individual catch trajectories (i.e., catch replicates) for scenario S1, median annual catch over 50 replicates, and the range containing the middle 90% of the annual catches for six management procedures representing the most conservative (Figure 8 a,c,e) and most aggressive (Figure 8 b,d,f) within each class. There is considerable variation among individual catch trajectories even though the underlying operating model dynamics use identical sequences of stochastic recruitment and observation errors. Differences arise only from applying the alternative management procedures where future catches for an individual replicate depend on the realized recruitments and consequent management actions taken during the earlier course of the projection. For example, although the three classes of procedure exhibit approximately synchronous responses to recruitment events (i.e., the main peaks of individual replicates reflect recruitment inputs), they each show unique lag effects, inter-annual variability, short-term responses (< 10 years), and long-term mean catches.

Median catch tends to change smoothly over time for the data-based procedures, and a bit more abruptly for the model-based procedures. Effects of the maximum 15% annual change constraint during the first 5 years varies among the procedures, with the model-based procedures showing a strong tendency to reduce annual catches by the maximum permitted 15%. In effect, the low initial depletion of the stock (0.18) for scenario S1 and low productivity means that in the absence of the 15% constraint, a large immediate reduction from current catches levels would result from application of most procedures. The onset of stock growth and increasing catch occurs earlier for the model-based procedures compared to the data-based procedures because catches were reduced further early in the management period.

Differences among management procedure outcomes were similar under the more productive scenario S4 (Figure 9) although the overall variability and mean catch levels were greater. Initial reductions in annual catch were less severe for all six procedures compared to the unproductive scenario S1. The more rapid onset of an increasing catch trend is in part due to higher stock productivity and a higher level of depletion at procedure implementation. For the data-based and catch-age model procedures, annual catch reductions over the first 5 years are often less than the 15% maximum as shown where the range containing 90% of the annual catches spreads away from the median early in the simulations.

5.2 Trade-off relationships between yield and conservation

Trade-off relationships between yield and conservation are an important way to identify "better performing" management procedures. Ideally, the best procedures provide high annual catches and while also maintaining the stock at highly productive levels. Here we examine the form of these trade-off relationships for the selected set of procedures (Figure 10). Construction of the trade-off figures is based on the concepts illustrated by Figure 5 of Section 4. Each panel of Figure 10 shows the trade-off relationship between median average catch and median average depletion over 50 simulations for each of the selected procedures and all four operating model scenarios. Outcomes for each specific management procedure are plotted using the same symbol over 1 -5, 6 – 10, 11 – 20, and 21 – 40 year time periods. Lines connecting procedure symbols are colour-coded to each scenario, where scenario S1 = black, S2 = red, S3 = green, and S4 = blue. For each scenario, the initial depletion prior to the start of the management procedure is indicated by an inverted triangle positioned along the upper *x*-axis using the appropriate scenario label and colour. This visual aid is intended to serve as a reference to allow easy judgment of the change in stock depletion under each procedure relative to conditions at the start of the simulation. Finally, different line styles are used to distinguish classes of procedures.

Over the range of procedures and scenarios examined, model-based and data-based management procedures tended to trade-off catch and conservation in similar ways. This is evident by examining the form of trade-off relationship between median average catch and median average stock depletion under each combination of management procedure and stock dynamics scenario (Figure 10). In the short-term, trade-offs were relatively steep regardless of procedure type or scenario indicating that large reductions in average catch would provide only small improvements in stock depletion (Figure 10 a,b). This reflects the fact that the 2006 operating model spawning biomasses are below the estimated maximum sustainable yield levels for all scenarios (Appendix Table C-4) and recent catches remove most of the surplus production available for stock growth. Therefore, even moderately higher levels of catch in the short term lead to stock declines and lower depletion, while lower catches lead to stock growth and higher depletion (i.e., larger stock size). In the medium- and long-term, however, trade-off relationships "flattened" indicating less severe trade-offs mainly because most management procedures caused increases in the stock under most scenarios (Figure 10 c,d).

As expected, perfect information procedures caused relatively rapid rebuilding to average depletion levels near B_{MSY} for high productivity scenarios S3-S4. Over the first 20 years, perfect information procedures obtained higher average annual catch, while maintaining lower average depletion levels than data-based and model-based procedures. Perfect information procedures generated the opposite pattern for low productivity scenarios S1-S2; that is, lower average annual catch and higher average depletion compared to other procedures. In all scenarios except S1, perfect information procedures rebuilt average depletion to B_{MSY} levels by the end of 40 years.

Scenario 1 – low productivity/low initial depletion

Under the low productivity/low depletion scenario (S1), all selected procedures result in stock declines over the first 10 years and increases over the last 20 years of the simulations. Declines in stock size during the first five years result from recent recruitment patterns (e.g., 2000 – 2006) prior to implementation of management procedures and the 15% constraint on catch changes, both of which are common to all procedures (Figure 10; Table 5).

Management procedure classes differed considerably in how catch-depletion trade-offs changed over time. For example, despite greater average catch and lower depletion during the first 10 years following procedure implementation (Figure 10 a,b), most data-based procedures resulted in 21 – 40 year average catch and depletion that were equivalent to, or greater than, model-based procedures (Figure 10d). Over 21 – 40 years, the catch-age model VHR procedure with $U^{ref} = 0.04$ produced the lowest average annual catch, but highest average depletion and was very similar to the perfect-information procedure (note that $U_{\text{MSY}} = 0.045$ for this scenario). The data-based VHR procedure with $\lambda_2 = 120$ produced slightly greater average annual catch at slightly lower depletion; however, this procedure was actually closer to perfect-information than the catch-age model with $U^{ref} = 0.04$. We did not examine a production model-based procedure with $U^{ref} = 0.04$.

Variable harvest rate rules helped to stabilize the stock trend or initiate stock increases during the 40-year projection period even for the high exploitation rate procedures. For example, despite the fact that the exploitation rate under the catch-age $U^{ref} = 0.10$ procedure is more than twice U_{MSY}, the spawning biomass reached the initial 2006 level by the 21 – 40 year period. Similarly, even though the production model $U^{ref} = 0.10$ VHR procedure used an aggressive exploitation rate that applied to an incorrectly defined stock, this procedure resulted in little change in spawning stock biomass as measured by average depletion over 21 – 40 years. Based on final depletion statistics, both of these aggressive strategies actually resulted in slight stock increases under VHR rules, although the catch-age procedure provided a higher probability of keeping the stock above 20% of the unfished level (Table 5). As mentioned above, VHR rules did not appear to hinder the progress of stock rebuilding and catch utilization for catch-age $U^{ref} = 0.04$ or data-based $\lambda_2 = 120$ procedures because both closely tracked the perfect-information procedure in terms of the catch-depletion trade-off (Figure 10 d).

Scenario 2 – low productivity/high initial depletion

Under the low productivity/high initial depletion scenario (S2), again, all procedures cause stock declines over the first 5 – 10 years (Figure 10 b); however, under this scenario two catch-age procedures left the stock in worse condition by 21 – 40 years compared to the initial depletion level. The catch-age procedures with $U^{ref}= \{0.08, 0.10\}$ both caused the stock to decline further from the 2006 level based on both the average and final depletion values (Table 6). Such declines reflect the fact that the equilibrium spawning biomass depletion in the operating model is lower than 30% for these exploitation rates. In fact, spawning biomass depletion levels under $U^{ref}= \{0.08, 0.10\}$ are 15% and 4%, respectively for scenario S2 (Appendix Figure C-1). Although both exploitation rates are excessive for this scenario, VHR rules act to stabilize the biomass somewhere between 1992 and 2006 levels (recall that VHR rules for catch-age procedures use 1992 and 2001 as high and low biomass reference points, respectively). All data-based procedures except $\lambda_2 = 240$ and all production model-based procedures result in stock increases beyond 2006 levels by 21 – 40 years.

The data-based $\lambda_2 = 150$, catch-age model with $U^{ref} = 0.04$, and production model with $U^{ref} = 0.06$ consistently tracked the perfect-information case in this scenario. All of these procedures left the stock within ± 5% of the MSY depletion level by 21 – 40 years. Interestingly, although the data-based procedure provided considerably greater average annual catch over 21 – 40 years, the production model procedure produced greater "guaranteed catch" compared to the others (Table 6; $C_{5\%}$). This pattern in which the production model procedures performed better in terms of $C_{5\%}$ despite lower average annual catches was repeated for other scenarios unless the average catch was much lower than other procedures.

Scenarios 3 and 4 – high productivity

Similar to the low productivity scenarios, the stock declined during the first five years for high productivity scenarios regardless of initial conditions (Figure 10a; Table 7; Table 8); however, the declines are less significant and do not last as long. The main difference is that under high productivity scenarios, most procedures allow for stock growth beyond 2006 levels within the 6 -10 year period (Figure 10b). Model-based procedures provide faster initial stock growth because initial catch levels are lower than for data-based procedures. Between 5 and 40 years, the model-based procedures also increased both catch and depletion at the same time, with catch-age procedures providing greater catches and production models providing greater stock growth. However, by 21 – 40 years, the trade-off lines for all procedure classes were very similar (Figure 10d).

Production models frequently under-estimated exploitable biomass during stock increases and therefore under-exploited the growing stock. For example, all production model procedures provided equivalent or higher depletion than the perfect-information procedures in both scenarios S3 and S4. The 21 – 40 year trade-offs for the catch-age procedures with $U^{ref} = 0.08$ (note $U_{\mathrm{MSY}} = 0.08$ in these scenarios) also provide higher depletion and slightly lower catch than the perfect information cases (Figure 10d). Similar to the production model, the catch-age model tends to under-estimate biomass during stock increases, which tends to cause under-exploit the stock relative to perfect-information. Both procedures also exploit the stock at slightly lower rates than indicated (e.g., $U^{ref} = 0.10$) when VHR rules reduce target exploitation rates. Causes of under- and over-estimation of spawning biomass by catch-age and production models are discussed in the next section.

5.3 Individual performance metrics for selected management procedures

Catch variability

Most management procedures met inter-annual catch variability objectives by maintaining fluctuations in catch at less than 15 – 20% per year over 11 – 20 and 21 – 40 year time horizons (Figure 11). The maximum 15% change constraint ensured that all procedures met *AAV* criteria over the first 5 years. Note that, although any particular simulation replicate might have exceeded 20% variability from year-to-year, the expected variability as represented by the median of *AAV* values over 50 replicates were always lower than 20%. However, even the most extreme *AAV* results for most procedures were less than 20%.

Approximately half of the inter-annual variability in catch is controlled by random variability in the surveys, age-composition data, and stock assessment models with the other half controlled by fluctuations in the stock biomass and recruitment. This is evident by comparing the *AAV* values that result from the perfect-information procedure (Figure 11 - "True") with model-based or data-based values. The median *AAV* values for perfect-information cases are approximately 5% per year on average whereas the candidate management procedures, in general, show *AAV* values ranging from 7 – 14%.

Inter-annual variability in catch decreased for larger, more productive operating model biomass scenarios (Figure 11). This is likely the case because at higher levels of biomass, variability in recruitment has less of an influence on survey variability and model biomass estimates. Also, changes in target exploitation rates resulting from VHR rules occurred less often. Production model procedures all showed greater inter-annual variability in catch than data-based procedures (Figure 11a, b).

Patterns of inter-annual variation in catch within-scenarios also differed among management procedure classes. In general, higher exploitation rates tended to lead to higher variability in catch compared to less aggressive policies under the same scenario and procedure class (Figure 11b).

Average annual catch

In general, average annual catches were higher for the large, more productive operating model biomass scenarios (Figure 12). Variations around the median average annual catch represent among-replicate variability, which is much larger than what is measured by *AAV* because replicates differed in both the timing and magnitude of simulated recruitment variability. For example, by chance, some replicates have runs of good recruitment years that lead to high biomass and high average catch, while other replicates have sequences of low recruitment, resulting in low biomass and catch. Therefore, whereas variability in catch limits among years (*AAV*) is driven approximately equally by recruitment variability and random observation errors within replicates, the variation in average annual catch among simulation replicates is driven by variation in biomass. This is particularly evident by similarities between the range of average catch variability for the "True" cases and the management procedures.

All procedures tended to over-exploit the stock under scenario S1 and under-exploit the stock under S4 when compared to the "True" procedures. Note that presentation of the average annual catches in Figure 12 mainly highlight the uncertainty in future average catch levels. Absolute differences among procedures cannot be interpreted in the absence of knowledge about the corresponding depletion trade-off. The only reasonable comparisons to make from Figure 12 are between each procedure and the perfect-information results, which apparently make the appropriate depletion trade-off for each scenario. For data-based procedures, these differences are caused by inappropriate initial choices of scaled exploitation rates λ_2 and VHR rules, whereas for model-based procedures, the differences arise from incorrect exploitation rates, persistent biomass estimation errors, and VHR decision rules (compared to CHR for "True"). The effective exploitation rates of data-based procedures depend upon the relationship between the true scenario biomass and the observed survey catch rates. It is evident from our earlier definition of the scaled exploitation rate for these data-based procedures that the effective exploitation rate is the product $q\lambda_2$. For the low initial depletion scenarios (S1 and S3), these effective exploitation rates are higher than for high initial depletion scenarios. For example, in scenario S1 the exploitation rates were $q\lambda_2 = \{0.043, 0.054, 0.065, 0.076, 0.087\}$ compared to $q\lambda_2 = \{0.030, 0.037, 0.045, 0.053, 0.060\}$ for the high initial depletion scenario S4. These scenarios also represent low and high productivity with corresponding optimal exploitation rates $U_{\mathrm{MSY}} = \{0.046, 0.083\}$, respectively. Therefore, it appears that $\lambda_2 = \{210, 240\}$ should over-exploit in S1 and under-exploit in S4.

The over-exploitation pattern by model-based procedures relative to the "True" cases in scenario S1 is caused mainly by the fact that reference exploitation rates are greater than U_{MSY} for this scenario. Under high productivity scenarios S3 and S4, however, procedures with $U^{ref} = 0.08$ or $U^{ref} = 0.10$ both under-exploited the stock compared to the "True" case even though $U_{\mathrm{MSY}} = 0.08$. In these cases, differences were caused mainly by persistent biomass estimation errors that arise as an interaction between the exploitation rate and the stock assessment model.

Procedures using exploitation rates greater than U_{MSY} sometimes generated long-term average catches greater than MSY. Such a counter-intuitive result arises mainly because catches were averaged over years 21-40, which represents slightly more than one sablefish generation (approx. 13-14 years). Averaging catches over much longer time frames would show more realistic consequences of long-term over-exploitation.

Although our catch-age model exhibited low uncertainty and overall bias in spawning biomass estimates, individual years tended to be systematically biased with respect to the true stock size. Analysis of retrospective biomass estimate traces revealed that catch-age stock assessment model biases depended upon (*i*) the initial depletion level and (*ii*) the direction of change in stock size (Figure 13). Under all scenarios, but particularly for high depletion scenarios S2 (Figure 13 d – f) and S4 (Figure 13 j – l), the catch-age model underestimated the operating model spawning biomass by a wide margin at the beginning of the simulation. This reflects the fact that the catch-age model interprets the existing 1992 – 2006 data differently than the operating model. Catch-age assessment model estimates of the biomass rely only on the commercial fishery and survey age composition and standardized survey index beginning in 1992, whereas the operating model used a larger data set that included fishery CPUE, tagging biomass, and all available age-proportions. Thus, it is not surprising that the assessment model provides a different interpretation of the stock trajectory. In fact, the catch-age model tends to agree more with the low depletion operating models as shown by the smaller initial biases for scenarios S1 and S3 (Figure 13 a – c and g – i). These differences are the main cause of very low catch limits under the catch-age model procedures during the first 5 years, and thus the necessity for the maximum 15% annual change constraint.

Regardless of the initial differences, the catch-age model showed retrospective bias patterns that are typical of stock assessment models. For example, during periods of stock stability or decline (Figure 13 a,d,j) the catch-age model tended to over-estimate spawning biomass, whereas biomass was under-estimated when the stock increased (Figure 13 b,e,h,k). Biases were greatest when stock increases or decreases were most rapid. When the stock increased and then declined, the catch-age model first under-estimated and then over-estimated the biomass trajectory, respectively (Figure 13 c,f,i,l). This retrospective bias pattern was repeated consistently, although to varying degrees, across all simulation replicates and scenarios. It arises because the average recruitment estimated by the catch-age model typically lags behind the operating model stock dynamics. For example, when the stock is at a low level, the catch-age model logically estimates low average recruitment; however, when recruitment increases, it takes several years for the catch-age model to detect the change. In fact, especially under productive scenarios, the continuously increasing biomass means that the catch-age model never actually "catches up" to the true stock dynamics. Therefore, even procedures that use reference exploitation rates that exceed U_{MSY} under-utilize the stock. Although over-estimation of biomass during stock declines leads to a potential positive feedback where declines in stock size cause further over-estimation biases and declines, the use of VHR rules essentially breaks this pattern and limits long-term stock declines even when the exploitation rates far exceeds U_{MSY}.

The production model showed retrospective behaviour that differed in some respects from the catch-age model (Figure 14). Note first that the production model does not provide specific estimates of spawning biomass of the type provided by the catch-age model. Instead, the production model estimates a single spawning-exploitable stock, which should be slightly larger than the spawning biomass shown in Figure 14 because in the operating model, fish recruit the exploitable stock before the spawning stock.

The production model shows a similar under-/over-estimation pattern with respect to initial depletion conditions as that of the catch-age model. Spawning biomass tends to be under-estimated in the high initial depletion scenarios (S2 and S4) and over-estimated under low initial depletion scenarios (S1 and S3). However, in contrast to the catch-age model, retrospective patterns in the most recent biomass estimates decreased in variability over time; that is, while catch-age model terminal biomass estimates varied considerably depending on stock dynamics, the production model terminal biomass estimates tended to become more precise over time.

Unlike the catch-age model, the production model accounts for the biomass effect on recruitment; therefore, some retrospective problems with using estimates of recent average recruitment are reduced. On the other hand, uncertainty in production model assessments tends to be reflected mainly in the unfished biomass estimate (the catch-age model does not estimate unfished biomass). Failure to follow major stock increases is caused in part from an under-estimation of unfished biomass, which limited the range of estimable stock increases (e.g., Figure 14 l). Therefore, increases in stock size under the high productivity scenarios usually accelerated the rate of increase because the production model could not allow high enough biomass estimates to cause full exploitation of the true stock. This explains why the production model procedures generally did not show depletion values as small as those obtained for equivalent exploitation rates under catch-age model procedures. Ultimately, these errors reduce yield performance of production model procedures relative to both data-based and catch-age model procedures, but improve conservation performance (Figure 15).

Average depletion

Levels of stock depletion over the 40-year period improved under most management procedures, although some procedures caused further stock declines under low productivity scenarios (Figure 15). All procedures avoided stock declines over the long-term for high productivity scenarios S3 and S4. The catch-age model procedure avoided stock declines under most scenarios except when $U^{ref} = 0.10$ for scenario S1 and $U^{ref} = \{0.08,\ 0.10\}$ for scenario S2. Data-based procedures produced long-term increases in the stock except for $\lambda_2 = 240$ under the low productivity scenarios S1 and S2.

6 Discussion

We utilized the management strategy evaluation approach to actively engage stakeholders in the process of developing fisheries management procedures for sablefish. Stakeholders provided compelling reasons for evaluating simple data-based methods for determining catch limits, as well as more elaborate methods based on modern catch-age analysis that use industry-funded fishery monitoring programs. Simulation testing of candidate management procedures indicated that both approaches could meet inter-annual catch variability criteria set by stakeholders. The range of candidate data-based procedures provided stable or increasing stock sizes over time under most circumstances, while model-based procedures tended to be more efficient in terms of catch, particularly in situations where the stock is productive. These comparisons showed that using a data-based or model-based procedure worked equally well provided that the exploitation rates were set close to the optimal value. Although this is not particularly surprising, it does point to the need for greater emphasis on setting better target exploitation rates rather than what is perhaps the current trend toward developing more sophisticated stock assessment models that attempt to better estimate stock biomass (Cotter et al. 2004; Kell et al. 2005).

Although the variable harvest rate rules we examined comply conceptually with DFO (2006), we did not conduct an extensive set of simulations based on a “factorial design” to identify an “optimal” choice of reference point parameters for model-based or data-based harvest control rules. Similarly, we only examined a limited range of stock assessment methods for estimating stock status, which is also required by those rules. As our analyses show, combinations of assessment methods and harvest control rules interact to affect the ability of management procedures to meet pre-defined objectives.

In our analysis, data-based procedures performed about as well as model-based procedures but substantial work remains to demonstrate that data-based procedures are robust to typical fisheries uncertainties. In particular, we have not challenged any candidate management procedures with more realistic sources of error such as density-dependent changes in survey selectivity, temporal patterns in survey catchability, or spatial stock structure. Changes in catchability and selectivity are not unrealistic given experiences in many fisheries (Harley et al. 2001; Parma 2002). The potential problem with data-based procedures is that changes in selectivity or catchability are translated directly into changes in fishery catch limits and thus exploitation rates on the stock. Undetected increases (decreases) in catchability will cause unwanted direct increases (decreases) in exploitation rate. Similarly, shifts in selectivity toward younger ages will cause increases in catch limits independent of actual changes in the stock.

Catch-age model-based procedures have at least two advantages over simple data-based methods for setting catch limits when catchability and selectivity are subject to change – provided that other model assumptions are not strongly violated. First, catch-age models treat survey catchability as a nuisance parameter that simply scales the average of the surveys to the average population biomass. Patterns of exploitation are inferred partly from changes in age composition independent of fishery surveys and partly from long-term trends in the surveys. Therefore, considerable time may pass before short-term changes in survey catchability affected catch limits. Furthermore, it is also possible that retrospective patterns would appear in the catch-age predictions (Mohn 1999) thus potentially providing early warning that survey catchability was potentially changing. Alternatively, very slow changes in survey catchability can possibly go undetected for very long periods (NRC 1998; Walters 2004). Second, catch-age models can use independent estimates of survey or fishery selectivity to account for potential changes in these processes over time (Myers and Hoenig 1997). Clearly, catch-age models are not without problems, especially where monitoring programs and data cannot target specific stocks; however, sablefish are possibly unique in this case among British Columbia groundfish species because the directed fisheries and surveys based on longline trap gear are very specific for sablefish. The point here is that a targeted science program aimed at providing high quality data is required for both data-based and model-based procedures. In such situations, catch-age modeling approaches may provide substantial benefits over simple data-based methods for setting catch limits.

The performance of management procedures that involved simple production model estimators was not particularly outstanding. For example, although production model procedures tended to generate the greatest increases in biomass under some scenarios, such increases were largely the result of persistent under-estimation errors for increasing stocks. Contrary to our expectation, the production model tended to react to variation in the survey data even more strongly than data-based procedures. This is partly a result of the particular errors-in-variables tuning we used where we specified the proportion of the total error owing to biomass dynamics (i.e., $1 - \rho = 0.20$). Smaller fractions frequently caused unstable estimator and management procedure behaviour because such observation error approaches deal poorly with dynamics that are driven by highly variable recruitment (Punt 2003).

Our results for data-based fishery management procedures reflect other simulation studies and empirical management experience. Where data- and model-based procedures have actually been compared against the same operating models, data-based management procedures generally give slightly lower average annual catch and higher interannual catch variability than model-based procedures (Rademeyer et al. 2007). Our results for data-based and model-based procedures agreed in general with Hilborn et al. (2002) who showed that a constant escapement form of data-based rule for long-lived west coast rockfish (*Sebastes spp*) gave comparable results to model-based procedures, although they did not simulate an actual stock assessment model. Their case, like ours, pointed to the difficulty of specifying exploitation rates or data-based targets for quota fisheries; that is, long-term declines are likely where initial choices for exploitation parameters are based on biased biomass estimates. Data-based procedures have been developed and adopted for Namibian hake and South African west coast rock lobster (Rademeyer et al. 2007) and New Zealand rock lobster (Bentley et al. 2005); however, procedures for hake and rock lobster in South Africa were both considered interim in the absence of better quality data for future model-based procedures.

6.1 A strategy for choosing a management procedure

Implementing the precautionary approach to fishery management involves three decision-making principles. The first integrates conservation, economic, and social objectives into decision making, although conservation of resource productivity takes precedence. Failure to follow this principle leads to development of narrowly focused management decisions that may ultimately lead to poor outcomes. The second principle requires that uncertainty and risk be taken into account in developing fishery management plans. Uncertainty has many layers in natural resource management systems such as fisheries, and therefore, management plans should be robust to failure of assumptions about the natural environment, data collection, fishery modeling, and implementation of regulations. Finally, the third principle involves the active participation of stakeholders in decision making. Involving stakeholders in the original development of management plans increases the likelihood that such plans will be adopted and followed faithfully.

Development of precautionary fishery management plans creates a complex decision environment where value-laden tradeoffs between conflicting (e.g., catch vs. conservation) objectives need to be made. The management strategy evaluation approach we describe is designed to expose these trade-offs, but managers and stakeholders must eventually choose only one particular management procedure. Such a choice is not easy because of the large and complex set of analyses and results that arise from the MSE process. Therefore, we suggest a straightforward strategy for choosing among candidate management procedures. The approach orders fishery management objectives linearly according to their level of priority under a precautionary fishery management policy. Alternatively, objectives may be arranged in a hierarchical structure if multiple objectives occur at the same levels of priority. In either structure, higher level objectives must be met before results related to lower level objectives are considered. Management procedures failing to meet an objective at any level are discarded as not being effective at generating desirable outcomes. Treatment of uncertainty is accomplished by stating certain objectives in probabilistic terms (e.g., probability of stock increase over 10 years) and measuring performance over replicate stochastic simulations and also over broad scenarios for fish stock dynamics. The procedure surviving to the lowest level of the hierarchy represents the choice that is most consistent with the objectives.

An example procedure for the simple case of three priority-ordered objectives related to conservation, catch variability, and average yield is:

1. Conservation
    1. Choose a limit reference point that defines an "undesirable outcome" for conservation – spawning biomass less than 20% of unfished;
    2. Specify a time horizon over which performance should be measured – 21 – 40 years;
    3. Specify the minimum probability that the limit reference point is avoided over the time horizon – 90%;
    4. Specify scenario weights – 100% scenario S1.
2. Catch variability
    1. Choose maximum permitted inter-annual variability in quota - 15%;
    2. Specify time horizon over which performance should be measured – 11 – 20 years;
    3. Specify the minimum probability that objective is attained over the time horizon – use median statistic;
    4. Specify scenario weights – all equal.
3. Yield
    1. Choose desired average annual catch – maximum possible;
    2. Specify time horizon over which performance should be measured – 11 – 20 years;
    3. Specify scenario weights – all equal.

Note that we are not recommending any of the specific objectives stated at each level; we merely include them to facilitate discussion on how a management procedure can be chosen based on this linearized objective structure.

Performance related to the conservation objective can be obtained graphically from Figure 15 or more specifically from Appendix Table F-1 (21 – 40 years) by selecting procedures for which performance measure $P_{cons} \geq 0.90$. Procedures meeting this objective are data-based with $\lambda_2 = \{120, 150\}$, and catch-at-age model with $U^{ref} = \{0.04, 0.06\}$; no production model-based procedures meet this objective and are thus eliminated. Next, we see that all of the procedures surviving to this point meet the catch variability objective as indicated in Figure 11a. Finally, based on Objective 3, we choose the procedure that provides the maximum average annual catch over 11 – 20 years according to the following table:

	Scenario				
	S1	S2	S3	S4	Average
Data ($\lambda_2 = 120$)	1025	1694	2357	2322	1849
Data ($\lambda_2 = 150$)	1192	1981	2796	2787	2189
CA $U^{ref} = 0.04$	1215	1978	2000	2466	1914
CA $U^{ref} = 0.06$	1574	2586	2688	3327	**2543**

Thus, CA $U^{ref} = 0.06$ provides the highest average annual catch over 11 – 20 years while still meeting conservation and catch variability constraints. The second best procedure, data-based with $\lambda_2 = 150$, does not perform quite as well, but might be preferred by industry stakeholders because of the transparent way in which quotas are calculated. At this point, any of the above procedures meet precautionary requirements and therefore, final choices may reflect social or economic objectives alone.

The proposed strategy can be easily adapted to more complex objective structures that will undoubtedly be common in commercial fisheries management applications. The key steps are in specifying the objectives to a sufficient level of detail. In particular, every objective should have (*i*) a measurable value or indicator from the simulation model; (*ii*) a time-horizon over which performance is to be measured; (*iii*) a probability statement of meeting the objective; and (*iv*) the simulation conditions or scenario to use in computing performance.

In executing the above strategy, we noted that only four procedures from the initial set of more than 30 reasonable candidates survived to be evaluated at the yield step. The 15% constraint on changes in catch during the first five years is one of the main reasons why so many candidate procedures failed to achieve the conservation objective. It is interesting to note that none of these failed procedures would have been considered unreasonable given historical stock assessment and management of sablefish in the northeast Pacific.

6.2 Limitations and future work

Scenarios considered in this paper focused on B.C. sablefish stock productivity and the level of depletion at the present time. Although these two uncertainties are amongst the most critical to evaluate in management procedure simulations, these scenarios do not capture the full range of uncertainties associated with the B.C. sablefish stock and fishery. Future management strategy evaluations will address uncertainties related to stock assessment data and discarding, the historical catch record, spatial stock structure and exchange with U.S. stocks, and spatial dynamics of the fishery.

The candidate management procedures we evaluated all utilized the standardized survey program data (SS; 1990-2006) simply because of the long time series of observations. A stratified random survey program (2003-2006) is also conducted annually and, therefore, a choice will need to be made between survey programs because resources available for gathering fishery-independent data will likely decline in the near future. In our opinion, the stratified random survey (StRS) will almost surely be the long-term choice for several reasons. First, the StRS encompasses a greater spatial and depth range of potential sablefish habitat, and thus individual localities cannot contribute disproportionately to the overall average survey catch rate. Furthermore, survey catch rates from random set locations should be less sensitive to moderate changes in the distribution of the stock, whereas catch rates from fixed locality surveys like the SS program can be very sensitive. The StRS is also a statistically-based sampling design that enjoys the advantages of randomization and replication, whereas the standardized survey is essentially *ad hoc*. Thus, incorporating the StRS program will be an important subject of management procedure development for sablefish in the short-term. Data-based procedures may be particularly sensitive to this development because, as noted earlier, effective exploitation rates for these procedures are totally dependent upon survey catchability, which will not be estimated particularly well given only five years of observations.

Both the operating model and management procedure simulations were based on sablefish landings data. We did not attempt to reconstruct the historical pattern of sablefish discards from the longline trap, longline hook (e.g., directed sablefish, directed rockfish, halibut) and trawl gear sectors that intercept B.C. sablefish. These data are poorly known with the exception of trawl discard estimates beginning in 1996 and, beginning in late 2006, for all fisheries with the advent of the Groundfish Integration Pilot Proposal. Discarding from the sablefish trap fishery has likely declined since the adoption of escape rings into commercial trap gear in 1998, but may have continued to some extent as higher prices are paid for fish that exceed about 60 cm fork length as compared to the 55 cm legal size limit. Non-directed longline hook fisheries were not permitted to retain sablefish, except on so-called combination trips, until Groundfish Integration provided non-quota holders the opportunity to access sablefish quota. The challenge will be to devise a realistic sablefish discard scenario that does not, for example, (*i*) simply add the higher quality discard data to the more recent landings, or (*ii*) prorates discards based the post-1996 trawl observer period and or recent Groundfish Integration period since the former coincided with a wholesale change to ITQs for the trawl fishery and both have strong potential for altering fishing behavior relative to past practices. The treatment of discard data will be examined in the next phase of MSE development for B.C. sablefish.

Continued evaluation of management procedures based on catch-age models will require greater attention to the sources of error in the ageing process and biological sampling. Sablefish are a notably difficult species to age reliably (Pearson and Shaw 2004). Alaskan stock assessments partially account for ageing errors by using an ageing-error bias correction matrix (Hanselman et al. 2006, Heifetz et al. 1998), which could be applied to B.C. ageing data for those fish similarly aged using the burnt otolith section method. Future management procedure research will experiment with aging error effects in the operating model to determine how important these errors have to be to reduce the performance of the catch-age model procedures.

The spatial structure of the coast-wide sablefish population is a significant uncertainty that must be addressed in future management strategy evaluations. Ignoring spatial stock structure can seriously degrade the performance of management procedures that otherwise perform adequately when the stock structure is well-known (de la Mare 1996; IWC 1992). As noted in our description of the operating models, we assumed the sablefish stock is closed to immigration and emigration, even though we know of evidence to the contrary. For example, the current hypothesis concerning coast-wide sablefish stock structure proposes two populations consisting of (*i*) a northern stock ranging from the Bering Sea southeast through the Gulf of Alaska to northwest Vancouver Island, and (*ii*) a west coast stock extending from southwest Vancouver Island to Baja, California (Kimura et al. 1998). The northern stock is thought to be more mobile than the west coast stock with strong mutual exchanges among all northern areas. Although long-term tagging studies show considerable long-range movement of some tagged sablefish, for the most part, remarkable site fidelity, particularly for adult fish, seems more common (Beamish and McFarlane 1988; Kimura et al. 1998). Ultimately, redistribution of sablefish throughout the northeast Pacific appears to be driven by both movement of adults (Kimura et al. 1998) and by emigration of juveniles from nearshore waters and inlets; however, specific movement rates have yet to be quantified. Therefore, available tagging information should be used to construct plausible stock structure and spatial dynamics scenarios. Specifically, these scenarios should include (*i*) two stocks (Alaska and B.C.) sharing only juvenile recruitment; (*ii*) two stocks (Alaska and B.C.) sharing adults and juveniles; and (*iii*) two coast-wide stocks as proposed by Kimura et al (1998) with age-structured movement within regions. Scenarios involving spatial exchange of adults are important for addressing immigration/emigration hypotheses, effects of adult movement on management procedures that involve catch-age modeling, and simulation of more realistic spatial variability in stock indices. For example, sampling variability of B.C. surveys may be influenced by inter-annual fluctuations in the fraction of the northern stock in the sampling frame at the time the survey is conducted (Cochrane et al. 1998, De Oliveira et al. 1988).

Scenarios aimed at testing management procedure performance under alternative sablefish stock structure and spatial dynamics hypotheses will need to consider the spatial dynamics of harvesting. Unlike many groundfish species in B.C., sablefish are managed under a single coast-wide quota; therefore, harvesters are free to take the quota from the most profitable locations. Areas in which fishing costs are lower may absorb more of the total catch than local production can support in the long-term, which may ultimately lead to failure of a management procedure. A potential solution in this case is to partition the management region by presumed stock boundaries and apply a separate catch limit to each. Such a tactic tends toward more conservative combined quotas, but also mediates the risk of overall stock decline (de la Mare 1996).

Industry stakeholders have suggested other uncertainties for evaluation. For example, some stakeholders have expressed concern over fishing spawning aggregations of sablefish during the January to March period when commercial trap fishery catch rates tend to be highest. Stakeholders have aligned on both sides of this issue, as profitability is obviously higher during periods when catch rates increase seasonally. There has also been a substantial industry investment in tagging programs since 1990 in British Columbia. Here we used the tag-recovery data to estimate gear selectivity independently of the operating model and assessment methods, and used a tagging index of relative biomass as in input to the operating model. We did not, however, use a tagging index in the candidate management procedures because conclusions ultimately depend on assumptions about tag reporting rates as well as (*i*) movement, (*ii*) contagiously distributed tag recoveries, (*iii*) tag retention, (*iv*) tagging mortality, (*v*) tag reporting rate, and (*vi*) sort/grading effects as fisherman sort through fish to be discarded to retain tags. Extensions of the work by Mathur (2007), which simulated the tag release-recovery process and estimation procedures, might be undertaken in future work.

7 Conclusion

Ultimately, choosing a specific fishery management procedure involves a compromise among possible candidates that perform differently under equally plausible, yet contrasting scenarios (i.e., operating models). Deliberate evaluation of management procedures in this way addresses national precautionary fishery management policy requirements for adjusting harvests in response to departures from operational objectives. The involvement of stakeholders in the evaluation process helped us to consider practical data-based and model-based fishery management procedures that addressed particular industry, as well as scientific concerns. Thus, industry stakeholders are in a better position to make the necessary compromises and trade-offs compared to situations where complex management procedures are defined outside the co-management arena. Furthermore, iterative refinement and testing of these procedures against known uncertainties provides a formal mechanism for fishery co-management in which stakeholders have a central role in decision-making, providing effective input into the process by which catch limit decision will be made.

Acknowledgments

We thank Steve Martell, Alan Sinclair and Andrew Edwards for their thoughtful reviews. We owe a great deal of thanks to the Canadian Sablefish Association (CSA) for their support and contributions to this work. Bill de la Mare and Alan Sinclair provided scientific input to the design and testing of our management strategy evaluation approach. Colleagues in the Groundfish Section at the Pacific Biological Station provided feedback on an earlier version of this paper. We thank an anonymous reviewer for many suggestions and improvements. We also thank students from the SFU-REM Fisheries Research and Management Group for their help in testing earlier versions of the software.

Literature Cited

Beamish, R.J. and McFarlane, G.A. 1988. Resident and dispersal behaviour of adult sablefish (*Anoplopoma fimbria*) in the slope waters off Canada's west coast. Can. J. Fish. Aquat. Sci. 4 45: 152 – 164.

Beddington, J.R. and Cooke, J.G. 1983. The potential yield of fish stocks. FAO Fisheries Technical Paper, 242, Rome FAO. 47pp.

Bentley, N., Breen, P.A., Kim, S.W., and Starr, P.J. 2005. Can additional abundance indices improve harvest control rules for New Zealand rock lobster (*Jasus edwardsii*) fisheries? New Zealand J. Mar. Fresh. Res. 39: 629 – 644.

Butterworth, D.S. 2007. Why a management procedure approach? Some positives and negatives. 11 ICES J. Mar. Sci. 64: 629 - 644.

Butterworth, D.S. and Punt, A.E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985-998.

Clark, W.G. and Kaimmer, S.M. 2006. Estimates of commercial longline selectivity for Pacific halibut (*Hippoglossus stenolepis*) from multiple tagging experiments. Fish. Bull. 104: 465 – 467.

Cochrane, K.L., Butterworth, D.S., De Oliveira, J.A.A., and Roel, B.A. 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Reviews in Fish Biology and Fisheries 8: 177-214.

Cooke, J.G. 1985. On the relationship between catch per unit effort and whale abundance. Rep. Int. Whal. Comm. 35: 511 – 519.

Cooke, J.G. 1995. The International Whaling Commission's revised management procedure as an example of a new approach to fishery management. In Whales, seals, fish and man. pp. 647-657. Ed. by A.S. Blix, L. Walloe, and O. Ulltang. Developments in Marine Biology, 4. 720 pp.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES J. of Mar. Sci., 56: 797-810.

Cotter, A.J.R., Burt, L., Paxton, C.G.M., Fernandez, C., Buckland, S.T., and Pan, J-X. 2004. Are stock assessment methods too complicated? Fish and Fisheries 5: 235 – 254.

Cox, S.P. and Kronlund, A.R. 2008. Practical stakeholder-driven harvest policies for groundfish in British Columbia, Canada. Fish. Res. 94(3): 224-237.

de la Mare, W.K. 1986. Simulation studies on management procedures. Rep. Int. Whal. Commn. 36: 429-450.

de la Mare, W.K. 1996. Some recent developments in the management of marine living resources. *In* Frontiers of Population Ecology. R.B. Floyd, A.W. Sheppard, and P.J. De Barro, eds. CSIRO Publishing, Melbourne. pp. 599-616.

de la Mare, W.K. 1998. Tidier fisheries management requires a new MOP (management-oriented paradigm). Rev. Fish. Biol. Fisheries 8: 349-356.

De Oliveira, J.A.A., Butterworth, D.S., Roel, B.A., Cochrane, K.L. and Brown, J.P. 1998. The application of a management procedure to regulate the directed and bycatch fishery of South African sardine, *Sardinaps sagax*. South African J. Mar. Sci. 19: 449-469.

Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37: 268-282.

DFO. 2002. Proceedings of the PSARC Groundfish Subcommittee Meeting January 29, 2002. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2002/001.

DFO. 2006. A harvest strategy compliant with the precautionary approach. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2006/023.

FAO. 1995. Precautionary approach to fisheries: Part 1: Guidelines on the precautionary approach to capture fisheries and species introductions. FAO Fisheries Technical Paper No. 10 350, Part 1. Rome FAO. 52 p.

Fisheries and Oceans Canada. 2002. Integrated fishery management plan: sablefish, August 1, 2002 to July 31, 2003. Fisheries and Oceans Canada. 50 p.

Fisheries and Oceans Canada. 2006. Integrated fishery management plan: sablefish, April 1, 2006 to March 31, 2007. Fisheries and Oceans Canada. 137 p.

Fisheries and Oceans Canada. 2007. Integrated fishery management plan: groundfish, March 10, 2007 to March 31, 2008. Fisheries and Oceans Canada. 134p.

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.

Haist, V., Kronlund, A.R., and Wyeth, M.R. 2005. Sablefish (*Anoplopoma fimbria*) in British Columbia, Canada: stock assessment for 2004 and advice to managers for 2005. DFO Can. Sci. Advis. Sec. Res. Doc. 2005/031. 182p.

Haist, V., Fournier, D., and Saunders, M.W. 1999. Reconstruction of B.C. sablefish stocks, 1966-1998, and catch projections for 1999, using an integrated catch-age mark-recapture model with area and depth movement. DFO Can. Stock Assess. Sec. Res. Doc. 99/79.

Haist, V., Hilborn, R., and Saunders, M.W. 1999. Sablefish stock assessment for 1999 and recommended yield options for 2000 and 2001. DFO Can. Stock Assess. Sec. Res. Doc. 99/195.

Haist, V., Saunders, M.W., Hilborn, R., and Maunder, M. 1997. Sablefish stock assessment for 1997 and recommended yield options for 1998. DFO Can. Stock Assess. Sec. Res. Doc. 97/146.

Hall, D.L., Hilborn, R., Stocker, M., and Walters, C. 1988. Alternative harvest strategies for Pacific herring (Clupea harengus pallasi). Can. J. Fish. Aquat. Sci. 45: 888-897.

Hanselman, D. H., Lunsford, C.R., Fujioka, J.T., and Rodgveller, C.J. 2006. Alaska sablefish assessment for 2007. *In* Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2007. November 2006, Plan Team Draft (http://www.afsc.noaa.gov/refm/stocks/assessments.htm).

Harley, S.J., Myers, R.A. and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. 58: 1760 – 1772.

Harvey, A.C. 1989. Forecasting, structural time-series models and the Kalman filter. Cambridge Univ. Press. Cambridge, UK. 554p.

Heifetz, J., Anderl, D., Maloney, N.E., and Rutecki, T.L. 1998. Age validation and analysis of ageing error from marked and recaptured sablefish, *Anoplopoma fimbria*. Fish. Bull. 97: 256-263.

Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. J. Fish. Res. Bd. Can. 33: 1 – 5.

Hilborn, R. 1997. Uncertainty, risk, and the precautionary approach. *In* Global trends: fisheries management, pp. 100-106. Ed. by E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine. Amer. Fish. Soc. Symp. 20, Bethesda, Maryland.

Hilborn, R., Parma, A., and Maunder, M. 2002. Exploitation rate reference points for west coast rockfish: are they robust and are there better alternatives? N. Amer. J. Fish. Mgmt. 22: 365-375.

Hilborn, R. and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York. 570p.

IWC. 1992. Report of the Third Comprehensive Assessment Workshop on Management Procedures. Reports of the International Whaling Commission, 42: 87-136.

Kimura, D.K., Shimade, A.M., and Shaw, F.R. 1998. Stock structure and movements of tagged sablefish, *Anoplopoma fimbria* in offshore northeastern Pacific waters and the effects of El Nino-southern oscillation on migration and growth. Fish. Bull. 96: 462-481.

Kell, L.T., O'Brien, C.M., Smith, M.T., Stokes, T.K., and Rackham, B.D. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North Sea plaice (*Pleuronectes platessa* L.). ICES J. Mar. Sci. 56: 834 – 845.

Koolman, J., Mose, B., Stanley, R.D., and Trager, D. 2007. Developing an integrated commercial groundfish strategy for British Columbia: insights gained about participatory management. In Biology, Assessment, and Management of North Pacific Rockfishes, Alaska Sea Grant College Program, AK-SG-07-01. p. 287-300.

Kronlund, A.R., Wyeth, M.R., and Hilborn, R. 2002. Review of survey, commercial fishery and tagging data for sablefish (Anoplopoma fimbria) in British Columbia: supplement to the November 2001 sablefish stock assessment. DFO Can. Sci. Adv. Res. Doc. 2002/074. 109 p.

Logan, G., de la Mare, W., King, J., and Haggarty, D. 2005. Management Framework for Strait of Georgia Lingcod. DFO Can. Sci. Advis. Sec. Res. Doc. 2005/048.

MacLellan, S.E. 1997. How to age rockfish (*Sebastes*) using *S. alutus* as an example – the otolith burnt section technique. Can. Tech. Rep. Fish. Aquat. Sci. 2146: 42 pp.

McFarlane, G.A. and R.J. Beamish. 1983. Overview of the fishery and management strategy for sablefish (*Anoplopoma fimbria*) in waters off the west coast of Canada. p. 13-35. *In* Proceedings of the Second Lowell Wakefield Fisheries Symposium, Anchorage, AK. Alaska Sea Grant Report 83-3.

Mathur, R. 2007. An evaluation of the Brownie tag-recapture model using simulation-estimation procedures, with application to British Columbia sablefish (*Anoplopoma fimbria*). M.R.M. Res. Project. School of Resource and Environmental Management, Simon Fraser University, Burnaby, B.C.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality in northern anchovy. Amer. Fish. Soc. Symp. **6**: 66-82.

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56: 473 – 488.

Myers, R.A. and Hoenig, J.M. 1997. Direct estimates of gear selectivity from multiple tagging experiments. Can. J. Fish. Aquat. Sci. 54: 1-9.

Myers, R.A., Bowen, K.G., and Barrowman, N.J. 1999. Maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56: 2404 – 2419.

NRC. 1998. Improving fish stock assessments. Washington, D.C., National Academy Press. 188 p.

Parma, A.M. 2002. Bayesian approaches to the analysis of uncertainty in the stock assessment of Pacific halibut. American Fisheries Society Symposium 27: 113-136.

Pearson, D.E. and Shaw, F.R. 2004. Sources of age determination errors for sablefish (*Anoplopoma fimbria*). Fish. Bull. 102:127-141.

Punt, A.E. 1995. The performance of a production-model management procedure. Fish. Res. 21: 349-374.

Punt, A.E. 1997. The performance of VPA-based management. Fish. Res. 29:217-243.

Punt, A.E. 2003. Extending production models to include process error in the population dynamics. Can. J. Fish. Aquat. Sci. 60: 1217-1228.

Punt, A.E., Smith, A.D.M., and Cui, G. 2002a. Evaluation of management tools for Australia's South East fishery. 1. Modelling the South East fishery taking account of technical interactions. Mar. Freshwater Res. 53: 615-629.

Punt, A.E., Smith, A.D.M., and Cui, G. 2002b. Evaluation of management tools for Australia's South East fishery. 2. How well can management quantities be estimated? Mar. Freshwater Res. 53: 615-629.

Punt, A.E., Smith, A.D.M., and Cui, G. 2002c. Evaluation of management tools for Australia's South East fishery. 3. Towards selecting appropriate harvest strategies. Mar. Freshwater Res. 53: 615-629.

Punt, A.E., Smith, A.D.M., and Cui, G. 2001. Review of progress in the introduction of management strategy evaluation (MSE) approaches in Australia's south east fishery. Mar. Freshwater Res. 52: 719-726.

Punt, A.E. and Smith, A.D.M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solnadri*). ICES J. Mar. Sci. 56: 860-875.

Quinn, T.J. and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York. 542 p.

Quinn, T.J., Fagen, R. and Zheng, J. 1990. Threshold management policies for exploited populations. Can. J. Fish. Aquat. Sci. 47: 2016-2029.

Rademeyer, R.A., Plagnáyi, É.E., and Butterworth, D.S. 2007. Tips and tricks in designing management procedures. ICES J. Mar. Sci. 64: 618-625.

Sainsbury, K.J., Punt, A.E., and Smith, A.D.M. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57: 731-741.

Saunders, M.W., B.M. Leaman, and McFarlane, G.A. 1995. Sablefish stock assessment for 1995 and recommended yield options for 1996. PSARC Working Paper G95-5.

Saunders, M.W., Leaman, B.M., Haist, V, Hilborn, R., and McFarlane, G.A. 1996. Sablefish stock assessment for 1996 and recommended yield options for 1997. PSARC Working Paper G96-5.

Schirripa, M.J. 2007. Status of the sablefish resource off the continental U.S. Pacific coast in 2007. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR 97220-1384.

Schirripa, M.J. 2002. Status of the sablefish resource off the continental U.S. Pacific coast in 2002. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR 97220-1384.

Schnute, J.T. 1985. A general theory for analysis of catch and effort data. Can. J. Fish. Aquat. Sci. 42: 419 – 429.

Schnute, J.T. and Richards, L.J. 1995. The influence of error on population estimates from catch age models. Can. J. Fish. Aquat. Sci. 52: 2063-2077.

Schnute, J.T. and Kronlund, A.R. 2002. Estimating salmon stock-recruitment relationships from catch and escapement data. Can. J. Fish. Aquat. Sci. 59(3): 433-449.

Sigler, M.F., Lunsford, C.R., Fujioka, J.T, and Lowe, S.A. 2004. Alaska sablefish assessment for 2004. *In* Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2005. November 2004, Plan Team Draft (http://www.afsc.noaa.gov/refm/stocks/assessments.htm).

Sinclair, A.F. 1993. Partial recruitment considerations in setting catch quotas. Can. J. Fish. Aquat. Sci. 50(4): 734-742.

Smith, A.D.M., Sainsbury, K.J. and Stevens, R.A. 1999. Implementing effective fisheries management systems – management strategy evaluation and the Australian partnership approach. ICES J. of Mar. Sci. 56: 967-979.

Stocker, M. 1994. Groundfish stock assessments for the west coast of Canada in 1993 recommended yield options for 1994. Can. Tech. Rep. Fish. Aquat. Sci. 1975: 352 p.

Stocker, M. and Fargo, J. 1995. Groundfish stock assessments for the west coast of Canada in 1994 recommended yield options for 1995. Can. Tech. Rep. Fish. Aquat. Sci. 2069: 440p.

Stock, M. and Saunders, M.W. 1997. Utility of an age- and sex-structured model for sablefish, *Anoplopoma fimbria*, stock assessment off the west coast of Canada. In Wilkins M.E. and Saunders, M.W. Biology and management of sablefish, *Anoplopoma fimbria*. NOAA Technical Report NMFS 130. 207-213.

Walters, C.J. 2004. Simple representation of the dynamics of biomass error propagation for stock assessment models. Can. J. Fish. Aquat. Sci. 61: 1061 – 1065.

Walters, C.J. 1986. Adaptive management of renewable resources. The Blackburn Press, NJ, 374p.

Walters, C.J. and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press, NJ. 399p.

Wyeth, M.R., Kronlund, A.R, and Elfert, M. 2007. Summary of the 2005 British Columbia sablefish (*Anoplopoma fimbria*) research and assessment survey. Can. Tech. Rep. Fish. Aquat. Sci. 2694. 105p.

Wyeth, M.R., Kronlund, A.R, and Elfert, M. 2006. Summary of the 2004 British Columbia sablefish (*Anoplopoma fimbria*) research and assessment survey. Can. Tech. Rep. Fish. Aquat. Sci. 2660. 74p.

Table 1 Summary of management procedures indicating the data, assessment method, and harvest control rule type and parameters.

MP Class	Data	Method Type	Assessment Method	Rule Type	Rule Parameters
1	Catch Survey index	Data-based	3-year running mean of survey	Constant harvest rate	$\lambda_1 = \{0.2, 0.5, 0.8\}$ $\lambda_2 = \{150, 180, 210, 240\}$
2	Catch Survey index	Data-based	3-year running mean of survey	Variable harvest rate	$\lambda_1 = \{0.5\}$ $\lambda_2 = \{150, 180, 210, 240\}$ $I_{low} = \{3, 4\}$, $I_{high} = \{10, 15\}$
3	Catch Survey index	Model-based	Production model	Constant harvest rate	$u_{ref} = \{0.06, 0.08, 0.10\}$
4	Catch Survey index	Model-based	Production model	Variable harvest rate with risk adjustment	$u_{ref} = \{0.06, 0.08, 0.10\}$ $Q = \{0.4, 0.5\}$ $D_{low} = \{0.1\}$, $D_{high} = \{0.4\}$
5	Catch Survey index Trap fishery ages Survey ages	Model-based	Catch age model	Constant harvest rate	$u_{ref} = \{0.06, 0.08, 0.10\}$
6	Catch Survey index Trap fishery ages Survey ages	Model-based	Catch age model	Variable harvest rate with risk adjustment	$u_{ref} = \{0.04, 0.06, 0.08, 0.10\}$ $Q = \{0.4, 0.5\}$ $D_{low} = \{0.25\}$, $D_{high} = \{1.0\}$

Table 2 Distinguishing features of operating model scenarios S1-S4. Parameters represent the steepness of the Beverton-Holt stock-recruitment function (h) and trap fishery hyperstability parameter ($q_{2,trap}$). The initial depletion of the stock as of 2006 (D_{init}) depends on the operating model parameters.

Scenario	Parameters	Description	Initial Depletion
S1	$\left\{h = 0.45,\ \hat{q}_{2,trap}\right\}$	Low stock-recruitment steepness Very low initial depletion	$D_{init} = 0.18$
S2	$\left\{h = 0.45,\ q_{2,trap} = 1\right\}$	Low stock-recruitment steepness Moderate initial depletion	$D_{init} = 0.30$
S3	$\left\{h = 0.65,\ \hat{q}_{2,trap}\right\}$	High stock-recruitment steepness Very low initial depletion	$D_{init} = 0.20$
S4	$\left\{h = 0.65,\ q_{2,trap} = 1\right\}$	High stock-recruitment steepness Moderate initial depletion	$D_{init} = 0.31$

Table 3 Definitions of performance statistics used for sablefish management strategy evaluation. Symbols t_1 and t_2 define the period over which performance measures are computed.

Symbol	Definition	Description
AAV	$AAV = \sum_{t=t_1}^{t_2} \left\| C_t - C_{t-1} \right\| \Big/ \sum_{t=t_1}^{t_2} C_t$	Average annual absolute change in the catch over the time interval, where C_t is the catch biomass in year t.
$\overline{C}$	$\overline{C} = \frac{1}{n}\sum_{t=t_1}^{t_2} C_t$	Arithmetic mean of annual catches over the time interval.
$C_{5\%}$	5th percentile of the catch from $t = t_1,\ldots,t_2$.	5th percentile of the catch distribution that represents a proxy for the minimum expected catch over the time interval.
$\overline{D}$	$\overline{D} = \frac{1}{t_2 - t_1 + 1}\sum_{t=t_1}^{t_2}\left(\frac{S_t}{B_0}\right)$	Arithmetic mean of annual depletion.
P_{cons}	$P\left(B > 0.2B_0\right)$	The probability that the true spawning biomass is greater than the conservation limit – 20% of unfished biomass.
$\overline{I}_{trap}$	$\overline{I}_{trap} = \frac{1}{t_2 - t_1 + 1}\sum_{t=t_1}^{t_2} I_{trap,t}$	Arithmetic mean of trap fishery CPUE.

Table 4 Summary of selected management procedures.

Class	Data	Method	Assessment Method	Rule Type	Rule Parameters
2	Catch Survey index	Data-based	3-year mean of survey	VHR	$\lambda_1 = 0.5$ $\lambda_2 = \{150, 180, 210, 240\}$ $I_{low} = 4$, $I_{high} = 15$
4	Catch Survey index	Model-based	Production model	VHR Risk-adjusted	$u_{ref} = \{0.06,\ 0.08,\ 0.10\}$ $Q = 0.4$ $D_{low} = 0.1$, $D_{high} = 0.4$
6	Catch Survey index Trap and survey ages	Model-based	Catch-age	VHR Risk-adjusted	$u_{ref} = \{0.06,\ 0.08,\ 0.10\}$ $Q = 0.4$ $D_{low} = 0.25$, $D_{high} = 1.0$

Table 5 Summary of performance statistics and ranks by management procedure for Scenario 1. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	**Statistic**
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.151	0.139	2081	12.4	0.00	7.5	11.5	4
	Data ($\lambda_2 = 150$)	2856	16.1	0.151	0.138	2081	12.4	0.00	7.5	11.5	6
	Data ($\lambda_2 = 180$)	3067	16.1	0.148	0.132	2234	12.3	0.00	3	8	10
	Data ($\lambda_2 = 210$)	3366	15.3	0.144	0.125	2473	12.2	0.00	2	3	11
	Data ($\lambda_2 = 240$)	3679	14.5	0.140	0.119	2722	12.1	0.00	1	2	12
	Prod $U^{ref} = 0.06$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	9.5	7
	Prod $U^{ref} = 0.08$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	9.5	8
	Prod $U^{ref} = 0.10$	2943	14.0	0.150	0.136	2272	12.4	0.00	4	1	9
	CA $U^{ref} = 0.04$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.06$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.08$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.10$	2856	16.1	0.151	0.138	2081	12.4	0.00	10.5	5.5	5
21-40	Data ($\lambda_2 = 120$)	2101	7.8	0.295	0.316	1314	16.2	1.00	11	4	2
	Data ($\lambda_2 = 150$)	2397	8.6	0.267	0.280	1413	15.6	1.00	3	5	3
	Data ($\lambda_2 = 180$)	2452	9.4	0.236	0.254	1375	14.9	0.85	1	7	6
	Data ($\lambda_2 = 210$)	2447	10.1	0.214	0.225	1302	14.4	0.63	2	8	8
	Data ($\lambda_2 = 240$)	2343	10.6	0.196	0.204	1208	13.9	0.50	4	9	10
	Prod $U^{ref} = 0.06$	2289	11.4	0.243	0.258	1750	15.0	0.38	9	10	5
	Prod $U^{ref} = 0.08$	2338	13.1	0.212	0.225	1735	14.3	0.05	6	11	9
	Prod $U^{ref} = 0.10$	2328	14.9	0.190	0.198	1651	13.7	0.00	7	12	12
	CA $U^{ref} = 0.04$	1832	4.8	0.316	0.364	1397	16.6	1.00	12	1	1
	CA $U^{ref} = 0.06$	2197	6.3	0.261	0.279	1696	15.5	1.00	10	2	4
	CA $U^{ref} = 0.08$	2322	7.6	0.222	0.227	1800	14.4	0.68	8	3	7
	CA $U^{ref} = 0.10$	2339	9.1	0.191	0.192	1773	13.6	0.30	5	6	11

Table 6 Summary of performance statistics and ranks by management procedure for Scenario 2. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	**Statistic**
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.271	0.261	2081	12.0	1.00	10	12	3
	Data ($\lambda_2 = 150$)	2892	15.3	0.271	0.260	2146	12.0	1.00	9	9	6
	Data ($\lambda_2 = 180$)	3208	13.5	0.267	0.253	2466	11.9	1.00	5	7	10
	Data ($\lambda_2 = 210$)	3579	12.2	0.264	0.245	2783	11.8	1.00	2	3	11
	Data ($\lambda_2 = 240$)	3948	11.1	0.260	0.237	3094	11.6	1.00	1	1	12
	Prod $U^{ref} = 0.06$	2926	14.2	0.271	0.260	2256	12.0	1.00	8	8	4
	Prod $U^{ref} = 0.08$	3059	12.6	0.270	0.257	2509	12.0	1.00	6	5	7
	Prod $U^{ref} = 0.10$	3236	11.6	0.270	0.255	2746	12.0	1.00	3	2	9
	CA $U^{ref} = 0.04$	2856	16.1	0.271	0.261	2081	12.0	1.00	11.5	10.5	1.5
	CA $U^{ref} = 0.06$	2856	16.1	0.271	0.261	2081	12.0	1.00	11.5	10.5	1.5
	CA $U^{ref} = 0.08$	2988	13.3	0.271	0.259	2436	12.0	1.00	7	6	5
	CA $U^{ref} = 0.10$	3231	12.5	0.270	0.257	2828	12.0	1.00	4	4	8
21-40	Data ($\lambda_2 = 120$)	2377	6.9	0.438	0.454	1741	19.5	1.00	12	3	1
	Data ($\lambda_2 = 150$)	2728	7.3	0.400	0.407	1991	18.0	1.00	9	4	3
	Data ($\lambda_2 = 180$)	2925	7.9	0.367	0.367	2090	16.6	1.00	7	6	5
	Data ($\lambda_2 = 210$)	2988	8.9	0.333	0.343	2007	15.3	1.00	5	7	8
	Data ($\lambda_2 = 240$)	2961	9.4	0.310	0.317	1928	14.1	1.00	6	9	10
	Prod $U^{ref} = 0.06$	2722	10.2	0.380	0.380	2245	16.9	1.00	10	10	4
	Prod $U^{ref} = 0.08$	2922	10.8	0.347	0.343	2340	15.5	1.00	8	11	6
	Prod $U^{ref} = 0.10$	3016	11.9	0.321	0.317	2320	14.5	1.00	4	12	9
	CA $U^{ref} = 0.04$	2469	4.5	0.416	0.437	2043	18.6	1.00	11	1	2
	CA $U^{ref} = 0.06$	3081	6.1	0.342	0.335	2500	15.5	1.00	3	2	7
	CA $U^{ref} = 0.08$	3339	7.6	0.286	0.269	2653	13.2	1.00	1	5	11
	CA $U^{ref} = 0.10$	3332	9.2	0.242	0.222	2672	11.2	0.80	2	8	12

Table 7 Summary of performance statistics and ranks by management procedure for Scenario 3. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	Statistic
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.197	0.201	2081	13.7	0.40	9	12	4
	Data ($\lambda_2 = 150$)	2886	15.3	0.197	0.200	2153	13.7	0.40	8	8	7
	Data ($\lambda_2 = 180$)	3178	13.7	0.193	0.193	2433	13.6	0.20	4	5	10
	Data ($\lambda_2 = 210$)	3537	12.6	0.189	0.185	2750	13.5	0.20	2	3	11
	Data ($\lambda_2 = 240$)	3889	11.6	0.185	0.177	3038	13.4	0.20	1	2	12
	Prod $U^{ref} = 0.06$	2905	14.2	0.197	0.200	2253	13.7	0.00	7	6	5
	Prod $U^{ref} = 0.08$	3042	12.7	0.196	0.198	2484	13.7	0.00	5	4	8
	Prod $U^{ref} = 0.10$	3217	11.5	0.195	0.195	2716	13.6	0.00	3	1	9
	CA $U^{ref} = 0.04$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.06$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.08$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.10$	2912	14.9	0.197	0.200	2266	13.7	0.40	6	7	6
21-40	Data ($\lambda_2 = 120$)	3405	6.0	0.483	0.510	2717	19.4	1.00	11	4	2
	Data ($\lambda_2 = 150$)	3891	6.1	0.435	0.450	3169	18.7	1.00	7	5	5
	Data ($\lambda_2 = 180$)	4218	6.2	0.392	0.398	3434	18.0	1.00	4	6	7
	Data ($\lambda_2 = 210$)	4462	6.5	0.351	0.345	3663	17.3	1.00	3	7	10
	Data ($\lambda_2 = 240$)	4643	7.0	0.317	0.305	3643	16.6	1.00	1	8	12
	Prod $U^{ref} = 0.06$	3547	8.2	0.436	0.455	3047	18.7	1.00	10	10	4
	Prod $U^{ref} = 0.08$	3806	8.2	0.409	0.419	3285	18.2	1.00	8	11	6
	Prod $U^{ref} = 0.10$	3969	8.5	0.387	0.390	3427	17.8	1.00	6	12	8
	CA $U^{ref} = 0.04$	2747	3.8	0.528	0.565	2279	20.2	1.00	12	1	1
	CA $U^{ref} = 0.06$	3576	4.6	0.453	0.471	3050	19.0	1.00	9	2	3
	CA $U^{ref} = 0.08$	4179	5.8	0.386	0.395	3460	18.0	1.00	5	3	9
	CA $U^{ref} = 0.10$	4534	7.2	0.337	0.339	3674	17.1	1.00	2	9	11

Table 8 Summary of performance statistics and ranks by management procedure for Scenario 4. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	Statistic
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2867	15.6	0.307	0.311	2123	13.6	1.00	10	10	3
	Data ($\lambda_2 = 150$)	2972	12.9	0.307	0.309	2373	13.6	1.00	9	7	6
	Data ($\lambda_2 = 180$)	3332	11.2	0.303	0.301	2701	13.5	1.00	4	4	10
	Data ($\lambda_2 = 210$)	3719	9.7	0.299	0.294	3101	13.3	1.00	2	2	11
	Data ($\lambda_2 = 240$)	4104	8.4	0.296	0.285	3499	13.2	1.00	1	1	12
	Prod $U^{ref} = 0.06$	2992	13.4	0.307	0.309	2432	13.6	1.00	8	9	5
	Prod $U^{ref} = 0.08$	3197	11.5	0.306	0.306	2730	13.6	1.00	6	5	8
	Prod $U^{ref} = 0.10$	3375	10.0	0.305	0.303	2957	13.5	1.00	3	3	9
	CA $U^{ref} = 0.04$	2856	16.1	0.307	0.312	2081	13.6	1.00	11.5	11.5	1.5
	CA $U^{ref} = 0.06$	2856	16.1	0.307	0.311	2081	13.6	1.00	11.5	11.5	1.5
	CA $U^{ref} = 0.08$	2999	13.0	0.307	0.310	2456	13.6	1.00	7	8	4
	CA $U^{ref} = 0.10$	3242	12.5	0.306	0.307	2850	13.6	1.00	5	6	7
21-40	Data ($\lambda_2 = 120$)	3015	6.3	0.580	0.603	2500	25.1	1.00	12	3	1
	Data ($\lambda_2 = 150$)	3544	6.4	0.536	0.550	2904	23.4	1.00	9	5	3
	Data ($\lambda_2 = 180$)	3975	6.5	0.496	0.503	3269	21.7	1.00	6	6	6
	Data ($\lambda_2 = 210$)	4311	6.8	0.458	0.459	3511	20.3	1.00	4	7	9
	Data ($\lambda_2 = 240$)	4539	7.1	0.425	0.421	3659	19.0	1.00	3	8	10
	Prod $U^{ref} = 0.06$	3423	8.6	0.532	0.545	2944	23.1	1.00	10	12	4
	Prod $U^{ref} = 0.08$	3698	8.5	0.506	0.509	3189	22.1	1.00	8	10	5
	Prod $U^{ref} = 0.10$	3893	8.6	0.486	0.485	3343	21.2	1.00	7	11	8
	CA $U^{ref} = 0.04$	3093	4.0	0.572	0.583	2646	24.7	1.00	11	1	2
	CA $U^{ref} = 0.06$	4088	4.9	0.488	0.482	3502	21.3	1.00	5	2	7
	CA $U^{ref} = 0.08$	4746	6.3	0.417	0.405	4019	18.7	1.00	2	4	11
	CA $U^{ref} = 0.10$	5122	7.6	0.358	0.345	4270	16.4	1.00	1	9	12

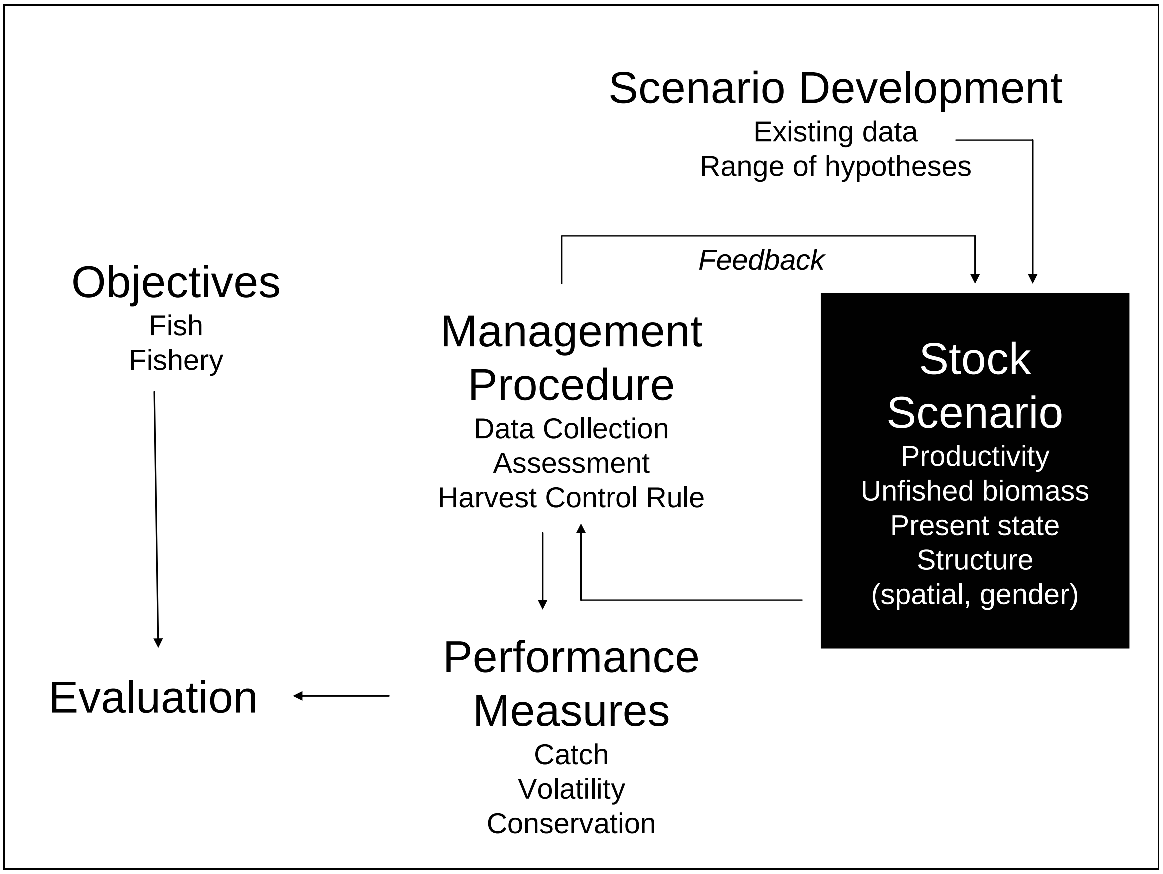


Figure 1 Schematic representation of the management strategy evaluation process including the simulation feedback loop between the state of the stock and the iterative application of management procedures.

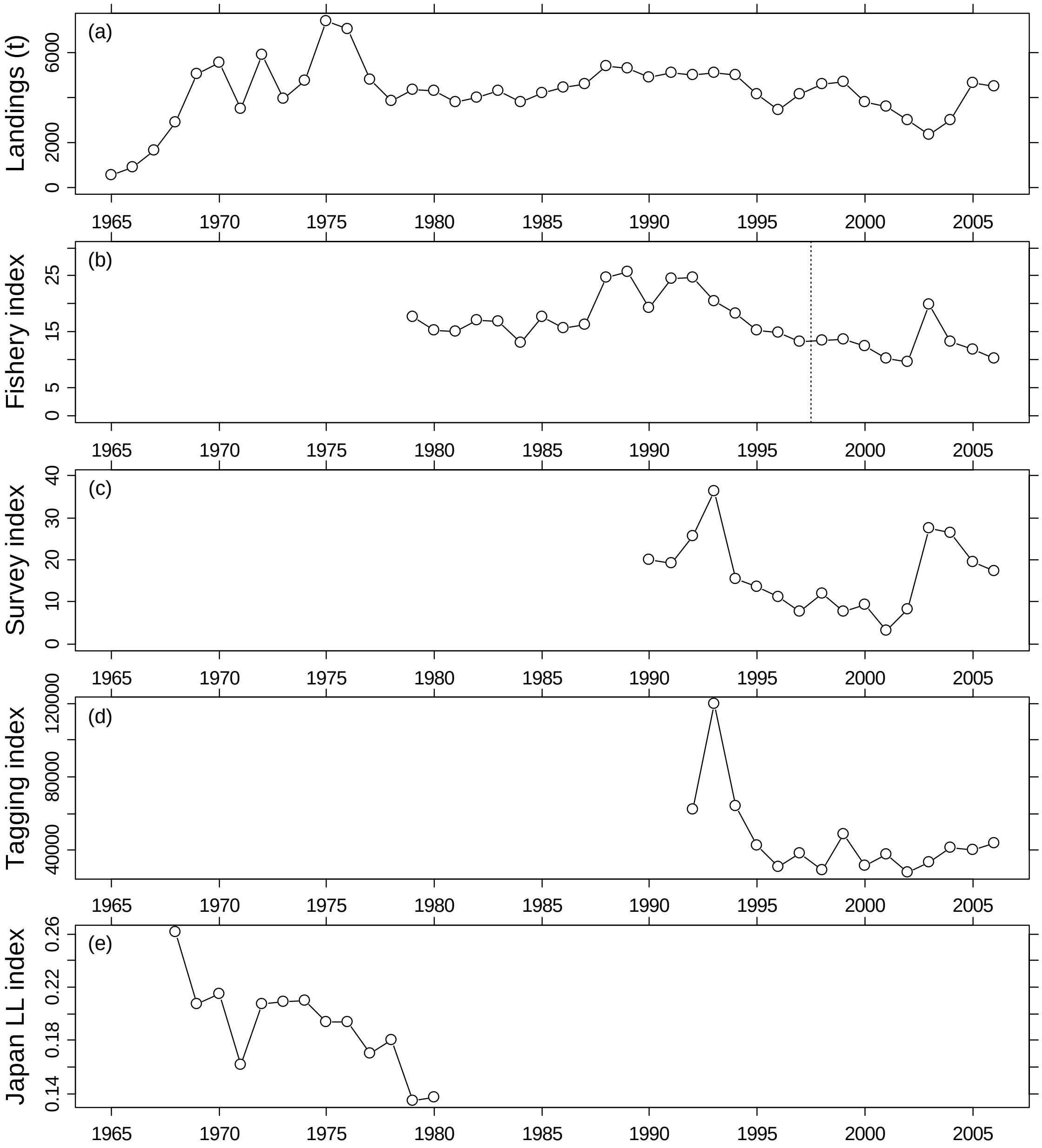


Figure 2 Annual sablefish landings (a) and stock abundance indices include (b) nominal fishery CPUE (kg/trap), (c) standardized survey mean CPUE (kg/trap), (d) tagging biomass index, and (e) Japanese LL CPUE index.

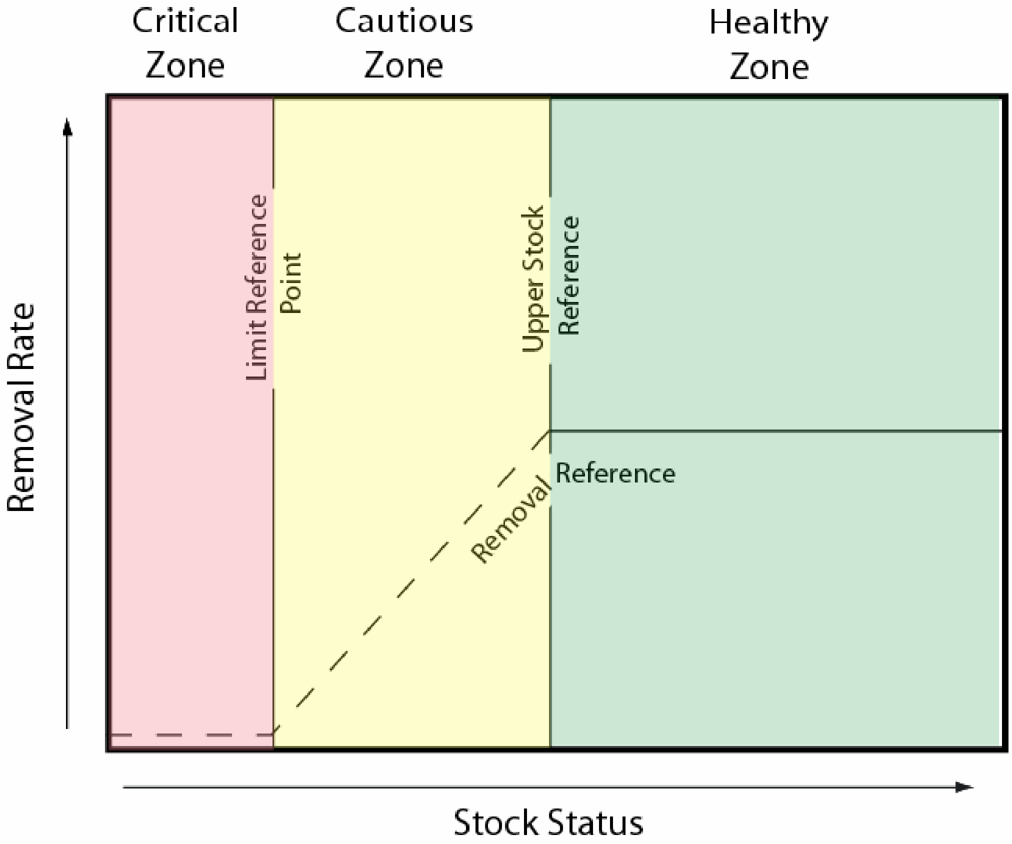


Figure 3 Conceptual diagram of a harvest strategy consistent with the precautionary approach (from DFO 2006).

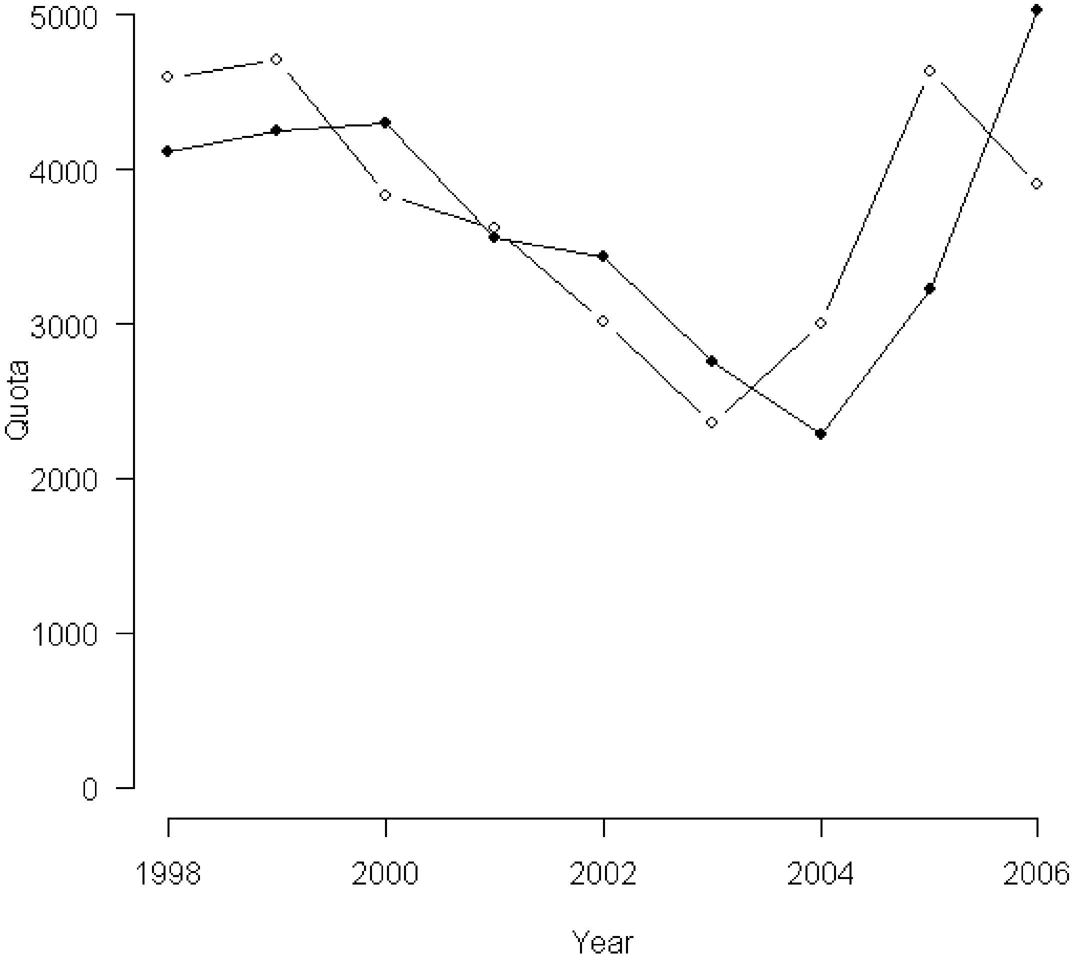


Figure 4 Observed annual quotas for B.C. sablefish (1998-2006; open circles) and predicted (solid circles) values based on equation (1) with $\lambda_1 = 0.75$ and $\lambda_2 = 299$.

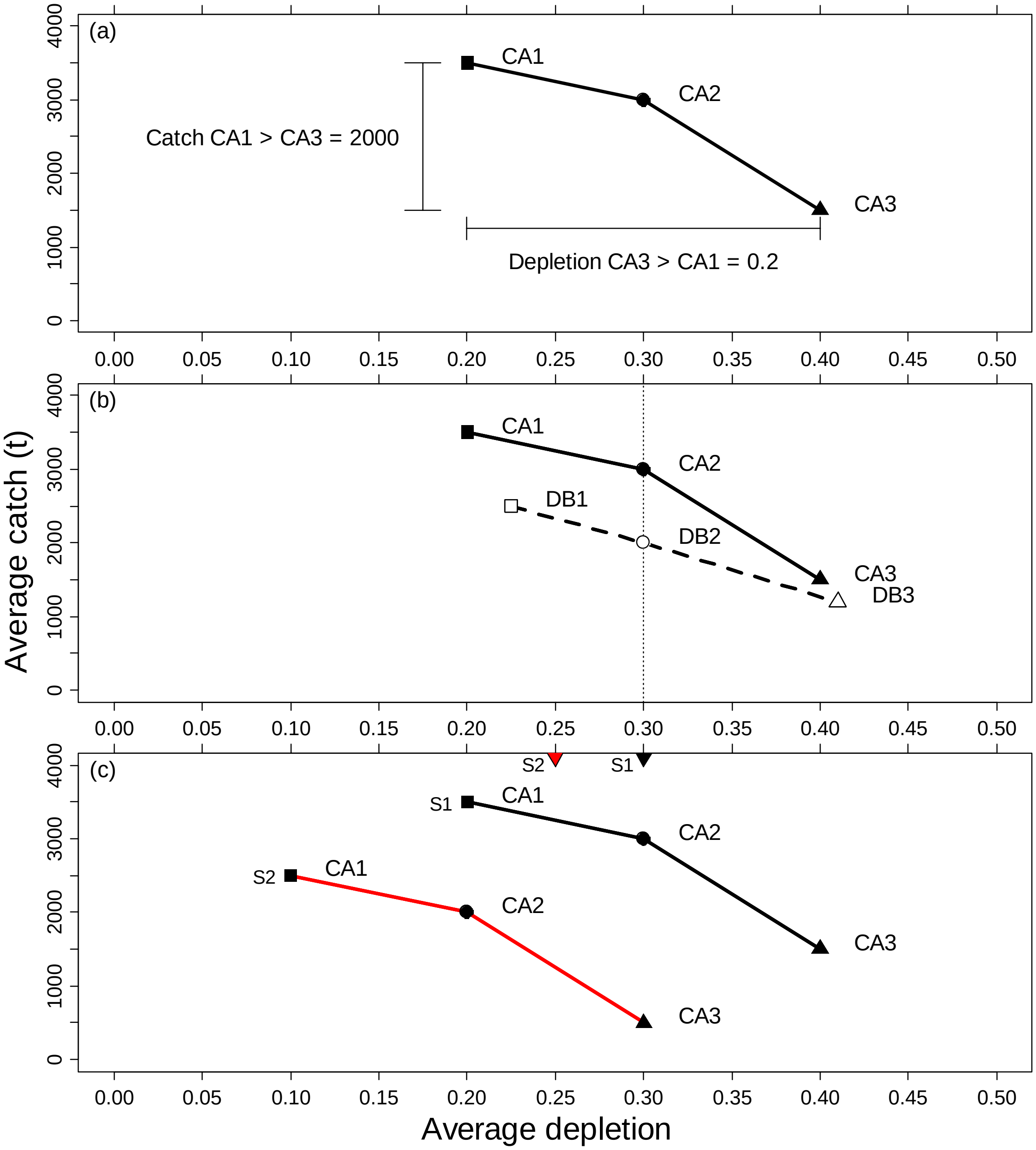


Figure 5 Example trade-offs between median average catch and median average depletion. This figure demonstrates how these plots are constructed and therefore does not represent actual results. Panel (a) compares catch-age procedures CA1 - CA3 with lines to indicate the difference in average catch (vertical) and depletion (horizontal) between CA1 and CA3; panel (b) compares three data-based (DB1 - DB3) and three catch-age model-based procedures; and panel (c) shows the same three catch-age procedures for two hypothetical stock scenarios S1 (red) and S2 (black). Note the colored and labeled indicators along the top of the plot showing the initial stock depletion (e.g. 2006) for each scenario.

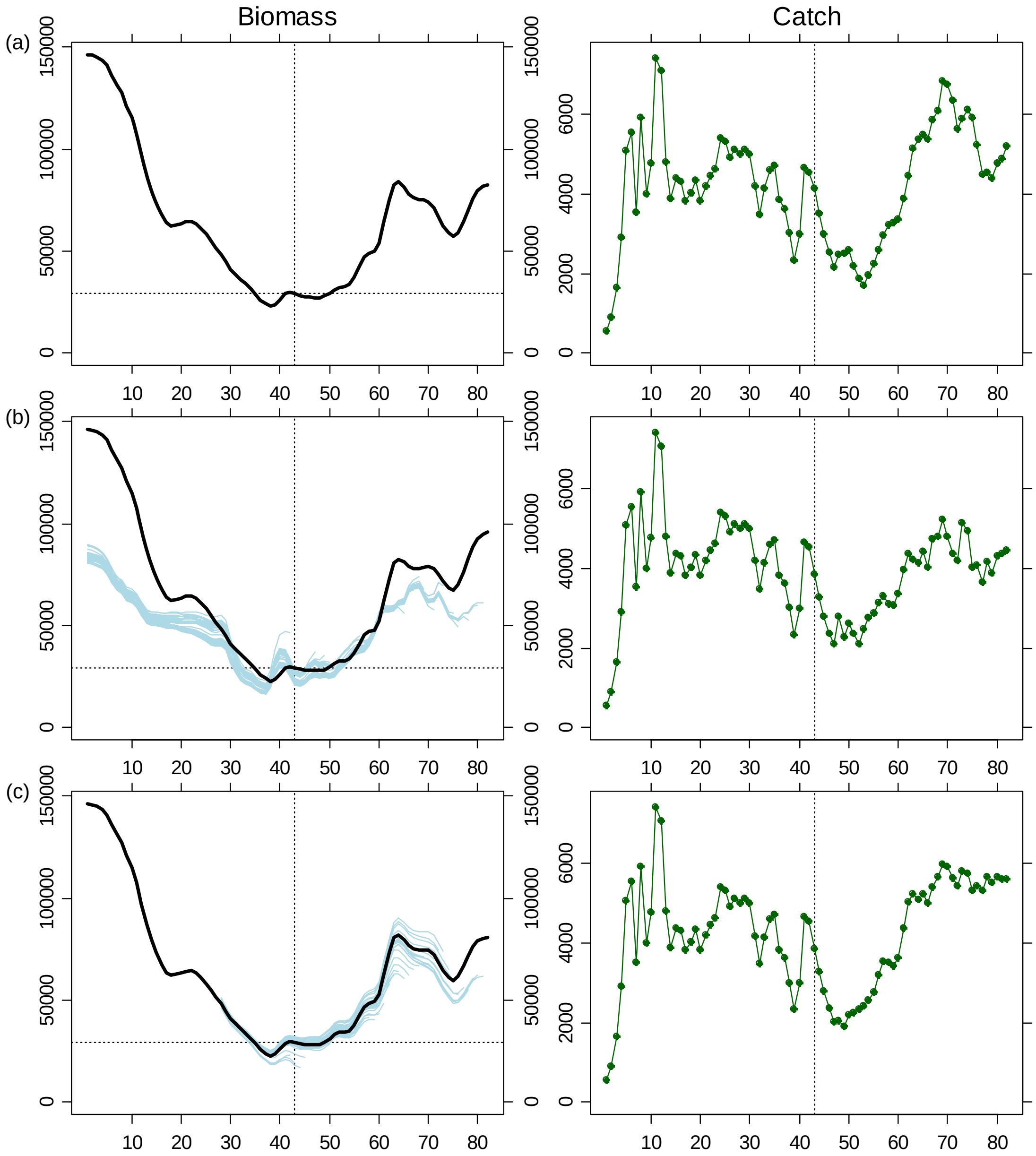


Figure 6 Results for a single replicate of scenario S3 when spawning biomass shows a strongly increasing trend under specific (a) data-based, (b) production model-based, and (c) CA model-based management procedures. Biomass and catch units in metric tons. All procedures use VHR decision rules with the first five years constrained to a maximum 15% change in catch. The vertical dotted line indicates the start of the simulation period and the horizontal dotted line indicates spawning biomass in the first year of procedure implementation.

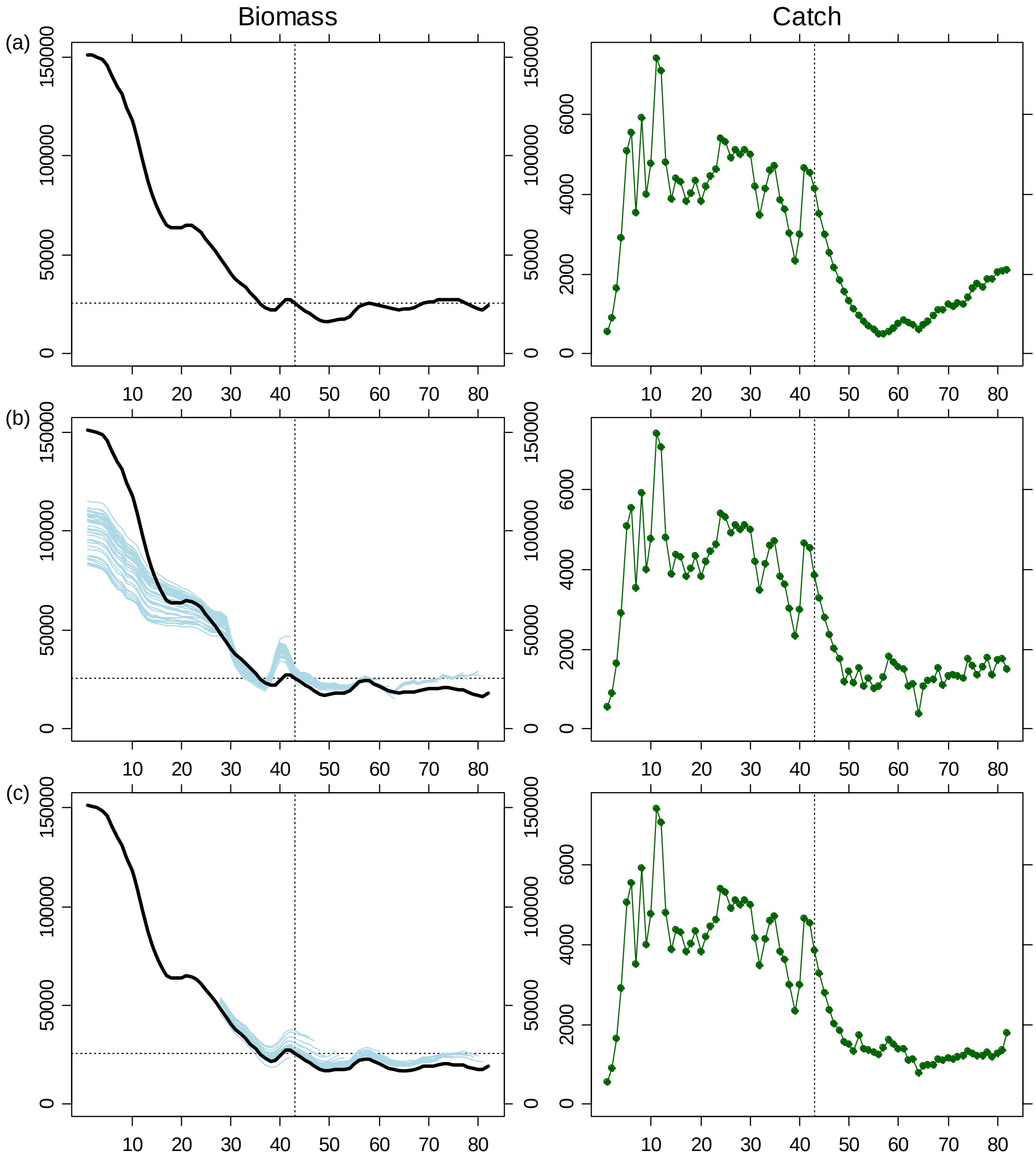


Figure 7 Results for a single replicate of scenario S1 when spawning biomass shows a stable future trend under specific (a) data-based, (b) production model-based, and (c) CA model-based management procedures. Biomass and catch units in metric tons. All procedures use VHR decision rules with the first five years constrained to a maximum 15% change in catch. The vertical dotted line indicates the start of the simulation period and the horizontal dotted line indicates spawning biomass in the first year of procedure implementation..

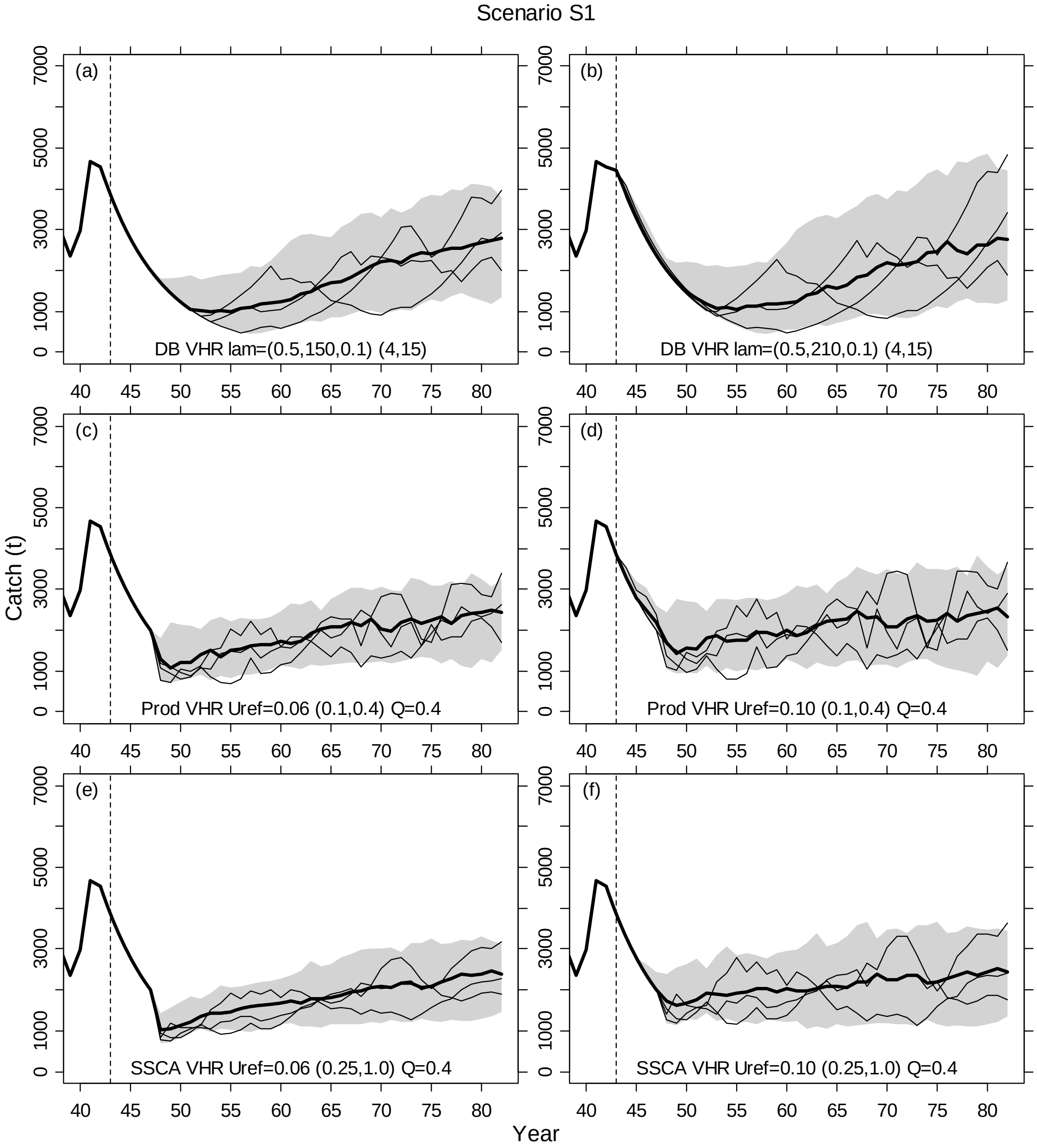


Figure 8 Trajectories of catch under scenario S1 using data-based procedures with (a) λ_2 = 150 and (b) λ_2 = 210; production model with (c) U^{ref} = 0.06 and (d) U^{ref} = 0.10; and catch-age model with (e) U^{ref} = 0.06 and (f) U^{ref} = 0.10. Vertical dashed lines indicate year 2006. Trajectories are summarized by the median (thick black line), three individual simulation replicates (thin black lines), and 5th to 95th percentiles (shaded area).

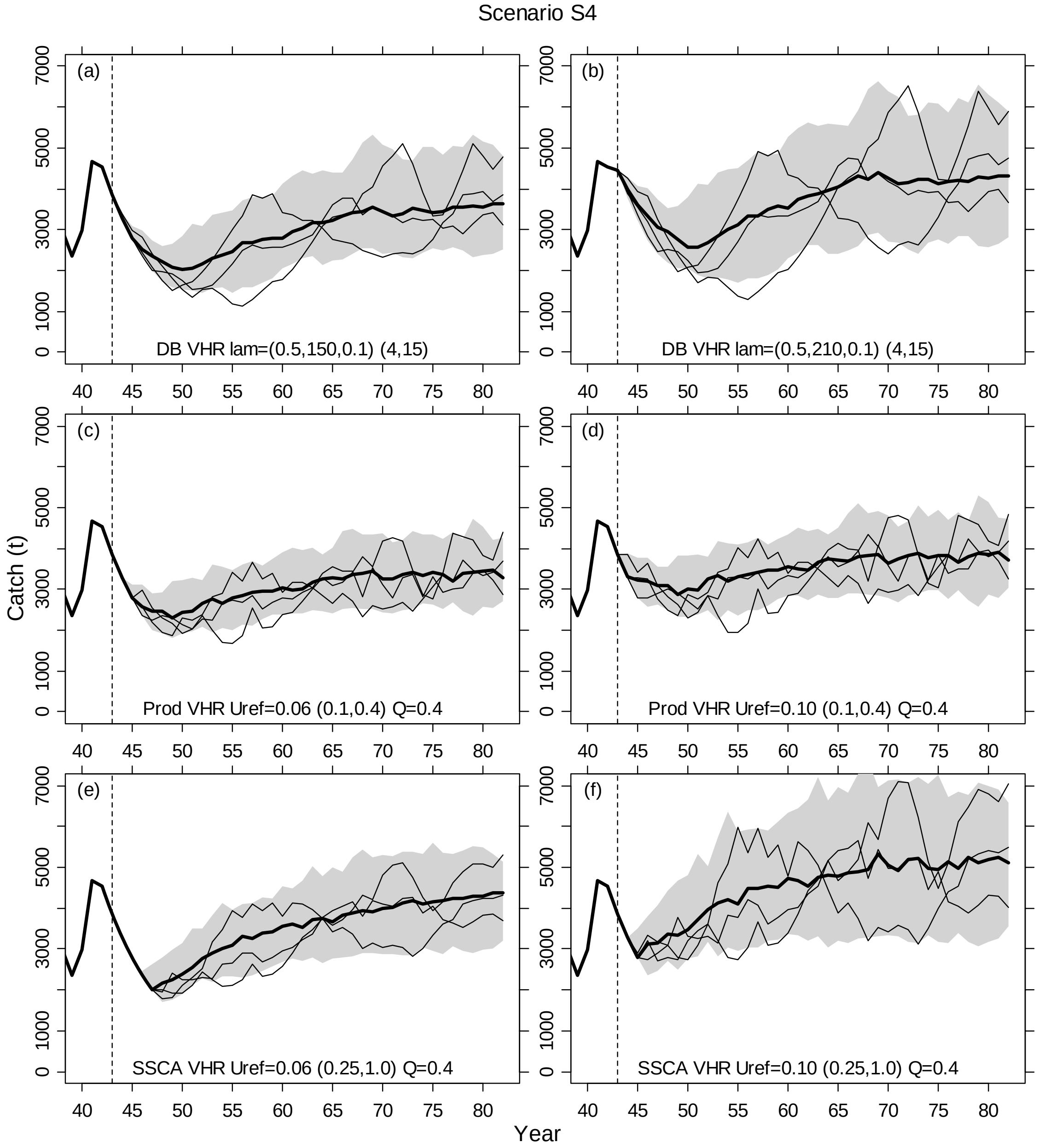


Figure 9 Trajectories of catch under scenario S4 using data-based procedures with (a) $\lambda_2 = 150$ and (b) $\lambda_2 = 210$; production model with (c) $U^{ref} = 0.06$ and (d) $U^{ref} = 0.10$; and catch-age model with (e) $U^{ref} = 0.06$ and (f) $U^{ref} = 0.10$. Vertical dashed lines indicate year 2006. Trajectories are summarized by the median (thick black line), three individual simulation replicates (thin black lines), and 5th to 95th percentiles (shaded area).

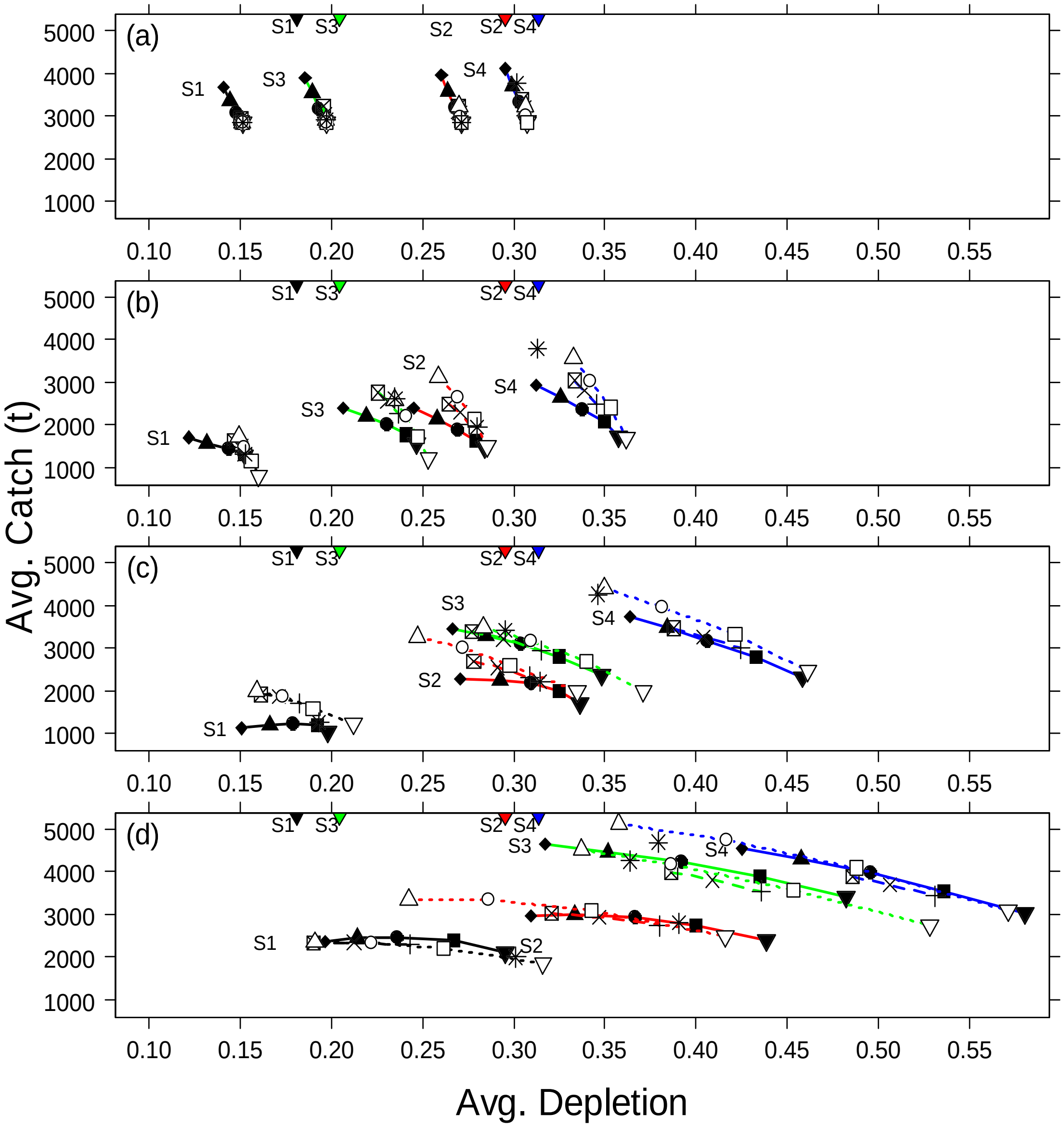


Figure 10 Trade-off relationship between median average catch (t) and median average spawning biomass depletion for scenarios S1-S4 measured over (a) 1-5, (b) 6-10, (c) 11-20 and (d) 21-40 year time intervals. Symbols are ordered from left to right corresponding to data-based with λ_2 = {240, 210, 180, 150, 120} (♦, ▲, ●, ■, ▼), production model-based with U^{ref} = {0.10, 0.08, 0.06} (☒, ×, +), and CA model-based with U^{ref} = {0.10, 0.08, 0.06, 0.04} (△, ○, □, ▽) procedures. Perfect-information procedures with U_{MSY} = {0.04, 0.06, 0.08, 0.10} are indicated by asterisks. Inverted triangles and labels along the upper *x*-axis indicate initial depletion values for each scenario.

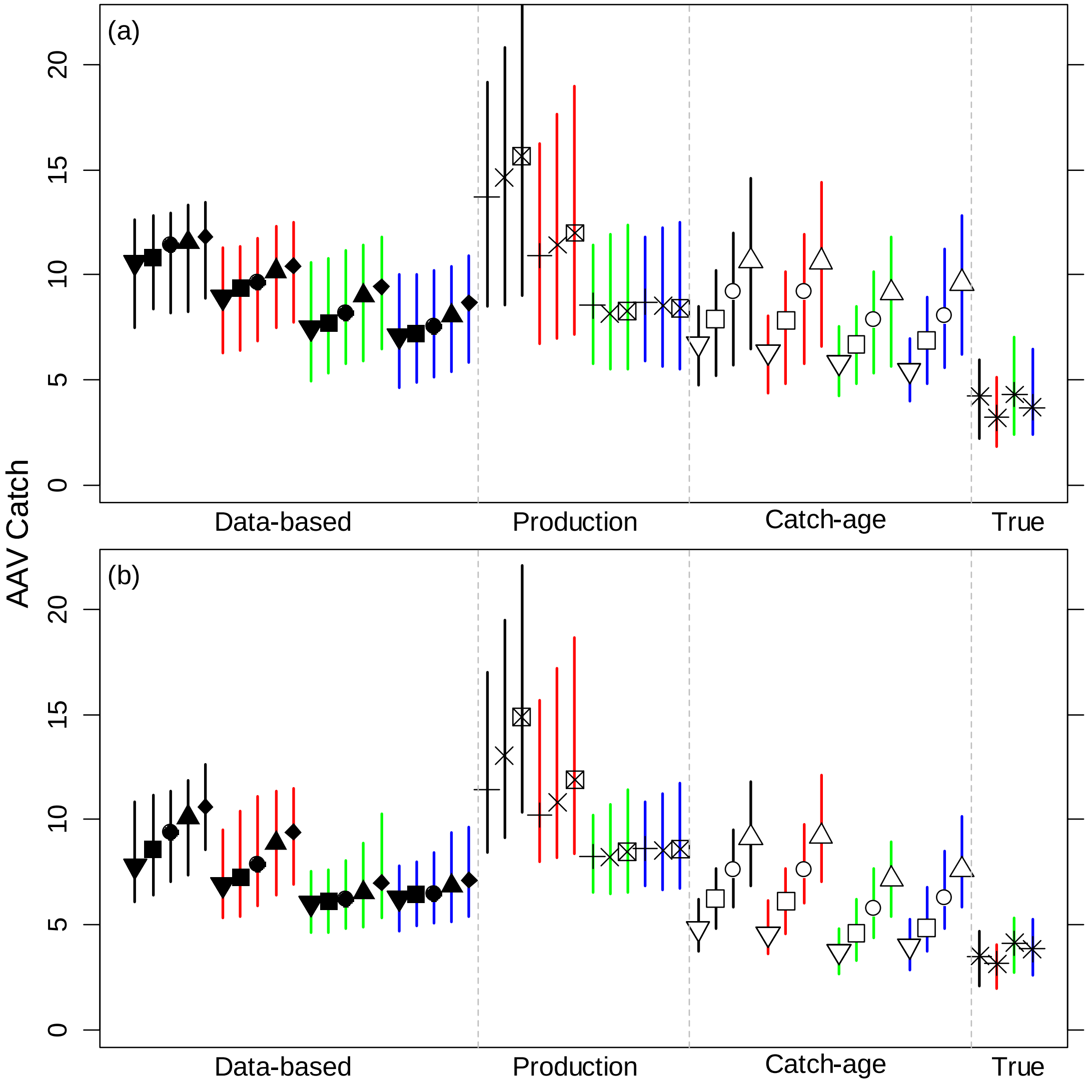


Figure 11 Summary of catch variability performance over projection years 11-20 (panel a) and 21-40 (panel b) for three selected classes of management procedures. Results are shown for scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Results for the quasi-perfect information procedures are shown as “True”. Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

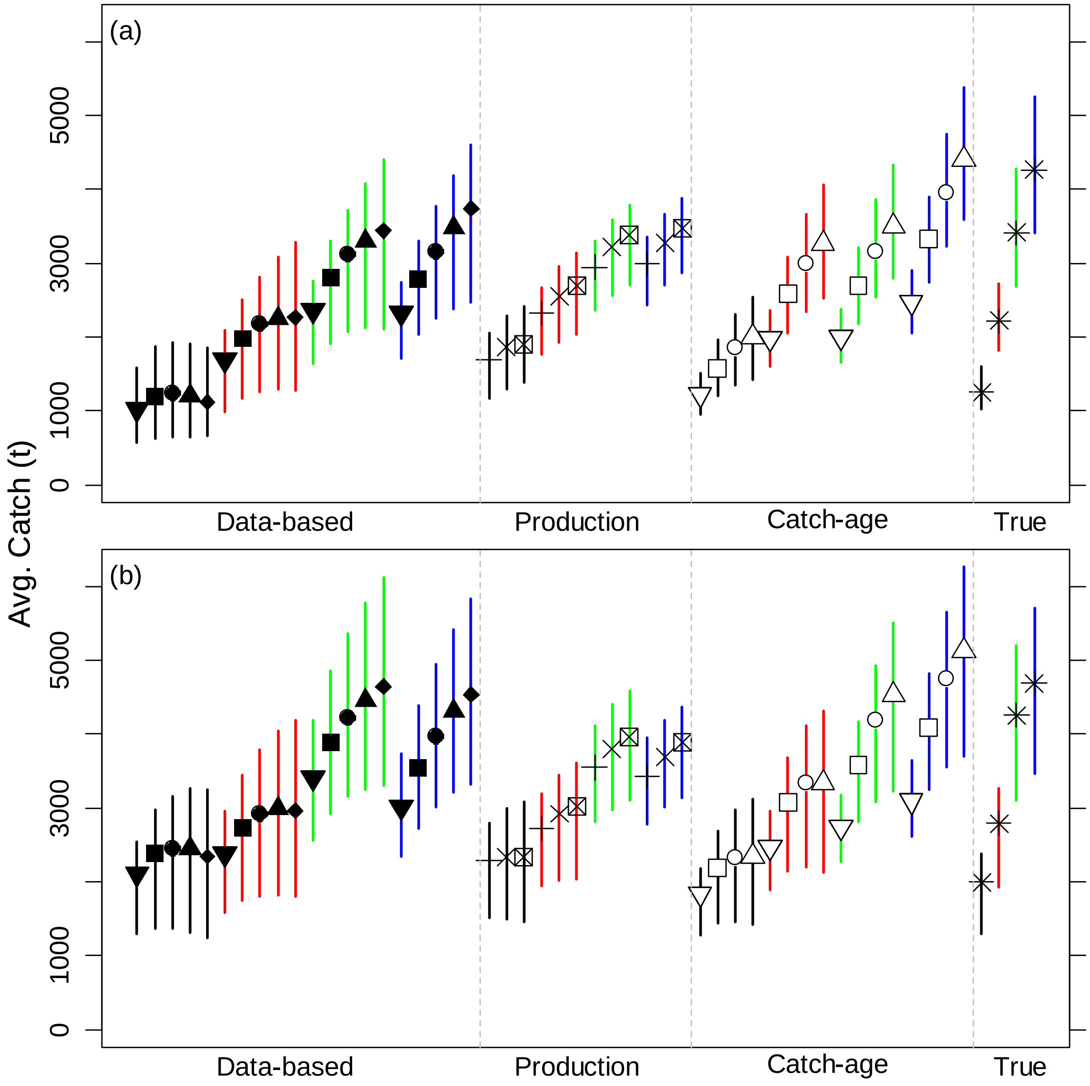


Figure 12 Average annual catch over projection years 11-20 (panel a) and 21-40 (panel b) for three selected classes of management procedures under scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Perfect-information procedures are shown as "True". Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

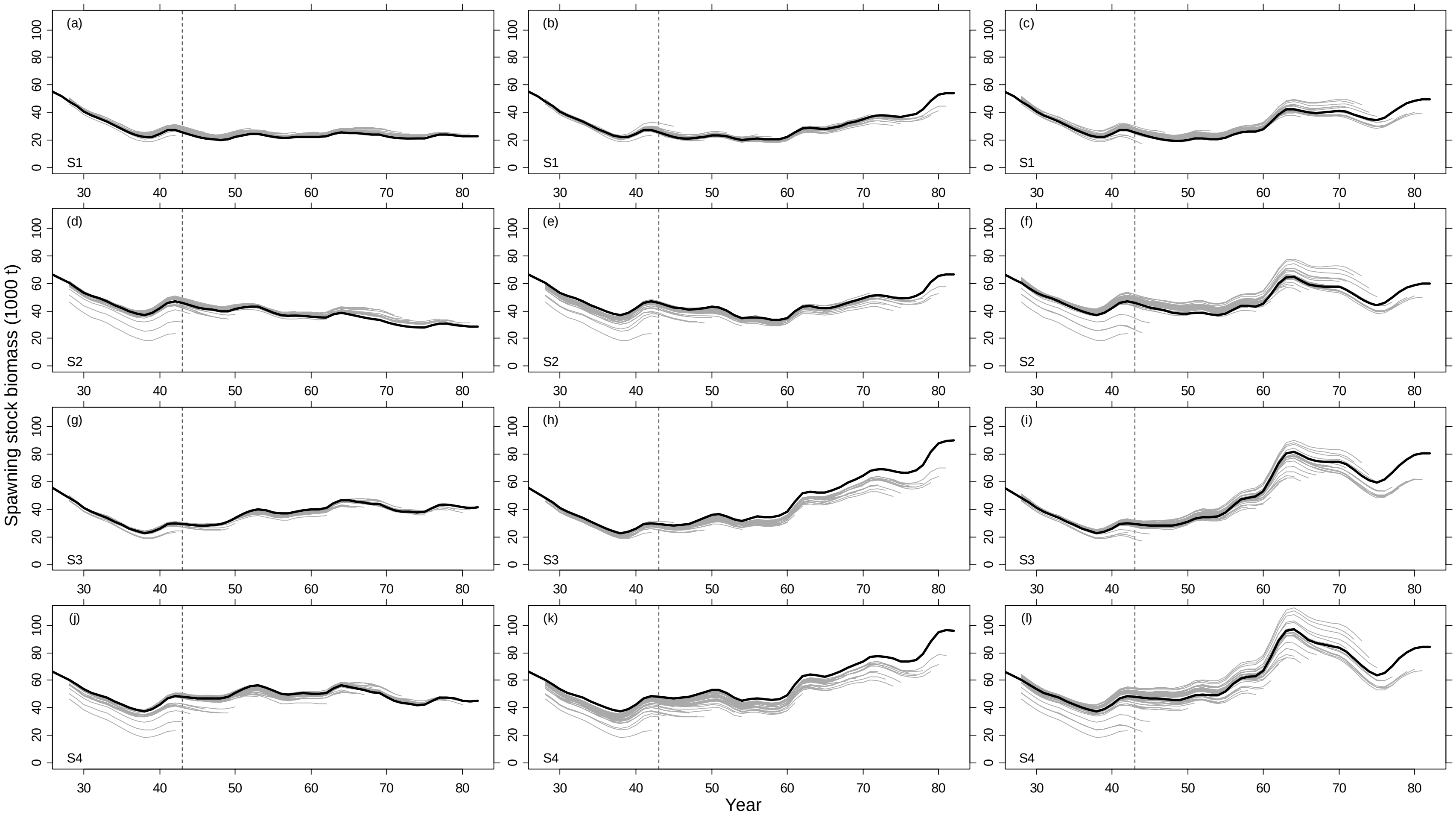


Figure 13 Retrospective patterns in estimated spawning biomass for declining/stable (left column), increasing (center column), and increasing/declining (right column) true spawning biomass trajectories (solid black lines). Solid gray lines represent successive spawning biomass trajectories estimated by the catch-age model under a VHR rule with U^{ref}=0.08.

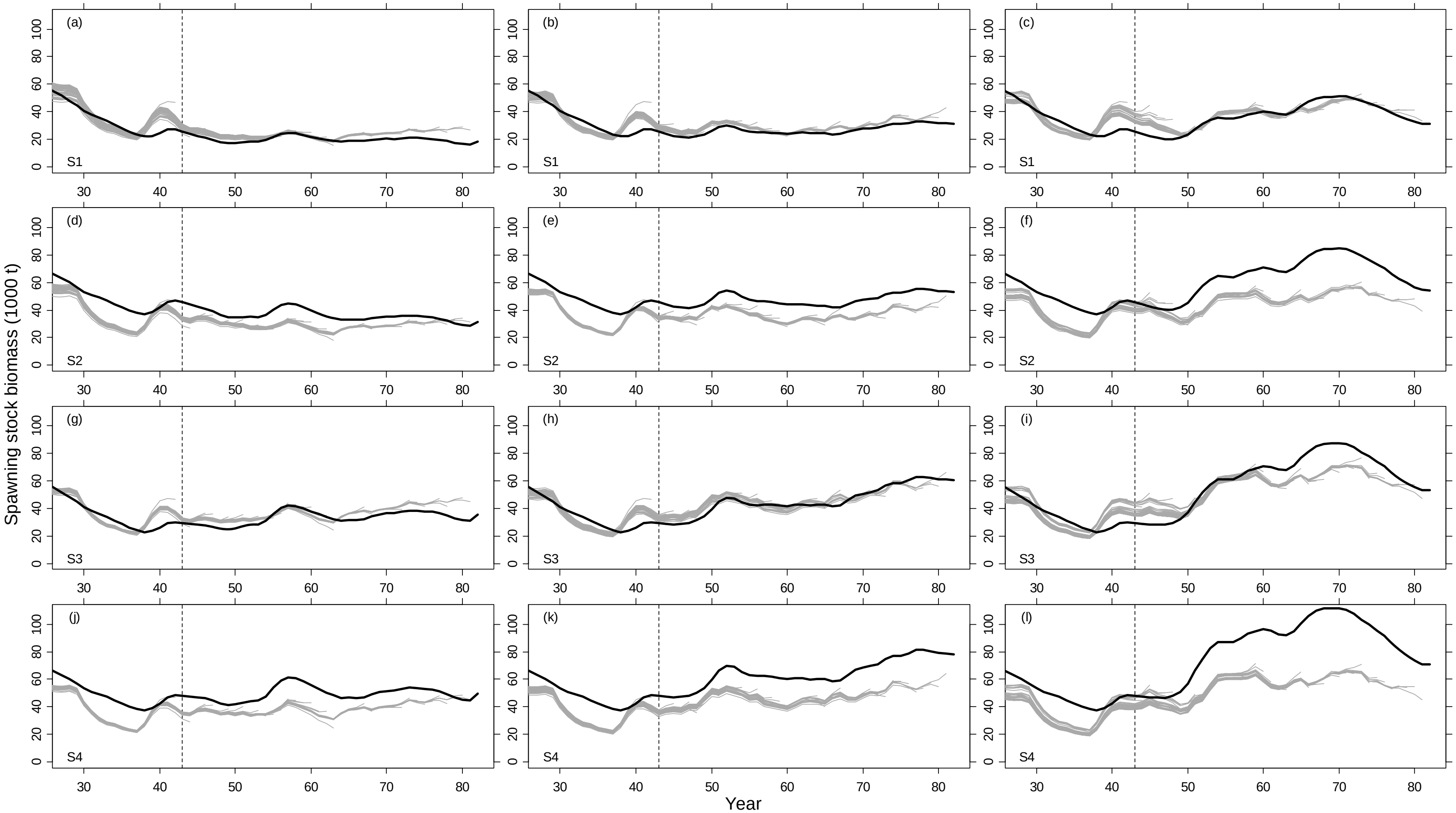


Figure 14 Retrospective patterns in estimated spawning biomass for declining/stable (left column), increasing (center column), and increasing/declining (right column) true spawning biomass trajectories (solid black lines). Solid gray lines represent successive spawning biomass trajectories estimated by the production model under a VHR rule with U^{ref}=0.08.

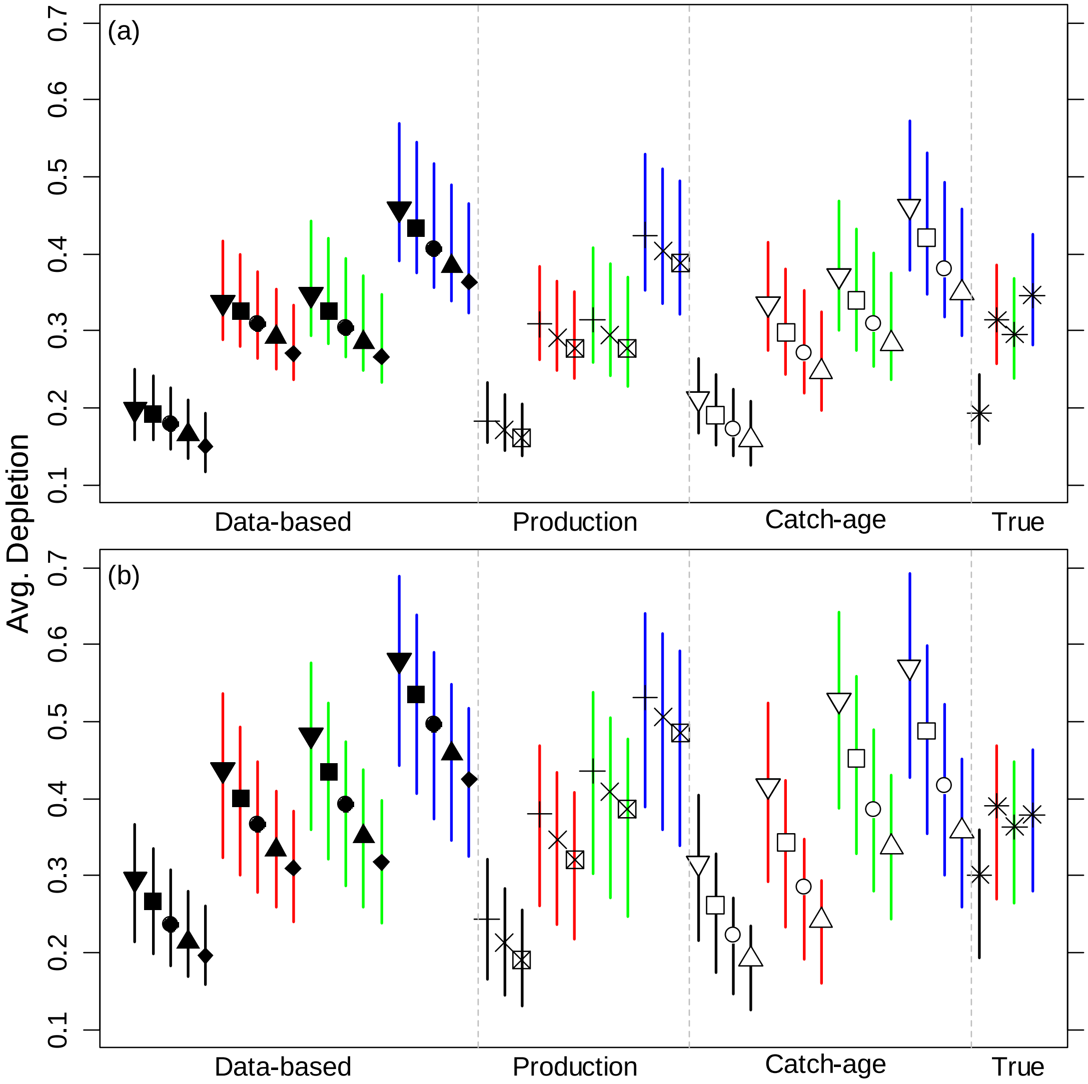


Figure 15 Average spawning biomass depletion over projection years 11-20 (panel a) and 21-40 (panel b) for three selected classes of management procedures under scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Perfect-information procedures are shown as “True”. Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

Appendix A Request for Science Information and/or Advice

REQUEST FOR SCIENCE INFORMATION AND/OR ADVICE
Centre of Science Advice - Pacific

PART 1: DESCRIPTION OF THE REQUEST (to be filled by the Branch requesting Information/Advice)

Date (Initial submission to Science): August 29, 2007

Directorate, Branch or group initiating the request and category of request

Directorate/Branch	Category of Request
☒ Fisheries and Aquaculture Management	
☐ Oceans and Habitat Management	☐ Stock Assessment
☐ Policy	☐ Species at Risk
☒ Science	☐ Habitat
☒ Other (please specify): Industry Supported	☐ Aquaculture
	☐ Ocean Action Plan
	☒ Other (please specify): Management procedures

Initiating Branch Contact:

Name:	**A.R. Kronlund**	Telephone Number: (250)756-7108
Email:	kronlunda@pac.dfo-mpo.bc.ca	Fax Number: (250)756-7053

Issue Requiring Science Advice (i.e., “the question”):
(Issue posed as a question for Science response)

Management strategy evaluation methodology review with application to sablefish.

Rationale for Advice Request:
(What is the issue, what will it address, importance, scope and breadth of interest, etc.)

Sablefish fishery objectives are not well-defined as noted in recent stock assessments and by PSARC. Using the context of management procedure evaluation, this work is required to solicit candidate objectives from fishery managers and fishery stakeholders and to illustrate the trade-offs that must be considered in the selection of a management procedure. The methodology has general application beyond sablefish and has seen application for Australian and South African fisheries, but has not been used for Canadian fisheries.

Intended Uses and Potential Impacts of Advice within DFO:
(Who will be the end user of the advice, e.g., DFO, another government agency or Industry? What impact could the advice have on other sectors?)

The methodology, which is consistent with the aims and requirements of the FAO Precautionary Approach, can be used by fishery managers, ENGOs and industry stakeholders to understand trade-offs among yield, inter-annual variability in yield, and conservation. It can easily incorporate the requirements of Canada’s policy on harvest strategies compliant with the Precautionary Approach (DFO 2006). It is important to note that this is a methodology paper and that there is no explicit advice for management decision making regarding harvest in 2008 or beyond. Harvest advice will be provided at a subsequent peer-review meeting, taking into account review recommendations and following consultation with fishery managers and stakeholders.

Potential Impacts of the Advice on the Public:
(Who will be impacted by the advice and to what extent?)

Potential impacts include the eventual transition to a long-term management procedure for sablefish in place of the current approach of traditional data-fitting assessments that provide short-term (e.g., 1-2 year) advice. The presentation of results can enhance discussion of fishery objectives and trade-offs between fishery managers and stakeholders, but requires science guidance during the discussions to explain and adapt to refinement of objectives.

Date Advice Required:

Latest possible date to receive Science advice: Not specified.

Rationale: N/A.

Appendix B Data

Overview

Landings data (retained catch) used for the simulation analysis were summarized for calendar years 1965 to 2006 from the GFCatch, PacHarvSable, and FOS databases maintained by Fisheries and Oceans Canada. Landings from seamount fishing were excluded where they could be identified since seamount harvest is not included in the coast-wide quota management area. Landings data prior to 1965 are available but averaged less than 1,000 t after 1920 prior to the ramping-up of the Canadian domestic sablefish fishery in the late 1960s (Figure B-1, McFarlane and Beamish 1983). Total annual landings as high as 5,956 metric tons (t) were realized during the 1910s, however landings remained modest from 1920 to 1965, ranging between 209 t and 1,895 t. Exploitation increased in the late 1960s with the arrival of foreign longline fleets from Japan, the US, the USSR and the Republic of Korea (McFarlane and Beamish 1983, Figure B-1). The largest annual landings of sablefish occurred during this period with a peak 7408 t removed in 1975. Declaration of the Canadian 200 mile Economic Exclusive Zone in 1977 ended unrestricted foreign fishing. However, some foreign fishing was allowed between 1977 and 1980 to utilize yield surplus to Canadian domestic fleet needs. Total landings have ranged from 2,345 t (2003) to 7,408 t (1975) since 1969 and averaged about 4741 t over the 1969 to 1999 period and 3570 t from 2000 to 2006 (Figure B-1).

The history of sablefish fishery management is summarized in Table B-1. The table contains a list of the annual total allowable catches (TACs) and quota allocations to the directed sablefish “K” fleet, the non-directed trawl “T” fleet, First Nations, and science projects. Landings by fishing year are also listed though note that the timing and duration of fishing years has changed over time, e.g., when an August 1 start date for the directed sablefish fishery was instituted in 1999 a fishing year of 19 months duration resulted. Also note that the trawl fishing year is defined as April 1 to March 31. Thus the “Total commercial allocation” does not apply to a 12 month period. For example, the 282 t trawl allocation for 2007/08 begins April 1, 2008 which is 8 months after the start of the 2007/08 fishing year for sablefish. Fishery landings data are incomplete for 2007.

Details of the 2001/2002 to 2006/2007 fishing year quotas and allocations are provided in Table B-2 to illustrate allocation and the carry-over provisions. The "carry-over" provision is a management tactic designed to allow individual quota holders the opportunity to delay taking current fishing quota until the following year, and to accommodate over-runs of quota in the current fishing year. The details of the overage/underage rules (i.e. "carry-over") have changed in two ways since their inception. First, the allowable percentages of overage and underage have been assigned various combinations of 5 and 10 percent over time (Haist et al. 2004). Second, the percentage overage was applied to the quota *remaining* to the vessel when the overage occurred until 1999, when the percentage was applied to the vessel's *total* quota (Haist et al. 2004). The point is that the TAC in a given year may not be met exactly for reasons of the carry-over provisions.

Canadian landings since 1951 have been reported by longline, trawl, and trap gear (Figure B-1, Table B-3). The fishery has been managed since 1981 under quotas allocated to the "K" licence (longline and trap gear) and "T" licence (trawl gear). Sablefish are caught incidentally in the halibut (*Hippoglossus stenolepis*) longline hook fishery, directed "Zn" rockfish longline hook, and there were allocations to research and to First Nations food fisheries (Table B-2). Since 1981, the trawl fishery has been allocated a fixed 8.75% of the total allowable catch based on historic average trawl landings.

Longline hook was the dominant gear type in the directed sablefish fishery for most years until 1973. At this time, the trap fishery began to develop and the proportion of the landings taken by longline gear declined (Figure B-1). By 1978, trap gear clearly accounted for the majority of landings. During the period from 1990 to 1992, the first three years of Individual Vessel Quota (IVQ) management, the proportion of landings attributed to longline hook gear was high (17 % to 28%) but then dropped to below 12% over the 1993 to 1998 period. The initial increase was due to large vessels that developed longline hook operations for other groundfish species that included sablefish caught under quota. In this way these vessels could fish most of the year. The subsequent decline in the proportion of longline hook landings was attributed to a move away from the multi-species longline hook approach in favor of dedicated trap gear fishing with transferable quota.

The transferable quota system allowed sablefish vessels to fish most of the year and traps were chosen as the most efficient gear. An increase in the proportion of the landings taken by longline hook gear from 1999 through 2004 may reflect a move back to a multiple target species approach, i.e., so-called "combination fishing" where halibut "L" or rockfish (*Sebastes*) "Zn" licenses may be fished in conjunction with a sablefish "K" license to avoid discarding imposed by license regulation. The increase in longline hook landings could also reflect reduced availability of sablefish to trap gear during the 1999 through 2002 period (Kronlund et al. 2002). In 2006 the proportion of landings by longline hook gear increased relative to trap gear, possibly as a result of the early effects of the Groundfish Integration Pilot Proposal as access to sablefish quota by non-"K" licensed vessels was permitted (Koolman et al. 2007).

Annual catches from Alaska and the west coast United States are plotted in Figure B-2 with landings from B.C. to illustrate the relative sizes of the respective sablefish fisheries. Data for Alaska are taken from Table 3.1a of Hanselman et al. (2007) and data for the US west coast from Table 2 of Schirripa (2007). Annual coastwide catches are dominated by the Gulf of Alaska fishery, which has ranged between 13,575 t (1999) and 17,782 t (2004) since 1996. During that period, Gulf of Alaska catches have been 8% (2000) to 26% (2006) lower than the total allowable catch (average 17% from 1996 to 2006) due in part to increasing interaction with marine mammals, restrictions on areas available to trawl, and the economics of fishing the Aleutian management area (Dana Hanselman, *pers. comm.*). In contrast to the catch trends in Alaska, sablefish catches in B.C. were not significantly increased during the late 1980s and early 1990s in response to increased abundance estimated for the Gulf of Alaska. Catches from the U.S. west coast have declined relatively smoothly since the early 1980s, with the exception of two downward excursions in the catch in 1998 and 2002 which were primarily due to parameter sensitivities in the stock assessment model (e.g., Schirripa 2002). Catches in all jurisdictions show a general decline after 1990 and increased following lows in 2001 and 2002 due to above average recruitments from the 1999/2000 year classes.

Data used for management procedure simulations

Landings were grouped by various sources to allow gear allocation during simulation experiments (Table B-4). In particular, the following data were combined:

- Foreign longline hook landings are the sum of Japanese and Republic of Korea longline hook landings;
- US landings from 1965 to 1980 are assumed to be taken by trawl gear;
- Trawl landings are the sum of U.S., U.S.S.R. and Canadian domestic trawl landings;
- Longline hook landings are the sum of domestic longline hook plus minor research longline retained catches (where they could be identified);
- Trap survey research catches were separated from commercial trap fishery catches;
- Landings attributed to “Other” were ignored (maximum 10 t in 1983).

Stock indices input to the operating model and assessment models (standardized trap survey CPUE only) are provided in Table B-5. Catch rate data for the Japanese longline fishery are available but were not used in the operating model for the analysis presented in this paper. Age proportions from commercial trap fishery and standardized trap survey sources are provided in Table B-6. Ages obtained using the burnt-otolith section method were pooled by sex and a minimum age class of age 3 and plus group of age 25 were selected.

Table B-1 Summary of management history. Note that the 1999/2000 fishing year was 19 months in duration to accommodate a shift in the fishing year from Jan 1 to August 1. Preliminary data for 2007/2008 current as of September 2007.

Year	Fishery	Assessment Yield Rec.	TAC	K Quota	T Quota	First Nations	Research	Landings FY	Date Open	Date Closed	Days Open	FY Days
1981	Derby		3500	3190	310			3830	01-Feb-81	04-Oct-81	245	245
1982	Derby		3500	3190	310			4028	01-Feb-82	22-Aug-82	202	202
1983	Derby		3500	3190	310			4346	01-May-83	26-Sep-83	148	148
1984	Derby		3500	3190	310			3827	01-Mar-84	22-Aug-84	174	174
1985	Derby		4000	3650	350			4193	01-Feb-85	08-Mar-85	35	92
									29-Mar-85	02-May-85	34	
									19-Jul-85	11-Aug-85	23	
1986	Derby		4000	3650	350			4449	17-Mar-86	21-Apr-86	35	63
									12-May-86	09-Jun-86	28	
1987	Derby		4100	3740	360			4630	16-Mar-87	10-Apr-87	25	45
									01-Sep-87	21-Sep-87	20	
1988	Derby		4400	4015	385			5403	06-Mar-88	26-Mar-88	20	140
									05-Apr-88	25-Apr-88	20	
									05-May-88	25-May-88	20	
									05-Jun-88	25-Jun-88	20	
									05-Jul-88	25-Jul-88	20	
									02-Aug-88	22-Aug-88	20	
									04-Sep-88	24-Sep-88	20	
1989	Derby		4400	4015	385			5324	14-Feb-89	28-Feb-89	14	112
									14-Mar-89	28-Mar-89	14	
									14-Apr-89	28-Apr-89	14	
									10-May-89	24-May-89	14	
									10-Jun-89	24-Jun-89	14	
									06-Jul-89	20-Jul-89	14	
									04-Aug-89	18-Aug-89	14	
									15-Sep-89	29-Sep-89	14	
1990	IVQ		4670	4260	410			4905	21-Apr-90	31-Dec-90	255	255
1991	IVQ	2,900-5,000	5000	4560	440			5112	01-Jan-91	31-Dec-91	365	365
1992	IVQ	2,900-5,000	5000	4560	440			5007	01-Jan-92	31-Dec-92	366	366

Year	Fishery	Assessment Yield Rec.	TAC	K Quota	T Quota	First Nations	Research	Landings FY	Date Open	Date Closed	Days Open	FY Days
1993	IVQ	2,900-5,000	5000	4560	440			5110	01-Jan-93	31-Dec-93	365	365
1994	IVQ	2,900-5,000	5000	4521	433			5002	01-Jan-94	31-Dec-94	365	365
1995	IVQ	2,725-5,550	4140	3709	356		29.48	4179	01-Jan-95	31-Dec-95	365	365
1996	IVQ	690-2,580	3600	3169	304		81.65	3471	01-Jan-96	31-Dec-96	366	366
1997	IVQ	6,227-16,285	4500	4023	386		45.36	4142	01-Jan-97	31-Dec-97	365	365
1998	IVQ	3,286-4,761	4500	4023	386		45.36	4592	01-Jan-98	31-Dec-98	365	365
1999/ 2000*	IVQ	2,977-5,052	4500	6395	386		45.36	7012	**01-Jan-99**	**31-Jul-00**	578	578
2000/ 2001	IVQ	3,375-5,625	4000	3555	350		45.36	3884	01-Aug-00	31-Jul-01	365	365
2001/ 2002	IVQ	4,000	2800	2657	342	45	45.36	3075	01-Aug-01	31-Jul-02	365	365
2002/ 2003	IVQ	4,000, revised to 2100-2800	2450	1883	206	45	45	2206	01-Aug-02	31-Jul-03	365	365
2003/ 2004	IVQ	Decision table	3000	2647	254	45	54	2983	01-Aug-03	31-Jul-04	365	365
2004/ 2005	IVQ	Decision table	4500	3995	384	45	75	4249	01-Aug-04	31-Jul-05	365	365
2005/ 2006	IVQ	Decision table	4600	4056	389	45	110	4498	01-Aug-05	31-Jul-06	365	365
2006/ 2007	IVQ	No Assessment	3900	3417	328	45	110	3950	01-Aug-06	31-Jul-07	365	365
2007/ 2008	IVQ	No Assessment	3300		384	45	35	-	01-Aug-07	31-Jul-08	365	365

Table B-2 TACs and allocations (nearest metric ton fresh round weight) for the sablefish 2001/02 fishing year to the 2007/08 fishing year. An in-season quota reduction in the 2001/02 fishing year of 910 t is shown as a carry-forward into 2002/03 designed to spread the quota reduction over two fishing years. Note that the fishing year for directed "K" sablefish is defined as August 1 to July 31, while the fishing year for trawl is defined as April 1 to March 31. Total "K" allocation, carryover, and IVQ available provided courtesy of the Groundfish Management Unit, DFO.

Allocation and Landings	**2001/02**	**2002/03**	**2003/04**	**2004/05**	**2005/06**	**2006/07**	**2007/08**
TAC	**2800**	**2450**	**3000**	**4500**	**4600**	**3900**	**3300**
Scientific purposes	25	45	54	65	110	110	35
First Nations allocation	45	45	45	45	45	45	45
Total commercial allocation	**3909**	**1179**	**2900**	**4389**	**4445**	**3745**	**3220**
Trawl "T" allocation (8.75%)	*342*	*206*	*254*	*384*	*389*	*328*	*282*
Sablefish "K" allocation	*3567*	*973*	*2647*	*3995*	*4056*	*3417*	*2938*
"K" scientific purposes	-	-	-	*10*	-	-	-
Sablefish "K" carry forward	153-(910)	79+910	35	87	231	329	-
Total "K" IVQ available	2806	1940	2669	4080	4284	3733	2938
Commercial landings (excluding seamounts)	3075	2206	2983	4249	4498	3950	*NA*

Table B-3 Annual sablefish landings (t) in Canadian waters by source from 1965-2006.

Year	Trapo	Res. Trap	Japan LL	ROK LL	Longline	Trawl	US (Trawl)	USSR Trawl	Other	Total
1965					193	262	92		0	547
1966			174		326	312	95		0	907
1967			1189		253	139	65		0	1646
1968			2390		292	167	65		15	2929
1969			4720		162	148	43		1	5074
1970			5142		142	166	104		1	5554
1971			3050		123	189	161			3523
1972			4236		400	688	582			5906
1973	746		2950		120	83	82	6		3986
1974	327		3866	129	41	122	227	65	2	4779
1975	469		4702	1263	152	280	541		1	7408
1976	303		3494	2335	89	382	473		0	7077
1977	215		2961	186	77	787	571		7	4803
1978	635		2103		57	131	948		8	3881
1979	1480		1112		277	276	1236		6	4387
1980	3211		199		249	335	317		3	4314
1981	3275				326	229				3830
1982	3438				344	246			0	4028
1983	3611				451	274			11	4347
1984	3275				365	187				3827
1985	3501				458	233				4193
1986	3277				619	552			1	4449
1987	2954				1269	407			1	4630
1988	3488				1274	637			3	5403
1989	3772				929	623			0	5324
1990	3072				1372	461				4905
1991	3494				1179	439			0	5112
1992	3710				849	449			0	5007
1993	4142				424	543			0	5110
1994	4051				468	483				5002
1995	3282				474	427			5	4189
1996	2984	15			279	191				3470
1997	3554	2			431	156				4142
1998	3772	0			444	376				4592
1999	3677	6			628	403				4714
2000	2745	13			752	326				3836
2001	2743	8			564	300				3614
2002	2159	20			564	267			0	3009
2003	1419	68			631	228				2345
2004	2129	48			465	345				2987
2005	3197	42			1145	277				4660
2006	2693	61			1327	445				4530

Table B-4 Landings (t) input to the operating model and assessment models, 1965-2006.

Year	Time Step	Trap	Research Trap	Foreign Longline	Longline	Trawl	Total
1965	1	0	0	0	193.2	353.9	547.1
1966	2	0	0	174	325.7	406.9	906.6
1967	3	0	0	1189	252.9	203.6	1645.5
1968	4	0	0	2390	292.3	232	2914.3
1969	5	0	0	4720	162.3	191.3	5073.6
1970	6	0	0	5142	142.1	269.9	5554
1971	7	0	0	3050	123	350.3	3523.3
1972	8	0	0	4236	399.7	1270.3	5906
1973	9	745.8	0	2950	119.8	170.8	3986.4
1974	10	327.1	0	3995	41.3	413.8	4777.2
1975	11	469.4	0	5965	152.2	820.8	7407.4
1976	12	303.4	0	5829	89.4	855	7076.8
1977	13	214.6	0	3147	77.1	1357.5	4796.2
1978	14	634.6	0	2103	57.2	1078.5	3873.3
1979	15	1480.1	0	1112	276.8	1512.1	4381
1980	16	3210.8	0	199	248.6	652.3	4310.7
1981	17	3275.3	0	0	326.1	228.8	3830.2
1982	18	3437.8	0	0	343.6	245.9	4027.3
1983	19	3610.5	0	0	451.4	274.1	4336
1984	20	3275.4	0	0	365.1	187	3827.5
1985	21	3501.3	0	0	458.3	233.1	4192.7
1986	22	3277.1	0	0	619.2	551.8	4448.1
1987	23	2954.3	0	0	1268.6	406.9	4629.8
1988	24	3488.5	0	0	1273.6	637.3	5399.4
1989	25	3772	0	0	928.6	623.4	5324
1990	26	3072.4	0	0	1371.8	460.7	4904.9
1991	27	3494.4	0	0	1179.2	438.8	5112.4
1992	28	3710.2	0	0	848.6	448.7	5007.5
1993	29	4142.4	0	0	424.2	543.1	5109.7
1994	30	4050.7	0	0	467.7	483.1	5001.5
1995	31	3282.2	0	0	474.3	427.4	4183.9
1996	32	2984.3	14.9	0	278.7	190.9	3468.8
1997	33	3553.6	1.5	0	430.6	156.3	4142
1998	34	3772	0	0	443.6	376.1	4591.7
1999	35	3677.3	5.7	0	627.9	403	4713.9
2000	36	2745.3	12.9	0	751.9	326.1	3836.2
2001	37	2742.8	7.5	0	564.4	299.6	3614.3
2002	38	2159	19.9	0	563.8	266.8	3009.5
2003	39	1419.2	67.5	0	630.8	227.6	2345.1
2004	40	2128.5	48.4	0	465.5	344.7	2987.1
2005	41	3196.5	41.6	0	1145.1	277.1	4660.3
2006	42	2693.3	61.1	0	1330.5	445.2	4530.1

Table B-5 Sablefish stock indices: nominal trap fishery CPUE, standardized survey CPUE, Japanese longline fishery CPUE and a tagging program biomass index.

Year	Trap Fishery CPUE (kg/trap)	Std. Trap Survey CPUE (kg/trap)	Japanese Longline CPUE (t/10 hachi)	Tagging Index
1965				
1966				
1967				
1968			0.261	
1969			0.207	
1970			0.215	
1971			0.162	
1972			0.207	
1973			0.209	
1974			0.21	
1975			0.194	
1976			0.194	
1977			0.17	
1978			0.18	
1979	17.661		0.135	
1980	15.312		0.137	
1981	15.056			
1982	16.973			
1983	16.819			
1984	13.059			
1985	17.687			
1986	15.602			
1987	16.16			
1988	24.736			
1989	25.695			
1990	19.222	20.017		
1991	24.562	19.336		
1992	24.73	25.569		62497
1993	20.421	36.509		119646
1994	18.3	15.571		64250
1995	15.255	13.665		42399
1996	14.928	11.258		30934
1997	13.314	7.721		38095
1998	13.387	12.037		28875
1999	13.711	7.72		48510
2000	12.456	9.296		31659
2001	10.139	3.092		37381
2002	9.659	8.206		28027
2003	19.813	27.59		33380
2004	13.194	26.415		41258
2005	11.852	19.432		40382
2006	10.261	17.382		43984

Table B-6 Proportions at age (sexes pooled) and sample size from commercial trap fishery and standardized survey samples.

	Age Class																						
Year	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**N**
Trap																								
1979	0.002	0.002	0.004	0.000	0.002	0.012	0.014	0.027	0.052	0.052	0.039	0.075	0.068	0.079	0.073	0.073	0.077	0.052	0.056	0.031	0.046	0.041	0.122	517
1980	0.229	0.077	0.020	0.024	0.033	0.028	0.032	0.042	0.046	0.060	0.047	0.041	0.043	0.049	0.044	0.032	0.018	0.035	0.020	0.022	0.007	0.008	0.043	1422
1981	0.047	0.160	0.053	0.027	0.027	0.007	0.007	0.020	0.007	0.027	0.020	0.027	0.060	0.033	0.027	0.047	0.040	0.080	0.033	0.040	0.013	0.007	0.193	150
1982	0.008	0.020	0.068	0.055	0.030	0.028	0.038	0.037	0.028	0.048	0.037	0.052	0.028	0.037	0.025	0.030	0.018	0.032	0.022	0.037	0.015	0.023	0.283	600
1983	0.026	0.083	0.071	0.242	0.086	0.043	0.025	0.040	0.022	0.035	0.035	0.033	0.050	0.019	0.025	0.017	0.018	0.022	0.019	0.013	0.006	0.016	0.055	1192
1984																								
1985																								
1986																								
1987	0.010	0.026	0.126	0.127	0.148	0.182	0.157	0.068	0.024	0.015	0.011	0.007	0.005	0.004	0.006	0.002	0.013	0.010	0.005	0.005	0.006	0.005	0.040	844
1988	0.021	0.049	0.047	0.091	0.184	0.131	0.126	0.100	0.079	0.022	0.012	0.010	0.009	0.005	0.005	0.014	0.009	0.009	0.006	0.014	0.005	0.006	0.043	770
1989	0.025	0.006	0.009	0.019	0.050	0.071	0.118	0.134	0.102	0.075	0.050	0.025	0.012	0.006	0.016	0.019	0.003	0.037	0.012	0.009	0.009	0.006	0.186	324
1990																								
1991	0.074	0.093	0.096	0.107	0.067	0.084	0.060	0.089	0.060	0.063	0.063	0.037	0.012	0.004	0.007	0.004	0.011	0.005	0.009	0.005	0.005	0.009	0.039	571
1992	0.024	0.010	0.024	0.047	0.064	0.137	0.086	0.069	0.095	0.096	0.068	0.061	0.052	0.041	0.037	0.010	0.007	0.003	0.003	0.003	0.003	0.008	0.051	596
1993	0.099	0.089	0.057	0.067	0.086	0.081	0.082	0.056	0.068	0.054	0.040	0.038	0.042	0.025	0.015	0.016	0.007	0.008	0.005	0.005	0.001	0.001	0.057	1398
1994	0.042	0.115	0.103	0.053	0.088	0.058	0.063	0.053	0.064	0.037	0.042	0.049	0.029	0.032	0.031	0.024	0.015	0.016	0.007	0.007	0.006	0.005	0.060	870
1995	0.008	0.045	0.152	0.066	0.033	0.053	0.065	0.079	0.054	0.051	0.037	0.042	0.038	0.032	0.030	0.037	0.020	0.024	0.004	0.011	0.004	0.005	0.111	837
1996	0.010	0.030	0.060	0.107	0.082	0.044	0.045	0.056	0.058	0.058	0.044	0.046	0.041	0.045	0.042	0.039	0.034	0.023	0.010	0.006	0.003	0.003	0.117	711
1997																								
1998	0.011	0.037	0.037	0.064	0.103	0.112	0.078	0.070	0.059	0.050	0.029	0.032	0.025	0.023	0.020	0.016	0.034	0.020	0.029	0.005	0.009	0.018	0.119	561
1999	0.000	0.051	0.063	0.071	0.090	0.101	0.099	0.080	0.054	0.039	0.037	0.023	0.031	0.017	0.014	0.019	0.012	0.025	0.014	0.020	0.014	0.008	0.118	647
2000	0.017	0.055	0.199	0.177	0.083	0.062	0.073	0.076	0.038	0.036	0.019	0.016	0.012	0.023	0.005	0.019	0.003	0.021	0.005	0.014	0.003	0.002	0.040	577
2001																								
2002	0.048	0.102	0.161	0.108	0.089	0.041	0.033	0.039	0.033	0.043	0.030	0.007	0.022	0.022	0.007	0.011	0.004	0.009	0.022	0.013	0.013	0.017	0.128	464

	Age Class																						
Year	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**N**
Survey																								
1988	0.114	0.079	0.076	0.085	0.106	0.112	0.088	0.057	0.034	0.012	0.014	0.011	0.013	0.011	0.015	0.004	0.011	0.013	0.013	0.011	0.009	0.004	0.107	1429
1989	0.068	0.067	0.072	0.088	0.111	0.110	0.103	0.071	0.046	0.027	0.016	0.009	0.011	0.008	0.010	0.007	0.009	0.010	0.013	0.008	0.008	0.007	0.122	3883
1990	0.081	0.097	0.068	0.039	0.042	0.046	0.031	0.038	0.038	0.016	0.022	0.012	0.005	0.011	0.005	0.001	0.008	0.015	0.008	0.009	0.011	0.015	0.381	756
1991	0.033	0.039	0.063	0.089	0.088	0.073	0.073	0.063	0.092	0.045	0.053	0.032	0.016	0.006	0.009	0.008	0.009	0.005	0.006	0.004	0.005	0.009	0.183	1053
1992	0.041	0.025	0.054	0.073	0.089	0.080	0.071	0.054	0.054	0.057	0.044	0.043	0.027	0.021	0.012	0.008	0.009	0.006	0.007	0.003	0.007	0.008	0.205	1848
1993	0.095	0.079	0.054	0.065	0.067	0.078	0.067	0.046	0.037	0.049	0.048	0.042	0.031	0.024	0.021	0.011	0.010	0.003	0.006	0.003	0.003	0.003	0.158	1782
1994	0.031	0.092	0.070	0.057	0.061	0.052	0.058	0.056	0.049	0.044	0.043	0.041	0.034	0.035	0.024	0.024	0.013	0.011	0.012	0.010	0.008	0.008	0.167	2064
1995	0.009	0.065	0.136	0.103	0.049	0.047	0.043	0.050	0.050	0.030	0.032	0.042	0.035	0.040	0.026	0.020	0.018	0.011	0.010	0.006	0.008	0.007	0.162	1708
1996	0.016	0.038	0.080	0.109	0.068	0.036	0.049	0.037	0.036	0.038	0.038	0.032	0.035	0.042	0.035	0.031	0.027	0.020	0.014	0.009	0.006	0.006	0.199	982
1997	0.055	0.044	0.066	0.126	0.192	0.055	0.077	0.055	0.044	0.027	0.011	0.022	0.022	0.011	0.022	0.022	0.022	0.022	0.005	0.000	0.005	0.005	0.088	187
1998																								
1999	0.025	0.057	0.085	0.074	0.068	0.045	0.085	0.030	0.025	0.036	0.045	0.021	0.025	0.023	0.030	0.034	0.023	0.026	0.019	0.008	0.008	0.002	0.208	534
2000	0.017	0.004	0.154	0.056	0.034	0.021	0.056	0.047	0.021	0.013	0.021	0.026	0.021	0.017	0.009	0.017	0.013	0.017	0.030	0.017	0.021	0.017	0.350	234
2001																								
2002	0.030	0.069	0.082	0.084	0.096	0.057	0.029	0.020	0.028	0.030	0.027	0.014	0.023	0.013	0.016	0.015	0.017	0.010	0.018	0.012	0.015	0.017	0.278	870
2003	0.095	0.116	0.147	0.104	0.039	0.064	0.056	0.017	0.031	0.017	0.015	0.010	0.025	0.010	0.010	0.010	0.015	0.012	0.012	0.012	0.010	0.015	0.156	487
2004	0.038	0.177	0.179	0.136	0.086	0.077	0.038	0.010	0.029	0.010	0.010	0.007	0.000	0.012	0.002	0.005	0.005	0.010	0.012	0.005	0.005	0.012	0.136	418

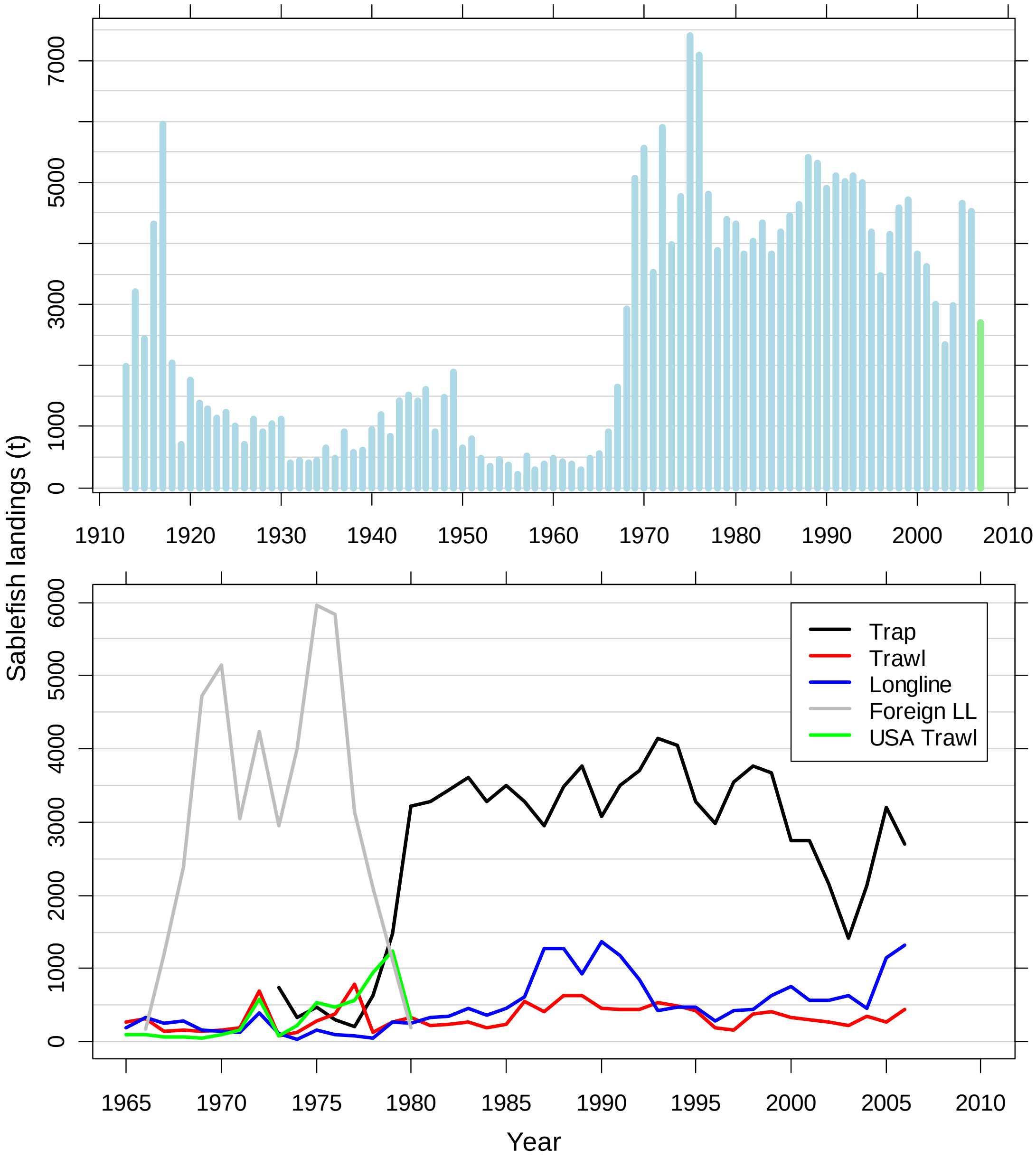


Figure B-1 Annual sablefish landings (t) from 1913 to 2007 from all sources (top panel). Annual landings by gear type for the period 1965 to 2006 are shown in the bottom panel. Preliminary data for 2007 represent the partial year to September 31, 2007.

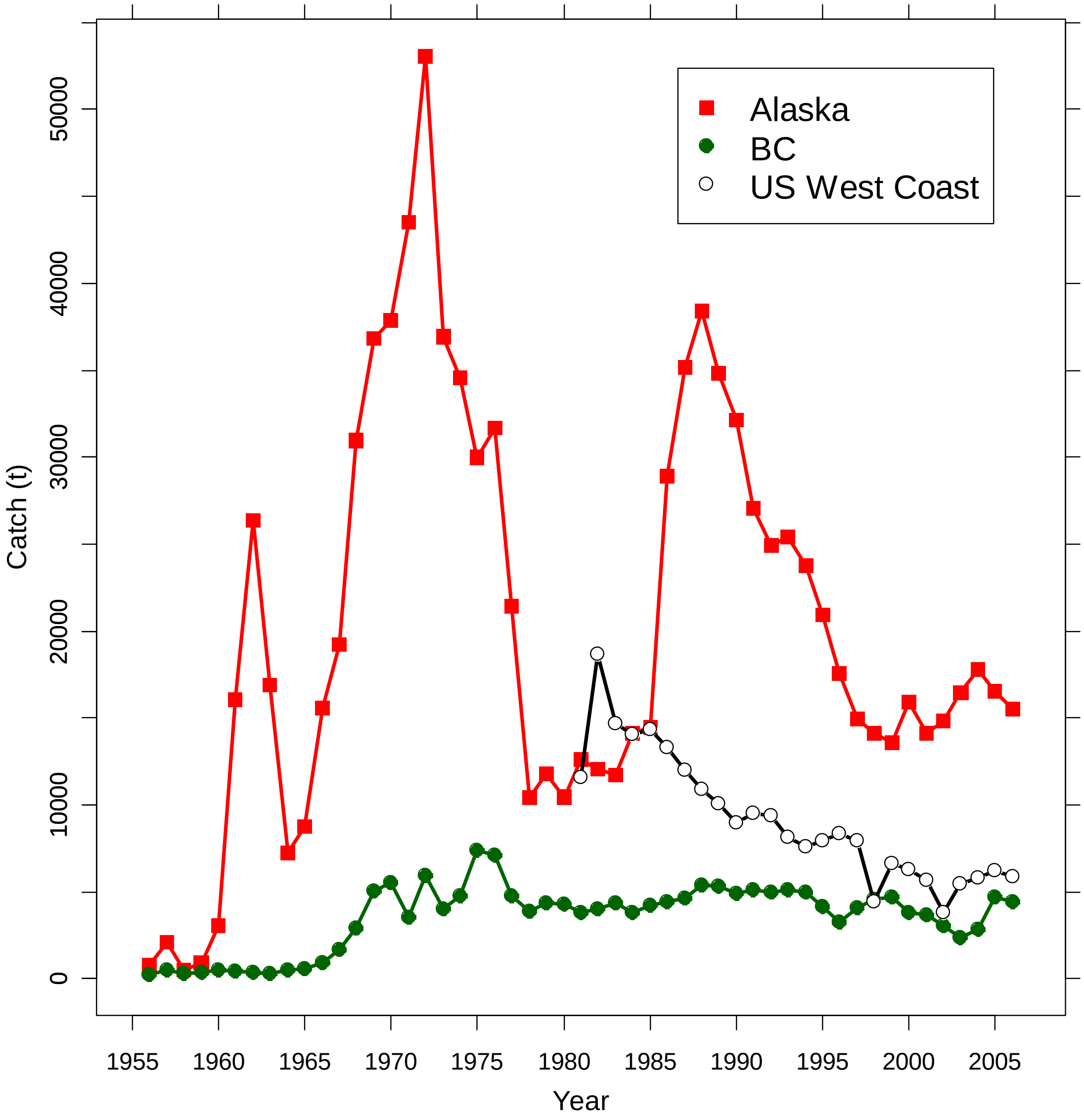


Figure B-2 Sablefish catch (t) for the Gulf of Alaska and U.S. west coast, and landings (t) for British Columbia.

Appendix C Sablefish Operating Model for Population and Fishery Dynamics

We used an age-structured population dynamics model to construct plausible scenarios for the "true" sablefish population in management procedure simulations. Here we describe the general structure of the operating model, with model notation provided in Table C-1 and the model equations listed in Table C-2. Model notation (and parameter settings) generally follows that used for the catch-age assessment model described in Appendix E. All operating models assume that the B.C. sablefish spawning stock was at unfished, deterministic equilibrium B_0 prior to directed fisheries in the mid-1960s. The models further assume that the B.C. population is closed to immigration and emigration. The unfished spawning biomass B_0 and steepness of the stock-recruitment relationship h determine the long-term yield and resilience of the stock, and are therefore among the most important uncertainties in management procedure evaluations (Butterworth and Punt 1999, Walters and Martell 2004). A Beverton-Holt stock-recruitment relationship based on $\{B_0,h\}$ is defined by the sequential calculation from C2.7 to C2.11, beginning with the unfished spawning biomass per recruit (C2.7) and ending with the initial recruitment slope and density-dependence parameters, respectively, in C2.10 and C2.11. Calculations C2.12 – C2.14 initialize the population age composition to the unfished equilibrium. Realized recruitment to age-1 is log-normally distributed with respect to the average stock-recruitment relationship and potentially auto-correlated in time (C2.17). The standard error $\sigma = 0.70$ controls the magnitude of independent recruitment process deviations while the first-order autoregression coefficient γ controls the degree of correlation among the annual deviations.

At each annual cycle, the operating model appends three observations to the existing sablefish monitoring dataset. The observations consist of a standardized survey index of relative abundance (kg/trap), fishery catch-age proportions, and survey catch-age proportions. The survey index in C2.23 is assumed proportional to the biomass available to the survey gear as defined by C2.21 with stochastic errors that are log-normal and corrected for bias by subtracting $0.5\tau_1^2$ from each observation. The bias correction is required here because simulation testing of data-based harvest policies requires that simulated future surveys have the same expected values as historical surveys for the same biomass levels. Note that hypotheses about hyperstability of fishery CPUE are addressed using the power parameter $q_{2,g}$, where g=1 (or $q_{2,\mathrm{trap}}$), in equation C2.23. Equations C2.24 and C2.25 model catch-at-age observations for the survey and trap fishery as multivariate-logistic random variables with gear-specific standard errors $\tau_{2,g}$ (Schnute and Richards 1995). The random variables $\delta_{g,t}$ in C2.23 and $\varepsilon_{g,a,t}$ in C2.24 are standard normal with mean zero and standard deviation equal to one.

Parameters in C2.1 were estimated by fitting the operating model to gear-specific catch (1965-2006), trap fishery catch-per-unit-effort (CPUE; 1979-2006), research survey CPUE (1990-2006), fishery catch-at-age (1979-2002), research survey catch-at-age (1988-2004), Japan longline CPUE (1965 - 1980), and tag-recovery estimates of biomass (1992-2006). Not all years were represented within the range of the two catch-at-age series. Natural mortality, length-at-age, maturity-at-age, and average selectivity-at-length parameters were all estimated external to the operating model. Selectivity in C2.5 is represented by a double-logistic function that, depending on parameter values, can take on a variety of shapes from uniform to dome-shaped. Parameters for each gear-type (Table C-3) were determined by fitting generalized linear models to annual tag release-recovery data using the method of Myers and Hoenig (1997). Model parameters were estimated using a penalized likelihood approach that was practically identical to the one used in the catch-age assessment model (Appendix E), except with additional likelihood components for fishery CPUE and tagging biomass, plus added catch-age observations.

Table C-4 provides a summary of operating model fits and derived management quantities for each scenario identified in Sections 3.2 and 3.3 of the main document.

Table C-1 Notation for a sablefish population and fishery operating model. Many parameters will have base values and then alternatives under different model configurations.

Symbol	Value	Description
Indices		
t		Time step $t = \{1,2,\ldots T\}$
a		Age-class in years $a = \{1,2,\ldots A\}$
g	Table C-3	Gear type index
Model parameters		
B_0		Unfished spawning biomass
h	{0.45,0.65}	Recruitment function steepness
δ_t	Normal(0,1)	Normally distributed log-recruitment deviation
$q_{i,g}$		Catchability coefficients $i = \{1,2\}$ for gear g
$\beta_{i,g}$	Table C-3	Double-logistic selectivity function parameters for gear g
$\tau_{1,g}$	Table C-3	Coefficient of variation for gear g abundance index
$\tau_{2,g}$	Table C-3	Standard error in observed proportions-at-age for gear g catch
σ	0.70	Standard error of log-recruitment deviations
γ	0.0	Lag-1 autocorrelation in log-recruitment deviations
M	0.08	Instantaneous natural mortality rate
L_∞	68.2	Asymptotic length (cm)
L_1	40.7	Length-at-age 1 (cm)
k	0.37	Von Bertalanffy growth constant
μ_2	5	Age-at-50% maturity

Symbol	Value	Description
μ_1	8	Maturity-at-age function steepness
Derived parameters		
R_0		Unfished recruitment
$s_{g,a}$		Selectivity-at-age in fishery g
m_a		Proportion mature-at-age
w_a		Body mass-at-age (tonnes)
ϕ		Unfished equilibrium spawning biomass per recruit
State variables		
$N_{a,t}$		Number of age a fish in year t
$B_{a,t}$		Biomass of age a fish in year t
$P_{g,t}$		Number of fish vulnerable to gear g in year t
$B^*_{g,t}$		Biomass of fish vulnerable to gear g in year t
$u_{g,a,t}$		Proportion of age a fish in harvestable population
S_t		Spawning biomass in year t
Observations		
$I_{g,t}$		Abundance index value for gear g in year t
$p_{g,a,t}$		Proportion of age a fish in gear g catch-age sample
Fishery controls		
$C_{g,t}$		Catch in fishery g (tonnes)
l_g		Minimum size limit in fishery g

Table C-2 Age-structured fish population and fishery operating model (simulation) used to evaluate management procedures. Beginning at the top, this table sequentially defines the population and time dynamics for a fixed set of input catches. The parameters in C2.1 were estimated using a similar penalized likelihood formulation as in Table E-2.

Parameters

C2.1 $\boldsymbol{\Theta} = \left(B_0, h, \boldsymbol{\delta}, q_{2,1}\right)$

Life history schedules

C2.2 $l_a = L_\infty + \left(L_1 - L_\infty\right) e^{(-k(a-1))}$

C2.3 $w_a = \exp[-25.9] l_a^{3.1}$

C2.4 $m_a = \dfrac{a^{\mu_1}}{a^{\mu_1} + \mu_2^{\mu_1}}$

Fishery selectivity

C2.5 $$\tilde{s}_{g,a} = \begin{cases} 0 & l_a < \bar{l}_g \\ \left(\dfrac{1}{1+e^{-\beta_{2,g}\left(l_a-\beta_{1,g}\right)}}\right)\left(1-\dfrac{1}{1+e^{-\beta_{4,g}\left(l_a-\beta_{1,g}-\beta_{3,g}\right)}}\right) & l_a \geq \bar{l}_g \end{cases}$$

C2.6 $s_{g,a} = \tilde{s}_{g,a} \Big/ \max\left[\tilde{s}_{g,1}, \tilde{s}_{g,2}, \ldots \tilde{s}_{g,A}\right]$

Stock-recruitment relationship

C2.7 $$\phi = \sum_{a=1}^{A-1} e^{-M(a-1)} m_a w_a + \frac{e^{-M(A-1)} m_A w_A}{1-e^{-M}}$$

C2.8 $R_0 = B_0 / \phi$

C2.9 $N_{1,1} = R_0$

C2.10 $a = \dfrac{4hR_0}{B_0\left(1-h\right)}$

C2.11 $b = \dfrac{5h-1}{B_0\left(1-h\right)}$

Initial population

C2.12 $N_{a,1} = R_0 e^{-M(a-1)}, \qquad 2 \leq a \leq A-1$

C2.13 $N_{A,1} = R_0 \dfrac{e^{-M(a-1)}}{1-e^{-M}}$

C2.14 $B_{a,1} = N_{a,1} w_a$

Age proportions in catch

C2.15 $u_{g,a,t} = s_{g,a} N_{a,t} \Big/ \sum_{a=1}^{A} s_{g,a} N_{a,t}$

State dynamics

C2.16 $\omega_t = \begin{cases} \dfrac{\sigma}{\sqrt{1-\gamma^2}} \delta_{1,t} & t = 1 \\ \gamma \omega_{t-1} + \sigma \delta_{1,t} & t > 1 \end{cases}$

C2.17 $N_{1,t} = \dfrac{a S_{t-1}}{1 + b S_{t-1}} \exp\left[\omega_t - 0.5 \sigma^2 \big/ \left(1 - \gamma^2\right) \right]$

C2.18 $N_{a,t} = e^{-M} \left[N_{a-1,t-1} - \sum_{g=1}^{G} u_{g,a-1,t-1} C_{g,t-1} / w_{a-1} \right] \quad 2 \le a \le A-1$

C2.19 $N_{A,t} = e^{-M} \left[N_{A-1,t-1} + N_{A,t-1} - \sum_{g=1}^{G} \left(u_{g,A-1,t-1} + u_{g,A,t-1} \right) C_{g,t-1} / w_A \right]$

C2.20 $B_{a,t} = w_a N_{a,t}$

C2.21 $B^{*}_{g,t} = \sum_{a=1}^{A} s_{g,a} B_{a,t}$

C2.22 $S_t = \sum_{a=1}^{A} m_a B_{a,t}$

Survey and catch-at-age observations

C2.23 $I_{g,t} = q_{1,g} {B^{*}_{g,t}}^{q_{2,g}} \exp\left[\tau_{1,g} \delta_{g,t} - 0.5 \tau_{1,g}^2 \right]$

C2.24 $x_{g,a,t} = \log u_{g,a,t} + \tau_{2,g} \varepsilon_{g,a,t} - \dfrac{1}{A} \sum_{a=1}^{A} \left[\log u_{g,a,t} + \tau_{2,g} \varepsilon_{g,a,t} \right]$

C2.25 $p_{g,a,t} = \exp\left[x_{g,a,t} \right] \Big/ \sum_{a=1}^{A} \exp\left[x_{g,a,t} \right]$

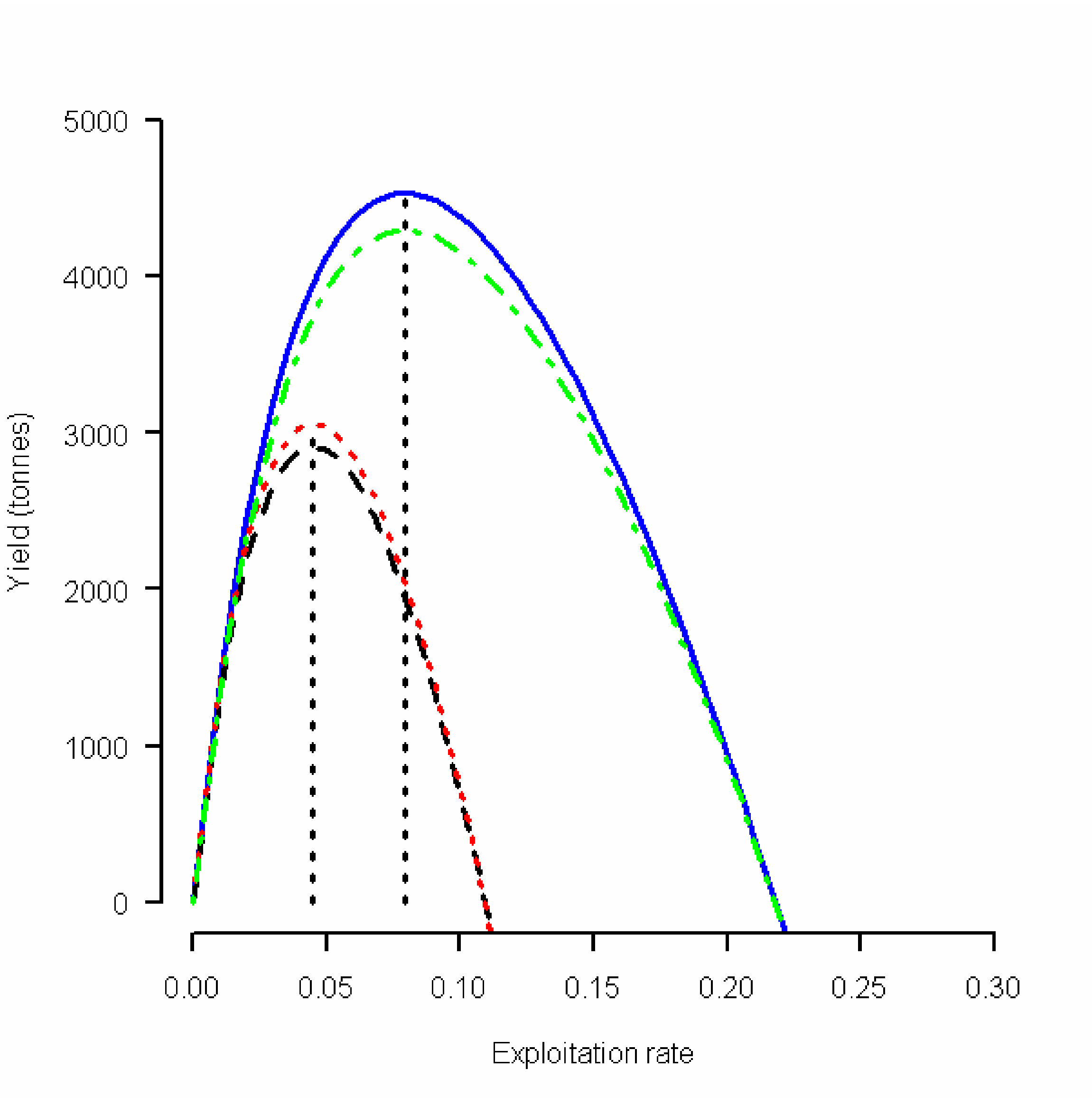
Table C-3 Fishery-specific parameters for the sablefish fishery operating model.

		Observation errors			Selectivity			
Fishery	*g*	$\tau_{1,g}$	$\tau_{2,g}$		$\beta_{1,g}$	$\beta_{2,g}$	$\beta_{3,g}$	$\beta_{4,g}$
Longline trap	1	0.06	0.66		57.94	0.21	14.10	0.19
Research survey	2	0.26	0.45		57.94	0.21	14.10	0.19
Japan longline	3	0.02	-		58.34	0.18	21.17	0.32
Tagging survey	4	0.07	-		57.94	0.21	14.10	0.19
Longline hook	5	-	-		58.34	0.18	21.17	0.32
Trawl	6	-	-		35.33	2.90	10.88	0.13

Table C-4 Summary of operating model characteristics that define scenarios for management procedure simulations. The first two symbols – recruitment steepness (*h*), initial depletion (D_{2006}) at the start of management procedure simulations – define a scenario. The next three columns show the negative log-likelihood (ℓ), number of parameters (*N*), and Akaike Information Criterion (*AIC*) values for operating model fits to existing data. The final three columns provide equilibrium quantities unfished spawning biomass (B_0), exploitation rate at the maximum sustainable yield (U_{MSY}), biomass at MSY (B_{MSY}), depletion at MSY (D_{MSY}), and the MSY.

Scenario	***h***	$\boldsymbol{D_{2006}}$	ℓ	***N***	***AIC***	$\boldsymbol{B_0}$	$\boldsymbol{U_{\text{MSY}}}$	$\boldsymbol{B_{\text{MSY}}}$	$\boldsymbol{D_{\text{MSY}}}$	**MSY**
S1	0.45	0.18	-189.46	43	103.46	151 288	0.045	56 747	0.375	2 914
S2	0.45	0.30	-184.62	42	100.62	158 830	0.045	59 576	0.375	3 059
S3	0.65	0.20	-188.80	43	102.80	146 428	0.079	44 709	0.305	4 294
S4	0.65	0.31	-184.80	42	100.80	154 465	0.079	47 162	0.305	4 529

Figure C-1 Equilibrium yield versus exploitation rate relationships for the four operating model scenarios. S1 (black, dashed), S2 (red, dotted), S3 (green, dot-dash), and S4 (blue, solid). Vertical dashed lines indicate U_{MSY} for each set of scenarios.



Appendix D Production models for sablefish stock assessment

Introduction

Although the catch sampling and ageing program for B.C. sablefish has generated catch-age samples from 1988 – 2000 (with some missing years), there is no catch sampling at the present time. Thus, it is possible that future management procedures for sablefish will involve only collecting catch-by-gear and abundance index data (e.g., survey catch rates). Therefore, stock assessments in the future may need to rely on aggregated (e.g., over age, size, etc.) production models for estimating biomass, exploitation, and production rates of B.C. sablefish.

Aggregated production models are stock assessment methods that rely only on fishery catch and survey data or, alternatively, fishery catch and effort data. These methods break into either (*i*) total biomass production or (*ii*) delay-difference model types. Delay-difference models (Deriso 1980; Schnute 1985; Hilborn and Walters 1992) offer two potential advantages in modeling catch-effort data. First, delay-difference models approximate age-structured fish dynamics without actually keeping track of individual cohorts. Therefore, key age-structured characteristics such as delays between birth and recruitment to the spawning or fishable stock and more realistic patterns of individual growth can be accounted for in stock assessments. Second, delay-difference approaches model the dynamics of average fish size by tracking numbers of fish N_y and the total biomass B_y in the population separately. This is important because data on the average body size of fish, which is often collected routinely for most fisheries, may exhibit signals of changes in population composition. For example, declines in average fish size during fishery development typically indicate fishery removals from a standing stock of old, large individuals. Potential overfishing may be identified where the average fish size approaches the average size of newly recruited fish; that is, indicating that there are few old, large fish remaining.

The advantages of delay-difference models are sometimes compensated by the potential pitfalls that accompany their restrictive assumptions, which are (Hilborn and Walters 1992): (*i*) average body weight – at – age follows a linear Ford-Brody growth model $w_{a+1} = \alpha + \rho w_a$; (*ii*) all individuals aged k and older are equally selected by the fishery; and (*iii*) all fish aged *k* and older have the same natural mortality rate. Difficulty here lies mainly in assumptions (*ii*) and (*iii*) because fish recruitment to fisheries and spawning stocks rarely coincide and are rarely "knife-edged" as implied by assumption (*ii*) in particular. Sometimes, small deviations in the age-at-recruitment *k* from can have profound implications for estimated management parameters such as optimal exploitation rates or yields. We discounted the role of delay-difference models for sablefish at this stage of management procedure development mainly because our independent analysis of tagging data suggested that selectivity was likely dome-shaped, at least for the trap fishery, which takes most of the catch. Dome-shaped selectivity clearly violates assumption (*ii*) of delay-difference models and leads to potentially severe biases in the interpretation of both catch-per-unit effort and average body size (Walters and Martell 2004). At our present stage of management procedure development, we viewed the added complexities of delay-difference assessment models as a disadvantage because of these known biases and the fact that added assessment model control variables (age-at-recruitment, fitting model to body size, etc.) would need to be evaluated.

In contrast to delay-difference approaches to stock assessment, total biomass production models have fewer biological assumptions and are thus potentially less "realistic". In fact, simple production models may even be more restrictive than delay-difference approaches because all process of growth, mortality, and recruitment are represented by a simple density-dependent production function. For example, the Schaefer form of production model, i.e.,

$$(1) \qquad B_{y+1} = B_y + rB_y\left(1 - B_y / K\right) - C_y \quad ,$$

assumes that annual net production depends only on biomass B_y in the previous year, a constant density-independent growth rate *r*, and a carrying capacity *K* or unfished biomass for the stock. Production models such as equation (1) therefore relate the cumulative effects of all population dynamics processes to population density. Furthermore, there are no lags between spawning stock and recruitment, which contrasts potentially long lags allowed by the delay-difference approach. Simple production models do have some advantages. First, provided that the catch – effort or catch – survey data accurately represent the fishery and stock dynamics, simple production models seem to provide reasonably robust assessments in both simulations and in practice (Hilborn 1979; Punt 2003). Another advantage is the relatively simple closed-form solutions for optimal exploitation and maximum sustainable yield, which allows for direct computation of Bayes posterior distributions for catch limits (Cooke 1999). Computing catch limits in this way provides the ability to examine harvest policies that have both precautionary and optimality objectives. This can also be done with more complex age-structured approaches, but such methods are more burdensome in quantitative simulations.

This appendix develops and evaluates production model estimators that could be applied to stock assessment for B.C. sablefish. Specifically, we formulated a Schaefer production model that employed an errors-in-variables paradigm (Schnute and Richards 1995; Punt 2003) to account for process and observation errors when attempting to estimate harvestable biomass. This estimator uses a tuning parameter, ρ, which represents the proportion of the total error assigned to the observations. Thus, ρ controls how much of the variability in survey catch rates is assigned to random observation errors and how much is assigned to random process error or unaccounted for changes in the stock biomass. Data for the production model always involved trap survey relative abundance indices (1992 – 2006) only combined with catches (1965 – 2006) that are aggregated over trap, longline hook, and trawl fisheries.

The sections below develop production model estimators using ρ = {0.70, 0.80}, which we apply to stock assessment analyses of B.C. sablefish to demonstrate how each interprets the historical dynamics and present status of the stock.

Methods

Production model for stock assessment

Our general stock production model derives inferences about management parameters from time-series observations of catch and survey CPUE. Notation for the model is provided in Table D-1 with sequential calculations listed in Table D-2. Here we elaborate on some of the model details.

A general stock production model predicts biomass in each year B_{y+1} based on four components: (1) the predicted stock present in the previous year B_y , (2) an average production function $f(B_y)$ that depends on biomass, (3) fishery catch C_y, and (4) a random deviation ω_y from the average production relationship (Punt 2003). These components can be written into a production model of the form

$$(2) \qquad B_{y+1} = \left(B_y + f\left(B_y\right) - C_y\right)e^{\omega_y} \quad ,$$

where B_y (tonnes) and C_y (tonnes) are the stock biomass at the start of year y $\left(y = 1,2,\ldots,n+1\right)$ and catch biomass during year y, respectively. The catch is assumed to be taken instantaneously and after production. The random production anomaly term ω_y is assumed independent of stock biomass and may represent, for example in the B.C. sablefish case, the net result of (*i*) immigration into the B.C. from Alaska or the lower west coast U.S., (*ii*) emigration out of the stock that is present in B.C. at any moment, (*iii*) and random deviations from the average production relationship within B.C. We assumed that these deviations, however they arise, are independent from year to year and are normally distributed with mean zero and coefficient of variation equal to σ. We adopt a Schaefer formulation for the production function, which gives the production model

$$(3) \qquad B_{y+1} = \left(B_y + rB_y\left(1 - B_y / K\right) - C_y\right)e^{\omega_y} \quad ,$$

where the middle term on the right-hand-side represents $f(B_y)$ in Equation (2), r is the intrinsic rate of biomass growth, and K (tonnes) is the long-term average carrying capacity or "unfished biomass". The Schaefer form assumes that fish production is a symmetric, dome-shaped function of existing stock biomass so that $U^{MSY} = r/2$ and $MSY = rK/4$ define the optimum exploitation rate and maximum sustainable yield, respectively. These quantities can be used by "passive adaptive" management strategies that attempt to steer fisheries exploitation toward theoretically optimal levels (c.f. Walters

1986 for full description of adaptive harvest policies). Calculating the apparent optimum quota C^*_{Y+1} for year Y+1 involves projecting the stock biomass one year into the future and multiplying by the estimated optimal exploitation rate U_Y^{MSY}, where the subscript Y recognizes that this estimate will change as new data are added.

Survey indices of relative abundance are used in estimating production model parameters for B.C. sablefish. We use the linear observation model

$$(4) \qquad I_y = qB_y e^{\xi_y} \quad ,$$

to relate predicted sablefish biomass to the index observations, where q is catchability and ξ_y is a normally distributed random observation error in year y. This observation model assumes that survey selectivity and catchability parameters remain constant over time.

Likelihood function

Different assumptions about how to allocate random deviations in the data to the stock dynamics (Equation 3) or the observations (Equation 4) give different production model estimators. For example, assigning all errors to the observations leads to an "observation-error" estimator in which the stock dynamics are assumed to be non-random and exactly equal to that predicted by Equation 3 with $\omega_y = 0$ for all values of y. Thus, observation-error models ignore inter-annual changes in stock biomass that occur via stochastic processes like natural mortality, immigration, emigration, or environmental influences on production. On the other hand, assigning all random error to the underlying stock dynamics by setting $\xi_{g,y} = 0$ in the observation model (Equation 4) for all values of g and y leads to a "process error" estimator in which the observations are assumed to be exact, i.e., $I_{g,y} = q_g B_y$, and thus inter-annual fluctuations in the data indicate changes in true stock biomass. For the process error estimator, the individual terms ω_y must be estimated along with the standard error σ.

Clearly, inferences about the dynamics of fish stocks depend upon uncertainty in both the observations and the underlying population dynamics processes. Admitting both observation and process errors in the stock assessment model leads to errors-in-variables estimators in which some proportion ρ of the total error variance is assigned to the random observations and the remainder 1- ρ is assigned to random changes in the underlying stock dynamics. Formally, errors-in-variables estimators define the total error variance, κ^2, as

$$(5) \qquad \kappa^2 = \tau^2 + \sigma^2 \quad .$$

This follows from the rules of variances. If the observation error proportion $\rho = \tau^2 / \left(\tau^2 + \sigma^2\right)$ is assumed known, the individual variance components can then be expressed as

$$(6) \qquad \tau^2 = \rho\kappa^2, \quad \sigma^2 = \left(1-\rho\right)\kappa^2 \quad ,$$

for observation and process errors, respectively. For our analysis, ρ is considered to act as a control or tuning parameter in the estimation procedure. As ρ approaches 0, the emphasis on process error will tend to allow for relatively large random changes in the estimated stock biomass from year to year, provided, of course, that multiple abundance indices suggest the same direction and magnitude of change. Conversely, values of ρ near 1 will cause the model biomass to change deterministically in response to changes in production and catch; that is, the stock will only increase if catches are less than deterministic surplus production. For application to B.C. sablefish, we examined two cases that involved either $\rho = 0.7$ or $\rho = 0.8$. Initial trials with an approximate observation-error estimator (i.e., $\rho = 1.0$) resulted in pathological convergence behaviour and were abandoned.

Prior distribution

Preliminary results from fitting the production model based on the likelihood alone and survey data for 1992 – 2006 suggested that the optimal exploitation rate for sablefish is approximately 18%, which is far greater than any previously reported estimate. Furthermore, biomass and quota estimates were very sensitive to the choice of ρ; for example, the estimated biomass in 2006 differed by 8,000 t between $\rho = 0.70$ and $\rho = 0.80$. We chose to stabilize model behaviour by including a relatively informative prior distribution on the intrinsic rate of biomass growth, *r*. The prior mean for *r* is a management procedure choice and should therefore be consistent with the reference exploitation rate choice U^{ref}, which is used in the same procedure. Consistency is needed because the two quantities are related, i.e., $U^{ref} = r/2$, if we assume $U^{ref} = U^{MSY}$. We established consistency by setting the prior mean on $r = \{0.12, 0.16, 0.20\}$ when we tested production model-based management procedures with $U^{ref} = \{0.06, 0.08, 0.10\}$, respectively. In each case, the prior standard error is set to 50% of the prior mean. This approach to specifying an informative prior on *r* to stabilize stock assessment estimator performance is similar to the approach taken in the IWC's Catch Limit Algorithm (Cooke 1999).

Figure D-1 shows the alternative model fits to standardized survey data (1992 – 2006). Although the fits look similar, changing from $\rho = 0.7$ to $\rho = 0.8$ changes the estimated 2006 biomass from 59,967 t to 53,053 t, estimated unfished biomass from 112,016 t to 106,351 t, and estimated *r* values from 0.150 to 0.152.

Table D-1 Notation for the aggregate production stock assessment model.

Symbol	**Description**
	Indices and Index Ranges
Y	Final year of modeled time horizon
y	Year, where $1 \le y \le Y$ and $y = 1$ corresponds to the first year
g	Gear index (fishery or survey), where $g = 1,\ldots,G$
n	Number of non-missing observations for the survey
i	Index for non-missing survey observations $i = 1,\ldots,n$
	Data
$C_{g,y}$	Catch biomass removed during year y by gear type g
I_y	Survey relative abundance observation for year y
	Model Parameters
r	Intrinsic rate of biomass growth
B_0	Unfished or pre-exploitation population biomass
q	Catchability coefficient for relative abundance survey
	State variables
B_y	Biomass at the beginning of year y
	Derived Management Quantities
D_y	Depletion level for year y
U_y^{MSY}	Optimal exploitation rate for year y
C^*	Quota based on optimal exploitation rate
	Statistical Errors
ω_y	Random process deviation in year y
σ^2	Process deviation squared coefficient of variation
ξ_y	Random log-survey observation error in year y
τ^2	Observation deviation squared coefficient of variation
ρ	Observation error proportion of total error variance
κ^2	Total error variance

Table D-2 Production model used in management procedure simulations and stock assessment analyses in this appendix. The table represents an errors-in-variables formulation of the Schaefer biomass dynamics stock assessment model for estimating biomass and management quantities each year.

Model parameters

D2.1 $$\Theta = \left(r, B_0, \left\{ \omega_y \right\}_{y=1}^{y=Y-1} \right)$$

Biomass dynamics model

D2.2 $$B_1 = B_0$$

D2.3 $$B_{y+1} = \begin{cases} \left(B_y + rB_y\left(1 - B_y / B_0\right) - \sum_{g=1}^{G} C_{g,y} \right) \mathrm{e}^{\omega_y} & 1 \le y \le Y-1 \\ B_y + rB_y\left(1 - B_y / B_0\right) - \sum_{g=1}^{G} C_{g,y} & y = Y \end{cases}$$

Residuals

D2.4 $$\xi_i = \log_e\left(I_i / B_i \right)$$

Conditional maximum likelihood estimates

D2.5 $$\log \hat{q} = \frac{1}{n} \sum_{i=1}^{n} \xi_i$$

D2.6 $$\hat{\kappa}^2 = \frac{1}{n+Y-1} \left(\frac{1}{\rho} \sum_{i=1}^{n} \left(\xi_i - \log \hat{q} \right)^2 + \frac{1}{1-\rho} \sum_{y=1}^{Y-1} \omega_y^2 \right)$$

Negative log-likelihood and objective function

D2.7 $$\ell\left(\mathbf{I} \mid \Theta\right) = \frac{n+Y-1}{2} \log_e \left(\frac{1}{\rho} \sum_{i=1}^{n} \left(\xi_i - \log \hat{q} \right)^2 + \frac{1}{1-\rho} \sum_{y=1}^{Y-1} \omega_y^2 \right)$$

D2.8 $$G\left(\Theta \mid \mathbf{I}\right) \propto \ell\left(\mathbf{I} \mid \Theta\right) + \frac{1}{2\sigma_r^2} \left(r - \mu_r \right)^2$$

Derived Management quantities

D2.9 $$D_Y = B_Y / B_0$$

D2.10 $$U_Y^{MSY} = r/2$$

D2.11 $$C_{Y+1}^{*} = U_Y^{MSY} B_{Y+1}$$

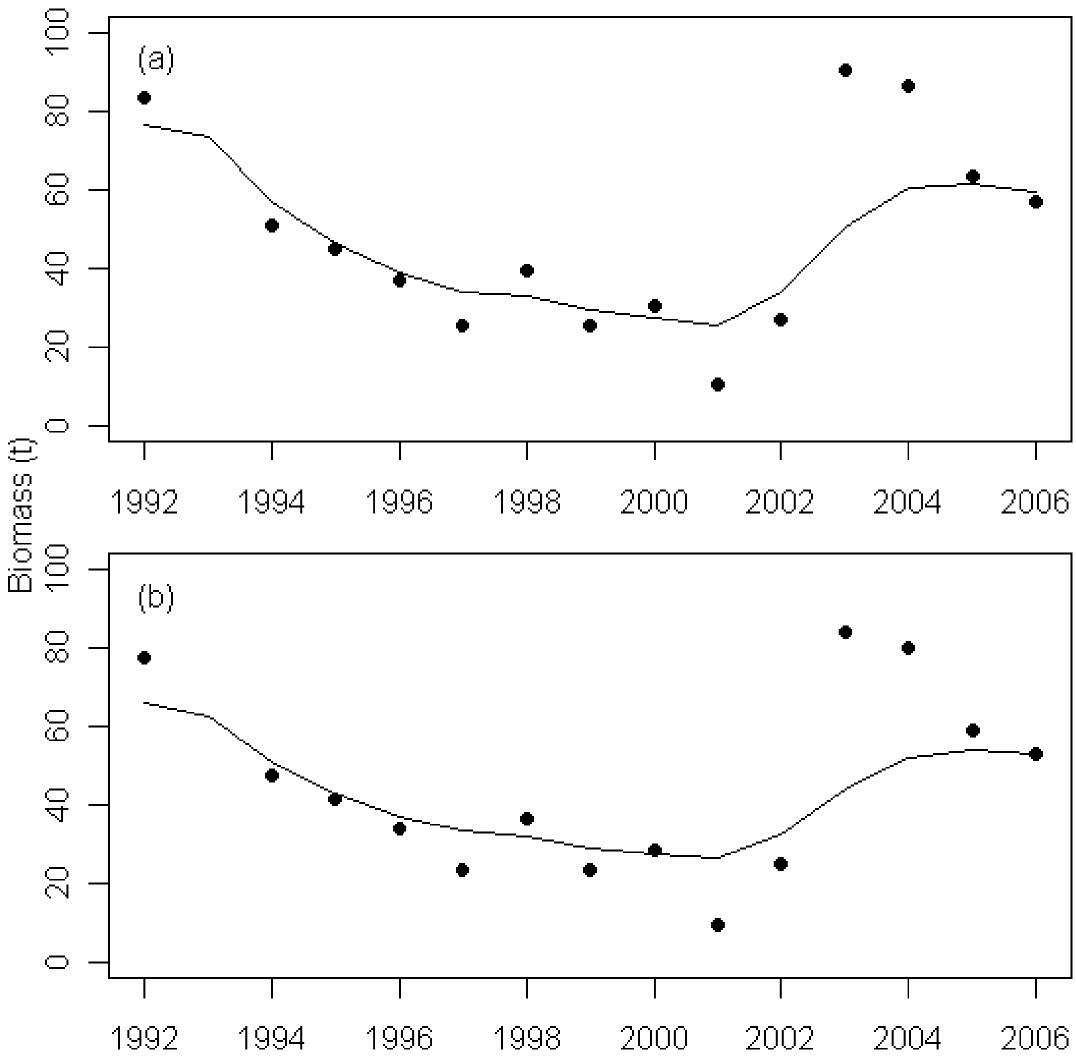


Figure D-1 Production model fits to standardized survey data for (a) ρ = 0.70 and (b) ρ = 0.80. There is an average biomass difference between (a) and (b) of 5,000t. Note that survey abundance indices are re-scaled to biomass units for comparison across models.

Appendix E Catch-age Modeling Approach to Sablefish Stock Assessments

This appendix provides a brief review of historical model-based approaches to stock assessment for B.C. sablefish followed by detailed specifications for the current catch-age model used in management procedure simulations. We developed this approach in response to industry stakeholder skepticism over stock assessment results based on older data from the 1970s and 1980s. The approach we describe is very similar in form and purpose to catch-age models employed by other sablefish management agencies as well as the International Pacific Halibut Commission. Basically, our catch-age approach reconstructs the abundance of individual cohorts that are present in fishery and survey age-proportion samples beginning in 1992 and ending in 2004 (for survey). This information is combined with the standardized sablefish survey (1992 – 2006).

Previous fisheries stock assessment modeling approaches for B.C. sablefish

Over the past 19 years stock assessments for sablefish in B.C. have taken a variety of "best assessment" approaches as reviewed by Haist et al. (2004). From 1989 through to 1994, abundance of BC sablefish was estimated using Virtual Population Analysis (VPA), and yield recommendations were based on forward projects and $F_{0.1}$ yield per recruit decision rules. Between 1994 and 1996, a stock synthesis (Methot 1989) approach was adopted and yield options were based on applying the $F_{40\%}$ rule (fishing mortality rate that would reduce the spawning stock biomass per recruit to 40% of its unfished state) to forward projections. During this time period, analyses of tagging data suggested northern and southern stock areas, and as a consequence yield options were provided for North and South coasts in addition to a coast wide yield option (e.g., Stocker and Fargo 1995). During the transition from VPA to stock synthesis, assumptions about M moved in a more conservative direction from M=0.1 to 0.15 selected for VPA assessments to M=0.05 to 0.1 for stock synthesis assessments. A new catch-age model replaced stock synthesis for the 1996 stock assessment and yield options were now derived using a fixed harvest rule of $F = 0.12$ corresponding roughly to the $F_{40\%}$ rule derived from spawning stock biomass per recruit arguments. During this assessment, the highest ever quota recommendation (6,227 – 16,285 t coastwide) was produced, representing almost a 7-fold increase over the previous year's recommend yield and overly optimistic in comparison to the previous stock synthesis model applied to the same data. In 1997, the assessment model increased in complexity with the development of a new catch-at-age mark-recapture model (Haist et al. 1997). For this model, more detailed information about age- and sex-specific selectivity was included and the number of estimated parameters continued to increase above that required for stock synthesis. This model was run in parallel with the previous assessment model with some conflicting results. These conflicts appeared partially due to the sensitivity of the assumed plus group age in the new age-structured model. In 1998, the assessment model continued to evolve in complexity in an attempt to reflect current thinking about sablefish movement and depth-related behavior, with further disaggregating of the data into 6 sub-regions (North-South shallow, mid and deep depths), movement among these regions, and movement into U.S. waters. Natural mortality rate was now fixed at $M = 0.08$. The integrated catch-age mark-recapture model was further refined and applied up until the 2000 stock assessment. In 2001, the assessment methods were radically simplified with the adoption of a Petersen-type mark-recapture estimate using fish recovered in the year following release from which exploitation rates were derived and analysis of CPUE trends. In January of 2002, the recommended quota option was revised downward to from 4,000 t to 2,800 t in response to declines in standardized survey CPUE (Kronlund et al. 2002). Since that time, a monthly tagging model that integrated available fishery and survey abundance indices was developed (e.g., Haist et al. 2005). These population dynamics models contained no information on age-structure, and biomass estimates were based on predicted numbers and observed mean body weights in surveys, in part as a result of the cessation of production ageing of sablefish in 1997.

Statistical catch-age stock assessment model

Catch-age stock assessment modeling options are potentially appealing for several scientific reasons. First, age-composition changes over time may contain information about temporal trends in fishing mortality and recruitment. Indeed, this particular capability is among the main reasons why so many fisheries agencies attempt to utilize ageing data. Second, in contrast to aggregate biomass production models, observed changes in fishery selectivity as measured by the annual sablefish tagging program can be accounted for in assessments as either fixed parameters or priors. Changes in fishery (and possibly survey) selectivity can have profound influences on abundance estimates from age-structured models, especially when there are few data to distinguish between dome-shaped and asymptotic selectivity functions. An extensive, industry-funded tag-recovery program for B.C. sablefish allows for direct estimation of length-based selectivity from tagging, and therefore potentially large improvements in age-structured assessment estimates. Finally, a catch-age assessment approach provides the ability to use shorter times-series (< 20 years) of fishery-independent data alone, which avoids the many of the potential biases associated with fishery-dependent abundance indices.

Model-based management procedures employed a catch-age model to project exploitable biomass one year into the future so that an upcoming catch limit in year T+1 can be calculated via the harvest rule, where T is the year in which the simulated assessment is performed. Generating this projection first involved estimating the initial population composition in 1992, $N_{3:A}$, and vector of annual age-1 recruitments $\mathbf{R}$ from 1992 to the year T-1, and then projecting exploitable biomass for year T+1 assuming that recruitment would be equal to the historical average. Notation for the catch-age model is provided in Table E-1 with the sequential calculation in Table E-2. This approach is a modification of the model described in Schnute and Richards (1995). The set of estimated model parameters defined in Equation E2.1 include $\hat{\mathbf{N}}_1$ - the initial numbers-at-age present in the first year (i.e., 1992), $\mathbf{R}$ - age-1 recruitment from 1992 to the year in which the assessment is performed, and $\overline{\log R}$ - the long-term average age-1 recruitment. The first section of calculations represents the age-specific lengths (E2.2), weights (E2.3), and proportions mature at each age (E2.4). Next, fishery selectivity (E2.5) is modeled as a double-logistic function to accommodate the estimated dome-shaped selectivity with respect to length for some gear types. Note that selectivity is set equal to zero for lengths smaller than the minimum size limit $\overline{l}_g$, which is 55 cm for all commercial gear and zero for surveys. Equation E2.6 normalizes the age-specific selectivity values so that the maximum over all ages is 1.0. True values for gear-specific selectivity parameters are provided to the simulated assessment and are assumed constant in the future for the purpose of this analysis. Clearly, full evaluation of sablefish management procedures should examine uncertainty in selectivity parameter values derived from tagging, further temporal changes perhaps due to density-dependent growth, and changes in selectivity as a function of sablefish abundance (e.g., changes in fishery targeting behaviour).

We developed a penalized maximum likelihood approach for fitting the catch-age model to simulated observations of relative abundance and catch-at-age. The residual function for the relative abundance survey assumes an observation model of the form $I_t = qB_t e^{v_t}$ where the random variable $v_t \sim N(0, \tau_1^2)$, q is survey catchability, and B_t is the fully selected biomass available to the survey. The latter state variable, which is calculated in E2.14 with $g = 2$, is the only relative abundance index used in this model. Equation E2.19 therefore represents the conditional maximum likelihood estimate (MLE) of log-survey catchability, where the conditioning is on the initial abundance and recruitment parameters of the model. Equation E2.20 is the conditional MLE of the survey variance τ_1^2. The likelihood function for the age proportions is a multivariate-logistic, which we adopted because it does not over-weight age-proportion data in the manner of traditional multinomial likelihoods (Schnute and Richards 1995). The age proportion residual calculation (E2.17) is done for trap fishery and trap survey ages proportions and involves only ages 3 and older to the plus group at age-25. Note that $p_{g,a,t}$ in E2.17 represents the observed age proportions. Equation E2.21 provides the conditional MLEs of the age proportion variances $\tau_{2,g}^2$ for the two gear types.

The final term in the total likelihood (E2.22) is the kernel of a $N(0, \sigma^2)$ prior on annual log-recruitment deviations from the long-term average. Note that we provide this prior with the true recruitment standard deviation (0.70) used in the operating model, because the maximum likelihood approach cannot estimate both process (σ^2) and observation (τ^2) error variances simultaneously. Although we could have chosen an errors-in-variables approach (Schnute and Richards 1995), this would involve making another assumption about the ratio of process to observation errors, similar to that made for the production model in Appendix D, which adds another management procedure option. Along with the requirement to investigate selectivity uncertainties, future management procedure evaluations should examine whether this catch-age model is robust to mis-specification of the process error variance. This is especially important for sablefish because recruitment variances can be poorly estimated for species that are difficult to age.

Estimates of selectivity based on tag-recovery data

As mentioned above we assumed that selectivity-at-length was known for each gear type that exploits sablefish. Selectivity functions were estimated by fitting the double-logistic model in equations E2.5-2.6 to tag release-recovery data using a binomial likelihood (Figure E1). The resulting selectivity-at-length relationships were then transformed to selectivity-at-age (Figure E-2) using the von Bertalanffy growth function in equation E2.2. One result of these gear-specific selectivity functions is that there are different exploitable biomasses available to each gear type. All of these exploitable biomasses are different from spawning biomass because fish tend to recruit to fisheries before the spawning stock.

Model fit to observed data, 1992 - 2006

As is common in many fisheries stock assessments that combine survey and catch-age data, this analysis is very strongly influenced by the age proportions. The shear quantity of age proportion observations (176 for commercial and 220 for surveys) means that this data set contributes about 10 times as much to the overall model likelihood as the survey, which has only 15 data points in total. In addition, the maximum likelihood estimate of the survey standard error is $\tau_1 = 0.32$, which allows considerable scope for alternative interpretation of the stock trajectory over the past 15 years from survey catch rates alone. Thus, the model attempts to fit the multiple complex patterns in the annual age proportion data (Figure E-3 and Figure E-4) more than the single time-series pattern of the survey (Figure E-5). This does not mean that the survey carries no weight in this analysis; on the contrary, the survey actually provides a strong constraint on the stock dynamics as implied by the age proportion data alone. For example, ignoring the survey altogether produces biomass estimates that decline much more rapidly than that shown in Figure E-5.

The survey age proportions appear to have lower variance than the commercial age proportions, which is expected given that commercial samples have not been collected in as highly regimented a manner as the surveys. On the other hand, the difference is not as great as expected because the survey shows a standard error of $\tau_{2,2} = 0.43$ compared to $\tau_{2,1} = 0.54$ for commercial ages. Neither set of age proportions are particularly precise, which, combined with the lack of strong differences between collections, suggests some difficulties that hinder precise age classification. We did not attempt a deconvolution of the aging errors because comprehensive aging error validation studies have not been performed for the B.C. sablefish aging program.

It is important to note that the fit to the survey shown in Figure E-5 (solid line) will remain the same across operating model scenarios used in the MPE exercise for the first year of each simulation. Biomass estimates remain the same because we do not alter the existing sablefish catch, survey, and age proportion data in any way; instead, we simply augment these data sets with new values simulated from the scenarios. Thus, the catch-age model described here will interpret the historical period 1992 – 2006 exactly the same under each scenario for assessment year 2006. This model eventually adapts, however, as new data are added that support some alternative scenario. Our philosophy here is that short-term patterns of catch resulting from this stock assessment method will be more accurately represented using this approach compared to simulating new observations for the historical period.

For this paper, we used standardized surveys combined with commercial and standardized survey age proportions because this combination appeared to provide the best overall performance based on general statistics of the fits (Table E-3). Thus, according to this catch-age analysis, the projected trap exploitable biomass for 2007 is 26 971 t and the spawning biomass in 2006 is estimated to be approximately 45% of the spawning biomass in 1992 (Table E-3). The recommended 2007 quota generated from these conditions will depend on the harvest control rule component of the management procedures.

Table E-1 Notation for the catch-age stock assessment model presented in Table E-2.

Symbol	Value	Description
Indices		
T		Time step $t = \{1,2,\ldots T\}$
A		Age-class in years $a = \{1,2,\ldots A\}$
G	Table E1	Fishery gear type index.
Model parameters		
q		Catchability coefficients $i = \{1,2\}$ for gear g
$\beta_{i,g}$	Table E1	Double-logistic selectivity function parameters for gear g
τ_1	Table E1	Coefficient of variation for survey abundance index
$\tau_{2,g}$	Table E1	Standard error in observed proportions-at-age for gear g
σ	0.70	Standard error of log-recruitment deviations
M	0.08	Instantaneous natural mortality rate (/yr)
L_∞	68.2	Asymptotic length (cm)
L_1	40.7	Length-at-age 1 (cm)
K	0.37	von Bertalanffy growth constant
μ_2	5	Age-at-50% maturity
μ_1	8	Maturity-at-age function slope
Derived parameters		
$s_{g,a}$		Selectivity-at-age in fishery g
m_a		Proportion mature-at-age
w_a		Body mass-at-age
State variables		
$N_{a,t}$		Number of age a fish in year t
$B_{a,t}$		Biomass of age a fish in year t
$B^*_{g,t}$		Biomass of fish vulnerable to gear g in year t
$u_{g,a,t}$		Proportion of age a fish in harvestable population
S_t		Spawning biomass in year t
Observations		
I_t		Survey abundance index value in year t
$p_{g,a,t}$		Proportion of age a fish in gear g catch-age sample
Fishery controls		
$C_{g,t}$		Catch in fishery g (tonnes)
l_g		Minimum size limit in fishery g

Table E-2 State-space catch-age model for estimating stock biomass from survey relative abundance and catch-age data. The table defines a calculation sequence from input parameter values through to the likelihood function. Note that the bold parameters in E2.1 are vectors.

Estimated parameters

E2.1 $$\boldsymbol{\Theta} = \left(\hat{\mathbf{N}}_1, \mathbf{R}, \overline{\log R}\right)$$

Life history schedules

E2.2 $$l_a = L_\infty + \left(L_1 - L_\infty\right) e^{\left(-k(a-1)\right)}$$

E2.3 $$w_a = \exp[-25.9] l_a^{3.1}$$

E2.4 $$m_a = \frac{a^{\mu_1}}{a^{\mu_1} + \mu_2^{\mu_1}}$$

Fishery selectivity

E2.5 $$\tilde{s}_{g,a} = \begin{cases} 0 & l_a < \overline{l}_g \\ \left(\dfrac{1}{1 + e^{-\beta_{2,g}\left(l_a - \beta_{1,g}\right)}}\right)\left(1 - \dfrac{1}{1 + e^{-\beta_{4,g}\left(l_a - \beta_{1,g} - \beta_{3,g}\right)}}\right) & l_a \geq \overline{l}_g \end{cases}$$

E2.6 $$s_{g,a} = \tilde{s}_{g,a} \Big/ \max\left[\tilde{s}_{g,1}, \tilde{s}_{g,2}, \ldots \tilde{s}_{g,A}\right]$$

Initial population $t = 1$

E2.7 $$N_{a,1} = \begin{cases} R_1 & a = 1 \\ \hat{N}_a & a \geq 2 \end{cases}$$

E2.8 $$B_{a,1} = w_a N_{a,1}$$

Age proportions in catch

E2.9 $$u_{g,a,t} = s_{g,a} N_{a,t} \Big/ \sum_a s_{g,a} N_{a,t}$$

State dynamics

E2.10 $$N_{1,t} = \begin{cases} R_t & 1 < t \leq T \\ \exp\left[\overline{\log R}\right] & t = T + 1 \end{cases}$$

E2.11 $$N_{a,t} = e^{-M}\left[N_{a-1,t-1} - \sum_{g=1}^{G} u_{g,a-1,t-1} C_{g,t-1} / w_{a-1}\right], \; 2 \leq a \leq A - 1$$

E2.12 $$N_{A,t} = e^{-M}\left[N_{A-1,t-1} + N_{A,t-1} - \sum_{g=1}^{G}\left(u_{g,A-1,t-1} + u_{g,A,t-1}\right)C_{g,t-1} / w_A \right] \quad a = A$$

E2.13 $$B_{a,t} = w_a N_{a,t}$$

E2.14 $$B_{g,t}^{*} = \sum_{a=1}^{A} s_{g,a} B_{a,t}$$

E2.15 $$S_t = \sum_{a=1}^{A} m_a B_{a,t}$$

Residuals

E2.16 $$\xi_t = \log\left(\frac{I_t}{B_{2,t}^{*}}\right) - \frac{1}{T}\sum_{t}\log\left(\frac{I_t}{B_{2,t}^{*}}\right)$$

E2.17 $$\eta_{g,a,t} = \log p_{g,a,t} - \log u_{g,a,t} - \frac{1}{A}\sum_{a=3}^{A}\left[\log p_{g,a,t} - \log u_{g,a,t}\right]$$

E2.18 $$\omega_t = \log R_t - \overline{\log R}$$

Conditional maximum likelihood estimates

E2.19 $$\widehat{\log q} = \frac{1}{T}\sum_{t}\log\left(\frac{I_t}{B_{2,t}^{*}}\right)$$

E2.20 $$\hat{\tau}_1^2 = \frac{1}{T}\sum_{t=1}^{T}\xi_t^2$$

E2.21 $$\hat{\tau}_{g,2}^2 = \frac{1}{\left(A-3\right)n_g}\sum_{a=3}^{A}\sum_{t=1}^{n_g}\eta_{g,a,t}^2$$

Log-likelihood

E2.22 $$\ell = \frac{T}{2}\log\hat{\tau}_1^2 + \sum_{g}\frac{\left(A-3\right)n_g}{2}\log\hat{\tau}_{2,g}^2 + \frac{1}{2\sigma^2}\sum_{t}\omega_t^2$$

Table E-3 Summary of catch-age model estimates based on alternative combinations of survey and age proportion datasets. Estimated quantities are projected 2007 trap exploitable biomass (B_{2007}), exploitation rate on trap exploitable stock in 2006 (U_{2006}), ratio of spawning biomass in 2006 relative to 1992 (S_{2006}/S_{1992}). Model statistics are negative-log-likelihoods (Likelihoods) for standardized survey (Survey) and age proportions (Ages; commercial trap on top, survey on bottom) and corresponding variances. Top row shown in bold is the configuration used in management procedure simulations to date.

				Likelihoods			Variances	
Dataset	B_{2007}	U_{2006}	S_{2006}/S_{1992}	Survey	Ages		Survey	Ages
Survey + All ages	**26 971**	**0.099**	**0.45**	**11.33**	**-109** **-185**		**0.32**	**0.54** **0.43**
Survey + Survey ages	28 890	0.093	0.44	11.11	-87 -196		0.31	0.61 0.41
Survey + Comm ages	37 584	0.073	0.67	12.03	-129 -129		0.35	0.48 0.55

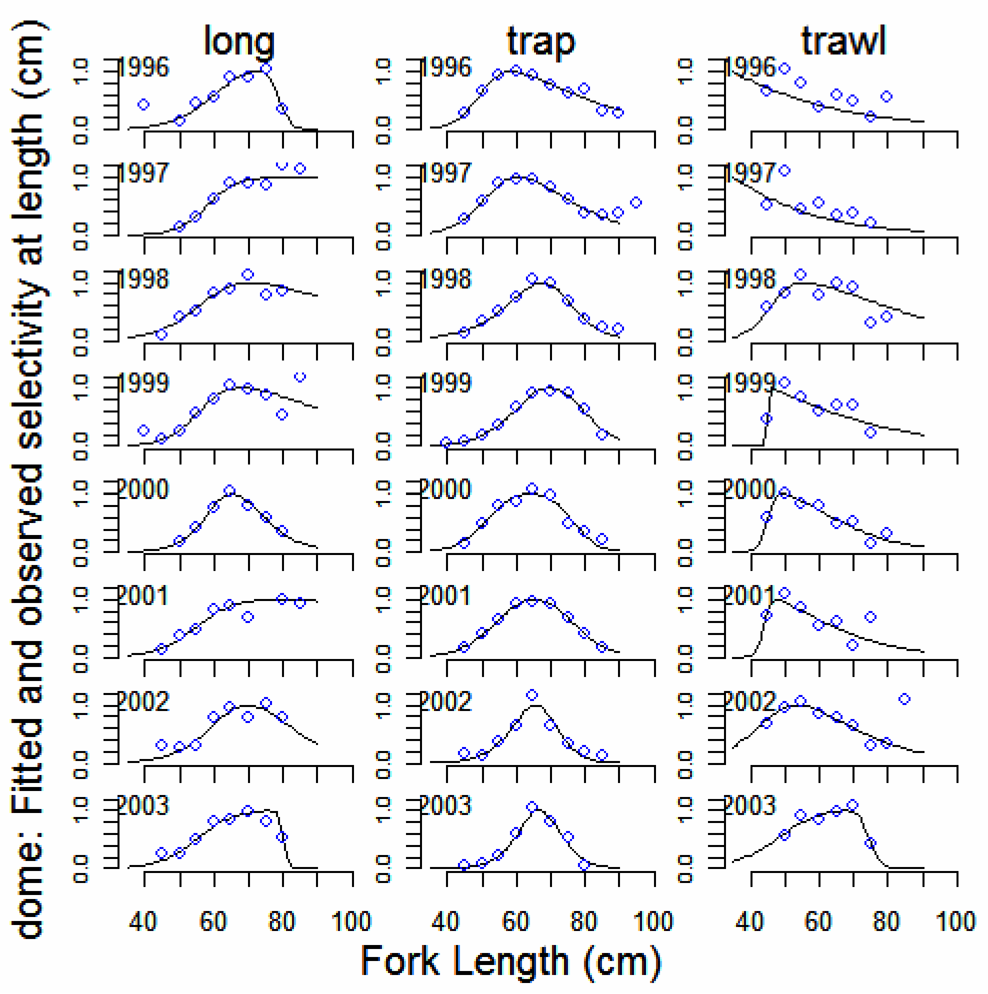


Figure E-1 Selectivity-at-length for sablefish released from traditional offshore tagging surveys and recaptured in longline hook (long), (b) longline trap, and (c) trawl fisheries from 1996 – 2003. Fitted double-logistic functions are shown as black lines and circles are empirical selectivity values calculated directly from release-recovery data.

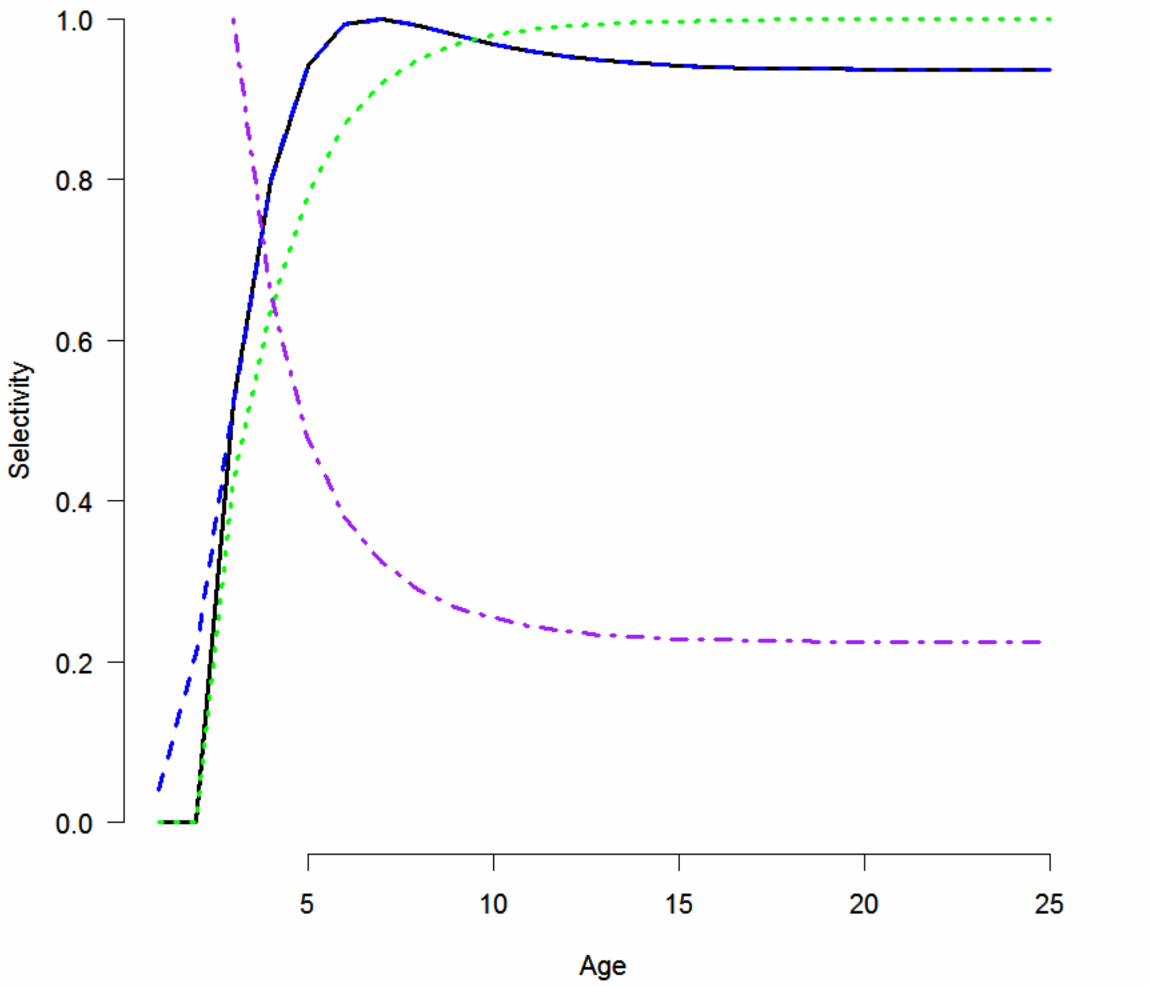


Figure E-2 Average selectivity-at-age functions for trap fishery (solid black), trap survey (dashed blue), longline hook (dotted green), and trawl (dot-dash purple). Actual selectivity functions were parameterized based on length and therefore the curves shown here involve a length-at-age model.

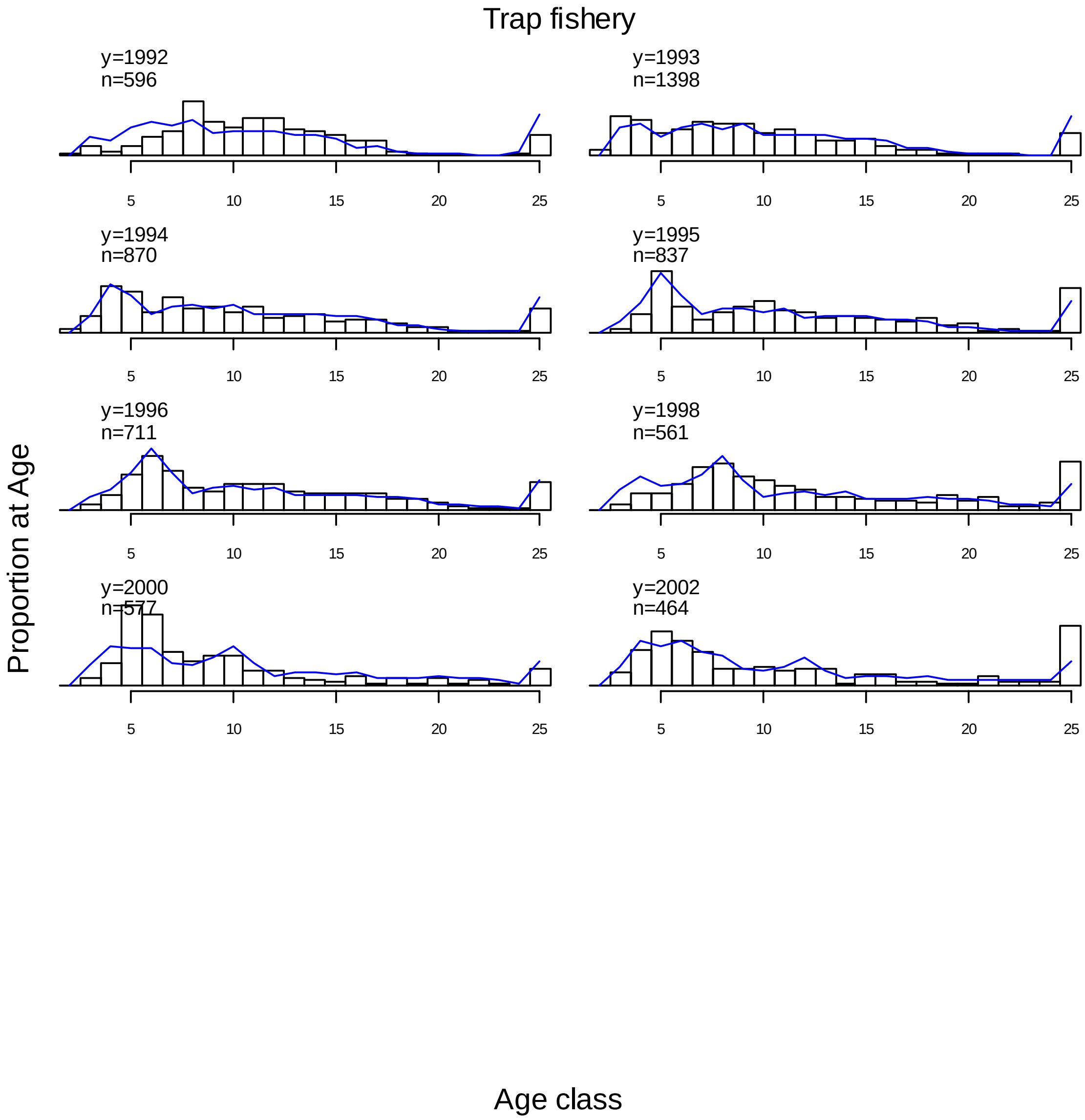


Figure E-3 Catch-age model fit to trap fishery age proportions. Indices y = year and n = sample size.

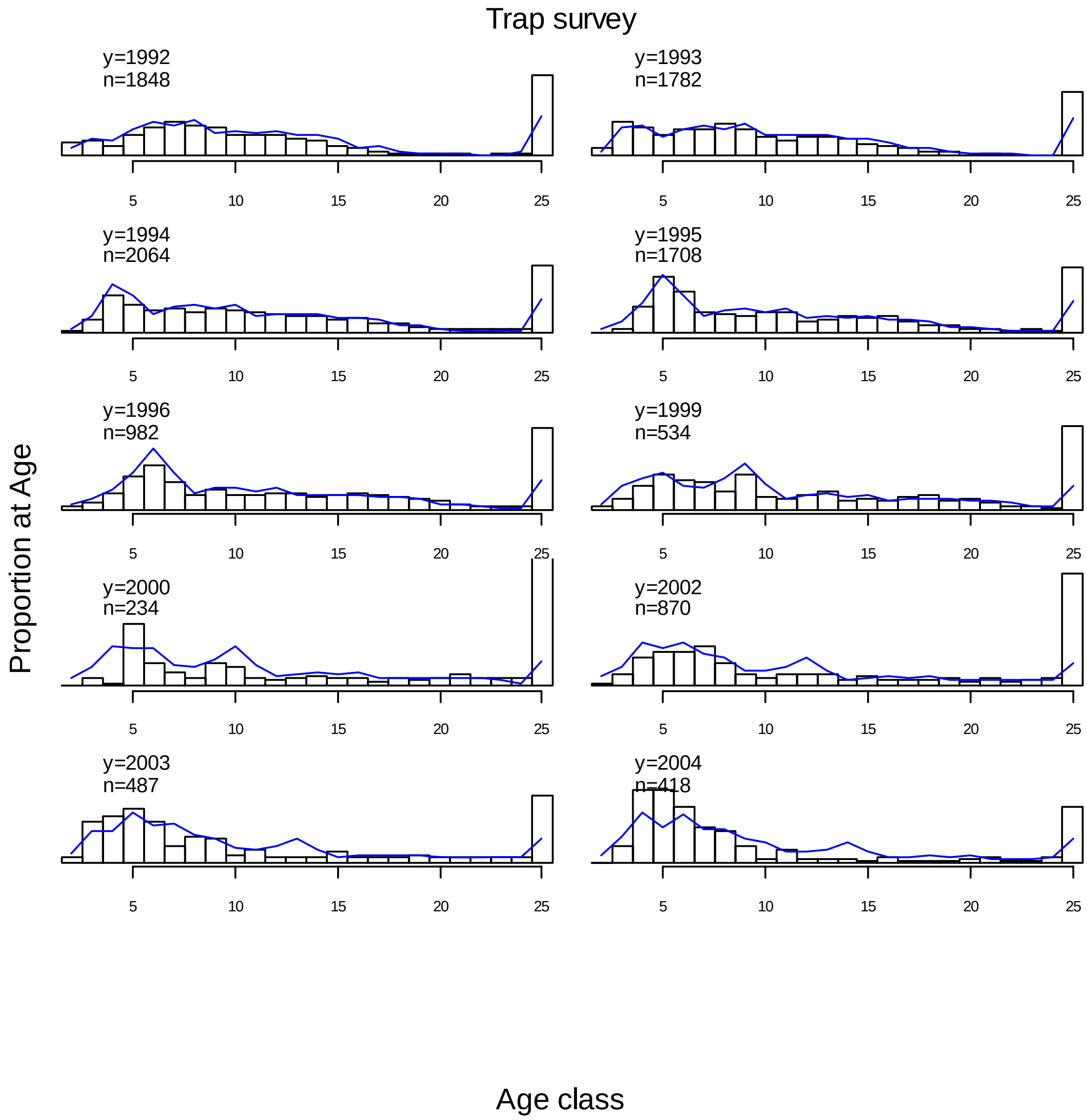


Figure E-4 Catch-age model fit to standardized survey age proportions. Indices y = year and n = sample size.

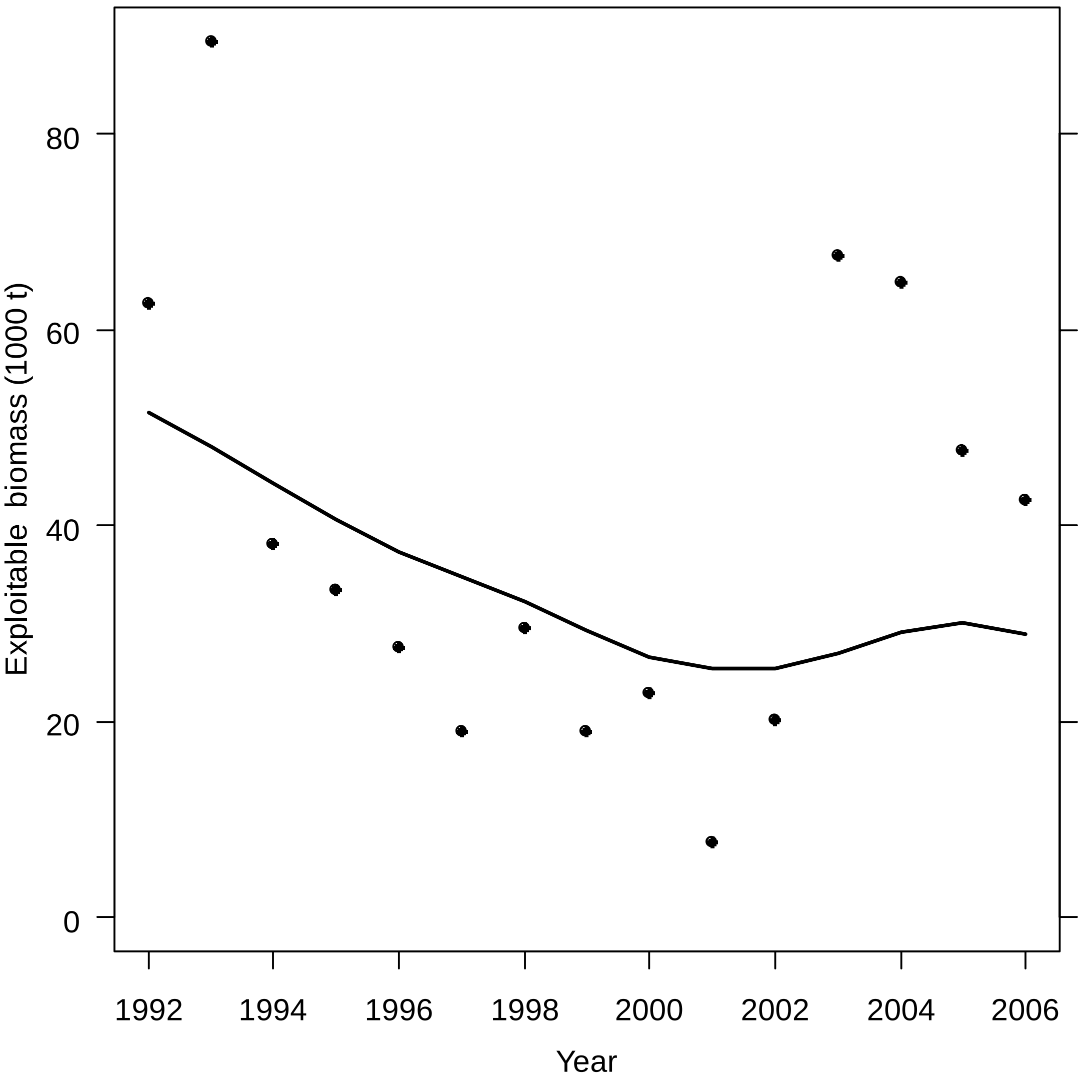


Figure E-5 Fit of catch-age model (line) to standardized surveys (circles) from 1992 to 2006. Note that survey relative abundance indices have been re-scaled to biomass units.

Appendix F Trade-off Relationships and Performance Statistics

This appendix provides performance summaries for all data- based, production model-based and catch-age model based procedures listed in Table 1. The results are presented in the form of the trade-off relationships between median average catch and median average depletion as presented for selected procedures in Section 5 of the main paper. For the selected procedures, full tabular listings of the performance measures are computed for non-overlapping time blocks corresponding to projection years 1-5, 6-10, 11-20 and 21-40 (Table F-1 through Table F-4).

We also include a listing of performance measures and figures for selected procedures calculated using cumulative time blocks corresponding to projection years 1-5, 1-10, 1-20 and 1-40 (Table F-5 through Table F-8). Corresponding cumulative trade-off relationships for all data-based, production model-based and catch-age model-based procedures are depicted in Figure F-1 through Figure F-3 with Figure 10 showing these trade-offs for selected procedures only. Graphical summaries of average catch, catch variability, and spawning biomass depletion are shown in Figure 11 through Figure 15.

Data-based procedures

As expected, increasing the value of the parameter λ_1 was effective at reducing inter-annual fluctuations in catch. Results from the data-based procedures show the expected negative trade-off relationship between median average catch and median average depletion for each scenario (Figure F-1); that is, higher levels of catch are associated with lower values of stock depletion. The short-term (5-10 year) results shown in Figure F-1a,b show a somewhat steep relationship with strong contrast in average catch among procedures associated with only minor differences in stock depletion; relatively large reductions in average annual catch are required to make modest gains along the depletion *x*-axis. Over longer time horizons (20-40 years, Figure F-1c,d), procedures that imposed reduced average catches during the first 10 years of the simulation tended to promote stock rebuilding and thus recovered some of the lost catch due as stock size increased and resulted in higher average catches. This can be seen by tracing particular symbols down and across the panels a-d in Figure F-1 and by reference to the initial depletion markers. The slope of the trade-off contour changed over time, favoring large gains in average depletion at the expense of relatively smaller decreases in catch as the time horizon increased.

Constant harvest rate procedures yield the highest average catches and correspondingly the lowest average depletion values in each scenario for the most aggressive tuning of the data-based decision rule $\left(\lambda_1 = 0.8,\ \lambda_2 = 210\right)$. The CHR tuning with $\left(\lambda_1 = 0.8,\ \lambda_2 = 240\right)$ was not included as scenario S1 depletion levels would be unacceptably small over the long-term. Differences between the CHR and variable harvest rate rules were not great during the first 5 to 10 years of the simulation, but the most aggressive CHR rules tended to deliver the highest average catch and VHR rules tended towards higher average depletion values at any given level of average catch. After 20 years (Figure F-1c) the average depletion levels of scenarios S1 and S3 converged, as the more productive scenario S3 $\left\{h = 0.65,\ \hat{q}_{2,trap}\right\}$ is depleted by procedures with higher effective exploitation rates, but driven to higher stock sizes by less aggressive tunings of the rule. Scenario S2 loses ground along the conservation axis over the first 20 years but reduced catches allow this less productive scenario $\left\{h = 0.45,\ q_{2,trap} = 1\right\}$ to rebuild as it does not start from a low initial depletion value. The converse occurs for scenario S1 as less aggressive procedures allow some stock building at the expense of reduced catches, but most lead to reduced average depletion values. Data-based rules with effectively constant exploitation rates do not lie on a smooth contour due to the tunings achieved by the various levels of λ_1 parameter. This effect is most notable in the long-term (Figure F-1d) and in particular for scenario S1 $\left\{h = 0.45,\ \hat{q}_{2,trap}\right\}$, with diminishing differences in outcomes as results progress to scenario S4.

Variable harvest rate rules tend to lower the effective harvest rate and therefore the average annual catches relative to the constant harvest rate tunings of the rule. The most extreme case occurs for scenario S1, where stock depletion is almost 23% higher than the initial depletion level for about the same average catch (Figure F-1d) whereas this effect becomes almost negligible for scenario S4. This result likely relates to that utility of a given catch difference at low versus high depletion levels, i.e., a given catch has a much greater impact on depletion at low stock sizes. The effect of the risk adjustment appeared to be to lower the effective harvest rate along the same trade-off contour as for a variable harvest rate rule without an adjustment. Nevertheless, this may be an important consideration for low productivity scenarios where VHR procedures with $Q = 0.4$ were most successful at maintaining the stock above the $0.2B_0$ limit over the long-term.

Catch-age procedures

Under catch-age model-based CHR rules, exploitation rate targets of $U^{ref} = \{0.06, 0.08, 0.10\}$ differed substantially in median average annual catch in the short-term and median average depletion in the long-term (Figure F-2). Under the two unproductive scenarios, CHR rules resulted in stock declines over the first 20 years (Figure F-2a-c) to levels lower than those at the time when the management procedure started (i.e., 18% for scenario S1 and 30% for scenario S2). Only the least aggressive CHR procedure with $U^{ref} = 0.06$ was able to essentially maintain the initial depletion value for scenario S1 over the long-term (40 years) and no CHR procedures were able to rebuild to the initial depletion value for scenario S2 . For scenario S4, the constant harvest rate procedure with $U^{ref} = 0.1$ was only able to maintain the stock at the initial depletion value over the long-term but resulted in the highest average catches of 4913 t. All CHR procedures lead to increases in stock depletion over all time periods (Figure F-2).

In contrast to CHR rules, almost all VHR rules increased stock size over 40 years (Figure F-2d). The most aggressive VHR procedures with $U^{ref} = 0.10$ produced a minor reduction under scenario S1 and a relatively large reduction under S2 from approximately 0.30 to 0.24. The risk adjustment factor Q had little effect on long-term recovery for model-based VHR rules. Increases in stock size were primarily due to much lower average catches under VHR rules in the short-term, by as much as 1 000 t. Over time, the differences between average catches for CHR and VHR procedures diminished because

VHR procedures increased stock biomass and thus increased catch, whereas CHR rules tended to cause declines in catch as a result of stock depletion.

Accounting for uncertainty in estimated biomass via the risk adjustment shifted outcomes along the trade-off contour toward higher average depletion and lower average catch, essentially reducing the average exploitation rate. Small effects of the risk adjustment mainly reflect a low degree of uncertainty in the assessment estimates of biomass, particularly for the catch-age model approach. Although our catch-age model exhibited low uncertainty in biomass estimates, the estimates themselves tended to be systematically biased with respect to the true stock size as noted in the main document. However, note the considerable overlap between CHR and VHR rules that occurred along the catch-depletion trade-off contour for the higher productivity scenarios, indicating that for at least some reference exploitations rates the CHR results can be achieved by an appropriately tuned VHR rule which is compliant with DFO policy.

Production model-based procedures

Production model procedures did not show as much range in both median average catch and median average depletion as the results for data-based or catch-age model-based procedures (Figure F-3). The trade-off relationship between median average catch and median average depletion showed the expected negative slope in the short-term which progressively flattened over time (Figure F-3a-d) in a manner similar to result reported above for the data-based and catch-age model-based procedures.

For productive scenarios (S3 and S4), variable harvest rate procedures without the risk adjustment (i.e., $Q = 0.5$) had little effect on the average catch-average depletion trade-off relationship compared to CHR procedures. This is because the estimated stock depletion never fell below the reference level $D_{high} = 0.4$ where the exploitation rate would be reduced, even though the true scenario stock biomass fell well below these levels. A negligible effect of VHR procedures resulted for scenario S2, but the difference was more pronounced for scenario S1 where VHR procedures tended to increase the depletion level at the expense of small reductions in median average catch. The risk adjustment ($Q = 0.4$), when coupled with a VHR rule, achieved significantly increased average depletion albeit at the expense of reductions in median average catch of about

184 to 150 t per year as the exploitation rate increased from $U^{ref} = 0.06$ to $U^{ref} = 0.10$ over the long-term (Figure F-3d). The effect of $Q = 0.4$ was much greater on production model-based procedures than on catch-age model procedures. This is because the production model has a greater number of effective parameters per observation than the catch-age model does. Thus, uncertainty is spread over more parameters, which results in greater uncertainty in biomass projections and quotas.

Cumulative Performance Measures for Selected Procedures

Measuring performance of management procedures using cumulative time blocks tends to mask some of the important differences in procedure performance over time. For example, less aggressive procedures that provided low catches during early time periods tended to promote stock growth, which resulted in higher catches in later time periods. In some cases, catches in later time periods exceeded those provided by more aggressive procedures. However, differences in catch may not appear when averaged over the full 1 – 40 year window. On the other hand, cumulative averaging provides a more realistic economic view of average fishery yield into the future. In our view, stock depletion is probably better measured using non-overlapping periods, while catch performance should be measured cumulatively.

Average Catch and Trap Fishery CPUE Trade-offs for Selected Procedures

Trap fishery CPUE was identified by industry stakeholders as an output of interest. Here we examine the trade-off relationships between median average catch (t) and predicted average trap fishery CPUE (kg/trap; Figure F-8). As was the case for the trade-off relationship between median average catch and median average depletion discussed in Section 5, the form of the trade-off relationship is steep over the short term (5-10 years), but flattens by 20 years, and by 40 years large increases in trap fishery CPUE are achieved with relatively small reductions in average annual catch. All procedures attain CPUE objectives identified by industry for scenarios S3 and S4, while only the least aggressive model-based procedures $U^{ref} = 0.06$ and data-based procedures with $\lambda_2 = \{150, 180\}$ meet the objectives for scenarios S1 and S2.

The trade-offs between average catch and CPUE can similarly be used to eliminate some procedures based on stakeholder objectives following inspection of Figure F-8. None of the procedures achieve stakeholder objectives for CPUE within 5-years, but by 10 years most of the procedures produce median average trap CPUE values greater than 14 kg/trap for scenarios S3 and S4. This outcome remains true for the 20 year time horizon, and it is not until the long-term that at least some procedures tested under scenarios S1 and S2 attain stakeholder objectives for fishery CPUE. Once again, only the least aggressive model-based procedures meet or exceed the objective for all scenarios, as do data-based procedures with $\lambda_2 = \{150, 180\}$ (Table F-5 to Table F-8).

Table F-1 Summary of performance statistics and ranks by management procedure for Scenario 1 $\left\{h = 0.45, \hat{q}_{2,trap}\right\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	**Statistic**
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.151	0.139	2081	12.4	0.00	7.5	11.5	4
	Data ($\lambda_2 = 150$)	2856	16.1	0.151	0.138	2081	12.4	0.00	7.5	11.5	6
	Data ($\lambda_2 = 180$)	3067	16.1	0.148	0.132	2234	12.3	0.00	3	8	10
	Data ($\lambda_2 = 210$)	3366	15.3	0.144	0.125	2473	12.2	0.00	2	3	11
	Data ($\lambda_2 = 240$)	3679	14.5	0.140	0.119	2722	12.1	0.00	1	2	12
	Prod $U^{ref} = 0.06$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	9.5	7
	Prod $U^{ref} = 0.08$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	9.5	8
	Prod $U^{ref} = 0.10$	2943	14.0	0.150	0.136	2272	12.4	0.00	4	1	9
	CA $U^{ref} = 0.04$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.06$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.08$	2856	16.1	0.151	0.139	2081	12.4	0.00	10.5	5.5	2
	CA $U^{ref} = 0.10$	2856	16.1	0.151	0.138	2081	12.4	0.00	10.5	5.5	5
6-10	Data ($\lambda_2 = 120$)	1271	15.7	0.152	0.163	938	12.6	0.00	10	7	4
	Data ($\lambda_2 = 150$)	1309	15.1	0.152	0.162	1021	12.6	0.00	8	5	5
	Data ($\lambda_2 = 180$)	1424	15.3	0.144	0.152	1134	12.3	0.00	7	6	10
	Data ($\lambda_2 = 210$)	1571	15.9	0.132	0.139	1218	12.0	0.00	4	9	11
	Data ($\lambda_2 = 240$)	1700	16.1	0.121	0.127	1281	11.6	0.00	2	11	12
	Prod $U^{ref} = 0.06$	1272	15.7	0.153	0.166	994	12.6	0.00	9	8	3
	Prod $U^{ref} = 0.08$	1497	16.0	0.150	0.160	1184	12.5	0.00	5	10	7
	Prod $U^{ref} = 0.10$	1620	17.3	0.146	0.154	1297	12.3	0.00	3	12	9
	CA $U^{ref} = 0.04$	806	12.8	0.160	0.175	674	12.8	0.00	12	3	1
	CA $U^{ref} = 0.06$	1155	12.5	0.156	0.168	981	12.7	0.00	11	1	2
	CA $U^{ref} = 0.08$	1470	12.9	0.152	0.162	1250	12.5	0.00	6	4	6
	CA $U^{ref} = 0.10$	1730	12.6	0.149	0.155	1438	12.4	0.00	1	2	8

									Rank Statistic		
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	$C_{5\%}$	**Avg. CPUE**	P_{cons}	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
11-20	Data ($\lambda_2 = 120$)	1025	10.6	0.198	0.225	758	13.8	0.50	12	4	2
	Data ($\lambda_2 = 150$)	1192	10.8	0.192	0.218	889	13.6	0.40	10	6	3
	Data ($\lambda_2 = 180$)	1233	11.5	0.179	0.203	924	13.2	0.20	7	7	6
	Data ($\lambda_2 = 210$)	1201	11.5	0.166	0.188	902	12.9	0.00	9	8	9
	Data ($\lambda_2 = 240$)	1116	11.8	0.150	0.173	829	12.4	0.00	11	9	12
	Prod $U^{ref} = 0.06$	1700	13.7	0.183	0.199	1216	13.3	0.00	5	10	5
	Prod $U^{ref} = 0.08$	1857	14.6	0.171	0.181	1353	13.0	0.00	4	11	8
	Prod $U^{ref} = 0.10$	1909	15.7	0.161	0.169	1381	12.7	0.00	2	12	10
	CA $U^{ref} = 0.04$	1215	6.7	0.212	0.231	1011	14.0	0.60	8	1	1
	CA $U^{ref} = 0.06$	1574	7.9	0.190	0.205	1319	13.5	0.30	6	2	4
	CA $U^{ref} = 0.08$	1858	9.2	0.173	0.185	1483	13.1	0.00	3	3	7
	CA $U^{ref} = 0.10$	1993	10.7	0.159	0.168	1545	12.7	0.00	1	5	11
21-40	Data ($\lambda_2 = 120$)	2101	7.8	0.295	0.316	1314	16.2	1.00	11	4	2
	Data ($\lambda_2 = 150$)	2397	8.6	0.267	0.280	1413	15.6	1.00	3	5	3
	Data ($\lambda_2 = 180$)	2452	9.4	0.236	0.254	1375	14.9	0.85	1	7	6
	Data ($\lambda_2 = 210$)	2447	10.1	0.214	0.225	1302	14.4	0.63	2	8	8
	Data ($\lambda_2 = 240$)	2343	10.6	0.196	0.204	1208	13.9	0.50	4	9	10
	Prod $U^{ref} = 0.06$	2289	11.4	0.243	0.258	1750	15.0	0.38	9	10	5
	Prod $U^{ref} = 0.08$	2338	13.1	0.212	0.225	1735	14.3	0.05	6	11	9
	Prod $U^{ref} = 0.10$	2328	14.9	0.190	0.198	1651	13.7	0.00	7	12	12
	CA $U^{ref} = 0.04$	1832	4.8	0.316	0.364	1397	16.6	1.00	12	1	1
	CA $U^{ref} = 0.06$	2197	6.3	0.261	0.279	1696	15.5	1.00	10	2	4
	CA $U^{ref} = 0.08$	2322	7.6	0.222	0.227	1800	14.4	0.68	8	3	7
	CA $U^{ref} = 0.10$	2339	9.1	0.191	0.192	1773	13.6	0.30	5	6	11

Table F-2 Summary of performance statistics and ranks by management procedure for Scenario 2 $\{h = 0.45,\ q_{2,trap} = 1\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

									Rank Statistic		
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.271	0.261	2081	12.0	1.00	10	12	3
	Data ($\lambda_2 = 150$)	2892	15.3	0.271	0.260	2146	12.0	1.00	9	9	6
	Data ($\lambda_2 = 180$)	3208	13.5	0.267	0.253	2466	11.9	1.00	5	7	10
	Data ($\lambda_2 = 210$)	3579	12.2	0.264	0.245	2783	11.8	1.00	2	3	11
	Data ($\lambda_2 = 240$)	3948	11.1	0.260	0.237	3094	11.6	1.00	1	1	12
	Prod $U^{ref} = 0.06$	2926	14.2	0.271	0.260	2256	12.0	1.00	8	8	4
	Prod $U^{ref} = 0.08$	3059	12.6	0.270	0.257	2509	12.0	1.00	6	5	7
	Prod $U^{ref} = 0.10$	3236	11.6	0.270	0.255	2746	12.0	1.00	3	2	9
	CA $U^{ref} = 0.04$	2856	16.1	0.271	0.261	2081	12.0	1.00	11.5	10.5	1.5
	CA $U^{ref} = 0.06$	2856	16.1	0.271	0.261	2081	12.0	1.00	11.5	10.5	1.5
	CA $U^{ref} = 0.08$	2988	13.3	0.271	0.259	2436	12.0	1.00	7	6	5
	CA $U^{ref} = 0.10$	3231	12.5	0.270	0.257	2828	12.0	1.00	4	4	8
6-10	Data ($\lambda_2 = 120$)	1461	12.2	0.284	0.299	1226	13.0	1.00	12	10	2
	Data ($\lambda_2 = 150$)	1631	11.3	0.279	0.294	1409	12.8	1.00	10	5	3
	Data ($\lambda_2 = 180$)	1880	12.0	0.269	0.281	1600	12.3	1.00	9	6	7
	Data ($\lambda_2 = 210$)	2130	12.1	0.258	0.266	1796	11.8	1.00	6	9	11
	Data ($\lambda_2 = 240$)	2380	12.1	0.245	0.252	1974	11.4	1.00	4	8	12
	Prod $U^{ref} = 0.06$	2002	12.0	0.276	0.290	1712	12.7	1.00	8	7	5
	Prod $U^{ref} = 0.08$	2293	12.4	0.270	0.280	1957	12.4	1.00	5	11	6
	Prod $U^{ref} = 0.10$	2480	13.3	0.264	0.271	2086	12.1	0.80	3	12	9
	CA $U^{ref} = 0.04$	1479	10.0	0.285	0.301	1291	13.0	1.00	11	4	1
	CA $U^{ref} = 0.06$	2123	9.8	0.278	0.288	1881	12.7	1.00	7	3	4
	CA $U^{ref} = 0.08$	2648	9.7	0.269	0.273	2320	12.4	1.00	2	1	8
	CA $U^{ref} = 0.10$	3132	9.8	0.258	0.259	2665	11.9	1.00	1	2	10

										Rank	Statistic
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
11-20	Data ($\lambda_2 = 120$)	1694	8.9	0.336	0.371	1262	15.0	1.00	12	3	1
	Data ($\lambda_2 = 150$)	1981	9.3	0.325	0.350	1466	14.5	1.00	10	5	3
	Data ($\lambda_2 = 180$)	2176	9.7	0.309	0.328	1600	13.7	1.00	9	6	4
	Data ($\lambda_2 = 210$)	2252	10.2	0.293	0.306	1635	12.9	1.00	8	7	7
	Data ($\lambda_2 = 240$)	2263	10.4	0.271	0.286	1672	12.2	1.00	7	8	11
	Prod $U^{ref} = 0.06$	2322	10.9	0.309	0.323	1812	13.8	1.00	6	10	5
	Prod $U^{ref} = 0.08$	2546	11.4	0.291	0.301	2004	13.1	1.00	5	11	8
	Prod $U^{ref} = 0.10$	2689	12.0	0.278	0.283	2063	12.5	0.85	3	12	9
	CA $U^{ref} = 0.04$	1978	6.3	0.335	0.349	1680	14.7	1.00	11	1	2
	CA $U^{ref} = 0.06$	2586	7.8	0.298	0.306	2171	13.5	1.00	4	2	6
	CA $U^{ref} = 0.08$	2997	9.2	0.272	0.272	2487	12.3	1.00	2	4	10
	CA $U^{ref} = 0.10$	3272	10.6	0.247	0.239	2616	11.3	1.00	1	9	12
21-40	Data ($\lambda_2 = 120$)	2377	6.9	0.438	0.454	1741	19.5	1.00	12	3	1
	Data ($\lambda_2 = 150$)	2728	7.3	0.400	0.407	1991	18.0	1.00	9	4	3
	Data ($\lambda_2 = 180$)	2925	7.9	0.367	0.367	2090	16.6	1.00	7	6	5
	Data ($\lambda_2 = 210$)	2988	8.9	0.333	0.343	2007	15.3	1.00	5	7	8
	Data ($\lambda_2 = 240$)	2961	9.4	0.310	0.317	1928	14.1	1.00	6	9	10
	Prod $U^{ref} = 0.06$	2722	10.2	0.380	0.380	2245	16.9	1.00	10	10	4
	Prod $U^{ref} = 0.08$	2922	10.8	0.347	0.343	2340	15.5	1.00	8	11	6
	Prod $U^{ref} = 0.10$	3016	11.9	0.321	0.317	2320	14.5	1.00	4	12	9
	CA $U^{ref} = 0.04$	2469	4.5	0.416	0.437	2043	18.6	1.00	11	1	2
	CA $U^{ref} = 0.06$	3081	6.1	0.342	0.335	2500	15.5	1.00	3	2	7
	CA $U^{ref} = 0.08$	3339	7.6	0.286	0.269	2653	13.2	1.00	1	5	11
	CA $U^{ref} = 0.10$	3332	9.2	0.242	0.222	2672	11.2	0.80	2	8	12

Table F-3 Summary of performance statistics and ranks by management procedure for Scenario 3 $\left\{h = 0.65,\ \hat{q}_{2,trap}\right\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	Statistic
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2856	16.1	0.197	0.201	2081	13.7	0.40	9	12	4
	Data ($\lambda_2 = 150$)	2886	15.3	0.197	0.200	2153	13.7	0.40	8	8	7
	Data ($\lambda_2 = 180$)	3178	13.7	0.193	0.193	2433	13.6	0.20	4	5	10
	Data ($\lambda_2 = 210$)	3537	12.6	0.189	0.185	2750	13.5	0.20	2	3	11
	Data ($\lambda_2 = 240$)	3889	11.6	0.185	0.177	3038	13.4	0.20	1	2	12
	Prod $U^{ref} = 0.06$	2905	14.2	0.197	0.200	2253	13.7	0.00	7	6	5
	Prod $U^{ref} = 0.08$	3042	12.7	0.196	0.198	2484	13.7	0.00	5	4	8
	Prod $U^{ref} = 0.10$	3217	11.5	0.195	0.195	2716	13.6	0.00	3	1	9
	CA $U^{ref} = 0.04$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.06$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.08$	2856	16.1	0.197	0.201	2081	13.7	0.40	11	10	2
	CA $U^{ref} = 0.10$	2912	14.9	0.197	0.200	2266	13.7	0.40	6	7	6
6-10	Data ($\lambda_2 = 120$)	1544	11.0	0.247	0.279	1375	15.1	1.00	11	7	3
	Data ($\lambda_2 = 150$)	1765	10.3	0.241	0.273	1569	15.0	1.00	9	4	5
	Data ($\lambda_2 = 180$)	1989	10.7	0.231	0.259	1758	14.7	1.00	8	5	9
	Data ($\lambda_2 = 210$)	2193	11.5	0.219	0.243	1902	14.4	0.80	7	9	11
	Data ($\lambda_2 = 240$)	2376	11.2	0.206	0.227	2010	14.1	0.60	4	8	12
	Prod $U^{ref} = 0.06$	2260	9.9	0.237	0.265	1922	14.9	0.20	5	3	6
	Prod $U^{ref} = 0.08$	2553	9.9	0.231	0.256	2240	14.7	0.20	3	2	8
	Prod $U^{ref} = 0.10$	2754	9.8	0.225	0.245	2448	14.5	0.00	1	1	10
	CA $U^{ref} = 0.04$	1198	13.0	0.253	0.290	990	15.2	1.00	12	12	1
	CA $U^{ref} = 0.06$	1724	11.7	0.247	0.278	1466	15.1	1.00	10	11	2
	CA $U^{ref} = 0.08$	2208	11.0	0.241	0.268	1880	15.0	1.00	6	6	4
	CA $U^{ref} = 0.10$	2586	11.5	0.234	0.257	2245	14.8	1.00	2	10	7

									Rank Statistic		
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
11-20	Data ($\lambda_2 = 120$)	2357	7.5	0.348	0.400	1767	17.0	1.00	11	3	2
	Data ($\lambda_2 = 150$)	2796	7.7	0.325	0.370	2097	16.7	1.00	9	4	4
	Data ($\lambda_2 = 180$)	3117	8.2	0.303	0.341	2322	16.3	1.00	7	7	7
	Data ($\lambda_2 = 210$)	3298	9.0	0.284	0.311	2463	15.9	1.00	4	10	9
	Data ($\lambda_2 = 240$)	3442	9.5	0.266	0.291	2523	15.5	1.00	2	12	12
	Prod $U^{ref} = 0.06$	2947	8.5	0.315	0.354	2359	16.5	1.00	8	9	5
	Prod $U^{ref} = 0.08$	3215	8.1	0.294	0.330	2616	16.1	0.95	5	6	8
	Prod $U^{ref} = 0.10$	3382	8.3	0.277	0.309	2810	15.8	0.80	3	8	11
	CA $U^{ref} = 0.04$	2000	5.8	0.371	0.423	1668	17.4	1.00	12	1	1
	CA $U^{ref} = 0.06$	2688	6.7	0.340	0.377	2237	16.8	1.00	10	2	3
	CA $U^{ref} = 0.08$	3161	7.9	0.309	0.345	2596	16.2	1.00	6	5	6
	CA $U^{ref} = 0.10$	3493	9.1	0.284	0.315	2874	15.7	1.00	1	11	10
21-40	Data ($\lambda_2 = 120$)	3405	6.0	0.483	0.510	2717	19.4	1.00	11	4	2
	Data ($\lambda_2 = 150$)	3891	6.1	0.435	0.450	3169	18.7	1.00	7	5	5
	Data ($\lambda_2 = 180$)	4218	6.2	0.392	0.398	3434	18.0	1.00	4	6	7
	Data ($\lambda_2 = 210$)	4462	6.5	0.351	0.345	3663	17.3	1.00	3	7	10
	Data ($\lambda_2 = 240$)	4643	7.0	0.317	0.305	3643	16.6	1.00	1	8	12
	Prod $U^{ref} = 0.06$	3547	8.2	0.436	0.455	3047	18.7	1.00	10	10	4
	Prod $U^{ref} = 0.08$	3806	8.2	0.409	0.419	3285	18.2	1.00	8	11	6
	Prod $U^{ref} = 0.10$	3969	8.5	0.387	0.390	3427	17.8	1.00	6	12	8
	CA $U^{ref} = 0.04$	2747	3.8	0.528	0.565	2279	20.2	1.00	12	1	1
	CA $U^{ref} = 0.06$	3576	4.6	0.453	0.471	3050	19.0	1.00	9	2	3
	CA $U^{ref} = 0.08$	4179	5.8	0.386	0.395	3460	18.0	1.00	5	3	9
	CA $U^{ref} = 0.10$	4534	7.2	0.337	0.339	3674	17.1	1.00	2	9	11

Table F-4 Summary of performance statistics and ranks by management procedure for Scenario 4 $\{h = 0.65,\ q_{2,trap} = 1\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

										Rank	**Statistic**
Years	**Management Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 120$)	2867	15.6	0.307	0.311	2123	13.6	1.00	10	10	3
	Data ($\lambda_2 = 150$)	2972	12.9	0.307	0.309	2373	13.6	1.00	9	7	6
	Data ($\lambda_2 = 180$)	3332	11.2	0.303	0.301	2701	13.5	1.00	4	4	10
	Data ($\lambda_2 = 210$)	3719	9.7	0.299	0.294	3101	13.3	1.00	2	2	11
	Data ($\lambda_2 = 240$)	4104	8.4	0.296	0.285	3499	13.2	1.00	1	1	12
	Prod $U^{ref} = 0.06$	2992	13.4	0.307	0.309	2432	13.6	1.00	8	9	5
	Prod $U^{ref} = 0.08$	3197	11.5	0.306	0.306	2730	13.6	1.00	6	5	8
	Prod $U^{ref} = 0.10$	3375	10.0	0.305	0.303	2957	13.5	1.00	3	3	9
	CA $U^{ref} = 0.04$	2856	16.1	0.307	0.312	2081	13.6	1.00	11.5	11.5	1.5
	CA $U^{ref} = 0.06$	2856	16.1	0.307	0.311	2081	13.6	1.00	11.5	11.5	1.5
	CA $U^{ref} = 0.08$	2999	13.0	0.307	0.310	2456	13.6	1.00	7	8	4
	CA $U^{ref} = 0.10$	3242	12.5	0.306	0.307	2850	13.6	1.00	5	6	7
6-10	Data ($\lambda_2 = 120$)	1721	9.8	0.357	0.391	1556	16.2	1.00	11	4	2
	Data ($\lambda_2 = 150$)	2068	10.1	0.350	0.383	1842	16.0	1.00	10	6	4
	Data ($\lambda_2 = 180$)	2367	10.3	0.337	0.368	2068	15.5	1.00	9	8	8
	Data ($\lambda_2 = 210$)	2655	10.5	0.326	0.350	2284	14.9	1.00	6	11	11
	Data ($\lambda_2 = 240$)	2928	10.8	0.312	0.333	2502	14.4	1.00	4	12	12
	Prod $U^{ref} = 0.06$	2486	9.6	0.346	0.374	2226	15.8	1.00	7	3	5
	Prod $U^{ref} = 0.08$	2805	9.1	0.339	0.364	2523	15.5	1.00	5	2	7
	Prod $U^{ref} = 0.10$	3035	8.7	0.333	0.354	2752	15.1	1.00	2	1	9
	CA $U^{ref} = 0.04$	1671	10.5	0.361	0.394	1426	16.4	1.00	12	10	1
	CA $U^{ref} = 0.06$	2401	9.9	0.353	0.378	2112	16.0	1.00	8	5	3
	CA $U^{ref} = 0.08$	3029	10.4	0.342	0.362	2638	15.6	1.00	3	9	6
	CA $U^{ref} = 0.10$	3575	10.2	0.332	0.345	3100	15.1	1.00	1	7	10

Years	Management Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Rank Statistic		
									Avg. Catch	AAV Catch	Avg. Depletion
11-20	Data ($\lambda_2 = 120$)	2322	7.1	0.459	0.506	1861	19.8	1.00	12	3	2
	Data ($\lambda_2 = 150$)	2787	7.2	0.433	0.472	2230	19.0	1.00	10	4	3
	Data ($\lambda_2 = 180$)	3160	7.5	0.406	0.442	2510	18.1	1.00	8	5	6
	Data ($\lambda_2 = 210$)	3485	8.1	0.384	0.415	2716	17.2	1.00	4	7	9
	Data ($\lambda_2 = 240$)	3744	8.7	0.364	0.391	2804	16.3	1.00	3	10	11
	Prod $U^{ref} = 0.06$	2992	8.7	0.425	0.455	2466	18.5	1.00	9	11	4
	Prod $U^{ref} = 0.08$	3270	8.5	0.404	0.432	2733	17.8	1.00	7	9	7
	Prod $U^{ref} = 0.10$	3466	8.4	0.388	0.414	2936	17.2	1.00	5	8	8
	CA $U^{ref} = 0.04$	2466	5.4	0.462	0.497	2135	19.7	1.00	11	1	1
	CA $U^{ref} = 0.06$	3327	6.9	0.421	0.441	2781	18.3	1.00	6	2	5
	CA $U^{ref} = 0.08$	3961	8.1	0.381	0.401	3281	16.9	1.00	2	6	10
	CA $U^{ref} = 0.10$	4409	9.6	0.350	0.362	3639	15.7	1.00	1	12	12
21-40	Data ($\lambda_2 = 120$)	3015	6.3	0.580	0.603	2500	25.1	1.00	12	3	1
	Data ($\lambda_2 = 150$)	3544	6.4	0.536	0.550	2904	23.4	1.00	9	5	3
	Data ($\lambda_2 = 180$)	3975	6.5	0.496	0.503	3269	21.7	1.00	6	6	6
	Data ($\lambda_2 = 210$)	4311	6.8	0.458	0.459	3511	20.3	1.00	4	7	9
	Data ($\lambda_2 = 240$)	4539	7.1	0.425	0.421	3659	19.0	1.00	3	8	10
	Prod $U^{ref} = 0.06$	3423	8.6	0.532	0.545	2944	23.1	1.00	10	12	4
	Prod $U^{ref} = 0.08$	3698	8.5	0.506	0.509	3189	22.1	1.00	8	10	5
	Prod $U^{ref} = 0.10$	3893	8.6	0.486	0.485	3343	21.2	1.00	7	11	8
	CA $U^{ref} = 0.04$	3093	4.0	0.572	0.583	2646	24.7	1.00	11	1	2
	CA $U^{ref} = 0.06$	4088	4.9	0.488	0.482	3502	21.3	1.00	5	2	7
	CA $U^{ref} = 0.08$	4746	6.3	0.417	0.405	4019	18.7	1.00	2	4	11
	CA $U^{ref} = 0.10$	5122	7.6	0.358	0.345	4270	16.4	1.00	1	9	12

Table F-5 Summary of performance statistics and ranks by management procedure for Scenario 1 $\left\{h = 0.45,\ \hat{q}_{2,trap}\right\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	**Procedure**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**	**Final Depletion**	**$C_{5\%}$**	**Avg. CPUE**	**P_{cons}**	**Avg. Catch**	**AAV Catch**	**Avg. Depletion**
1-5	Data ($\lambda_2 = 150$)	2856	16.1	0.151	0.138	2081	12.4	0.00	7	10	4
	Data ($\lambda_2 = 180$)	3067	16.1	0.148	0.132	2234	12.3	0.00	3	7	8
	Data ($\lambda_2 = 210$)	3366	15.3	0.144	0.125	2473	12.2	0.00	2	3	9
	Data ($\lambda_2 = 240$)	3679	14.5	0.140	0.119	2722	12.1	0.00	1	2	10
	Prod $U^{ref} = 0.06$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	8.5	5
	Prod $U^{ref} = 0.08$	2856	16.1	0.151	0.138	2081	12.4	0.00	5.5	8.5	6
	Prod $U^{ref} = 0.10$	2943	14.0	0.150	0.136	2272	12.4	0.00	4	1	7
	CA $U^{ref} = 0.06$	2856	16.1	0.151	0.139	2081	12.4	0.00	9	5	1.5
	CA $U^{ref} = 0.08$	2856	16.1	0.151	0.139	2081	12.4	0.00	9	5	1.5
	CA $U^{ref} = 0.10$	2856	16.1	0.151	0.138	2081	12.4	0.00	9	5	3
1-10	Data ($\lambda_2 = 150$)	2083	15.6	0.152	0.162	1048	12.5	0.00	8	5	4
	Data ($\lambda_2 = 180$)	2262	15.2	0.146	0.152	1158	12.3	0.00	5	3	8
	Data ($\lambda_2 = 210$)	2473	14.6	0.139	0.139	1234	12.0	0.00	2	2	9
	Data ($\lambda_2 = 240$)	2697	14.4	0.132	0.127	1332	11.8	0.00	1	1	10
	Prod $U^{ref} = 0.06$	2064	18.9	0.153	0.166	1025	12.5	0.00	9	10	2
	Prod $U^{ref} = 0.08$	2197	17.3	0.151	0.160	1207	12.5	0.00	6	8	5
	Prod $U^{ref} = 0.10$	2324	16.3	0.149	0.154	1336	12.3	0.00	3	6	7
	CA $U^{ref} = 0.06$	2005	18.5	0.154	0.168	1007	12.6	0.00	10	9	1
	CA $U^{ref} = 0.08$	2163	16.6	0.152	0.162	1274	12.5	0.00	7	7	3
	CA $U^{ref} = 0.10$	2308	15.2	0.150	0.155	1481	12.4	0.00	4	4	6
1-20	Data ($\lambda_2 = 150$)	1681	13.2	0.173	0.218	901	13.1	0.23	10	5	2
	Data ($\lambda_2 = 180$)	1791	13.2	0.164	0.203	940	12.8	0.15	9	6	5
	Data ($\lambda_2 = 210$)	1885	13.2	0.154	0.188	924	12.5	0.00	6	4	9
	Data ($\lambda_2 = 240$)	1966	13.2	0.145	0.173	870	12.2	0.00	5	3	10

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
	Prod $U^{ref}=0.06$	1880	16.2	0.170	0.199	1034	13.0	0.00	7	9	3
	Prod $U^{ref}=0.08$	2038	15.9	0.163	0.181	1186	12.7	0.00	3	8	6
	Prod $U^{ref}=0.10$	2162	16.5	0.156	0.169	1260	12.6	0.00	2	10	7
	CA $U^{ref}=0.06$	1811	13.4	0.174	0.205	1008	13.1	0.20	8	7	1
	CA $U^{ref}=0.08$	2022	12.8	0.165	0.185	1262	12.8	0.00	4	2	4
	CA $U^{ref}=0.10$	2177	12.6	0.156	0.168	1447	12.6	0.00	1	1	8
1-40	Data ($\lambda_2=150$)	1994	10.6	0.222	0.280	910	14.4	0.59	9	3	1
	Data ($\lambda_2=180$)	2106	11.0	0.206	0.254	939	14.0	0.50	7	5	4
	Data ($\lambda_2=210$)	2185	11.5	0.190	0.225	907	13.5	0.38	4	6	6
	Data ($\lambda_2=240$)	2174	11.8	0.174	0.204	858	13.1	0.25	5	7	9
	Prod $U^{ref}=0.06$	2041	13.8	0.207	0.258	1098	14.1	0.21	8	8	3
	Prod $U^{ref}=0.08$	2190	14.7	0.188	0.225	1223	13.6	0.03	3	9	7
	Prod $U^{ref}=0.10$	2258	15.7	0.174	0.198	1263	13.2	0.00	2	10	10
	CA $U^{ref}=0.06$	1984	9.5	0.219	0.279	1078	14.3	0.53	10	1	2
	CA $U^{ref}=0.08$	2172	10.0	0.196	0.227	1296	13.7	0.39	6	2	5
	CA $U^{ref}=0.10$	2271	10.9	0.177	0.192	1359	13.2	0.20	1	4	8

Table F-6 Summary of performance statistics and ranks by management procedure for Scenario 2 $\{h = 0.45,\ q_{2,trap} = 1\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
1-5	Data ($\lambda_2 = 150$)	2892	15.3	0.271	0.260	2146	12.0	1.00	9	9	4
	Data ($\lambda_2 = 180$)	3208	13.5	0.267	0.253	2466	11.9	1.00	5	7	8
	Data ($\lambda_2 = 210$)	3579	12.2	0.264	0.245	2783	11.8	1.00	2	3	9
	Data ($\lambda_2 = 240$)	3948	11.1	0.260	0.237	3094	11.6	1.00	1	1	10
	Prod $U^{ref} = 0.06$	2926	14.2	0.271	0.260	2256	12.0	1.00	8	8	2
	Prod $U^{ref} = 0.08$	3059	12.6	0.270	0.257	2509	12.0	1.00	6	5	5
	Prod $U^{ref} = 0.10$	3236	11.6	0.270	0.255	2746	12.0	1.00	3	2	7
	CA $U^{ref} = 0.06$	2856	16.1	0.271	0.261	2081	12.0	1.00	10	10	1
	CA $U^{ref} = 0.08$	2988	13.3	0.271	0.259	2436	12.0	1.00	7	6	3
	CA $U^{ref} = 0.10$	3231	12.5	0.270	0.257	2828	12.0	1.00	4	4	6
1-10	Data ($\lambda_2 = 150$)	2320	12.9	0.276	0.294	1435	12.4	1.00	10	9	1
	Data ($\lambda_2 = 180$)	2592	12.5	0.268	0.281	1634	12.1	1.00	7	8	6
	Data ($\lambda_2 = 210$)	2885	11.7	0.261	0.266	1830	11.8	1.00	3	4	9
	Data ($\lambda_2 = 240$)	3199	11.2	0.252	0.252	1994	11.4	1.00	1	1	10
	Prod $U^{ref} = 0.06$	2467	13.1	0.274	0.290	1755	12.3	1.00	9	10	3
	Prod $U^{ref} = 0.08$	2681	12.2	0.271	0.280	2006	12.2	1.00	6	6	4
	Prod $U^{ref} = 0.10$	2837	11.9	0.267	0.271	2170	12.0	0.90	5	5	7
	CA $U^{ref} = 0.06$	2496	12.3	0.276	0.288	1900	12.4	1.00	8	7	2
	CA $U^{ref} = 0.08$	2840	11.7	0.270	0.273	2293	12.2	1.00	4	3	5
	CA $U^{ref} = 0.10$	3173	11.3	0.266	0.259	2548	11.9	1.00	2	2	8
1-20	Data ($\lambda_2 = 150$)	2151	11.4	0.305	0.350	1382	13.6	1.00	10	7	1
	Data ($\lambda_2 = 180$)	2379	11.1	0.293	0.328	1557	13.0	1.00	9	5	3
	Data ($\lambda_2 = 210$)	2584	10.9	0.278	0.306	1633	12.6	1.00	6	3	6
	Data ($\lambda_2 = 240$)	2776	11.0	0.264	0.286	1671	12.0	1.00	3	4	9

Time Horizon (years)	Management Procedure			Median	Statistic					Rank	Statistic
		Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
	Prod $U^{ref} = 0.06$	2385	12.6	0.293	0.323	1656	13.3	1.00	8	10	2
	Prod $U^{ref} = 0.08$	2600	12.0	0.284	0.301	1849	12.8	0.95	5	8	5
	Prod $U^{ref} = 0.10$	2765	12.2	0.277	0.283	1990	12.5	0.85	4	9	7
	CA $U^{ref} = 0.06$	2573	10.1	0.290	0.306	1897	13.2	1.00	7	1	4
	CA $U^{ref} = 0.08$	2965	10.4	0.274	0.272	2225	12.5	1.00	2	2	8
	CA $U^{ref} = 0.10$	3274	11.3	0.258	0.239	2444	11.8	1.00	1	6	10
1-40	Data ($\lambda_2 = 150$)	2397	8.8	0.356	0.407	1348	16.0	1.00	10	2	1
	Data ($\lambda_2 = 180$)	2643	9.3	0.333	0.367	1407	14.9	1.00	8	4	3
	Data ($\lambda_2 = 210$)	2799	9.6	0.310	0.343	1479	14.0	1.00	5	5	6
	Data ($\lambda_2 = 240$)	2906	10.0	0.291	0.317	1560	13.2	1.00	3	6	8
	Prod $U^{ref} = 0.06$	2482	11.3	0.338	0.380	1699	15.4	1.00	9	8	2
	Prod $U^{ref} = 0.08$	2694	11.6	0.318	0.343	1833	14.4	0.94	7	9	5
	Prod $U^{ref} = 0.10$	2847	12.1	0.301	0.317	1947	13.7	0.85	4	10	7
	CA $U^{ref} = 0.06$	2770	7.8	0.318	0.335	1912	14.4	1.00	6	1	4
	CA $U^{ref} = 0.08$	3120	8.9	0.280	0.269	2210	12.8	1.00	2	3	9
	CA $U^{ref} = 0.10$	3304	10.2	0.250	0.222	2337	11.5	0.88	1	7	10

Table F-7 Summary of performance statistics and ranks by management procedure for Scenario 3 $\left\{h = 0.65,\ \hat{q}_{2,trap}\right\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
1-5	Data ($\lambda_2 = 150$)	2886	15.3	0.197	0.200	2153	13.7	0.40	8	8	5
	Data ($\lambda_2 = 180$)	3178	13.7	0.193	0.193	2433	13.6	0.20	4	5	8
	Data ($\lambda_2 = 210$)	3537	12.6	0.189	0.185	2750	13.5	0.20	2	3	9
	Data ($\lambda_2 = 240$)	3889	11.6	0.185	0.177	3038	13.4	0.20	1	2	10
	Prod $u_{ref} = 0.06$	2905	14.2	0.197	0.200	2253	13.7	0.00	7	6	3
	Prod $u_{ref} = 0.08$	3042	12.7	0.196	0.198	2484	13.7	0.00	5	4	6
	Prod $u_{ref} = 0.10$	3217	11.5	0.195	0.195	2716	13.6	0.00	3	1	7
	CA $u_{ref} = 0.06$	2856	16.1	0.197	0.201	2081	13.7	0.40	9.5	9.5	1.5
	CA $u_{ref} = 0.08$	2856	16.1	0.197	0.201	2081	13.7	0.40	9.5	9.5	1.5
	CA $u_{ref} = 0.10$	2912	14.9	0.197	0.200	2266	13.7	0.40	6	7	4
1-10	Data ($\lambda_2 = 150$)	2371	12.8	0.220	0.273	1583	14.4	0.70	9	8	3
	Data ($\lambda_2 = 180$)	2679	12.2	0.213	0.259	1767	14.1	0.60	6	6	7
	Data ($\lambda_2 = 210$)	2945	11.7	0.205	0.243	1925	13.9	0.50	3	4	9
	Data ($\lambda_2 = 240$)	3189	11.4	0.197	0.227	2057	13.7	0.40	1	3	10
	Prod $U^{ref} = 0.06$	2596	12.1	0.218	0.265	1994	14.3	0.10	7	5	4
	Prod $U^{ref} = 0.08$	2796	10.7	0.214	0.256	2265	14.2	0.10	4	2	6
	Prod $U^{ref} = 0.10$	2961	10.3	0.210	0.245	2456	14.1	0.00	2	1	8
	CA $U^{ref} = 0.06$	2290	15.5	0.223	0.278	1492	14.4	0.70	10	10	1
	CA $U^{ref} = 0.08$	2536	13.1	0.220	0.268	1915	14.4	0.70	8	9	2
	CA $U^{ref} = 0.10$	2784	12.5	0.217	0.257	2164	14.2	0.70	5	7	5
1-20	Data ($\lambda_2 = 150$)	2599	9.9	0.278	0.370	1611	15.6	0.85	9	5	2
	Data ($\lambda_2 = 180$)	2915	9.7	0.263	0.341	1800	15.2	0.80	6	2	5
	Data ($\lambda_2 = 210$)	3188	9.8	0.249	0.311	1983	14.9	0.70	4	3	8
	Data ($\lambda_2 = 240$)	3429	10.0	0.234	0.291	2133	14.6	0.65	1	6	10

Time Horizon (years)	Management Procedure	Median Statistic							Rank Statistic		
		Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
	Prod $U^{ref}=0.06$	2745	10.6	0.270	0.354	2032	15.5	0.55	8	9	3
	Prod $U^{ref}=0.08$	2998	9.8	0.258	0.330	2286	15.2	0.50	5	4	6
	Prod $U^{ref}=0.10$	3194	9.5	0.248	0.309	2413	15.0	0.48	3	1	9
	CA $U^{ref}=0.06$	2524	10.6	0.282	0.377	1551	15.8	0.85	10	10	1
	CA $U^{ref}=0.08$	2911	10.1	0.267	0.345	1969	15.5	0.85	7	7	4
	CA $U^{ref}=0.10$	3207	10.5	0.254	0.315	2230	15.2	0.80	2	8	7
1-40	Data ($\lambda_2=150$)	3219	7.5	0.359	0.450	1727	17.3	0.93	8	2	3
	Data ($\lambda_2=180$)	3569	7.5	0.329	0.398	1949	16.7	0.90	4	3	6
	Data ($\lambda_2=210$)	3848	7.9	0.300	0.345	2097	16.2	0.85	3	5	9
	Data ($\lambda_2=240$)	4081	8.3	0.277	0.305	2167	15.6	0.83	1	6	10
	Prod $U^{ref}=0.06$	3140	8.7	0.361	0.455	2140	17.2	0.78	9	10	2
	Prod $U^{ref}=0.08$	3388	8.5	0.336	0.419	2387	16.8	0.74	7	8	4
	Prod $U^{ref}=0.10$	3564	8.5	0.317	0.390	2580	16.5	0.71	5	7	7
	CA $U^{ref}=0.06$	3032	7.1	0.372	0.471	1677	17.5	0.93	10	1	1
	CA $U^{ref}=0.08$	3535	7.7	0.333	0.395	2010	16.9	0.93	6	4	5
	CA $U^{ref}=0.10$	3905	8.6	0.302	0.339	2355	16.3	0.90	2	9	8

Table F-8 Summary of performance statistics and ranks by management procedure for Scenario 4 $\left\{h = 0.65,\ q_{2,trap} = 1\right\}$. Table values represent the median performance statistic or rank over 50 replicates of the projection period within the procedure and scenario.

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
1-5	Data ($\lambda_2 = 150$)	2972	12.9	0.307	0.309	2373	13.6	1.00	9	7	4
	Data ($\lambda_2 = 180$)	3332	11.2	0.303	0.301	2701	13.5	1.00	4	4	8
	Data ($\lambda_2 = 210$)	3719	9.7	0.299	0.294	3101	13.3	1.00	2	2	9
	Data ($\lambda_2 = 240$)	4104	8.4	0.296	0.285	3499	13.2	1.00	1	1	10
	Prod $U^{ref} = 0.06$	2992	13.4	0.307	0.309	2432	13.6	1.00	8	9	3
	Prod $U^{ref} = 0.08$	3197	11.5	0.306	0.306	2730	13.6	1.00	6	5	6
	Prod $U^{ref} = 0.10$	3375	10.0	0.305	0.303	2957	13.5	1.00	3	3	7
	CA $U^{ref} = 0.06$	2856	16.1	0.307	0.311	2081	13.6	1.00	10	10	1
	CA $U^{ref} = 0.08$	2999	13.0	0.307	0.310	2456	13.6	1.00	7	8	2
	CA $U^{ref} = 0.10$	3242	12.5	0.306	0.307	2850	13.6	1.00	5	6	5
1-10	Data ($\lambda_2 = 150$)	2557	11.2	0.330	0.383	1854	14.7	1.00	10	7	2
	Data ($\lambda_2 = 180$)	2894	10.3	0.322	0.368	2102	14.4	1.00	7	5	6
	Data ($\lambda_2 = 210$)	3241	9.8	0.313	0.350	2320	14.1	1.00	3	3	9
	Data ($\lambda_2 = 240$)	3573	9.6	0.304	0.333	2526	13.7	1.00	1	2	10
	Prod $U^{ref} = 0.06$	2744	10.8	0.328	0.374	2241	14.7	1.00	8	6	3
	Prod $U^{ref} = 0.08$	2971	10.2	0.323	0.364	2516	14.5	1.00	6	4	5
	Prod $U^{ref} = 0.10$	3168	9.1	0.320	0.354	2749	14.3	1.00	4	1	8
	CA $U^{ref} = 0.06$	2640	12.1	0.332	0.378	2083	14.8	1.00	9	10	1
	CA $U^{ref} = 0.08$	3038	11.3	0.326	0.362	2458	14.6	1.00	5	8	4
	CA $U^{ref} = 0.10$	3423	11.6	0.320	0.345	2777	14.4	1.00	2	9	7
1-20	Data ($\lambda_2 = 150$)	2685	9.1	0.386	0.472	1870	17.1	1.00	10	4	1
	Data ($\lambda_2 = 180$)	3047	9.1	0.373	0.442	2110	16.4	1.00	7	5	4
	Data ($\lambda_2 = 210$)	3385	9.0	0.357	0.415	2347	15.8	1.00	4	1	7
	Data ($\lambda_2 = 240$)	3697	9.0	0.340	0.391	2553	15.2	1.00	2	2	9

Time	Management			Median	Statistic					Rank	Statistic
Horizon (years)	Procedure	Avg. Catch	AAV Catch	Avg. Depletion	Final Depletion	$C_{5\%}$	Avg. CPUE	P_{cons}	Avg. Catch	AAV Catch	Avg. Depletion
	Prod $U^{ref} = 0.06$	2861	9.9	0.379	0.455	2230	16.8	1.00	9	9	2
	Prod $U^{ref} = 0.08$	3148	9.3	0.367	0.432	2488	16.4	1.00	6	6	5
	Prod $U^{ref} = 0.10$	3361	9.1	0.358	0.414	2697	15.9	1.00	5	3	6
	CA $U^{ref} = 0.06$	3016	9.3	0.377	0.441	2107	16.9	1.00	8	7	3
	CA $U^{ref} = 0.08$	3566	9.7	0.357	0.401	2521	16.1	1.00	3	8	8
	CA $U^{ref} = 0.10$	4000	10.4	0.339	0.362	2782	15.4	1.00	1	10	10
1-40	Data ($\lambda_2 = 150$)	3062	7.3	0.465	0.550	1970	20.6	1.00	10	2	1
	Data ($\lambda_2 = 180$)	3459	7.4	0.435	0.503	2236	19.4	1.00	7	3	5
	Data ($\lambda_2 = 210$)	3808	7.5	0.410	0.459	2325	18.2	1.00	4	4	7
	Data ($\lambda_2 = 240$)	4107	7.7	0.384	0.421	2455	17.3	1.00	3	5	9
	Prod $U^{ref} = 0.06$	3125	8.9	0.463	0.545	2324	20.4	1.00	9	9	2
	Prod $U^{ref} = 0.08$	3393	8.6	0.442	0.509	2575	19.6	1.00	8	8	3
	Prod $U^{ref} = 0.10$	3584	8.4	0.426	0.485	2754	19.0	1.00	5	7	6
	CA $U^{ref} = 0.06$	3522	6.9	0.438	0.482	2223	19.4	1.00	6	1	4
	CA $U^{ref} = 0.08$	4139	7.8	0.393	0.405	2657	17.7	1.00	2	6	8
	CA $U^{ref} = 0.10$	4586	9.1	0.357	0.345	2941	16.3	1.00	1	10	10

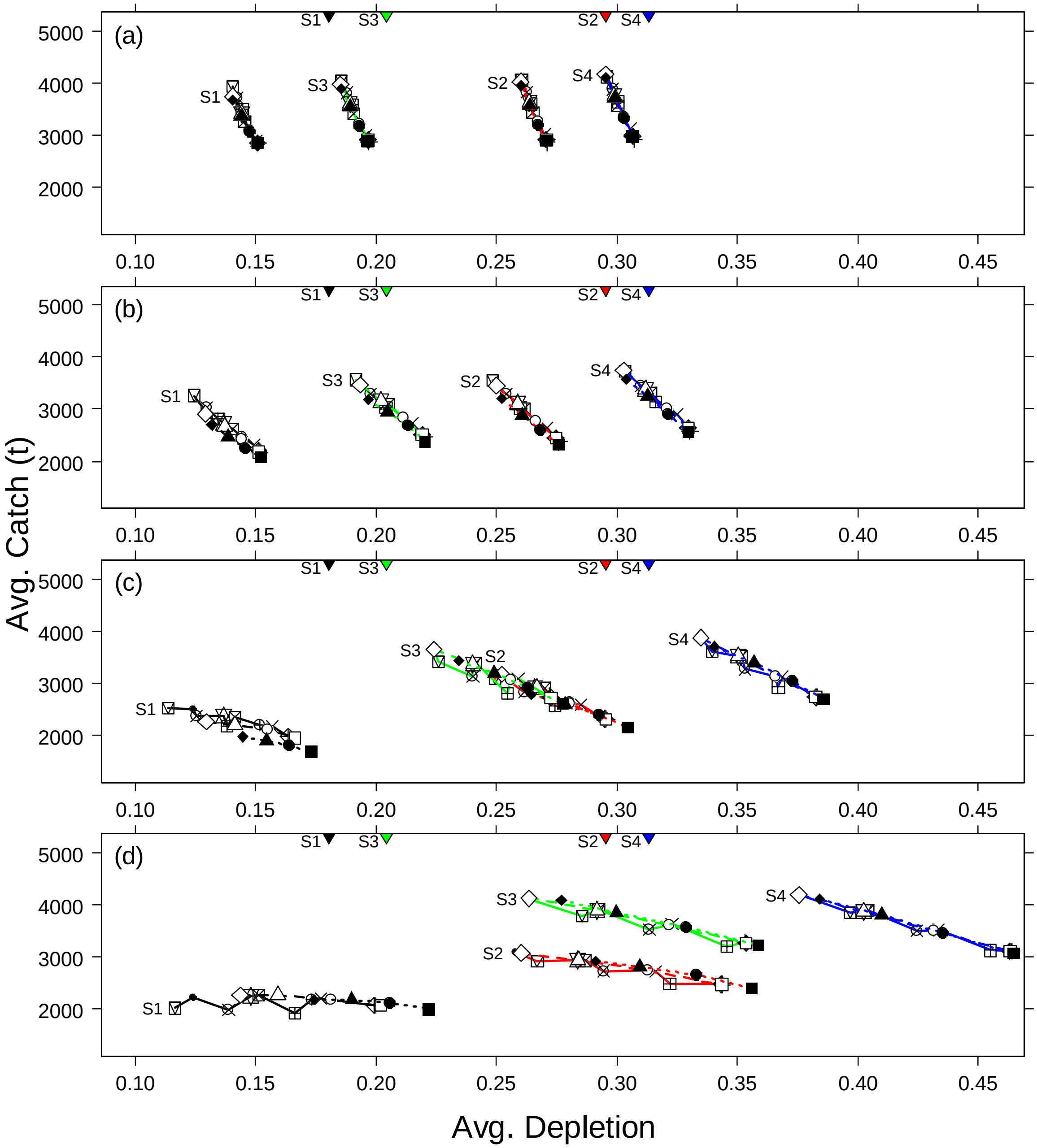


Figure F-1 Trade-off relationship between median average catch (t) and median average depletion for data-based procedures for scenarios S1-S4 and 5, 10, 20 and 40 year time horizons (panels a-d). Results show CHR procedures for $\left(\lambda_1 = 0.2,\ \lambda_2 = \{210,180,150\}\right)$ (⊠,×,+), $\left(\lambda_1 = 0.5,\ \lambda_2 = \{240,\ 210,\ 180,\ 150\}\right)$ (•, star,⊕, ⊞), and $\left(\lambda_1 = 0.8,\ \lambda_2 = \{210,180,150\}\right)$ (square+triangle,⊗, ⊞). Results for VHR procedures correspond to $\lambda_2 = \{240,\ 210,\ 180,\ 150\}$ with $\left\{\lambda_1 = 0.5,\ I_{low} = 3,\ I_{high} = 10\right\}$ (◇, △, ○, □) and $\left\{\lambda_1 = 0.5,\ I_{low} = 4,\ I_{high} = 15\right\}$ (♦, ▲, ●, ■). Initial depletion values for each scenario are indicated by inverted triangles along the upper *x*-axis.

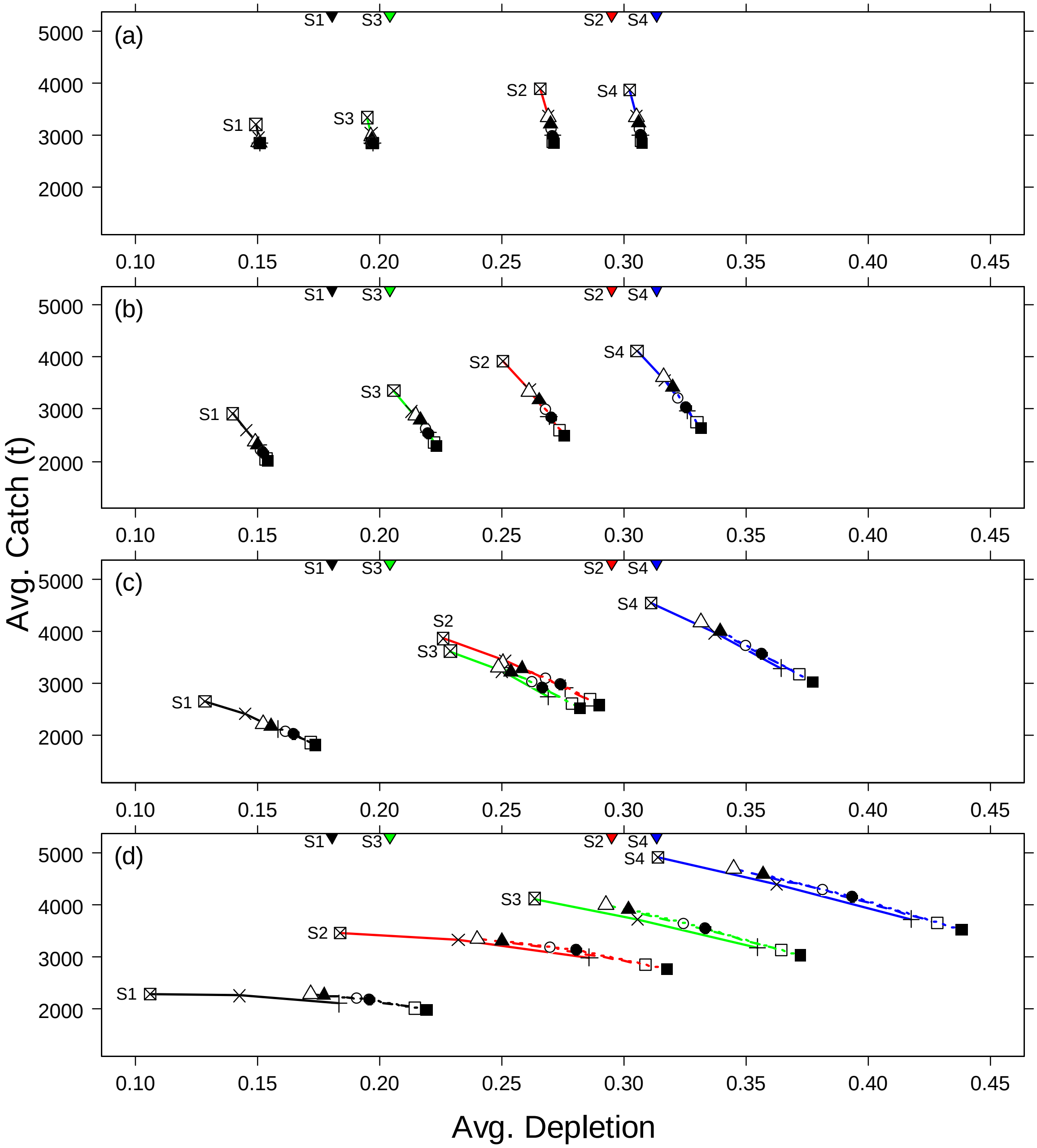


Figure F-2 Trade-off relationship between median average catch (t) and median average depletion for nine catch-age model-based procedures for scenarios S1-S4 and 5, 10, 20 and 40 year time horizons (panels a-d). Procedures are ordered from left to right with $U^{ref} = \{0.10, 0.08, 0.06\}$ for CHR procedures (solid lines, ⊠, ×, +), VHR procedures with $\{D_{low} = 0.25, D_{high} = 1.0\}$ (dotted lines, △, ○, □) and the same VHR procedures with $Q = 0.4$ (dashed lines, ▲, ●, ■). Initial depletion values for each scenario are indicated by inverted triangles along the upper *x*-axis.

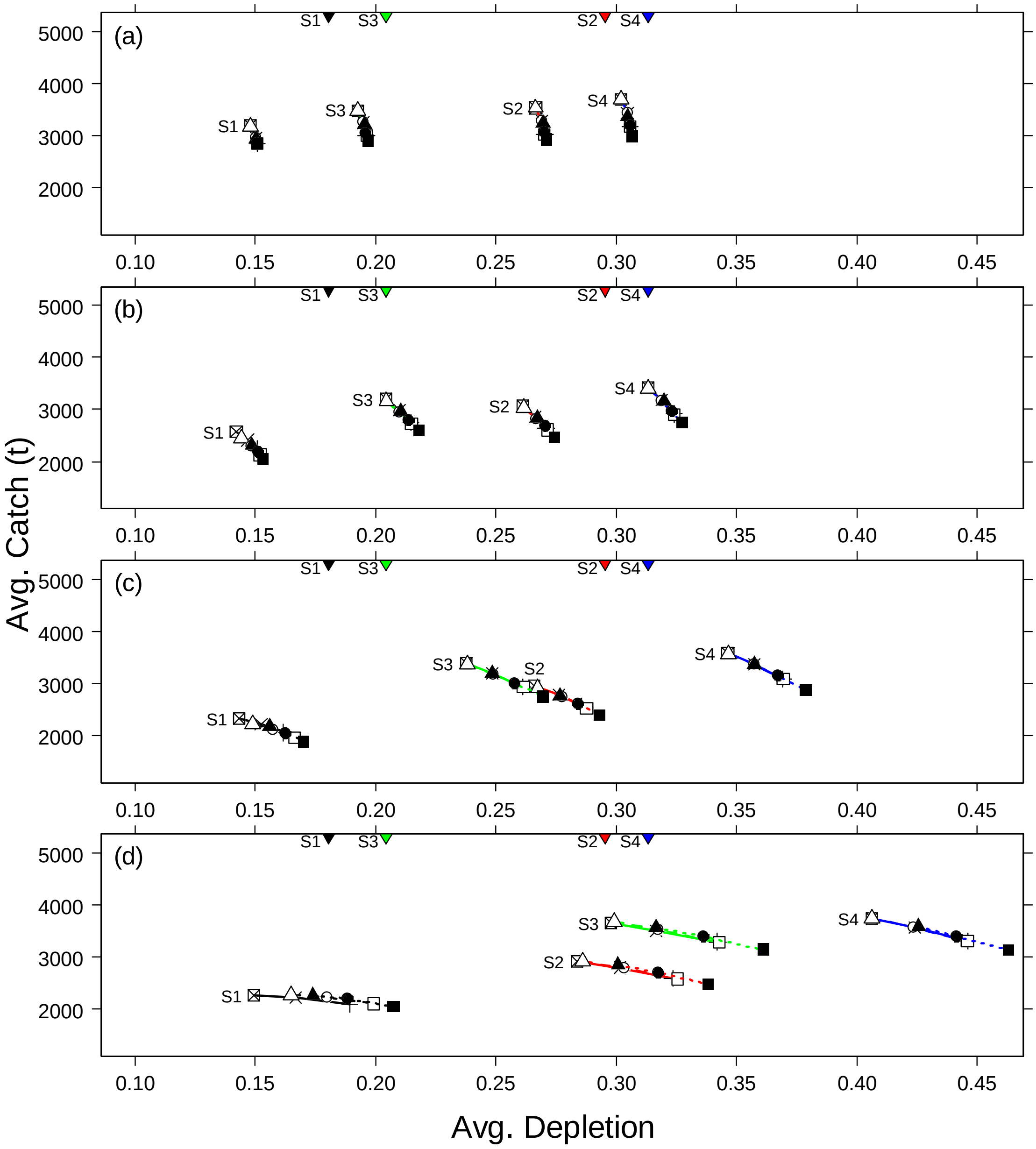


Figure F-3 Trade-off plot of median average catch against median average depletion for nine production model-based management procedures for scenarios S1 to S4 and 5, 10, 20 and 40 year time horizons (panels a-d). Procedures are ordered from left to right with $U^{ref} = \{0.10, 0.08, 0.06\}$ for CHR procedures (solid lines, ⊠, ×, +), VHR procedures with $\{D_{low} = 0.1, D_{high} = 0.4\}$ (dotted lines, △, ○, □) and the same VHR procedures with $Q = 0.4$ (dashed lines, ▲, ●, ■). . Initial depletion values for each scenario are indicated by inverted triangles along the upper *x*-axis.

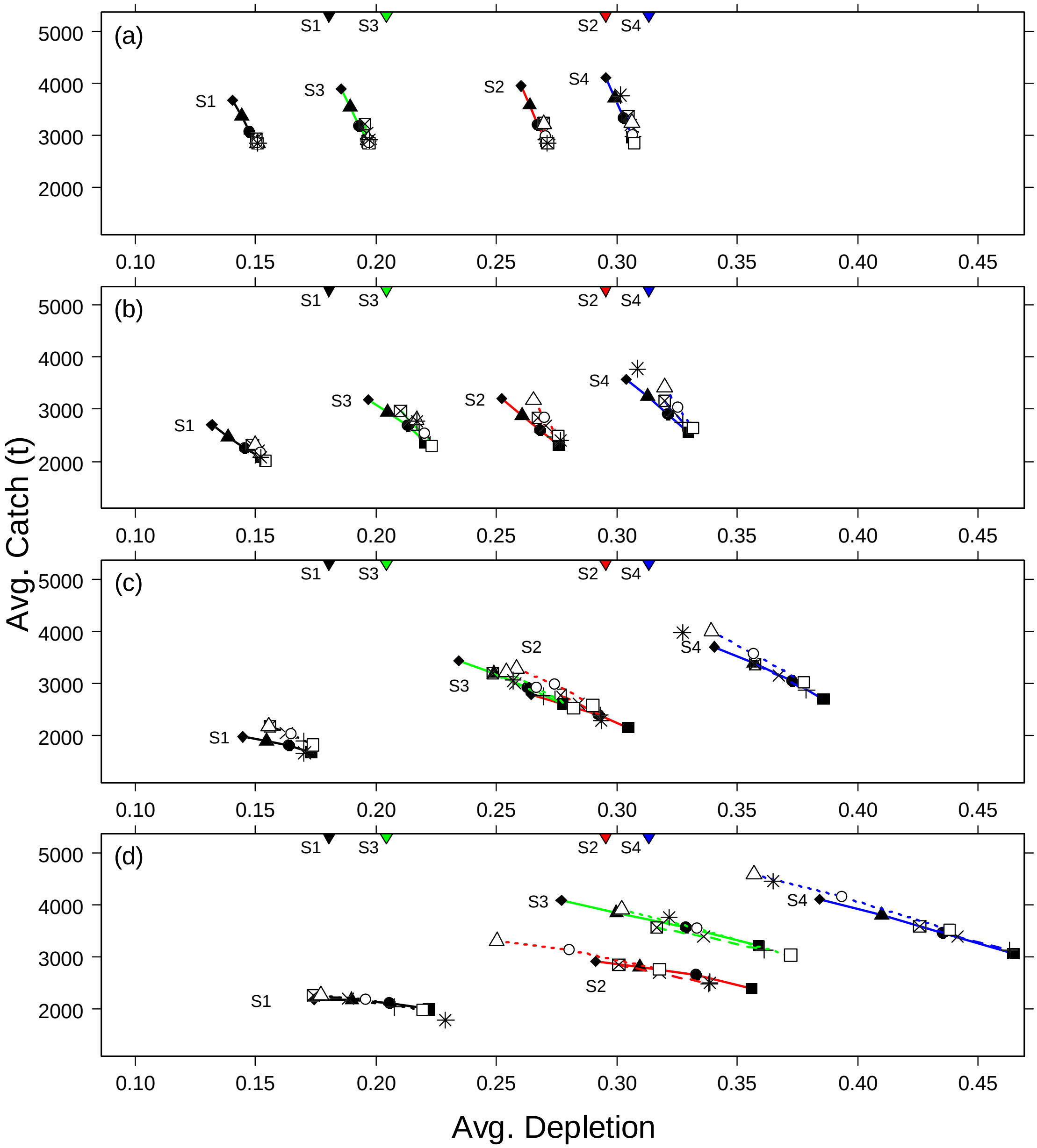


Figure F-4 Trade-off relationship between median average catch (t) and median average depletion for scenarios S1-S4 and 5, 10, 20 and 40 year time horizons (panels a-d). Symbols are ordered from left to right corresponding to data-based with $\lambda_2 = \{240, 210, 180, 150\}$ (♦, ▲, ●, ■), production model-based with U^{ref}={0.10, 0.08, 0.06} (☒, ×, +), and CA model-based with U^{ref}={0.10, 0.08, 0.06} (△, ○, □) procedures. Perfect-information procedures with U_{MSY} = {0.06, 0.08, 0.10} are indicated by asterisks. Inverted triangles and labels along the upper *x*-axis indicate initial depletion values for each scenario.

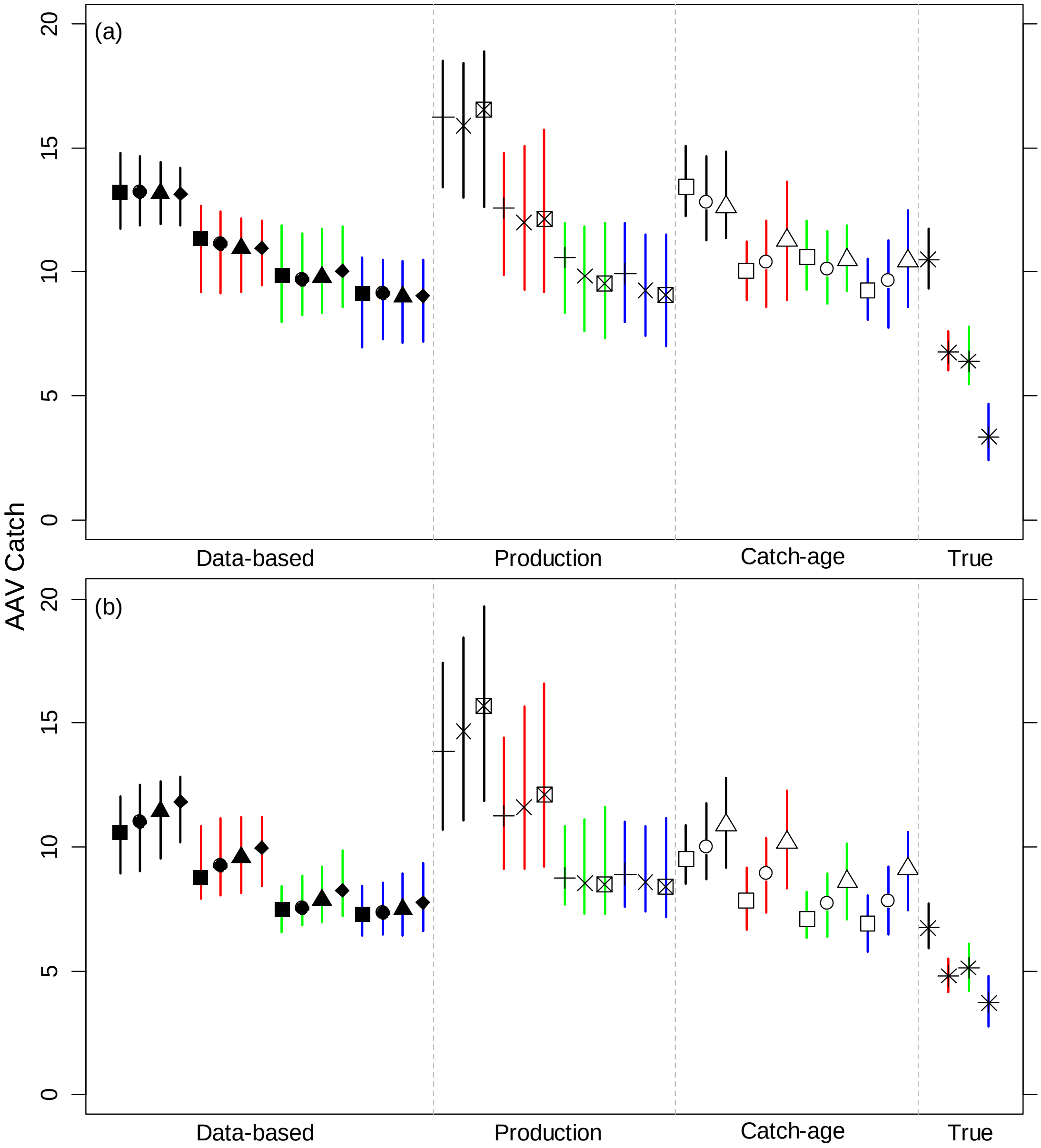


Figure F-5 Summary of catch variability performance over 20 years (panel a) and 40 years (panel b) for three selected classes of management procedures. Results are shown for scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Results for the quasi-perfect information procedures are shown as “True”. Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

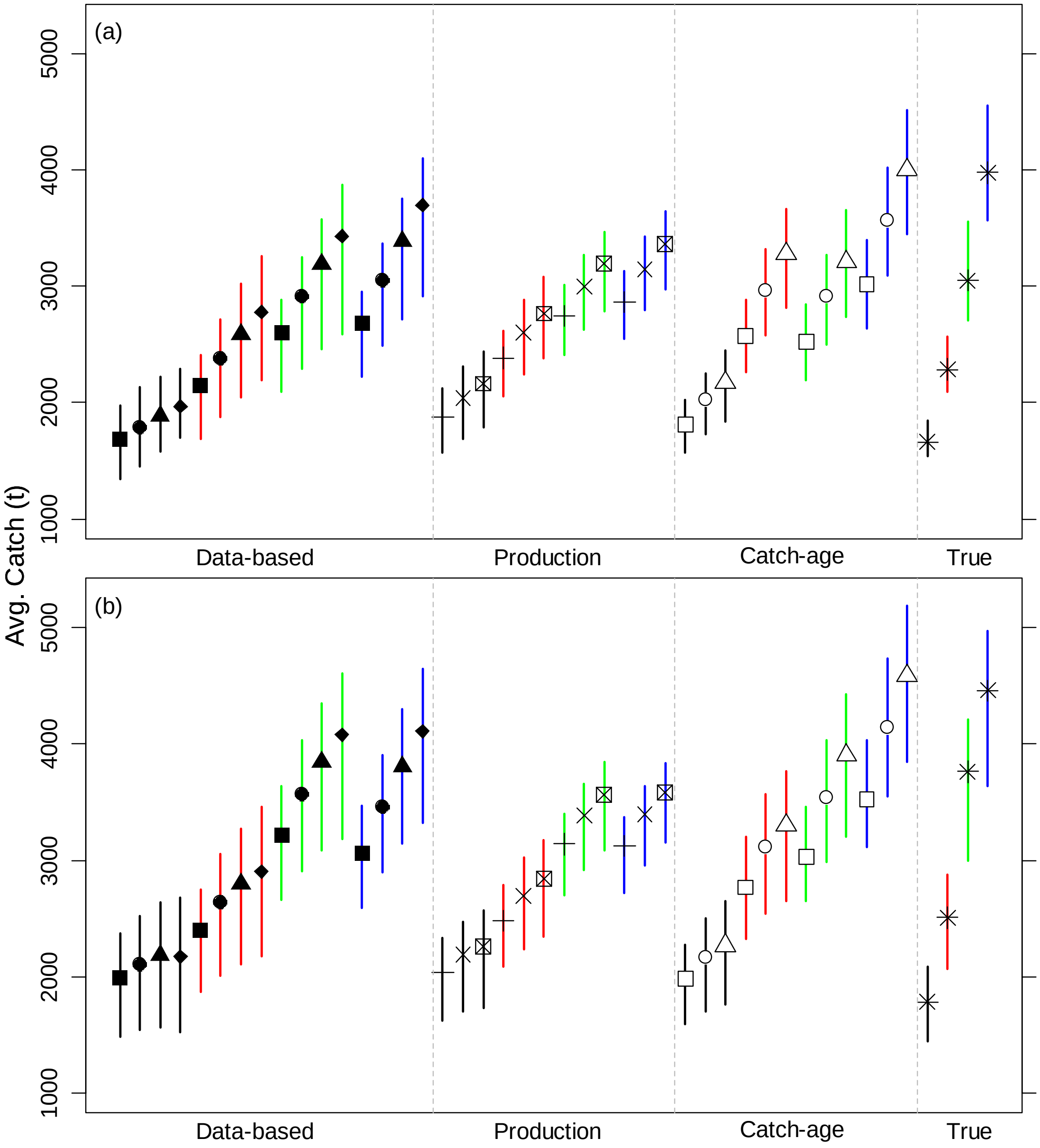


Figure F-6 Average annual catch over (a) 20 years and (b) 40 years for three selected management procedure classes under scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Perfect-information procedures are shown as “True”. Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

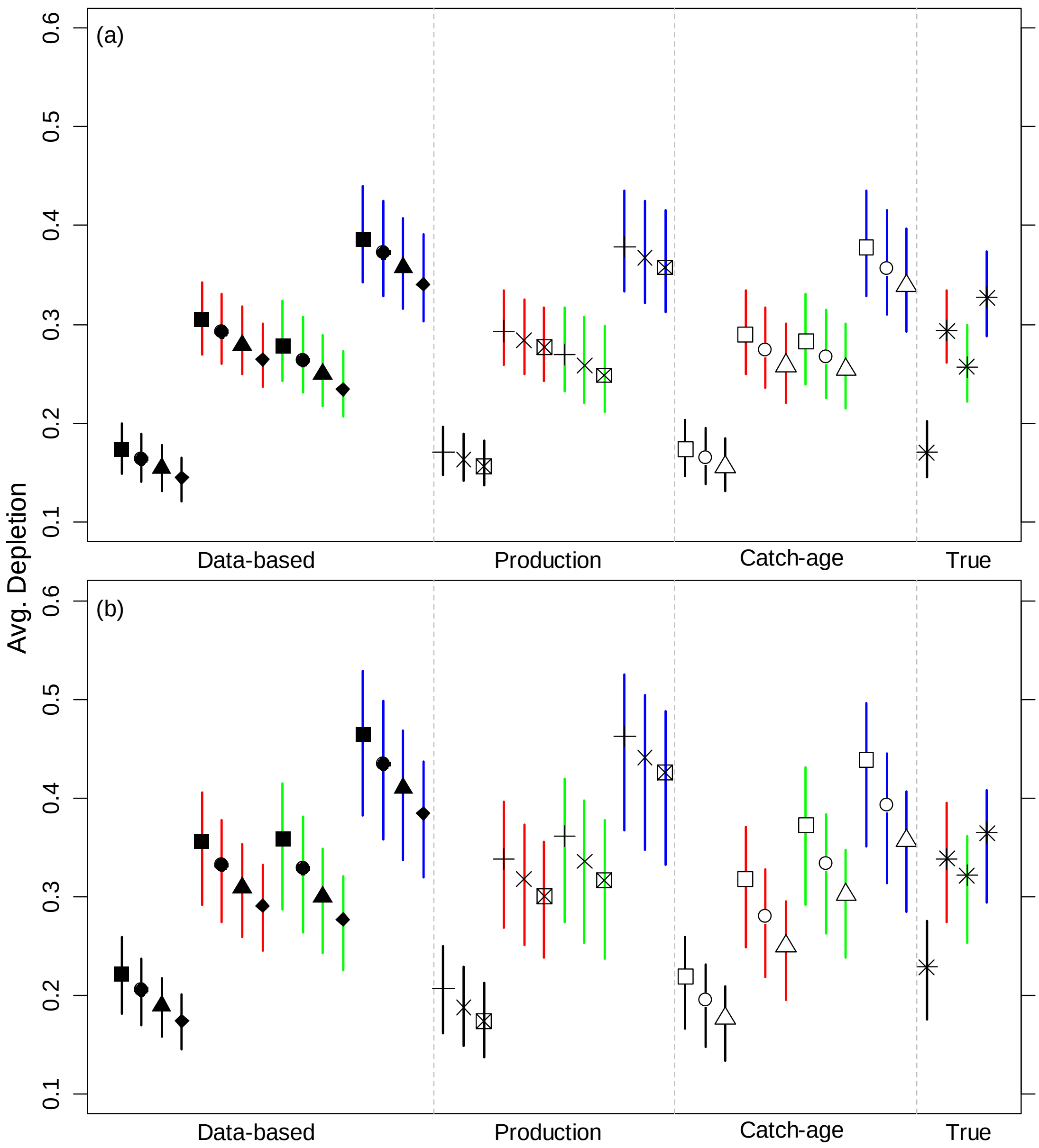


Figure F-7 Average spawning biomass depletion over (a) 20 years and (b) 40 years for three selected management procedure classes under scenarios S1 (black), S2 (red), S3 (green) and S4 (blue). Perfect-information procedures are shown as “True”. Symbols indicate the median value and the vertical bars indicate the 5th and 95th percentiles over 50 replicates.

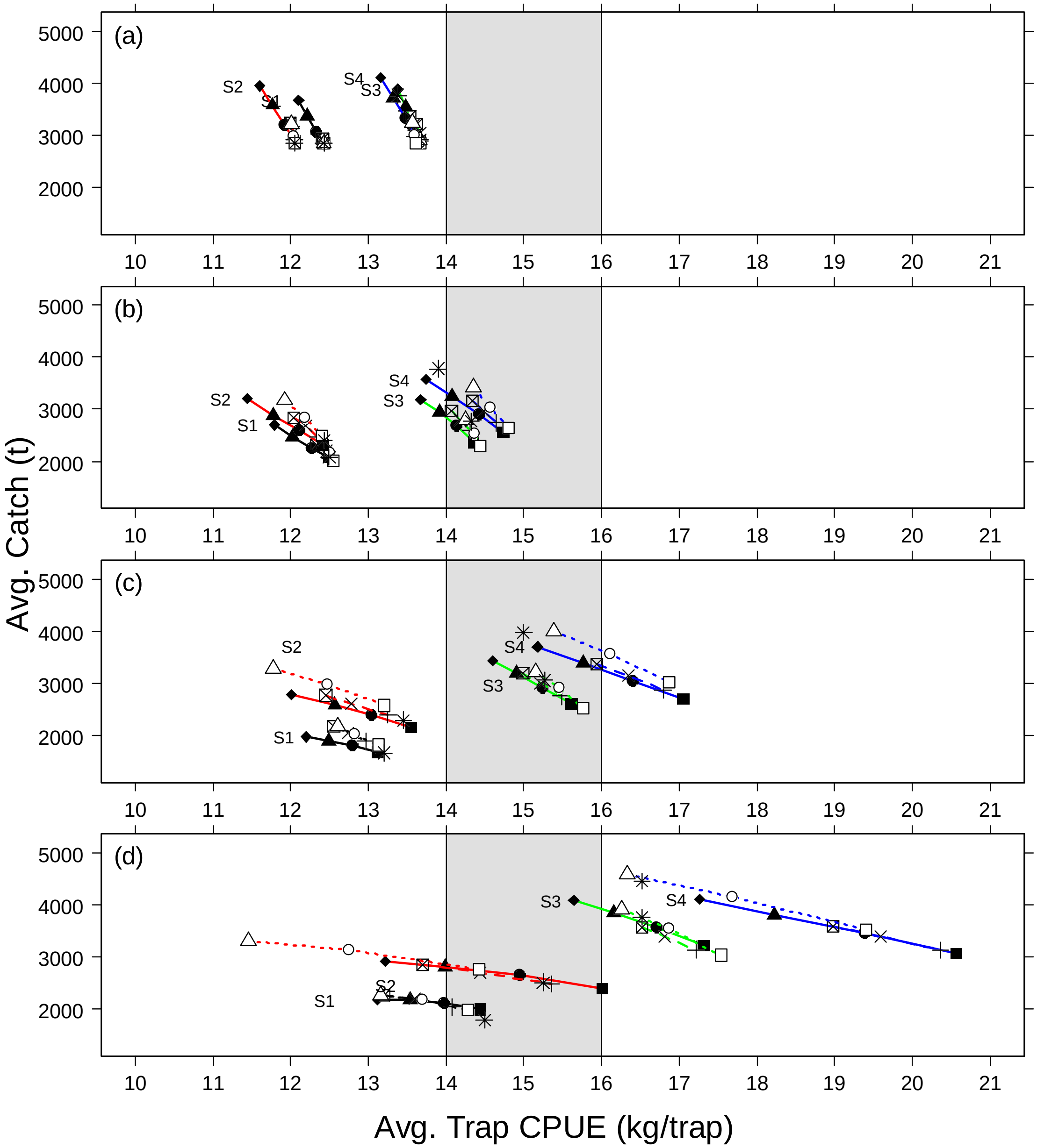


Figure F-8 Trade-off relationship between median average catch (t) and median average trap CPUE for scenarios S1-S4 and 5, 10, 20 and 40 year time horizons (panels a-d). Results are shown for selected data-based (solid lines, ♦, ▲, ●, ■), production model-based (dashed lines, ⊠, ×, +) and CA model-based (dotted lines, △, ○, □). The gray shaded region indicates CPUE levels identified by industry stakeholders as a consensus objective.