INTERNATIONAL PACIFIC
HALIBUT COMMISSION

IPHC–2025–SRB026–00
Last Update: 9 May 2025

26th Session of the IPHC Scientific Review Board (SRB026) – *Compendium of meeting documents*

10-12 June 2025, Seattle, WA, USA

Commissioners

Canada	United States of America
Mark Waddell	Jon Kurland
Neil Davis	Robert Alverson
Peter DeGreef	Richard Yamada

Executive Director

David T. Wilson, Ph.D.

DISTRIBUTION: IPHC WEBSITE
LAST UPDATE: 9 May 2025

BIBLIOGRAPHIC ENTRY
IPHC 2025. 26th Session of the IPHC Scientific Review Board (SRB026) - Compendium of meeting documents. Int. Pac. Halibut Comm.

Contact details:

International Pacific Halibut Commission
2320 W. Commodore Way, Suite 300
Seattle, WA, 98199-1287, U.S.A.
Phone: +1 206 634 1838
Fax: +1 206 632 2983
Email: secretariat@iphc.int
Website: http://iphc.int/

Last updated: 12 March 2025

PROVISIONAL: AGENDA & SCHEDULE FOR THE 26th SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB02)

Date: 10-12 June 2025
Location: Seattle, WA, USA
Venue: IPHC HQ (for SRB only) & Adobe Connect
Time: 09:00-17:00 (10-11th), 09:00-12:00 (12th) PDT
Chairperson: Dr Sean Cox (Simon Fraser University)
Vice-Chairperson: Nil

1. OPENING OF THE SESSION

2. ADOPTION OF THE AGENDA AND ARRANGEMENTS FOR THE SESSION

3. IPHC PROCESS

3.1. SRB annual workflow (D. Wilson)
3.2. Update on the actions arising from the 25th Session of the SRB (SRB025) (D. Wilson)
3.3. Outcomes of the 101st Session of the IPHC Annual Meeting (AM101) (D. Wilson)
3.4. Observer updates (e.g. Science Advisors)

4. INTERNATIONAL PACIFIC HALIBUT COMMISSION 5-YEAR PROGRAM OF INTEGRATED RESEARCH AND MONITORING (2026-31)

4.1. RESEARCH

4.1.1. Biology and ecology
4.1.2. Pacific halibut stock assessment
4.1.3. Management strategy evaluation

4.2. MONITORING

4.2.1. Fishery-dependent data
4.2.2. Fishery-independent data
- IPHC Fishery-Independent Setline Survey (FISS)
  - 2026 FISS design evaluation (R. Webster)
  - Updates to space-time modelling (R. Webster)

4.2.3. Age composition data (both fishery-dependent and fishery-independent)
- Ageing methods update

5. MANAGEMENT SUPPORTING INFORMATION

6. REVIEW OF THE DRAFT AND ADOPTION OF THE REPORT OF THE 26th SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB026)

SCHEDULE FOR THE 24th SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB024)

Tuesday, 10 June 2025		
Time	**Agenda item**	**Lead**
09:00-09:15	1. OPENING OF THE SESSION 2. ADOPTION OF THE AGENDA AND ARRANGEMENTS FOR THE SESSION	S. Cox & D. Wilson
09:15-10:00	3. IPHC PROCESS 3.1 SRB annual workflow (D. Wilson) 3.2 Update on the actions arising from the 25th Session of the SRB (SRB025) 3.3 Outcomes of the 101st Session of the IPHC Annual Meeting (AM101) 3.4 Observer updates (e.g. Science Advisors)	D. Wilson
10:00-11:00	4. INTERNATIONAL PACIFIC HALIBUT COMMISSION 5-YEAR PROGRAM OF INTEGRATED RESEARCH AND MONITORING (2026-31)	D. Wilson
11:00-12:30	4.1 RESEARCH 4.1.1 Biology and ecology	J. Planas
12:30-13:30	Lunch	
13:30-16:00	4.1.2 Pacific halibut stock assessment	I. Stewart
16:00-17:00	SRB drafting session	SRB members
18:30-21:00	SRB Function (Location TBA)	SRB

Wednesday, 11 June 2025		
Time	**Agenda item**	**Lead**
09:00-09:30	Review of Day 1 and discussion of SRB Recommendations from Day 1	Chairperson
09:30-12:30	4.1.3 Management Strategy Evaluation	A. Hicks
12:30-13:30	Lunch	
13:30-16:00	4.2 MONITORING 4.2.1. Fishery-dependent data 4.2.2. Fishery-independent data • IPHC Fishery-Independent Setline Survey (FISS) ○ 2026 FISS design evaluation (R. Webster) ○ Updates to space-time modelling (R. Webster) 4.2.3. Age composition data (both fishery-dependent and fishery-independent) • Ageing methods update ○ Using artificial intelligence (AI) for supplementing Pacific halibut age determination from collected otoliths (B. Hutniczak)	D. Wilson R. Webster K. Ualesi B. Hutniczak
16:00-17:00	SRB drafting session	SRB members
Thursday, 12 June 2025		
Time	**Agenda item**	**Lead**
09:00-09:30	5. MANAGEMENT SUPPORTING INFORMATION	All
09:30-10:30	SRB drafting session	SRB members
10:30-11:30	Time for all participants to review the draft report	All
11:30-12:30	6. ADOPTION OF THE REPORT OF THE 26th SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB026)	S. Cox
12:30-13:30	Lunch and departure	

Last updated: 9 May 2025

LIST OF DOCUMENTS FOR THE 26th SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB026)

Document	Title	Availability
IPHC-2025-SRB026-01	Agenda & Schedule for the 26th Session of the Scientific Review Board (SRB026)	✓ 12 Mar 2025
IPHC-2025-SRB026-02	List of Documents for the 26th Session of the Scientific Review Board (SRB026)	✓ 12 Mar 2025 ✓ 9 May 2025
IPHC-2025-SRB026-03	Update on the actions arising from the 25th Session of the SRB (SRB025) (IPHC Secretariat)	✓ 9 May 2025
IPHC-2025-SRB026-04	Outcomes of the 101st Session of the IPHC Annual Meeting (AM101) (D. Wilson)	✓ 9 May 2025
IPHC-2025-SRB026-05	International Pacific Halibut Commission 5-Year program of integrated research and monitoring (2022-26) (D. Wilson, J. Planas, I. Stewart, A. Hicks, R. Webster, & B. Hutniczak)	✓ 9 May 2025
IPHC-2025-SRB026-06	Report on current and future biological and ecosystem science research activities (J. Planas, C. Dykstra, A. Jasonowicz, & C. Jones)	✓ 7 May 2025
IPHC-2025-SRB026-07	Development of the 2025 Pacific halibut (*Hippoglossus stenolepis*) stock assessment (I. Stewart & A. Hicks)	✓ 9 May 2025
IPHC-2025-SRB026-08	IPHC Secretariat MSE Program of Work (2025) and an update on development of a Harvest Strategy Policy (A. Hicks & I. Stewart)	✓ 8 May 2025
IPHC-2025-SRB026-09	2026-28 FISS design evaluation (R. Webster, I. Stewart, K. Ualesi, T. Jack, & D. Wilson)	✓ 9 May 2025
IPHC-2025-SRB026-10	Using artificial intelligence (AI) for supplementing Pacific halibut age determination from collected otoliths (B. Hutniczak, J. Forsberg, K. Sawyer Van Vleck, & K. Magrane)	✓ 5 May 2025
Information papers		
IPHC-2025-SRB026-INF01	Nil to date	Due: 9 Jun 2025

UPDATE ON THE ACTIONS ARISING FROM THE 25TH SESSION OF THE IPHC SCIENTIFIC REVIEW BOARD (SRB025)

PREPARED BY: IPHC SECRETARIAT (9 MAY 2025)

PURPOSE

To provide the Scientific Review Board (SRB) with an opportunity to consider the progress made during the intersessional period, on the recommendations/requests arising from the SRB025.

BACKGROUND

At the SRB025, the members recommended/requested a series of actions to be taken by the IPHC Secretariat, as detailed in the SRB024 meeting report (IPHC–2024–SRB025–R) available from the IPHC website, and as provided in Appendix A.

DISCUSSION

During the 26th Session of the SRB (SRB026), efforts will be made to ensure that any recommendations/requests for action are carefully constructed so that each contains the following elements:

1) a specific action to be undertaken (deliverable);
2) clear responsibility for the action to be undertaken (such as the IPHC Staff or SRB officers);
3) a desired time frame for delivery of the action (such as by the next session of the SRB or by some other specified date).

RECOMMENDATIONS

That the SRB:

1) **NOTE** paper IPHC-2025-SRB026-03, that provided the SRB with an opportunity to consider the progress made during the inter-sessional period, in relation to the consolidated list of recommendations/requests arising from the previous SRB meeting (SRB025).
2) **AGREE** to consider and revise the actions as necessary, and to combine them with any new actions arising from SRB026.

APPENDICES

Appendix A: Update on actions arising from the 25th Session of the IPHC Scientific Review Board (SRB025).

APPENDIX A

Update on actions arising from the 25th Session of the IPHC Scientific Review Board (SRB025)

RECOMMENDATIONS

Action No.	Description	Update
SRB025–Rec.01 (para. 14)	***IPHC 5-Year Program of Integrated Research and Monitoring (2022-26)*** The SRB **RECOMMENDED** that the IPHC 5-year Program of Integrated Research and Monitoring be revised by SRB026 to reflect changing priorities in light of major progress on biological research and ongoing monitoring challenges.	***In Progress*** **Update**: The 5YPIRM is currently being revised and a draft for the next 5-year period is expected to be provided to the SRB027.
SRB025–Rec.02 (para. 15)	The SRB **RECOMMENDED** incorporating evaluation of new technologies into the 5-year Program of Integrated Research and Monitoring. Initial examples include: a) testing samples of AI-generated age compositions in the assessment model as soon as is practicable to determine their potential value for that purpose; b) using AI to support ageing requirements for gene-tagging and/or CKMR methods to estimate abundance. These ages would be required beyond ageing workloads for normal assessment purposes; c) epigenetic ageing (a new project beginning 2025), which could provide more reliable and unbiased ages than AI and perhaps comparable in precision to human-read ages.	***In Progress*** **Update**: The 5YPIRM is currently being revised and a draft for the next 5-year period is expected to be provided to the SRB027.
SRB025–Rec.03 (para. 18)	***Pacific halibut stock assessment*** The SRB **RECOMMENDED** considering the impact of alternative FISS designs not only on the coast-wide abundance estimates but also on our understanding of the stock distribution across regions.	***In Progress*** **Update**: Additional information on uncertainty due to reduced FISS designs was added to presentations and documents for AM101. Further, the uncertainty in stock distribution was propagated into projected TCEYs for 2025.

SRB025–Rec.04 (para. 24)	***Management strategy evaluation*** **NOTING** the analysis of depensation, the SRB **RECOMMENDED** redoing this analysis in the future whenever estimated spawning stock biomass falls below the minimum level previously observed within the corresponding PDO regime.	***Completed*** The estimates from the stock assessment will be monitored to determine if the analysis should be repeated.
SRB025–Rec.05 (para. 26)	The SRB strongly **RECOMMENDED** against using MSE (a strategic tool) in the annual TCEY setting process. Exceptional circumstances checks (on WPUE and CATCH) are used to judge whether management procedures are generating appropriate recommendations in a given year.	***Completed*** This recommendation was noted by the Commission.
SRB025–Rec.06 (para. 27)	The SRB **RECOMMENDED** including performance metrics expressing impacts of alternative FISS designs and MP options in terms of the dollar value of foregone yield to more directly capture economic outputs. The SRB **RECOGNISED** that there is long-term price uncertainty and complicated economics. Nevertheless, it is not unreasonable to present economic performance for the short-term projections.	***In Progress*** **Update**: Specific performance metrics are being developed, but general economic consequences have been communicated.
SRB025–Rec.07 (para. 30)	The SRB **RECOMMENDED** adopting realised coastwide catch as a fishery-dependent indicator for testing exceptional circumstances. Realised coastwide catch each year can be compared to the projected distribution of future TCEY for that year to determine whether biological or management processes (e.g. decision variability) are leading to unexpected TCEY.	***Completed*** This has been added to the draft Harvest Strategy Policy and is reported in document **IPHC-2025-SRB026-08**.
SRB025–Rec.08 (para. 31)	The SRB **RECOMMENDED** adding a measurable objective related to absolute spawning biomass under the general objective 2.1 “maintain spawning biomass at or above a level that optimises fishing activities” to be included in the priority Commission objectives after, or in place of, the current relative biomass threshold objective.	***In Progress*** **Update**: Objectives related to absolute spawning biomass are being discussed by the Commission and MSAB.
SRB025–Rec.09 (para. 35)	***Biology and ecology*** The SRB **RECOMMENDED** that when incorporating the new maturity ogive derived from the use of generalised additive models into	***Completed*** This recommendation is addressed in Section 2 of document **IPHC-SRB026-**

	the stock assessment, that the Secretariat consider using annual calculation of a regionally weighted ogive for years where FISS regional abundance estimates are available rather than one weighted by the 2023 FISS relative abundances by biological region.	**06** and results will be presented at the meeting.
SRB025–Rec.10 (para. 36)	The SRB **NOTED** a decrease in the coastwide A50, driven largely by changes in Biological Region 2 from 2022 to 2023 and **RECOMMENDED**: a) not to pool years to inspect potential decreasing trends in the age at maturity; b) investigating separately the maturity ogives and the age at the first maturity by determining, where possible, whether an individual has spawned previously.	***Completed*** This recommendation is addressed in Section 2 of document IPHC-SRB026-06 and results will be presented at the meeting.
SRB025–Rec.11 (para. 44)	***2025 FISS design evaluation*** The SRB **RECOMMENDED** a preliminary analysis of potential alternative approaches to generating Pacific halibut abundance estimates in the future. For example, the MSE simulations could be used to generate projected survey deficits over the next 3-5 yrs to estimate the distribution of cumulative "supplemental funding" (CSF) required over that time. The CSF can then be compared to the estimated cost of developing and executing alternative abundance estimators such as gene-tagging and/or CKMR, which partially rely on less expensive commercial catch sampling. Genetic methods require up-front development costs that may look more reasonable against the prospect of the CSF. Annual CKMR costs could be substantially less than annual FISS costs, while providing reliable absolute biomass estimates regardless of stock status.	***In Progress*** **Update**: Supplemental funding and FISS design needs are highly uncertain and rapidly evolving. The Secretariat will need to propose this type of work as part of the next research plan and receive feedback from the Commission on which alternative abundance estimators to explore.
SRB025–Rec.12 (para. 47)	***Age composition data (both fishery-dependent and fishery-independent)*** The SRB **RECOMMENDED** that the Secretariat investigate using the AI to identify region of collection. Otolith shape is sometimes used as a tool for understanding mixing and stock structure and the AI may have skill in identifying region of origin (and thus mixing and migration rates) from otolith images.	***In Progress*** **Update**: An update will be provided in-session. See also paper **IPHC-2025-SRB026-10**

REQUESTS

Action No.	Description	Update
SRB025–Req.01 (para. 20)	***Pacific halibut stock assessment*** The SRB **REQUESTED** an analysis of the relationship between commercial CPUE and the FISS WPUE at the coastwide and regional levels to investigate the strength of hyperstability/hyperdepletion in CPUE for the stock assessment in 2025. This analysis should include two scenarios: (i) the historical FISS WPUE estimates and (ii) FISS WPUE estimates calculated from reduced designs (i.e. subset the historical FISS data and recalculate WPUE from the reduced data set). The statistical model used for the analysis should account for uncertainty in the FISS index (the X-axis variable) using, for example, an error-in-variables approach like that in Harley et al. 2001 (CJFAS). This analysis represents a first step in including presumed hyperstability in scenarios that investigate the impacts of reduced FISS designs.	***Pending*** **Update**: This analysis was placed on hold while the full stock assessment was developed. It can be prioritized for SRB027 depending on other topics arising.
SRB025–Req.02 (para. 22)	**RECALLING** previous discussions at SRB020 (IPHC-2022-SRB020-R) and SRB021 (IPHC-2022-SRB021-R) regarding stock assessment research priorities and that several of the smaller topics have been addressed, the SRB **REQUESTED** an update on the list of larger topics larger topics that may require moving to a three-year schedule for stock assessment. Examples of such topics include the following: a) Exploration of alternative stock assessment model frameworks, e.g. state-space models like the Woods Holde Assessment Model (WHAM), Bayesian models, and spatially structured models beyond the Areas as Fleets model.	***Completed*** **Update**: An updated list of research topics is included in the preliminary assessment for 2025.
SRB025–Req.03 (para. 32)	***Management strategy evaluation*** **NOTING** that the definitions of “overfished” and “overfishing” are consistent with the use of these terms in the USA federal fishery management systems under the Magnuson-Stevens Act, but differ from the terms and definitions elsewhere, the SRB **REQUESTED** a broader investigating of terms and definitions related to B and F	***In Progress*** **Update**: This is being addressed with the Commission. See document **IPHC-2025-SRB026-08**.

	reference points used by fishery managements organisations throughout the world.	
SRB025–Req.04 (para. 37)	***Biology and ecology*** The SRB **REQUESTED** a preliminary evaluation of the feasibility for using information on the genetic differentiation of Pacific halibut parasites as a possible stock structure marker.	***In Progress*** The IPHC Secretariat has conducted literature searches on the types and prevalence of parasites in Pacific halibut and their outcomes will be discussed at the SRB026 meeting.

OUTCOMES OF THE 101ST SESSION OF THE IPHC ANNUAL MEETING (AM101)

PREPARED BY: IPHC SECRETARIAT (D. WILSON; 9 MAY 2025)

PURPOSE

To provide the SRB with the outcomes of the 101st Session of the IPHC Annual Meeting (AM10`), relevant to the mandate of the SRB.

BACKGROUND

Nil

DISCUSSION

During the course of the 101st Session of the IPHC Annual Meeting (AM101) the Commission made a number of specific recommendations and requests for action regarding the stock assessment, MSE process, and 5-year research program. Relevant sections from the report of the meeting are provided in Appendix A for the SRB's consideration.

RECOMMENDATION

That the SRB:

1) **NOTE** paper IPHC-2025-SRB026-04 which details the outcomes of the 101st Session of the IPHC Annual Meeting (AM101), relevant to the mandate of the SRB.

APPENDICES

Appendix A: Excerpts from the 101st Session of the IPHC Annual Meeting (AM101) Report (IPHC-2025-AM101-R).

APPENDIX A

Excerpts from the 101st Session of the IPHC Annual Meeting (AM101) Report

(IPHC-2025-AM101-R)

RECOMMENDATIONS

Nil

REQUESTS

Management Strategy Evaluation

AM101–Req.04 (para. 53) The Commission **REQUESTED** that the Secretariat facilitate informal intersessional workshops, consisting of Commissioners and key advisors, to review and consider the draft Harvest Strategy Policy, for adoption in mid-to-late 2025.

OTHER

Para. 23. The Commission **NOTED** that at the request of the SRB (see below), the IPHC Secretariat will be updating the 5YPIRM throughout the course of 2025 with the intention of presenting a draft of the next 5YPIRM (2026-31) to the Commission at IM101 in November 2025.

> *SRB025–Rec.01 (para. 14) The SRB RECOMMENDED that the IPHC 5-year Program of Integrated Research and Monitoring be revised by SRB026 to reflect changing priorities in light of major progress on biological research and ongoing monitoring challenges.*
>
> *SRB025–Rec.02 (para. 15) The SRB RECOMMENDED incorporating evaluation of new technologies into the 5-year Program of Integrated Research and Monitoring. Initial examples include:*
>
> > *a) testing samples of AI-generated age compositions in the assessment model as soon as is practicable to determine their potential value for that purpose;*
> >
> > *b) using AI to support ageing requirements for gene-tagging and/or CKMR* [Close Kin Mark Recapture] *methods to estimate abundance. These ages would be required beyond ageing workloads for normal assessment purposes;*
> >
> > *c) epigenetic ageing (a new project beginning 2025), which could provide more reliable and unbiased ages than AI and perhaps comparable in precision to human-read ages.*

Para. 24. The Commission **NOTED** paper IPHC-2025-AM101-INF03 that summarizes the information available on the use of artificial intelligence (AI) for determining the age of fish from images of collected otoliths and provides an update on the exploratory work of implementing an AI-based age determination model for Pacific halibut.

INTERNATIONAL PACIFIC HALIBUT COMMISSION

INTERNATIONAL PACIFIC HALIBUT COMMISSION 5-YEAR PROGRAM OF INTEGRATED RESEARCH AND MONITORING (2022-26): UPDATES

PREPARED BY: IPHC SECRETARIAT (D. WILSON, J. PLANAS, I. STEWART, A. HICKS, B. HUTNICZAK, AND R. WEBSTER; 9 MAY 2025)

PURPOSE

To provide the Commission with an annual opportunity to comment and amend the IPHC's 5-year Program of Integrated Research and Monitoring (2022-26) (the Plan).

BACKGROUND

Recalling that:

a) the IPHC Secretariat conducts activities to address key issues identified by the Commission, its subsidiary bodies, the broader stakeholder community, and the IPHC Secretariat;

b) the process of identifying, developing, and implementing the IPHC's science-based activities involves several steps that are circular and iterative in nature, but result in clear project activities and associated deliverables;

c) the process includes developing and proposing projects based on direct input from the Commission, the experience of the IPHC Secretariat given its broad understanding of the resource and its associated fisheries, and concurrent consideration by relevant IPHC subsidiary bodies, and where deemed necessary, including by the Commission, additional external peer review;

d) the IPHC Secretariat commenced implementation of the new Plan in 2022 and will keep the Plan under review on an ongoing basis.

Also recalling that an overarching goal of the IPHC 5-year Program of Integrated Research and Monitoring (2022-26) is to promote integration and synergies among the various research and monitoring activities of the IPHC Secretariat in order to improve knowledge of key inputs into the Pacific halibut stock assessment, and Management Strategy Evaluation (MSE) processes, thereby providing the best possible advice for management decision making processes.

The 1st iteration of the Plan was formally presented to the Commission at IM097 in November 2021 (IPHC-2021-IM097-12) for general awareness of the documents ongoing development. At the 98th Session of the IPHC Annual Meeting (AM098) in January 2022, the Commission requested a number of amendments which were subsequently incorporated.

In 2023 and 2024, the plan went through two cycles of review and improvement with the SRB, with amendments being suggested and incorporated accordingly.

DISCUSSION

The SRB should note that:

a) the intention is to ensure that the new integrated plan is kept as a '*living plan*', and is reviewed and updated annually based on the resources available to undertake the work of the Commission (e.g. internal and external fiscal resources, collaborations, internal expertise);

b) the plan focuses on core responsibilities of the Commission; and any redirection provided by the Commission;

c) each year the SRB may choose to recommend modifications to the current Plan, and that any modifications subsequently made would be documented both in the Plan itself, and through reporting back to the SRB and then the Commission.

Updates: The Secretariat is currently in the process of updating the Plan to meet the request of the SRB at its 24th Session, as per the below text:

> ***International Pacific Halibut Commission 5-year program of integrated research and monitoring (2022-26)***
>
> SRB024–Req.01 (para. 14) *The SRB **REQUESTED** that the IPHC 5-year Program of Integrated Research and Monitoring be revised by SRB026 to reflect changing priorities in light of major progress on biological research and ongoing monitoring challenges.*

RECOMMENDATION

That the SRB:

1) **NOTE** paper IPHC-2025-SRB026-05 that provides the latest iteration of the IPHC 5-year program of Integrated Research and Monitoring (2022-26).

APPENDICES

Appendix A: Updated: IPHC 5-year program of Integrated Research and Monitoring (2022-26)

INTERNATIONAL PACIFIC HALIBUT COMMISSION

5-YEAR PROGRAM OF INTEGRATED RESEARCH AND MONITORING

(2022 - 2026)

INTERNATIONAL PACIFIC

HALIBUT COMMISSION

Commissioners

Canada	United States of America
Paul Ryall	Jon Kurland
Neil Davis	Robert Alverson
Peter DeGreef	Richard Yamada

Executive Director

David T. Wilson, Ph.D.

BIBLIOGRAPHIC ENTRY

IPHC 2023. International Pacific Halibut Commission 5-Year program of integrated research and monitoring (2022-26). Seattle, WA, U.S.A. *IPHC–2023-5YPIRM, 58 pp.*

INTERNATIONAL PACIFIC HALIBUT COMMISSION

Contact details:

International Pacific Halibut Commission

2320 W. Commodore Way, Suite 300

Seattle, WA, 98199-1287, U.S.A.

Phone: +1 206 634 1838

Fax: +1 206 632 2983

Email: secretariat@iphc.int

Website: http://www.iphc.int/

ACRONYMS

AM	Annual Meeting
CB	Conference Board
DMR	Discard Mortality Rate
FAC	Finance and Administration Committee
FISS	Fishery-Independent Setline Survey
FSC	First Nations Food, Social, and Ceremonial [fishery]
IM	Interim Meeting
IPHC	International Pacific Halibut Commission
MSAB	Management Strategy Advisory Board
MSE	Management Strategy Evaluation
OM	Operating Model
PAB	Processor Advisory Board
PDO	Pacific Decadal Oscillation
PHMEIA	Pacific halibut multiregional economic impact assessment [model]
QAQC	Quality assurance/quality control
RAB	Research Advisory Board
SHARC	Subsistence Halibut Registration Certificates
SRB	Scientific Review Board
TCEY	Total Constant Exploitation Yield
U.S.A.	United States of America
WM	Work Meeting

DEFINITIONS

A set of working definitions are provided in the IPHC Glossary of Terms and abbreviations: https://iphc.int/the-commission/glossary-of-terms-and-abbreviations

TABLE OF CONTENTS

EXECUTIVE SUMMARY

An overarching goal of the *IPHC 5-Year Program of Integrated Research and Monitoring (2022-26)* is to promote integration and synergies among the various research and support activities of the IPHC Secretariat in order to improve our knowledge of key inputs into the Pacific halibut stock assessment and Management Strategy Evaluation (MSE) processes, and to provide the best possible advice for management decision-making processes.

Along with the implementation of the short- and medium-term activities contemplated in this *IPHC 5-Year Program of Integrated Research and Monitoring (2022-26)*, and in pursuit of the overarching objective, the IPHC Secretariat will also aim to:

1) undertake cutting-edge research programs in fisheries research in support of Pacific halibut fisheries management;
2) undertake groundbreaking methodological research;
3) undertake applied research;
4) establish new collaborative agreements and interactions with research agencies and academic institutions;
5) promote the international involvement of the IPHC by continued and new participation in international scientific organizations and by leading international science and research collaborations;
6) effectively communicate IPHC research outcomes;
7) incorporate talented students and early researchers in research activities contemplated.

The research and monitoring activities conducted by the IPHC Secretariat are directed towards fulfilling the following four (4) objectives within areas of data collection, biological and ecological research, stock assessment, and Management Strategy Evaluation (MSE). In addition, the IPHC responds to Commission requests for additional inputs to management and policy development which are classified under management support.

The Secretariat's success in implementing the *IPHC 5-Year Program of Integrated Research and Monitoring (2022-26)* will be measured according to the following criteria relevant to the stock assessment, the MSE and for all inputs to IPHC management:

1) Timeliness – was the research conducted, analyzed, published, and provided to the Commission at the appropriate points to be included in annual management decisions?
2) Accessibility – was the research published and presented in such a way that it was available to other scientists, stakeholders, and decision-makers?
3) Relevance – did the research improve the perceived accuracy of the stock assessment, MSE, or decisions made by the Commission?
4) Impact – did the research allow for more precision or a better estimate of the uncertainty associated with information for use in management?
5) Reliability – has the research resulted in more consistent information provided to the Commission for decision-making?

1. Introduction

The International Pacific Halibut Commission (IPHC) is a public international organization so designated via Presidential Executive Order 11059 and established by a Convention between Canada and the United States of America. The IPHC Convention was signed on 2 March 1923, ratified on 21 July 1924, and came into effect on 21 October 1924 upon exchange. The Convention has been revised several times since, to extend the Commission's authority and meet new conditions in the fishery. The most recent change occurred in 1979 and involved an amendment to the 1953 Halibut Convention. The 1979 amendment, termed a "protocol", was precipitated in 1976 by Canada and the United States of America extending their jurisdiction over fisheries resources to 200 miles. The 1979 Protocol along with the U.S. legislation that gave effect to the Protocol (Northern Pacific Halibut Act of 1982) has affected the way the fisheries are conducted, and redefined the role of IPHC in the management of the fishery. Canada does not require specific enabling legislation to implement the protocol.

The basic texts of the Commission are available on the IPHC website: https://www.iphc.int/the-commission, and prescribe the mission of the organization as:

*"..... **to develop the stocks of [Pacific] halibut in the Convention waters to those levels which will permit the optimum yield from the fishery and to maintain the stocks at those levels**. ....." IPHC Convention*, Article I, sub-article I, para. 2). The IPHC Convention Area is detailed in Fig. 1.

The IPHC Secretariat, formed in support the Commission's activities, is based in Seattle, WA, U.S.A. As its shared vision, ***the IPHC Secretariat aims to deliver positive economic, environmental, and social outcomes for the Pacific halibut resource for Canada and the U.S.A. through the application of rigorous science, innovation, and the implementation of international best practice.***

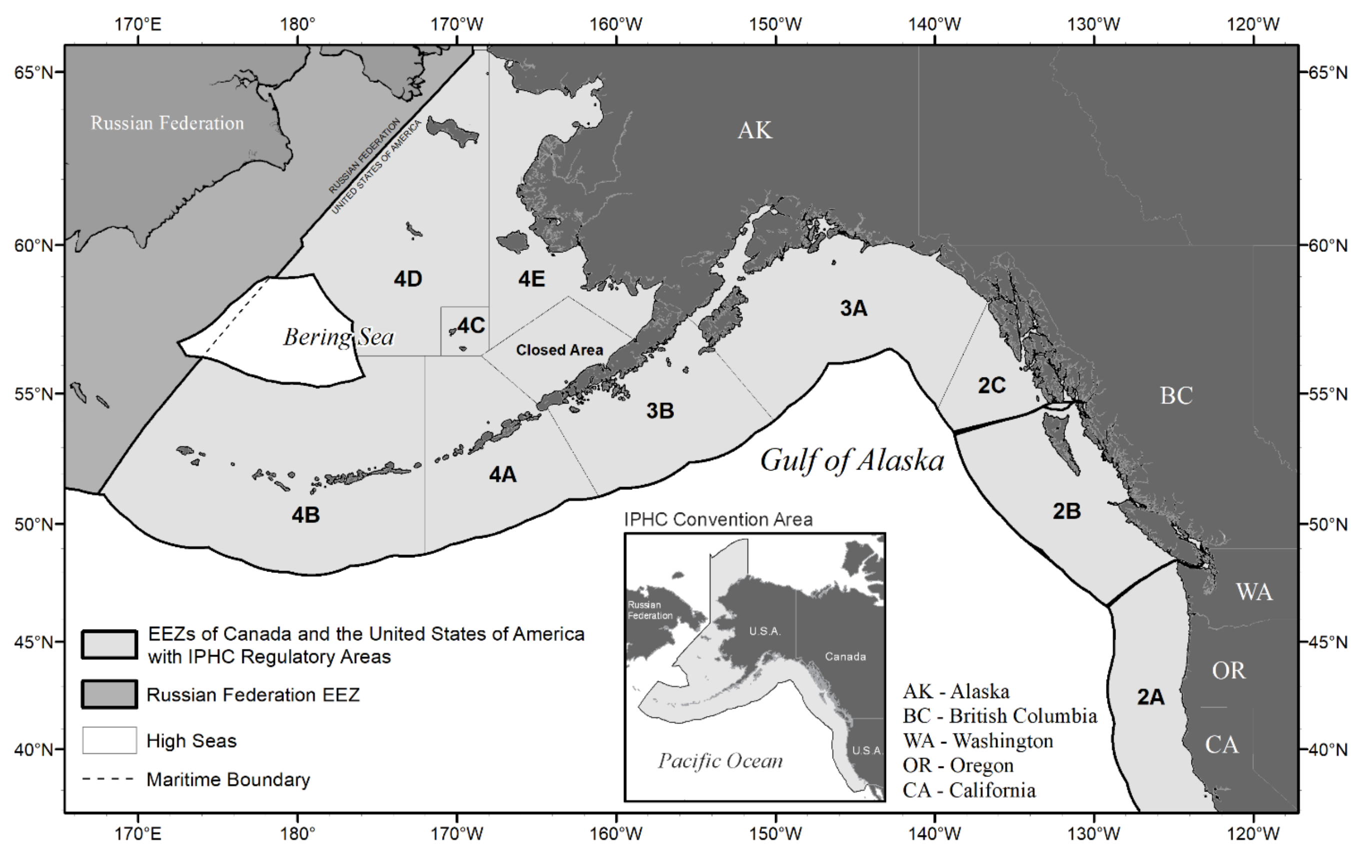


Figure 1. Map of the IPHC Convention Area (map insert) and IPHC Regulatory Areas.

2. Objectives

The IPHC has a long-standing history (since 1923) of collecting data, undertaking research, and stock assessment, devoted to describing and understanding the Pacific halibut (*Hippoglossus stenolepis*) stock and the fisheries that interact with it.

The IPHC Secretariat conducts activities to address key issues identified by the Commission, its subsidiary bodies, the broader stakeholder community, and of course, the IPHC Secretariat itself. The process of identifying, developing, and implementing our science-based activities involves several steps that are circular in nature, but result in clear research activities and associated deliverables. The process includes developing and proposing projects based on direct input from the Commission, the experience of the IPHC Secretariat given our broad understanding of the resource and its associated fisheries, and concurrent consideration by relevant IPHC subsidiary bodies, and where deemed necessary, additional external peer review.

Over the last five years (2017-2021), the research conducted by the IPHC Secretariat has been guided by a 5-Year Biological and Ecosystem Science Research Plan (IPHC–2019–BESRP-5YP) that aimed at improving knowledge on the biology of Pacific halibut in order to improve the accuracy of the stock assessment and in the management strategy evaluation (MSE) process. The IPHC-2019-BESRP-5YP contemplated research activities in five focal areas, namely Migration and Distribution, Reproduction, Growth and Physiological Condition, Discard Mortality Rates and Survival, and Genetics and Genomics. Research activities were highly integrated with the needs of stock assessment and MSE by their careful alignment with biological uncertainties and parameters, and the resulting prioritization (Appendix I). The outcomes of the IPHC-2019-BESRP-5YP have provided key inputs into stock assessment and the MSE process and, importantly, have provided foundational information for the successful pursuit of continuing and novel objectives within the new 5-Year Program of Integrated Research and Monitoring (2022-2026) (5YPIRM) (Appendix I).

The 2nd Performance Review of the IPHC (IPHC-2019-PRIPHC02-R), carried out over the course of 2019, also provided a range of recommendations to the Commission on ways in which it could continue to improve on the quality of scientific advice being provided to the Commission. There were nine (9) specific recommendations as provided below:

Science: Status of living marine resources

PRIPHC02–Rec.03 (para. 44) The PRIPHC02 ***RECOMMENDED*** *that opportunities to engage with western Pacific halibut science and management agencies be sought, to strengthen science links and data exchange. Specifically, consider options to investigate pan-Pacific stock structure and migration of Pacific halibut.*

PRIPHC02–Rec.04 (para. 45) The PRIPHC02 ***RECOMMENDED*** *that:*

a) further efforts be made to lead and collaborate on research to assess the ecosystem impacts of Pacific halibut fisheries on incidentally caught species (retained and/or discarded);
b) where feasible, this research be incorporated within the IPHC's 5-Year Research Plan (https://www.iphc.int/uploads/pdf/besrp/2019/iphc-2019-besrp-5yp.pdf);
c) findings from the IPHC Secretariat research and that of the Contracting Parties be readily accessible via the IPHC website.

Science: Quality and provision of scientific advice

PRIPHC02–Rec.05 (para. 63) The PRIPHC02 ***RECOMMENDED*** *that simplified materials be developed for RAB and especially MSAB use, including training/induction materials.*

PRIPHC02–Rec.06 (para. 64) The PRIPHC02 ***RECOMMENDED*** *that consideration be given to amending the Rules of Procedure to include appropriate fixed terms of service to ensure SRB peer review remains independent and fresh; a fixed term of three years seems appropriate, with no more than one renewal.*

PRIPHC02–Rec.07 (para. 65) The PRIPHC02 ***RECOMMENDED*** *that the peer review process be strengthened through expanded subject specific independent reviews including data quality and standards, the FISS, MSE, and biological/ecological research; as well as conversion of "grey literature" to primary literature publications. The latter considered important to ongoing information outreach efforts given the cutting-edge nature of the Commission's scientific work.*

PRIPHC02–Rec.08 (para. 66) The PRIPHC02 ***RECOMMENDED*** *that the IPHC Secretariat develop options for simple graphical summaries (i.e. phase plot equivalents) of fishing intensity and spawning stock biomass for provision to the Commission.*

Conservation and Management: Data collection and sharing

PRIPHC02–Rec.09 (para. 73) The PRIPHC02 ***RECOMMENDED*** *that observer coverage be adjusted to be commensurate with the level of fishing intensity in each IPHC Regulatory Area.*

Conservation and Management: Consistency between scientific advice and fishery Regulations adopted

PRIPHC02–Rec.10 (para. 82) The PRIPHC02 ***RECOMMENDED*** *that the development of MSE to underpin multi-year (strategic) decision-making be continued, and as multi-year decision making is implemented, current Secretariat capacity usage for annual stock assessments should be refocused on research to investigate MSE operating model development (including consideration of biological and fishery uncertainties) for future MSE iterations and regularised multi-year stock assessments.*

PRIPHC02–Rec.11 (para. 83) The PRIPHC02 ***RECOMMENDED*** *that ongoing work on the MSE process be prioritised to ensure there is a management framework/procedure with minimal room for ambiguous interpretation, and robust pre-agreed mortality limit setting frameworks.*

The work outlined in this document builds on the previous a 5-Year Biological and Ecosystem Science Research Plan (IPHC–2019–BESRP-5YP), closing completed projects, extending efforts where needed, and adding new avenues in response to new information. Appendix I provides a detailed summary of the previous plan and the status of the work specifically undertaken. Key highlights relevant to the stock assessment and MSE include:

- Completion of the genetic assay for determining sex from tissue samples, processing of commercial fishery samples collected during 2017-2020, inclusion of this information in the 2019 and subsequent stock assessments, and transfer of this effort from research to ongoing monitoring.
- Incremental progress toward population-level sampling and analysis of maturity and fecundity.
- Continued development of the understanding of physiological and environmental mechanisms determining growth for future field application.
- Published estimates of discard mortality rates for use in data processing and management accounting.
- Collection of genetic samples and genome sequencing to provide a basis for ongoing evaluation of stock structure at population-level and finer scales.

All previously described research areas continue to represent critical areas of uncertainty in the stock assessment and thus are closely linked to management performance. The previous 5-year plan was successful in either providing direct new information to the stock assessment or building the foundation for the collection/analysis of such information in this updated plan. As noted below, some new priorities have emerged, and others have evolved based on the work completed to date. The incorporation of research objectives in the 5YPIRM that address climate change as a factor influencing Pacific halibut biology and ecology as well as fishery performance and dynamics constitutes a timely and relevant contribution towards advancing IPHC-led research to the forefront of fisheries science.

An **overarching goal** of the *IPHC 5-Year Program of integrated research and monitoring (2022-26)* is therefore to promote integration and synergies among the various research and support activities of the IPHC Secretariat in order to improve our knowledge of key inputs into the Pacific halibut stock assessment and MSE processes, in order to provide the best possible advice for management decision-making processes.

Along with the implementation of the short- and medium-term activities contemplated in this *IPHC 5-Year Program of Integrated Research and monitoring (2022-26)*, and in pursuit of the overarching objective, the IPHC Secretariat will also aim to:

1) undertake cutting-edge research programs in fisheries research in support of fisheries management of Pacific halibut;
2) undertake groundbreaking methodological research;
3) undertake applied research;
4) establish new collaborative agreements and interactions with research agencies and academic institutions;
5) promote the international involvement of the IPHC by continued and new participation in international scientific organizations and by leading international science and research collaborations.
6) effectively communicate IPHC research outcomes
7) incorporate talented students and early researchers in research activities contemplated.

The research and monitoring activities conducted by the IPHC Secretariat are directed towards fulfilling the following four (4) **objectives** within areas of data collection, biological and ecological research, stock assessment, and MSE. In addition, the IPHC responds to Commission requests for additional inputs to management and policy development which are classified under management support. The overall aim is to provide a program of integrated research and monitoring (Fig 2):

Research

1) **Stock assessment**: apply the resulting knowledge to improve the accuracy and reliability of the current stock assessment and the characterization of uncertainty in the resultant stock management advice provided to the Commission;
2) **Management Strategy Evaluation (MSE)**: to develop an accurate, reliable, and informative MSE process to appropriately characterize uncertainty and provide for the robust evaluation of the consequences of alternative management options, known as harvest strategies, using defined conservation and fishery objectives;
3) **Biology and Ecology**: identify and assess critical knowledge gaps in the biology and ecology of Pacific halibut within its known range, including the influence of environmental conditions on population and fishery dynamics;

Monitoring

4) **Monitoring**: collect representative fishery dependent and fishery-independent data on the distribution, abundance, biology, and demographics of Pacific halibut through ongoing monitoring activities;

Integrated management support

5) **Additional management-supporting inputs**: respond to Commission requests for any additional information supporting management and policy development.

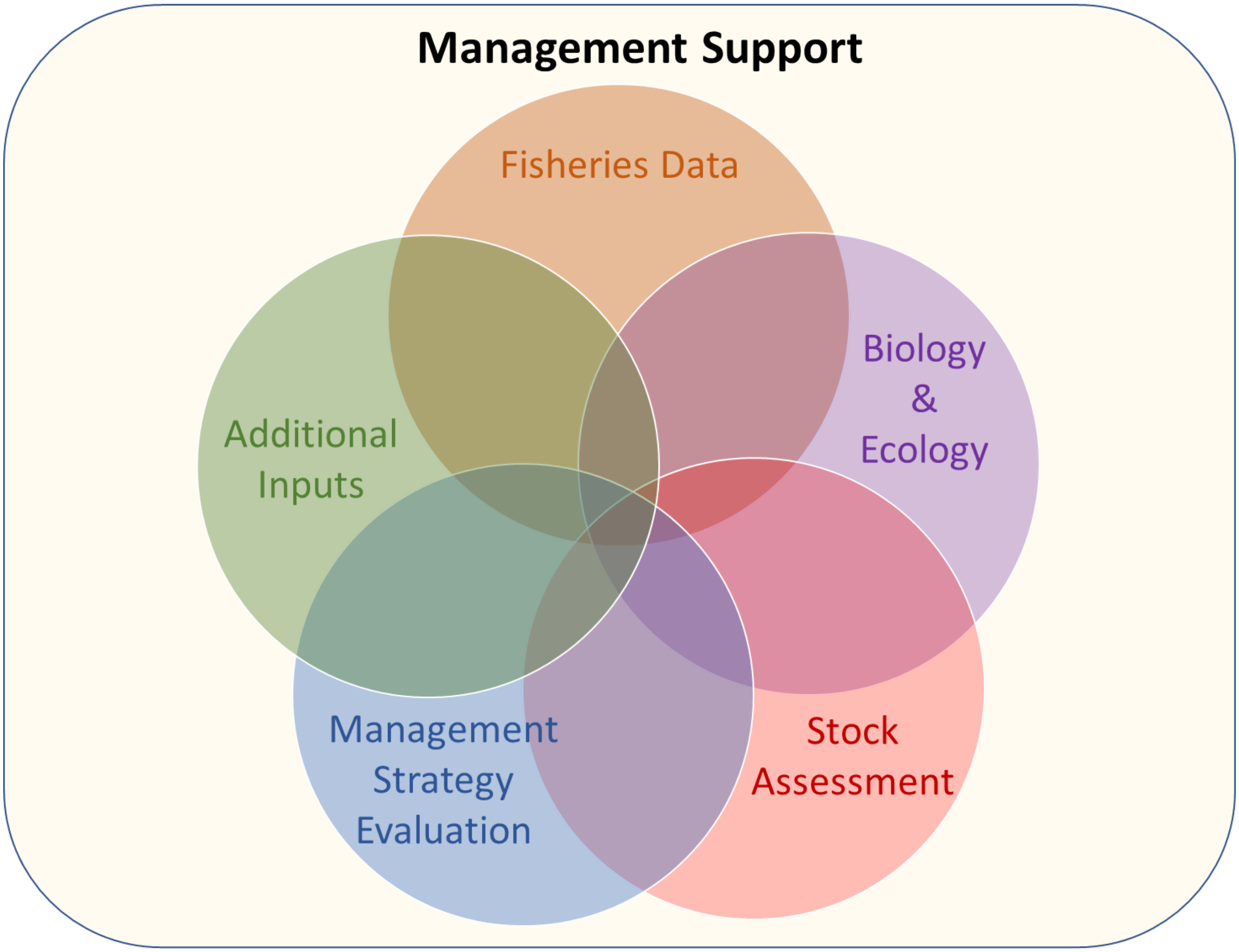


Figure 2. Core areas of the IPHC's program of integrated research and monitoring providing management support.

3. Strategy

The IPHC Secretariat has five (5) enduring strategic goals in executing our mission, including our overarching goal and associated science and research objectives, as articulated in our Strategic Plan (IPHC Strategic Plan (2019-23)): 1) To operate in accordance with international best practice; 2) Be a world leader in scientific excellence and science-based decision making; 3) To foster collaboration (within Contracting Parties and internationally) to enhance our science and management advice; 4) Create a vibrant IPHC culture; and 5) Set the standard for fisheries commissions globally.

Although priorities and tasking will change over time in response to events and developments, the Strategic Plan provides a framework to standardise our approach when revising or setting new priorities and tasking. The Strategic goals as they apply to the science and research activities of the IPHC Secretariat, will be operationalised through a multi-year tactical activity matrix at the organisational and management unit (Branch) level (Fig. 3). The tactical activity matrix is described in the sections below and has been developed based on the core needs of the Commission, in developing and implementing robust, scientifically-based management decisions on an annual, and multi-year level. Relevant IPHC subsidiary bodies will be involved in project development and ongoing review.

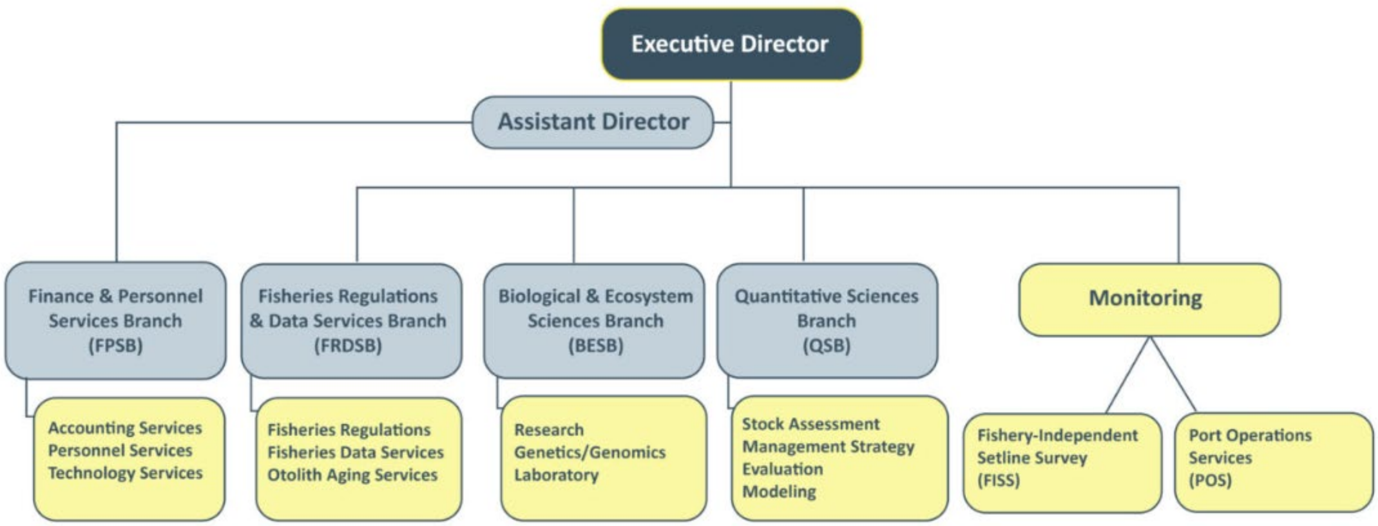


Figure 3. IPHC Secretariat organisation chart (2023).

4. Measures of Success

The Secretariat's success in implementing the *IPHC 5-Year Program of Integrated Research and Monitoring (2022-26)* will be measured according to the following criteria relevant to the stock assessment, the MSE and for all inputs to IPHC management:

1) Timeliness – was the research conducted, analyzed, published, and provided to the Commission at the appropriate points to be included in annual management decisions?
2) Accessibility – was the research published and presented in such a way that it was available to other scientists, stakeholders, and decision-makers?
3) Relevance - did the research improve the perceived accuracy of the stock assessment, MSE or decisions made by the commission?
4) Impact – did the research allow for more precision or a better estimate of the uncertainty associated with information for use in management?
5) Reliability - has research resulted in more consistent information provided to the Commission for decision-making.

4.1 Delivery of specified products

Each project line item will contain specific deliverables that constitute useful inputs into the stock assessment and the management strategy evaluation process, as well as support their implementation in the decision-making process at the level of the Commission.

4.2 Communication

The IPHC Secretariat will disseminate information about the activities contemplated in the IPHC 5-Year Program of Integrated Research and Monitoring (2022-2026) and the resulting products to Contracting Parties, stakeholders, the scientific community, and the general public through a variety of channels:

1) IPHC website (www.iphc.int);
2) Formal documentation provided for IPHC meetings (Interim and Annual Meetings, Subsidiary Body meetings, etc.);
3) Presentations at national and international scientific conferences;
4) Published reports and peer-reviewed publications (section 4.4);
5) Outreach events;
6) Social media outlets (e.g. Facebook, Twitter, LinkedIn, etc.);
7) Informal presentations and interactions with partners, stakeholders, and decision-makers at varied times and venues when needed.

4.3 External research funding

The Secretariat has set a funding goal of at least 20% of the funds for this program to be sourced from external funding bodies on an annual basis. Continuing the successful funding-recruitment strategy adopted during the previous 5-yr research plan (IPHC–2019–BESRP-5YP) (Appendix I), the Secretariat will identify and select external funding opportunities that are timely and that aim at addressing key research objectives (as outlined in Appendix II and summarized in Appendix V) that have important implications for stock assessment and the MSE process. The IPHC Secretariat has the necessary expertise to propose novel and important research questions to funding agencies and to recruit external collaborators from research agencies and universities as deemed necessary. The IPHC Secretariat will continue to capitalize on the strong analytical contributions of quantitative scientists to the development of biological research questions within the framework of research projects funded by external as well as internal funding sources.

4.4 Peer-reviewed journal publication

Publication of research outcomes in peer-reviewed journals will be clearly documented and monitored as a measure of success. This may include single publications at the completion of a particular project, or a series of publications throughout the project as well as at its completion. Each sub-project shall be published in a timely manner and shall be submitted no later than 12 months after the end of the research. In the sections that follow, the expected publications from each research stream and cross-stream are defined.

5. Core focal areas – Background

The goals of the main activities of the *5-Year program of integrated research and monitoring (2022-26)* are integrated across the organisation, involving 1) monitoring (fisheries-dependent and –independent data collection), and 2) research (biological, ecological), modelling (FISS and stock assessment), and MSE, as outlined in the following sub-sections. These components are closely linked to one another, and all feed into management decision-making (Fig. 4). Additionally, management-supporting information constitute a range of additional decision-making drivers within and beyond IPHC's current research and monitoring programs. The current program builds on the outcomes and experiences of the Commission arising from the implementation of the 2017-21 5-Year Biological and Ecosystem Science Research Plan (IPHC–2019–BESRP-5YP), and which is summarized in Appendix I.

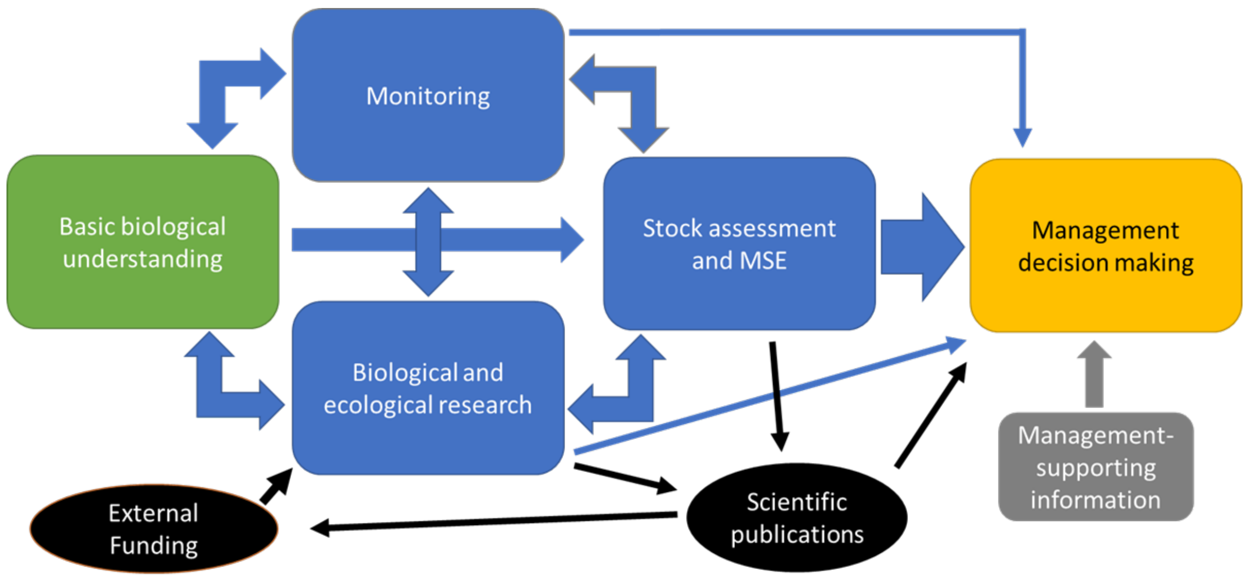


Figure 4. Flow of information from basic biological understanding of the Pacific halibut resource, through IPHC research components (monitoring, biological and ecological research, stock assessment, and MSE) to management decision-making. Management-supporting information (grey) constitute a range of additional decision-making drivers within and beyond IPHC's current research and monitoring programs. Arrows indicate the strength (size of the arrow) and direction of information exchange. Also identified (in black) are the external links from funding and scientific publications which supplement the IPHC's internal process.

5.1 Research

5.1.1 Stock Assessment

Focal Area Objective	To improve accuracy and reliability of the current stock assessment and the characterization of uncertainty in the resultant stock management advice provided to the Commission.
IPHC Website portal	https://www.iphc.int/management/science-and-research/stock-assessment

The IPHC conducts an annual stock assessment, using data from the fishery-independent setline survey (FISS), the commercial Pacific halibut and other fisheries, as well biological information from its research program. The assessment includes the Pacific halibut resource in the IPHC Convention Area, covering the Exclusive Economic Zones of Canada and the United States of America. Data sources are updated each year to reflect the most recent scientific information available for use in management decision-making.

The 2021 stock assessment relied on an ensemble of four population dynamics models to estimate the probability distributions describing the current stock size, trend, and demographics. The ensemble is designed to capture both uncertainty related to the data and stock dynamics (due to estimation) as well as uncertainty related to our understanding of the way in which the Pacific halibut stock functions and is best approximated by a statistical model (structural uncertainty).

Stock assessment results are used as inputs for harvest strategy calculations, including mortality projection tables for the upcoming year that reflect the IPHC's harvest strategy policy and other considerations, as well as the harvest decision table which provides a direct tool for the management process. The harvest decision table uses the probability distributions from short-term (three year) assessment projections to evaluate the trade-offs between alternative levels of potential yield (catch) and the associated risks to the stock and fishery.

The stock assessment research priorities have been subdivided into four categories:

1) Assessment data collection and processing;
2) technical development;
3) biological inputs; and
4) fishery yield.

It is important to note that ongoing monitoring, including the annual FISS and directed commercial landings sampling programs is not considered research and is therefore not included in this research priority list despite the critical importance of these collections. These are described in the sections below.

5.1.2 Management Strategy Evaluation (MSE)

Focal Area Objective	To develop an accurate, reliable, and informative MSE process to appropriately characterize uncertainty and provide for the robust evaluation of the consequences of alternative management options, known as harvest strategies, using defined conservation and fishery objectives.
IPHC Website portal	https://www.iphc.int/management/science-and-research/management-strategy-evaluation

Management Strategy Evaluation (MSE) is a process to evaluate the consequences of alternative management options, known as harvest strategies. MSE uses a simulation tool to determine how alternative harvest strategies perform given a set of pre-defined fishery and conservation objectives, taking into account the uncertainties in the system and how likely candidate harvest strategies are to achieve the chosen management objectives.

MSE is a simulation technique based on modelling each part of a management cycle. The MSE uses an operating model to simulate the entire population and all fisheries, factoring in management decisions, the monitoring program, the estimation model, and potential ecosystem effects using a closed-loop simulation.

Undertaking an MSE has the advantage of being able to reveal the trade-offs among a range of possible management decisions. Specifically, to provide the information on which to base a rational decision, given harvest strategies, preferences, and attitudes to risk. The MSE is an essential part of the process of developing, evaluating and agreeing to a harvest strategy.

The MSE process involves:

- Defining fishery and conservation objectives with the involvement of stakeholders and managers;
- Identifying harvest strategies (a.k.a. management procedures) to evaluate;
- Simulating a Pacific halibut population using those harvest strategies;
- Evaluating and presenting the results in a way that examines trade-offs between objectives;
- Applying a chosen harvest strategy for the management of Pacific halibut;
- Repeating this process in the future in case of changes in objectives, assumptions, or expectations.

There are many tasks that would continue to improve the MSE framework and the presentation of future results to the Commission. The tasks can be divided into five general categories, which are common to MSE in general:

1. **Objectives**: The goals and objectives that are used in the evaluation.
2. **Management Procedures (MPs)**: Specific, well-defined management procedures that can be coded in the MSE framework to produce simulated Total Constant Exploitation Yields (TCEY) for each IPHC Regulatory Area.
3. **Framework**: The specifications and computer code for the closed-loop simulations including the operating model and how it interacts with the MP.
4. **Evaluation**: The performance metrics and presentation of results. This includes how the performance metrics are evaluated (e.g. tables, figures, and rankings), presented to the Commission and its subsidiary bodies, and disseminated for outreach.
5. **Application**: Specifications of how an MP may be applied in practice and re-evaluated in the future, including responses to exceptional circumstances.

All these categories provide inputs and outputs of the MSE process, but the Framework category benefits most from the integration of biological and ecosystem research because the operating model, the simulation of the monitoring program, the estimation model, and potential ecosystem effects are determined from this knowledge.

Outcomes of the MSE process will not only inform the Commission on trade-offs between harvest strategies and assist in choosing an optimal strategy for management of the Pacific halibut resource but will inform the prioritization of research activities related to fisheries monitoring, biological and ecological research, stock assessment, and fishery socioeconomics.

5.1.3 Biology and Ecology

Focal Area Objective	To identify and assess critical knowledge gaps in the biology and ecology of Pacific halibut within its known range, including the influence of environmental conditions on population and fishery dynamics.
IPHC Website portal	https://www.iphc.int/management/science-and-research/biological-and-ecosystem-science-research-program-bandesrp

Since its inception, the IPHC has had a long history of research activities devoted to describe and understand the biology of the Pacific halibut. At present, the main objectives of the Biological and Ecosystem Science Research Program at IPHC are to: 1) identify and assess critical knowledge gaps in the biology of the Pacific halibut; 2) understand the influence of environmental conditions in the biology of the Pacific halibut and its fishery; and 3) apply the resulting knowledge to reduce uncertainty in current stock assessment models.

The primary biological research activities at the IPHC that follow Commission objectives and that are selected for their important management implications are identified and described in the proposed IPHC 5-Year Program of Integrated Research and Monitoring (2022-2026). An overarching goal of the 5-Year Program of Integrated Research and Monitoring (2022-2026) is to promote integration and synergies among the various research activities led by the IPHC to improve our knowledge of key biological inputs that feed into the stock assessment and MSE process. The goals of the main research activities of the 5-Year Program of Integrated Research and Monitoring (2022-2026) are therefore aligned and integrated with the IPHC stock assessment and MSE processes. The IPHC Secretariat conducts research activities to address key biological issues based on the IPHC Secretariat's own input as well as input from the IPHC Commissioners, stakeholders and particularly from specific subsidiary bodies to the IPHC, including the Scientific Review Board (SRB) and the Research Advisory Board (RAB).

The biological research activities contemplated in the 5-Year Program of Integrated Research and Monitoring (2022-2026) and their specific aims are detailed in Section 6. Overall, the biological research activities at the IPHC aim to provide information on 1) factors that influence the biomass of the Pacific halibut population (e.g. distribution and movement of fish among IPHC Regulatory Areas, growth patterns and environmental influences on growth in larval, juvenile and adult fish, drivers of changes in size-at-age); 2) the spawning (female) population (e.g. reproductive maturity, skipped spawning, reproductive migrations); and 3) resulting changes in population dynamics. Furthermore, the research activities of IPHC also aim to provide information on the survival of regulatory-discarded Pacific halibut in the directed fisheries with the objective to refine current estimates of discard mortality rates and develop best handling practices, and reduce whale depredation and Pacific halibut bycatch through gear modifications and through a better understanding of behavioral and physiological responses of Pacific halibut to fishing gear. The proposed timeline of activities and of staffing and funding indicators are provided in Appendix VI and Appendix VII, respectively.

5.2 Monitoring

Focal Area Objective	To collect fishery-dependent and fishery-independent data on the distribution, abundance, and demographics of Pacific halibut, as well as other key biological data, through ongoing monitoring activities.
IPHC Website portal	***Fishery-dependent data:*** • https://www.iphc.int/datatest/commercial-fisheries • https://www.iphc.int/data/datatest/pacific-halibut-recreational-fisheries-data • https://www.iphc.int/datatest/subsistence-fisheries • https://www.iphc.int/data/time-series-datasets ***Fishery-independent data:*** • https://www.iphc.int/management/science-and-research/fishery-independent-setline-survey-fiss • https://www.iphc.int/data/datatest/fishery-independent-setline-survey-fiss • https://www.iphc.int/datatest/data/water-column-profiler-data

5.2.1 Fishery-dependent data

The IPHC estimates all Pacific halibut removals taken in the IPHC Convention Area and uses this information in its yearly stock assessment and other analyses. The data are compiled by the IPHC Secretariat and include data from Federal and State agencies of each Contracting Party. Specific activities in this area are described below.

5.2.1.1 Directed commercial fisheries data

The IPHC Secretariat collects logbooks, otoliths, tissue samples, and associated sex-length-weight data from directed commercial landings coastwide (Fig. 5). A sampling rate is determined for each port by IPHC Regulatory Area. The applicable rate is calculated from the current year's mortality limits and estimated percentages of weight of fish landed, and estimated percentages of weight sampled in that port to allow for collection of the target number of biological samples by IPHC Regulatory Area. An example of the data collected and the methods used are provided in the annually updated directed commercial sampling manual (e.g. IPHC Directed Commercial Landings Sampling Manual 2022). Directed commercial fishery landings are recorded by the Federal and State agencies of each Contracting Party and summarized each year by the IPHC. Discard mortality for the directed commercial fishery is currently estimated using a combination of research survey (U.S.A.) and observer data (Canada).

5.2.1.2 Non-directed commercial discard mortality data

The IPHC accounts for non-directed commercial discard mortality by IPHC Regulatory Area and sector. Non-directed commercial discard mortality estimates are provided by State and Federal agencies of each Contracting Party and compiled annually for use in the stock assessment and other analyses.

Non-directed commercial discard mortality of Pacific halibut is estimated because not all fisheries have 100% monitoring and not all Pacific halibut that are discarded are assumed to die. The IPHC relies upon information supplied by observer programs run by Contracting Party agencies for non-directed commercial discard mortality estimates in most fisheries. Non-IPHC research survey information or other sources are used to generate estimates of non-directed commercial discard mortality in the few cases where fishery observations are unavailable. Non-directed fisheries off Canada British Columbia are monitored and discard mortality information is provided to IPHC by DFO. NOAA Fisheries operates observer programs off the USA West Coast and Alaska, which monitor the major groundfish fisheries. Data collected by those programs are used to estimate non-directed commercial discard mortality.

5.2.1.3 Subsistence fisheries data

Subsistence fisheries are non-commercial, customary, and traditional use of Pacific halibut for direct personal, family, or community consumption or sharing as food, or customary trade. The primary subsistence fisheries are the treaty Indian Ceremonial and Subsistence fishery in IPHC Regulatory Area 2A off northwest Washington State (USA), the First Nations Food, Social, and Ceremonial (FSC) fishery in British Columbia (Canada), and the subsistence fishery by rural residents and federally recognized native tribes in Alaska (USA) documented via Subsistence Halibut Registration Certificates (SHARC). Subsistence fishery removals of Pacific halibut, including estimated subsistence discard mortality, are provided by State and Federal agencies of each Contracting Party, estimated, and compiled annually for use in the stock assessment and other analysis.

5.2.1.4 Recreational fisheries data

Recreational removals of Pacific halibut, including estimated recreational discard mortality, are provided by National/State agencies of each Contracting Party, estimated, and compiled annually for use in the stock assessment and other analysis.

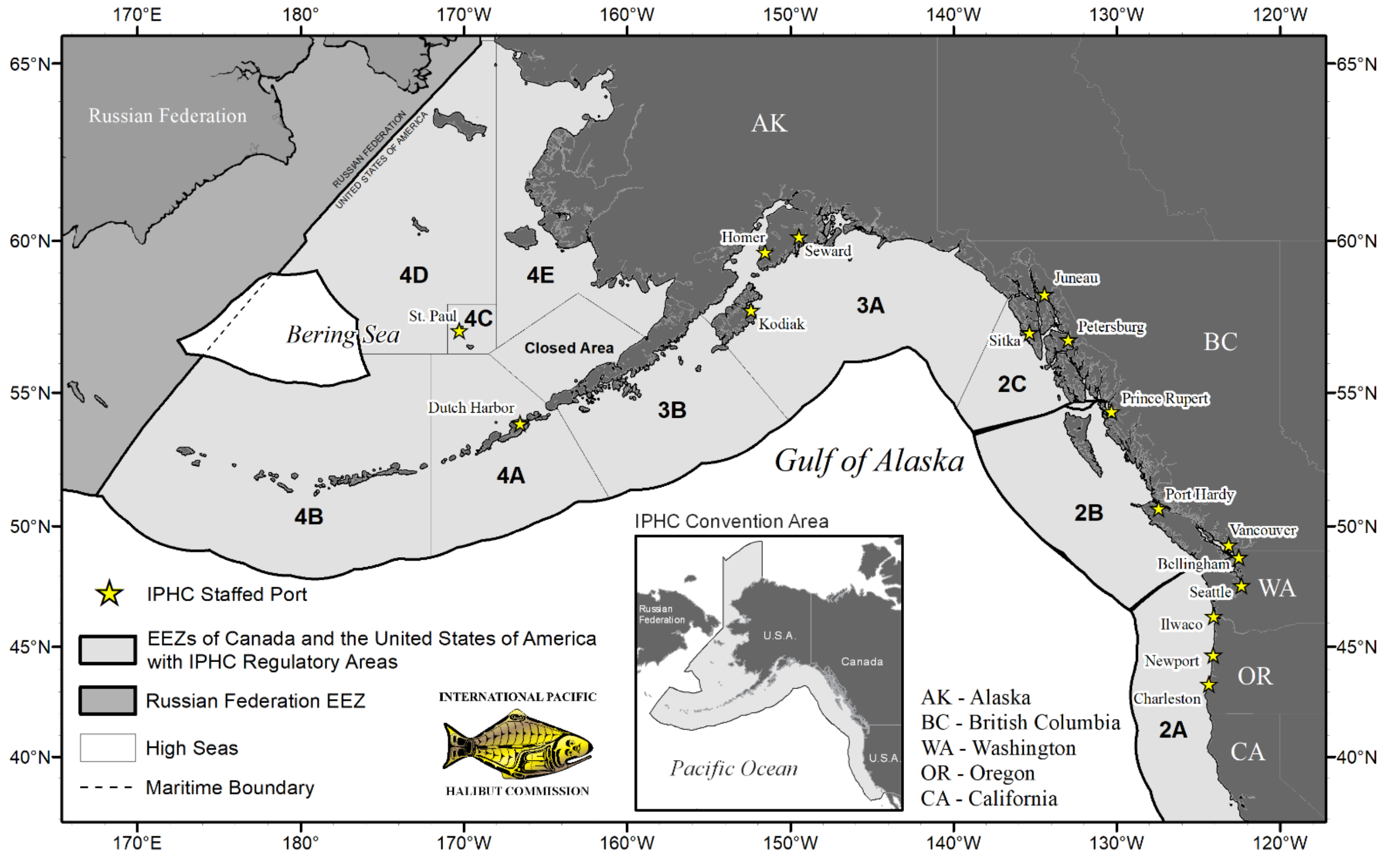


Figure 5. Ports where the IPHC has sampled directed commercial landings throughout the fishing period in recent years (note: ports sampled may change from year-to-year for operational reasons).

5.2.2 Fishery-independent data

Data collection and monitoring activities aimed at providing a standardised time-series of biological and ecological data that is independent of the fishing fleet.

5.2.2.1 Fishery-independent setline survey (FISS)

The IPHC Fishery-Independent Setline Survey (FISS) provides catch-rate information and biological data on Pacific halibut that are independent of the fishery. These data, collected using standardized methods, bait, and gear, are used to estimate the primary index of population abundance used in the stock assessment. The FISS is restricted to the summer months but encompasses the commercial fishing grounds in the Pacific halibut fishery, and almost all known Pacific halibut habitat in Convention waters outside the Bering Sea. The standard FISS grid totals 1,890 stations (Fig. 6). Biological data collected on the FISS (e.g. the length, weight, age, and sex of Pacific halibut) are used to monitor changes in biomass, growth, and mortality. In addition, records of non-target species caught during FISS operations provide insight into bait competition, and serve as an index of abundance over time, making them valuable to the potential management and avoidance of non-target species. Environmental data are also collected including water column temperature, salinity, dissolved oxygen, pH, and chlorophyll concentration to help identify the conditions in which the fish were caught, and these data can serve as co-variates in space-time modeling used in the stock assessment. An example of the data collected and the methods used are provided in the annually updated FISS sampling manual (e.g. IPHC FISS Sampling Manual 2022).

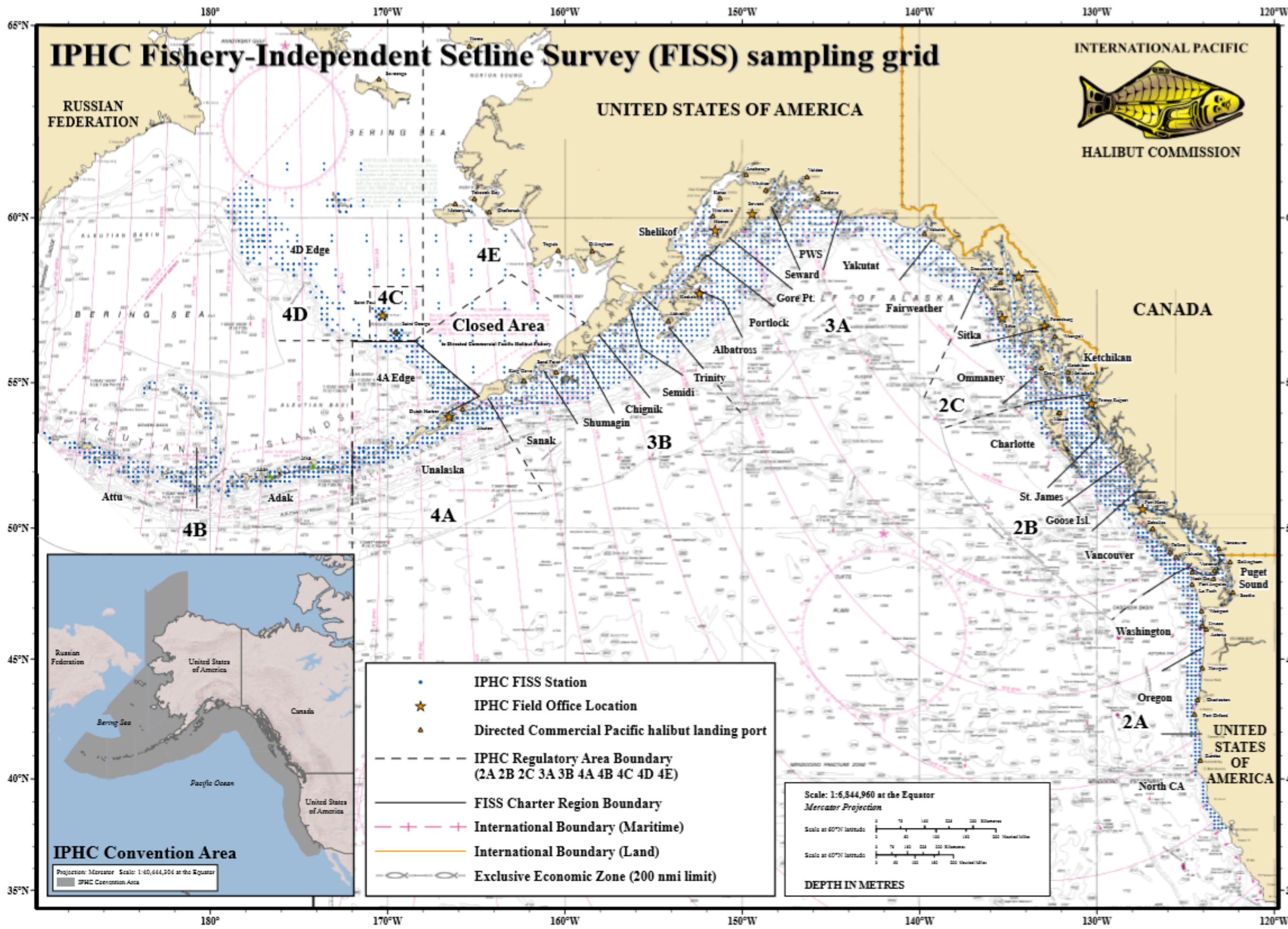


Figure 6. IPHC Fishery-Independent Setline Survey (FISS) with full sampling grid shown.

Quality control and sampling rate estimations: Following a program of planned FISS expansions from 2014-19, a process of rationialisation of the FISS was undertaken. The goal was to ensure that, given constraints on resources available for implementing the FISS, station selection was such that density indices would be estimated with high precision and low potential for bias. An annual design review process has been developed during which potential FISS designs for the subsequent three years are evaluated according to precision and bias criteria. The resulting proposed designs and their evaluation are presented for review at the June Scientific Review Board (SRB) meetings and potentially modified following SRB input before presentation to the Commissioners at the Work Meeting and Interim Meeting. Annual biological sampling rates for each IPHC Regulatory Area are calculated based on the previous year's catch rates and an annual target of 2000 sampled fish (with 100 additional archive samples).

5.2.2.2 Fishery-independent Trawl Survey (FITS)

The IPHC has participated routinely in the NOAA Fisheries trawl surveys operating in the Bering Sea (Fig. 7, annually since 1998), Aleutian Islands (intermittently since 1997) and Gulf of Alaska (since 1996). The information collected from Pacific halibut caught on these surveys, together with data from the IPHC Fishery-Independent Setline Survey (FISS) and commercial Pacific halibut data, are used directly in estimating indices of abundance and in the stock assessment and to monitor population trends, growth/size, and to supplement understanding of recruitment, distribution, and age composition of young Pacific halibut.

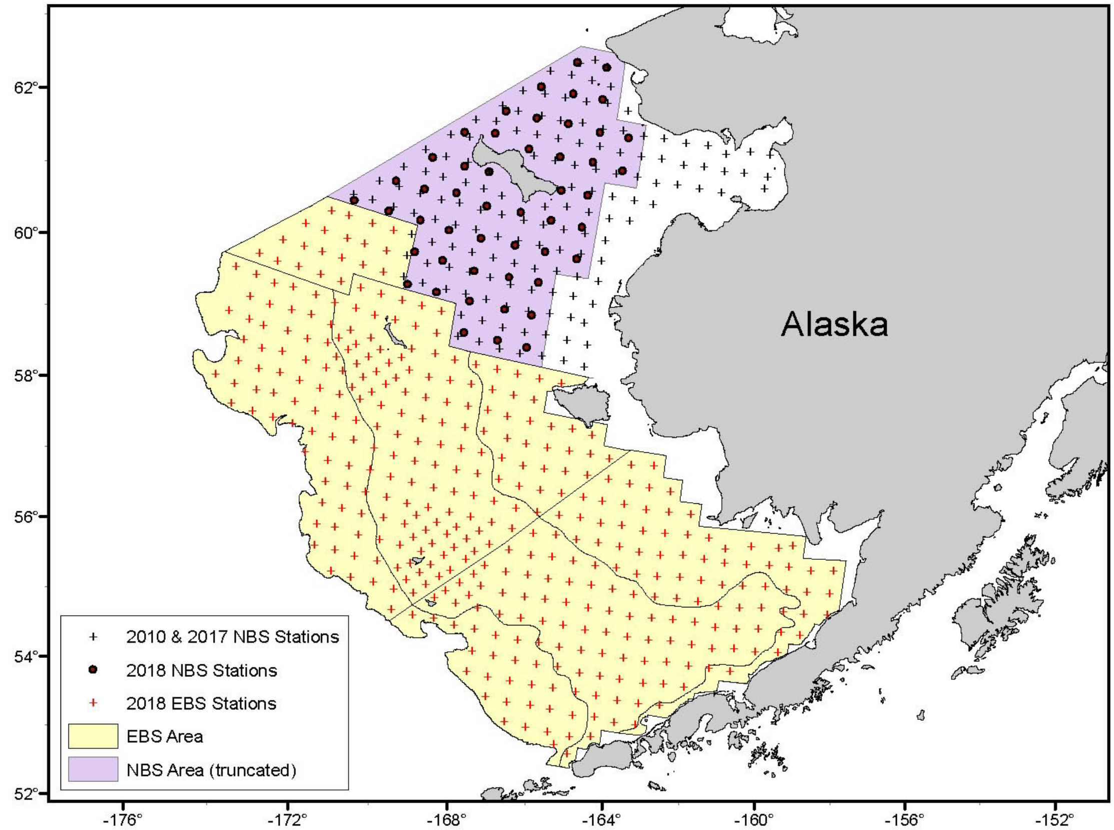


Figure 7. Sampling station design for the 2018 NOAA Bering Sea bottom trawl survey. Black dots are stations sampled in the 2018 "rapid-response" Northern Bering Sea trawl survey and black plus signs are stations sampled in standardized Northern Bering Sea trawl survey.

5.2.2.3 Norton Sound trawl survey

The Alaska Department of Fish and Game's annual Norton Sound trawl survey data contribute to the estimation of Pacific halibut indices of abundance in IPHC Regulatory Area 4E.

5.2.3 Age composition data (both fishery-dependent and fishery-independent)

The annually collected biological samples from commercial fisheries and FISS include otoliths, a crystalline calcium carbonate structure found in the inner ear of fish which growth patterns can be analyzed to estimate the age of fish. Fish age is a key input to stock assessment models that inform management decisions related to fish exploitation. Since inception, the IPHC aged over 1.5 million otoliths manually by trained readers under the stereoscopic microscope.

5.3 Management-supporting information

Successful fisheries management requires rigorous application of the scientific method of problem solving in the development of strategic alternatives and their evaluation on the basis of objectives that integrate ecosystem and human dynamics across space and time into management decision-making (Lane and Stephenson, 1995). This underscores the importance of a holistic understanding of a broad range of factors to deliver on the Commission's objective to develop the stocks of Pacific halibut to the levels that permit the optimum yield from the fishery over time. Management-supporting information beyond IPHC's current research and monitoring programs relate to, among others, socioeconomic considerations, community development, political constraints, and operational limitations.

Responding to the Commission's "*desire for more comprehensive economic information to support the overall management of the Pacific halibut resource in fulfillment of its mandate*" (economic study terms of reference adopted at FAC095 and endorsed at AM095 in 2019), between 2019 and 2021 the IPHC conducted a socioeconomic study. The study's core product, Pacific halibut multiregional economic impact assessment (PHMEIA) model, describes economic interdependencies between sectors and regions to bring a better understanding of the role and importance of the Pacific halibut resource to regional economies of Canada and the United States of America (see project report). The model details the within-region production structure of the Pacific halibut sectors (fishing, processing, charter) and cross-regional flows of economic benefits. The model also accounts for economic activity generated through sectors that supply fishing vessels, processing plants, and charter businesses with inputs to production, by embedding Pacific halibut sectors into the model of the entire economy of Canada and the USA. The PHMEIA model fosters stakeholders' better understanding of a broad scope of regional impacts of the Pacific halibut resource. The results highlight that the harvest stage accounts for only a fraction of economic activity that would be forgone if the resource was not available to fishers in the Pacific Northwest. Moreover, the study informs on the vulnerability of communities to changes in the state of the Pacific halibut stock throughout its range, highlighting regions particularly dependent on economic activities that rely on Pacific halibut. Leveraging multiple sources of socioeconomic data, the project provides complementary input for designing policies with desired effects depending on regulators' priorities which may involve balancing multiple conflicting objectives. A good understanding of the localized effects is pivotal to policymakers who are often concerned about community impacts, particularly in terms of impact on employment opportunities and households' welfare.

The economic impact assessment is supplemented by an analysis of the formation of the price paid for Pacific halibut products by final consumers (end-users) that is intended to provide a better picture of Pacific halibut contribution to the gross domestic product (GDP) along the entire value chain, from the hook-to-plate. This supplemental material is available in IPHC's Pacific halibut market analysis.

6. Core focal areas – Planned and opportunistic activities (2022-2026)

Research at IPHC can be classified as "use-inspired basic research" (Stokes 1997) which combines knowledge building with the application of existing and emerging knowledge to provide for the management of Pacific halibut. The four core focal areas: stock assessment, management strategy evaluation, management supporting information, and biology & ecology, all interact with each other as well as with fisheries monitoring activities in the IPHC program of integrated research and monitoring. Progress and knowledge building in one focal area influences and informs application in other core focal areas, also providing insight into future research priorities. The circular feedback loop is similar to the scientific method of observing a problem, creating a hypothesis, testing that hypothesis through research and analysis, drawing conclusions, and refining the hypothesis.

The IPHC Secretariat has been working with IPHC advisory bodies, such as the Scientific Review Board (SRB), and the Commission to conduct scientific research in a way that utilizes the scientific method. Problems are often identified by an advisory body or Commission and hypotheses are developed by the IPHC Secretariat. Research is reviewed by the SRB and refined hypotheses are presented to advisory bodies and the Commission. This process occurs via an annual schedule of meetings, as shown in Fig. 8. In May, an MSE informational session may be held if there is significant progress in the MSE such that it would be useful to prepare stakeholders for the Management Strategy Advisory Board (MSAB) meeting in October. Recommendations related to the MSE, and development of a harvest strategy directed to the Commission are a result of the MSAB meeting. The SRB holds two meetings each year: one in June where requests are typically directed to IPHC Secretariat, and one in September where recommendations are made to the Commission. The June SRB meeting has a focus on research;

the September meeting represents a final check of science products to be presented to the Commission for use in management. The Research Advisory Board (RAB) meets in November to discuss ongoing research, provide guidance and recommend new research projects. The Work Meeting (WM) is held in September and is a working session with IPHC Secretariat and the Commission to prepare for the Interim Meeting (IM) held in November and the Annual Meeting (AM) held in January. Outcomes from the AM include mortality limits (coastwide and by IPHC Regulatory Area), directed fishery season dates, domestic regulations, and requests and recommendations for the IPHC Secretariat. In conjunction with the AM are meetings of the Finance and Administration Committee (FAC), the Conference Board (CB), and the Processor Advisory Board (PAB). The Commission may also hold Special Sessions (SS) throughout the year to take up and make decisions on specific topics.

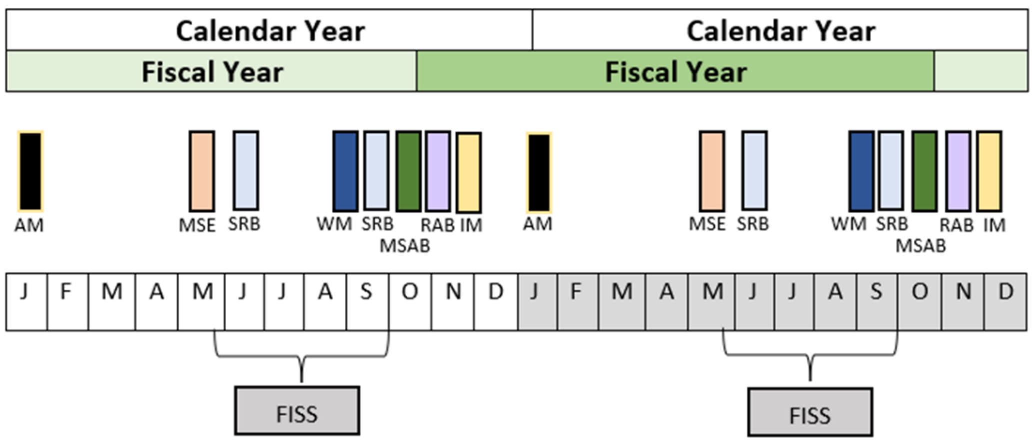


Figure 8. The typical IPHC annual meeting schedule with the calendar year and fiscal year shown. The meetings, shown in the middle row are: Annual Meeting where the Commission makes many final decisions for that year (AM), an MSE informational session (MSE), Scientific Review Board meetings (SRB), the Commission Work Meeting (WM), the Management Strategy Advisory Board meeting (MSAB), the Research Advisory Board Meeting (RAB), and the Interim Meeting (IM). The annual FISS schedule is also shown.

In addition to the annual meeting process at IPHC, individual core focal areas of research may identify and prioritize research for other core focal areas. For example, stock assessment research often identifies gaps in the knowledge of Pacific halibut biology and ecology, which then identifies priority research for the Biology and Ecology core area. Vice versa, basic biological and ecological research can identify concepts that could be better understood and result in improved implementation in any of the core areas. Furthermore, Management Strategy Evaluation can often be used to identify priority research topics for any core areas by simulation testing to identify research that may have the largest benefit to improving the management of Pacific halibut.

The top priorities of research for various categories in each of the core focal areas are provided below. The top priorities are a subset of the potential research topics in each core focal area. More exhaustive and up-to-date lists of research topics, that may extend beyond a five-year timeframe, can be found in recent meeting documents related to each core focal area.

6.1 Research

6.1.1 Stock Assessment

Within the four assessment research categories, the following topics have been identified as top priorities in order to focus attention on their importance for the stock assessment and management of Pacific halibut. A brief narrative is provided here to highlight the specific use of products from these studies in the stock assessment.

6.1.1.1 Stock Assessment data collection and processing

6.1.1.1.1 Commercial fishery sex-ratio-at-age via genetics

Commercial fishery sex-ratio information has been found to be closely correlated with the absolute scale of the population estimates in the stock assessment and has been identified as the greatest source of uncertainty since 2013. With only four years (2017-20) of commercial sex-ratio-at-age information available for the 2021 stock assessment, the annual genetic assay of fin clips sampled from the landings remains critically important. When the time series grows longer, it may be advantageous to determine the ideal frequency at which these assays need to be conducted. Development of approaches to use archived otoliths, scales or other samples to derive historical estimates (if possible) could provide valuable information on earlier time-periods (with differing fishery and biological properties), and therefore potentially reconcile some of the considerable historical uncertainty in the present stock assessment. This assessment priority directly informs *6.1.3.2 Reproduction* as described below.

6.1.1.1.2 Whale depredation accounting and tools for avoidance

Whale depredation currently represents a source of unobserved and unaccounted-for mortality in the assessment and management of Pacific halibut. A logbook program has been phased in over the last several years, in order to record whale interactions observed by commercial harvesters. Estimation of depredation mortality, from logbook records and supplemented with more detailed data and analysis from the FISS represents a first step in accounting for this source of mortality; however, such estimates will likely come with considerable uncertainty. Reduction of depredation mortality through improved fishery avoidance and/or catch protection would be a preferable extension and/or solution to basic estimation. As such, research to provide the fishery with tools to reduce depredation is considered a closely-related high priority. This assessment priority directly informs *6.1.3.4 Mortality and Survival Assessment* as described below.

6.1.1.2 Stock Assessment technical development

6.1.1.2.1 Maintaining coordination with the MSE

The stock assessment and MSE operating models have been developed in close coordination, in order to identify plausible hypotheses regarding the processes governing Pacific halibut population dynamics. Important aspects of Pacific halibut dynamics include recruitment (possibly related to extrinsic environmental factors in addition to spawning biomass), size-at-age, movement/migration and spatial patterns in fishery catchability and selectivity. Many approaches developed as part of the tactical stock assessment have been explored in the MSE operating model, and conversely, the MSE operating model has highlighted areas of data uncertainty or alternative hypotheses for exploration in the assessment (e.g. movement rates). Although these two modelling efforts target differing objectives (tactical vs. strategic) continued coordination is essential to ensure that the stock assessment and the MSE represent the Pacific halibut similarly and provide consistent and useful advice for tactical and strategic decision-making.

6.1.1.2.2 Data weighting

The stock assessment currently relies on iterative "Francis" weighting of the age compositional data using a multinomial likelihood formulation (Francis 2011) based on the number of samples available in each year. Exploration of a stronger basis for input sample sizes through analysis of sampling design, estimation of sample weighting and alternative likelihoods may all provide for a more stable approach and a better description of the associated uncertainty.

6.1.1.2.3 Environmental covariates to recruitment

The two long time-series models included in the stock assessment ensemble allow for the Pacific Decadal

Oscillation (PDO; Mantua et al. 1997) to be a binary covariate indicating periods of higher or lower average recruitment. This relationship has been observed to be consistent since its development over 20 years ago (Clark et al 1999) and is re-estimated in each year's stock assessment models. With additional years of data, evaluation of the strength of this relationship, as compared to other metrics of the PDO (e.g., annual deviations, running averages) or other indicators of NE Pacific Ocean productivity should be undertaken in order to provide the best estimates and projections of Pacific halibut recruitment and to provide for alternative hypotheses for use in the MSE. This assessment priority partially informs *6.1.3.2 Reproduction* as described below.

6.1.1.2.4 'Leading' parameter estimation

Stock assessments are generally very sensitive to the estimates of leading parameters (stock-recruitment parameters, natural mortality, sex-specific dynamics, etc.). For Pacific halibut some of these are fully integrated into the estimation uncertainty (average unexploited recruitment), or partially integrated (e.g. estimation of natural mortality in two of the four models). As time-series of critically informative data sources like the FISS and the sex-ratio of the commercial landings grow longer it may be possible to integrate additional leading parameters directly in the assessment models and/or include them as nested models within the ensemble.

6.1.1.3 Stock Assessment biological inputs

6.1.1.3.1 Maturity, skip-spawning, and fecundity

Management of Pacific halibut is currently based on reference points that rely on relative female spawning biomass. Therefore, any changes to the understanding of reproductive output – either across age/size (maturity), over time (skip spawning) or as a function of body mass (fecundity) are crucially important. Each of these components directly affects the annual reproductive output estimated in the assessment. Ideally, the IPHC would have a program in place to monitor each of these three reproductive processes over time and use that information in the estimation of the stock-recruitment relationship, and the annual reproductive output relative to reference points. This would reduce the potential for biased time-series estimates created by non-stationarity in these traits (illustrated via sensitivity analyses in several of the recent assessments). However, at present we have only historical time-aggregated estimates of maturity and fecundity schedules. Therefore, the current research priority is to first update our estimates for each of these traits to reflect current environmental and biological conditions. After current stock-wide estimates have been achieved, a program for extending this information to a time-series via transition from research to monitoring can be developed. This assessment priority directly informs *6.1.3.2 Reproduction* as described below.

6.1.1.3.2 Stock structure of IPHC Regulatory Area 4B relative to the rest of the convention area

The current stock assessment and management of Pacific halibut assume that IPHC Regulatory Area 4B is functionally connected with the rest of the stock, i.e., that recruitment from other areas can support harvest in Area 4B and that biomass in Area 4B can produce recruits that may contribute to other Areas. Tagging (Webster et al. 2013) and genetic (Drinan et al. 2016) analyses have indicated the potential for Area 4B to be demographically isolated. An alternative to current assessment and management structure would be to treat Area 4B separately from the rest of the coast. This would not likely have a large effect on the coastwide stock assessment as Area 4B represents only approximately 5% of the surveyed stock (Stewart and Webster 2022). However, it would imply that the specific mortality limits for Area 4B could be very important to local dynamics and should be separated from stock-wide trends. Therefore, information on the stock structure for Area 4B has been identified as a top priority. This assessment priority directly informs *6.1.3.1 Migration and Population Dynamics* as described below.

6.1.1.3.3 Meta-population dynamics (connectivity) of larvae, juveniles, and adults

The stock assessment and current management procedure treat spawning output, juvenile Pacific halibut abundance, and fish contributing to the fishery yield as equivalent across all parts of the Convention Area. Information on the connectivity of these life-history stages could be used for a variety of improvements to the assessment and current management procedure, including: investigating recruitment covariates, structuring spatial assessment models, identifying minimum or target spawning biomass levels in each Biological Region, refining the stock-recruitment relationship to better reflect source-sink dynamics and many others. Spatial dynamics have been highlighted as a major source of uncertainty in the Pacific halibut assessment for decades and will continue to be of high priority until they are better understood. This assessment priority directly informs *6.1.3.1 Migration and Population Dynamics* as described below.

6.1.1.4 Stock Assessment fishery yield

6.1.1.4.1 Biological interactions with fishing gear

In 2020, 16% of the total fishing mortality of Pacific halibut was discarded (Stewart et al. 2021). Discard mortality rates can vary from less than 5% to 100% depending on the fishery, treatment of the catch and other factors (Leaman and Stewart 2017). A better understanding of the biological underpinnings for discard mortality could lead to increased precision in these estimates, avoiding potential bias in the stock assessment. Further, improved biological understanding of discard mortality mechanisms could allow for reductions in this source of fishing mortality, and thereby increased yield available to the fisheries. This assessment priority directly informs *6.1.3.4 Mortality and Survival Assessment* as described below.

6.1.1.4.2 Guidelines for reducing discard mortality

Much is already known about methods to reduce discard mortality, in non-directed fisheries as well as the directed commercial and recreational sectors. Promotion and adoption of best handling practices could reduce discard mortality, lead to greater retained yield, and reduce the potential uncertainty associated with large quantities of estimated mortality due to discarding. This assessment priority directly informs *6.1.3.4 Mortality and Survival Assessment* as described below.

Outside of the four general assessment categories, the IPHC has recently considered adding close-kin genetics (e.g., Bravington et al. 2016) to its ongoing research program (see section 6.1.3.1). Close-kin mark-recapture can potentially provide estimates of the absolute scale of the spawning output from the Pacific halibut population. This type of information can be fit directly into the stock assessment, and if estimated with a reasonable amount of precision, even a single data point could substantially reduce the uncertainty in the scale of total population estimates. Further, close-kin genetics may provide independent estimates of total mortality (and therefore natural mortality conditioned on catch-at-age), relative fecundity-at-age, and the spatial dynamics of spawning and recruitment. All of these quantities could substantially improve the structure of the current assessment and reduce uncertainty. Data collection of genetic samples from 100% of the sampled commercial landings has been in place since 2017 (as part of the sex-ratio monitoring) and from the FISS since 2021. The genetic analysis required to produce data allowing the estimation of reproductive output and other population parameters from close-kin mark-recapture modelling is both complex and expensive, and it could take several years for this project to get fully underway. This five-year plan should consider a pilot evaluation, such that a broader study could be undertaken in the future, providing the likely results would meet the Commission's objectives and prove possible given financial constraints. Research related to close-kin genetics would be pursued under *6.1.3.1 Migration and Population Dynamics* as described below.

6.1.2 Management Strategy Evaluation

MSE priorities have been subdivided into three categories: 1) biological parameterisation, 2) fishery parameterization, and 3) technical development. Research provides specifications for the MSE simulations, such as inputs to the Operating Model (OM), but another important outcome of the research is to define the range of plausibility to include in the MSE simulations as a measure of uncertainty. The following topics have been identified as top priorities.

6.1.2.1 MSE Biological and population parameterization

6.1.2.1.1 Distribution of life stages and stock connectivity

Research topics in this category will mainly inform parameterization of movement in the OM, but will also provide further understanding of Pacific halibut movement, connectivity, and the temporal variability. This knowledge may also be used to refine specific MSE objectives to reflect reality and plausible outcomes. Research under Section 6.1.3.1 will inform this MSE priority.

This research includes examining larval and juvenile distribution which is a main source of uncertainty in the OM that is currently not fully incorporated. Outcomes will assist with conditioning the OM, verify patterns simulated from the OM, and provide information to develop reasonable sensitivity scenarios to test the robustness of MPs.

Also included in this number one priority is stock structure research, especially regarding IPHC Regulatory Area 4B. The dynamics of this IPHC Regulatory Area are not fully understood and it is useful to continue research on the connectivity of IPHC Regulatory Area 4B with other IPHC Regulatory Areas.

Finally, genomic analysis of population size is also included in this ranked category because that would help inform development of the OM as well as the biological sustainability objective related to maintaining a minimum spawning biomass in each IPHC Regulatory Area. An understanding of the spatial distribution of population size will help to inform this objective as well as the OM conditioning process.

6.1.2.1.2 Spatial spawning patterns and connectivity between spawning populations

An important parameter that can influence simulation outcomes is the distribution of recruitment across Biological Regions. Continued research in this area will improve the OM and provide justification for parameterising temporal variability. Research includes assigning individuals to spawning areas and establishing temporal and spatial spawning patterns. Outcomes may also provide information on recruitment strength and the relationship with environmental factors. For example, recent work by Sadorus et al (2020) used a biophysical and spatio-temporal models to examine connectivity across the Bering Sea and Gulf of Alaska. Furthermore, close-kin mark-recapture (Bravington et al. 2016) may provide insights into spatial relationships between juveniles and adults as well as abundance in specific regions. Research under Sections 6.1.3.1 and 6.1.3.2 will inform this MSE priority.

6.1.2.1.3 Understanding growth variation

Changes in the average weight-at-age of Pacific halibut is one of the major drivers of changes in biomass over time. The OM currently simulates temporal changes in weight-at-age via a random autocorrelated process which is unrelated to population size or environmental factors. Ongoing research in drivers related to growth in Pacific halibut will help to improve the simulation of weight-at-age. Research under Section 6.1.3.3 will inform this MSE priority.

6.1.2.1.4 MSE fishery parameterization

The specifications of fisheries and their parameterizations involved consultation with Pacific halibut stakeholders but some aspects of those parameterizations benefit from targeted research. One specific example is knowledge of discarding and discard mortality rates in directed and non-directed fisheries. Discard mortality can be a significant source of fishing mortality in some IPHC Regulatory Areas and appropriately modelling that mortality will provide a more robust evaluation of MPs. Research under Sections 6.1.3.4 and 6.1.3.5 will inform this MSE priority.

6.1.2.2 MSE technical development

Technical improvements to the MSE framework will allow for rapid development of alternative operating models and efficient simulation of management strategies for future evaluation. Coordination with the technical development of the stock assessment (Section 6.1.1.2.1) is necessary to ensure consistent assumptions and hypotheses for tactical (i.e. stock assessment) and strategic (i.e. MSE) models. Investigations done in the stock assessment will inform the stock assessment, which will then be informed by investigations using the closed-loop simulation framework. Multi-year assessments may allow for additional opportunity to coordinate between stock assessment and MSE.

6.1.2.2.1 Alternative migration scenarios

Including alternative migration hypotheses in the MSE simulations will assist in identifying management procedures that are robust to this uncertainty. This exploration will draw on general research on the movement and migration of Pacific halibut, observations from FISS and fisheries data, and outcomes of the stock assessment. Identification of reasonable hypotheses for the movement of Pacific halibut is essential to the robust investigation of management procedures. Research under Section 6.1.3.1 will inform this MSE priority.

6.1.2.2.2 Realistic simulations of estimation error

Closed loop simulation uses feedback from the management procedure to update the population in the projections. The management procedure consists of data collection, an estimation model, and harvest rules; currently IPHC uses a stock assessment as the estimation model. Future development of an efficient simulation process to mimic the stock assessment will more realistically represent the current management process. This involves using multiple estimation models to represent the ensemble and appropriately adding data and updating those models in the simulated projections. Improvements to the current MSE framework include adding additional estimation models to better represent the ensemble stock assessment, ensuring that the simulated estimation accurately represent the stock assessment now and, in the future, and speeding up the simulation process.

6.1.2.2.3 Incorporate additional sources of implementation uncertainty

Implementation uncertainty consists of three subcategories: 1) decision-making uncertainty, 2) realized uncertainty, and 3) perceived uncertainty. Decision-making uncertainty is the difference between mortality limits determined from the management procedure and those adopted by the Commission. This uncertainty is currently not implemented in the MSE framework but has been requested by the SRB and the independent peer review of the MSE. Realized uncertainty is the difference between the mortality limit set by the Commission and the actual mortality realized by the various fisheries. This type of uncertainty is currently partially implemented in the MSE framework. Finally, perceived uncertainty is the difference between the realized mortality and the estimated mortality limits from the various fisheries, which would be used in the estimation model. This third type of implementation uncertainty has not been implemented in the MSE framework. Implementing decision-making uncertainty is a priority for the MSE and will assist in understanding the performance of management procedures when they may not be followed exactly.

6.1.2.3 MSE Program of Work for 2021–2023

Following the 11th Special Session of the IPHC, an MSE program of work for 2021–2023 was developed. Seven tasks were identified that pertained to further developments of the MSE framework, evaluation of alternative MPs, and improvements in evaluation and presentation of results. Table 1 lists these tasks and provides a brief description. Additional details can be found in the program of work available on the MSE webpage.

Table 1. Tasks recommended by the Commission at SS011 (IPHC-2021-SS011-R para 7) for inclusion in the IPHC Secretariat MSE Program of Work for 2021–23.

ID	Category	Task	Deliverable
F.1	Framework	Develop migration scenarios	Develop OMs with alternative migration scenarios
F.2	Framework	Implementation variability	Incorporate additional sources of implementation variability in the framework
F.3	Framework	Develop more realistic simulations of estimation error	Improve the estimation model to more adequately mimic the ensemble stock assessment
F.5	Framework	Develop alternative OMs	Code alternative OMs in addition to the one already under evaluation.
M.1	MPs	Size limits	Identification, evaluation of size limits
M.3	MPs	Multi-year assessments	Evaluation of multi-year assessments
E.3	Evaluation	Presentation of results	Develop methods and outputs that are useful for presenting outcomes to stakeholders and Commissioners

6.1.2.4 Potential Future MSE projects

Management Strategy Evaluation is an iterative process where new management procedures may be evaluated, current management procedures may be re-evaluated under different assumptions, and the understanding of the population, environment, and fisheries may be updated with new information stemming from the stock assessment and biological/ecological research. The current Program of Work (Table 1) focuses on two elements of Management Procedures, but in the future other elements may be of interest, such as distribution procedures. The research being done now will inform the development of the MSE in the future to ensure a robust evaluation of any management procedure.

6.1.3 Biology and Ecology

Capitalizing on the outcomes of the previous 5-year plan (IPHC–2019–BESRP-5YP) (Appendix I), the IPHC Secretariat has identified five research areas that will provide key inputs for stock assessment and the MSE process. In addition to linking genetics and genomics with migration and distribution studies in the newly coined area of Migration and Population Dynamics, the IPHC Secretariat has incorporated a novel research area on Fishing Technology. A series of key objectives for each of the five research areas have been identified that integrate with specific needs for stock assessment and MSE processes and that are ranked according to their relevance (Appendix II). To further describe the IPHC Secretariat's rationale for establishing research priorities, a ranked list of biological uncertainties and parameters for stock assessment and the MSE process and their links to research activities and outcomes derived from the IPHC 5-Year Program of Integrated Research and Monitoring (2022-2026) are provided in Appendix III and Appendix IV.

6.1.3.1 Migration and Population Dynamics

Genetic and genomic studies aimed at improving current knowledge of Pacific halibut migration and population dynamics throughout all life stages in order to achieve a complete understanding of stock structure and distribution across the entire distribution range of Pacific halibut in the North Pacific Ocean and the biotic and abiotic factors that influence it (specifically excluding satellite tagging). Specific objectives in this area include:

- Improve current knowledge of the genetic structure of the Pacific halibut population through the use of state-of-the-art low-coverage whole genome resequencing approaches. Establishment of genetic signatures of spawning sites.
- Improve our understanding of the mechanisms and magnitude of larval connectivity in the North Pacific Ocean. Identification of environmental and biological predictors of larval abundance and recruitment.
- Improve our understanding of spawning site contributions to nursery/settlement areas in relation to year-class, recruit survival and strength, and environmental conditions in the North Pacific Ocean. Measure of genetic diversity of Pacific halibut juveniles from the eastern Bering Sea and the Gulf of Alaska.
- Improve our understanding of the relationship between nursery/settlement origin and adult distribution and abundance over temporal and spatial scales. Genomic assignment of individuals to source populations and assessment of distribution changes.
- Integrate analyses of Pacific halibut connectivity and distribution changes by incorporating genomic approaches.
- Improve estimates of population size, migration rates among geographical regions, and demographic parameters (e.g. fecundity-at-age, survival rate), through the application of close-kin mark-recapture-based approaches.
- Improve our understanding of the influences of oceanographic and environmental variation on connectivity, population structure and adaptation at a genomic level using seascape genomics approaches.
- Exploration and development of alternative methods for aging Pacific halibut based on genetic analyses of DNA methylation patterns in tissues (fin clips).
- Exploration of methods for individual identification based on computer-assisted tail image matching systems as an alternative for traditional mark and recapture tagging.

6.1.3.2 Reproduction

Studies aimed primarily at addressing two critical issues for stock assessment analysis based on estimates of female spawning biomass: 1) the sex ratio of the commercial catch and 2) maturity estimations. Specific objectives in this area include:

- Continued improvement of genetic methods for accurate sex identification of commercial landings from fin clips and otoliths in order to incorporate recent and historical sex-at-age information into the stock assessment process.
- Improve our understanding of the temporal progression of reproductive development and gamete production during an entire annual reproductive cycle in female and male Pacific halibut.
- Update current maturity-at-age estimates.
- Provide estimates of fecundity-at-age and fecundity-at-size.
- Investigate the possible presence of skip spawning in Pacific halibut females.

- Improve accuracy in current staging criteria of maturity status used in the field.
- Investigate possible environmental effects on the ontogenetic establishment of the phenotypic sex and their influence on sex ratios in the adult Pacific halibut population.
- Improve our understanding of potential temporal and spatial changes in maturity schedules and spawning patterns in female Pacific halibut and possible environmental influences.
- Improve our understanding of the genetic basis of variation in age and/or size-at-maturity, fecundity, and spawning timing, by conducting genome-wide association studies.

6.1.3.3 Growth

Studies aimed at describing the role of factors responsible for the observed changes in size-at-age and at evaluating growth and physiological condition in Pacific halibut. Specific objectives in this area include:

- Evaluate possible variation in somatic growth patterns in Pacific halibut as informed by physiological growth markers, physiological condition, energy content and dietary influences.
- Investigate the effects of environmental and ecological conditions that may influence somatic growth in Pacific halibut. Evaluate the relationship between somatic growth and temperature and trophic histories in Pacific halibut through the integrated use of physiological growth markers.
- Improve our understanding of the genetic basis of variation in somatic growth and size-at-age by conducting genome-wide association studies.

6.1.3.4 Mortality and Survival Assessment

Studies aimed at providing updated estimates of discard mortality rates (DMRs) for Pacific halibut in the guided recreational fisheries and at evaluating methods for reducing mortality of Pacific halibut. Specific objectives in this area include:

- Provide information on the types of fishing gear and fish handling practices used in the Pacific halibut recreational (charter) fishery as well as on the number and size composition of discarded Pacific halibut in this fishery.
- Establish best handling practices for reducing discard mortality of Pacific halibut in recreational fisheries.
- Investigate new methods for improved estimation of depredation mortality from marine mammals.

6.1.3.5 Fishing Technology

Studies aimed at developing methods that involve modifications of fishing gear with the purpose of reducing Pacific halibut depredation and bycatch. Specific objectives in this area include:

- Investigate new methods for whale avoidance and/or deterrence for the reduction of Pacific halibut depredation by whales (e.g. catch protection methods).
- Investigate physiological and behavioral responses of Pacific halibut to fishing gear in order to reduce bycatch.

6.2 Monitoring

The Commission's extensive monitoring programs include both direct data collection and coordination with domestic agencies to produce both fishery-dependent and fishery-independent information on the stock and fishery trends, and other information. These critical sources include estimates of fishing mortality from all fisheries encountering Pacific halibut, biological sampling from these fisheries as well as catch-rates and biological sampling from longline and trawl surveys. Monitoring data provide the basis for stock assessment and MSE analysis, many biological research studies, and some inputs directly to the decision-making process (Figure 4). While not the primary focus of this 5-year plan, a basic summary of the components led by the IPHC and those that are provided by domestic agencies is provided below.

6.2.1 Fishery-dependent data

Data collection and monitoring activities aimed at providing standardised time-series of mortality, fishery, and biological data from both direct target fisheries as well as fisheries that incidentally catch Pacific halibut. Directed commercial fisheries data are managed by IPHC. Non-directed commercial discard mortality data, subsistence fisheries data, and recreational fisheries data are managed by Contracting Party domestic agencies.

6.2.1.1 Directed commercial fisheries data

6.2.1.2 Annually review the spatial distribution of sampling effort among ports, data collection methods, sampling rates, and quality assurance/quality control (QAQC) processes, including in-season review of port sampling activities

Ensure current data collection efforts meet current and future needs of stock assessment, MSE and management. Collaborate and coordinate with other Secretariat functions to develop methods and procedures for incorporating promising research results into long-term monitoring program. The IPHC relies on domestic and Tribal agency programs to report annual mortality from incidental catches in non-directed commercial fisheries, catches from subsistence fisheries, and catches from recreational fisheries. Non-directed commercial discard mortality data

Annually collaborate with observer programs and other partners to ensure robust data collection and sampling, QAQC processes, and reporting of incidental catch and mortality, as well as biological sampling.

6.2.1.3 Subsistence fisheries data

Annually collaborate with Tribal, State and Federal agencies of each Contracting Party to ensure high quality data collection, sampling, and reporting in the subsistence fisheries in Canada and the United States of America.

6.2.1.4 Recreational fisheries data

Annually collaborate with National/State agencies of each Contracting Party to ensure and validate high quality data and reporting of recreational fishery mortality estimates and biological data.

6.2.2 Fishery-independent data

Data collection and monitoring activities aimed at providing a standardised time-series of biological and ecological data that is independent of the fishing fleet.

6.2.2.1 Fishery-independent setline survey (FISS)

An annual review process for the FISS station design has been developed (Fig. 9) and is expected to continue in coming years. This process involves scientific review of proposed FISS designs by the Scientific Review Board and includes input from stakeholders prior to review and approval of designs by the Commissioners.

Direct weighing of Pacific halibut has been integrated into the annual FISS sampling since 2019 and will continue into the future to ensure accurate estimation of WPUE and other weight-derived quantities. Sample rates for genetic monitoring will need to be determined for future sampling. Sampling rates of otoliths for aging, archive otoliths and tagged fish will continue to be reviewed annually to ensure the data needs of the IPHC stock assessment and research program are met. Annual FISS sampler training and data QAQC (including at point of data collection and during post-sampling review) will ensure high quality data from the FISS program. Procedures are reviewed annually.

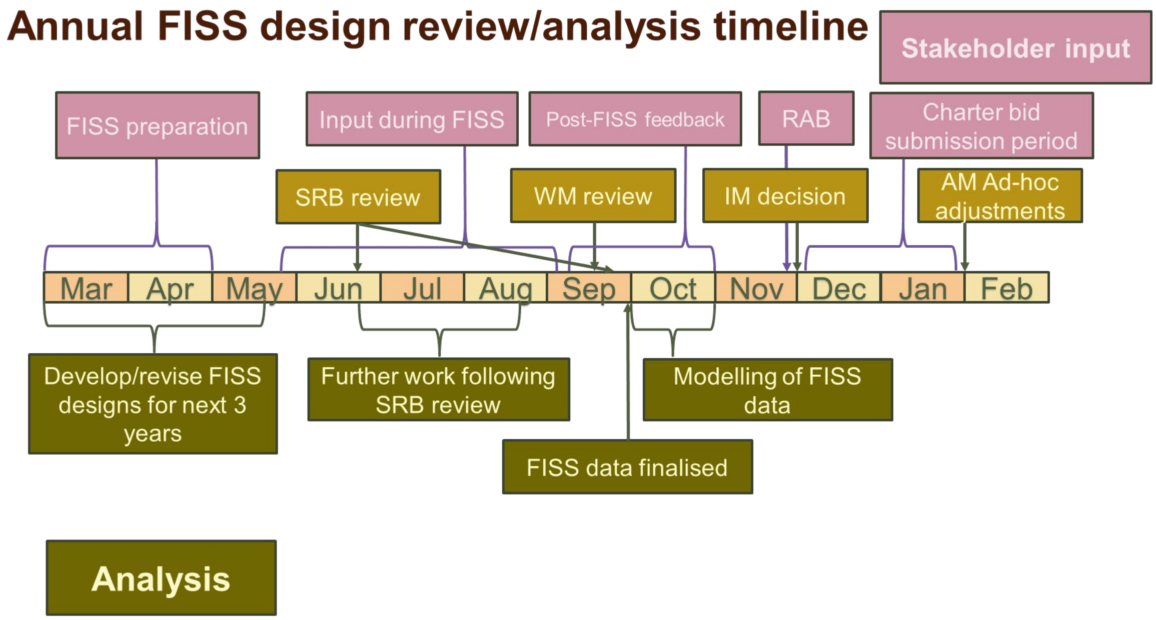


Figure 9. Timeline of annual FISS design review process.

6.2.2.2 Fishery-independent Trawl Survey (FITS)

The IPHC will continue to collaborate with NMFS on sampling procedures for Pacific halibut on the placement of an IPHC sampler onboard a survey vessel for the collection of biological data.

6.2.3 Age composition data (both fishery-dependent and fishery-independent)

The IPHC Secretariat is looking at options for supplementing current Pacific halibut ageing protocol with automatized ageing that does not require extensive otolith-reader training. The IPHC is investigating the potential use of artificial intelligence (AI) for determining the age of Pacific halibut from images of collected otoliths. The Secretariat is in the process of initializing creation of a database of pictures with expert-provided labels, utilizing previously aged otoliths, and assessing the option for the development of a Convolutional Neural Network (CNN) model specifically designed for image classification to determine Pacific halibut age. The goal is to create an AI-based age determination system that complements traditional methods for reliable fish stock assessment and management advice.

6.3 Potential of integrating human dynamics into management decision-making

The evolution of modern fisheries management is taking a transformative turn, emphasizing the integration of human dynamics into decision-making processes. As our world becomes more interconnected through globalization, understanding the intricate human dimension of the fisheries sector is emerging as a critical aspect of sustainable resource management. This forward-looking approach seeks to proactively address challenges while capitalizing on new opportunities.

In a global marketplace where local and imported products compete for consumer attention, vulnerability to disruptions, as evidenced by the COVID-19 pandemic (OECD 2020), has highlighted the need for adaptable strategies embracing the broader picture encompassing external influences. Recent IPHC's socioeconomic study underlines the far-reaching impacts of such dynamics, showcasing the income fluctuations experienced by households dependent on Pacific halibut during the pandemic. Acknowledging these complexities, there is a growing realization of the need for expanding the scope of management-supporting information the IPHC provides beyond stock condition.

The question of how small remote communities can capitalize on the high prices that the final customers are paying for premium seafood products demands innovative thinking. In 2021, fresh Alaskan Pacific halibut fillets routinely sold for USD 24-28 a pound, and often more, in downtown Seattle (e.g. USD 38 at Pike Place Market). Pacific halibut dishes at the restaurants typically sell for USD 37-43 for a dish including a 6oz fish portion. The IPHC's socioeconomic study detailed the geography of impacts of the Pacific halibut fisheries, providing a coherent picture of the exposure of fisheries-dependent households by location to changes in resource availability, but paying closer attention to quantifying leakage of economic benefits from communities strongly involved in fisheries, highlighted that the local earnings often do not align with how much fishing occurs within the community. This suggests the need for research focused on how to operationalize social equity in the context of the globalized market dynamics and the pursuit of stock sustainability.

In parallel, the accelerating impacts of climate change is placing fisheries at the forefront of environmental challenges. The rapid increase in water temperature off the coast of Alaska in 2014-16, termed *the blob*, exemplifies the changes that disrupt ecosystems and fisheries (Cheung and Frölicher 2020), and may have a long-term impact on Pacific halibut distribution. The consequences may include shifts in the distribution of benefits, but possibly go further, affecting the stability of agreements over allocation of a shared resource. Research on decision quality under fast-progressing climate-induced changes to stock distribution emerges as an avenue for impactful work.

Conflicting objectives among stakeholders regarding the use of limited resource in the context of globalization, calls for social equity and climate change are a major challenge of decision-making in fisheries management. Integrating approaches aimed at understanding the human dynamics and external factors with stock assessment and MSE can assist fisheries in bridging the gap between the current and the optimal performance without compromising the stock biological sustainability. For example, socioeconomic performance metrics presented alongside already developed biological/ecological performance metrics would supplement IPHC's portfolio of tools for assessing policy-oriented issues (as requested by the Commission, IPHC-2021-AM097-R, AM097-Req.02) and support decision-making. Moreover, continuing investment in understanding the human dimension of Pacific halibut fishing can also inform on other drivers such as human behavior or human organization that affect the dynamics of fisheries, and thus contribute to improved accuracy of the stock assessment and the MSE (Lynch et al.2018). As such, it can contribute to research integration at the IPHC and provide a complementary resource for the development of harvest control rules.

Lastly, Pacific halibut value is also in its contribution to the diet through subsistence fisheries and importance to the traditional users of the resource. To native people, traditional fisheries constitute a vital aspect of local identity and a major factor in cohesion. One can also consider the Pacific halibut's existence value as an iconic fish of the Pacific Northwest. Recognizing and adopting such an all-encompassing definition of the Pacific halibut resource contribution, the IPHC echoes a broader call to include the human dimension into the research on the impact of management decisions, as well as changes in environmental or stock conditions.

7. Amendment

The intention is to ensure the plan is kept as a '*living plan'*, that is reviewed and updated annually based on the resources available to undertake the work of the Commission (e.g. internal and external fiscal resources, collaborations, internal expertise). The IPHC Secretariat is committed to ensuring an exceptional level of transparency and commitment to the principles of open science.

8. *References*

Bravington, M.V., Skaug, H.J., and Anderson, E.C. 2016. Close-Kin Mark-Recapture. Statistical Science 31(2): 259-274. doi:10.1214/16-sts552.

Cheung, William W.L., and Thomas L. Frölicher. 2020. "Marine Heatwaves Exacerbate Climate Change Impacts for Fisheries in the Northeast Pacific." Scientific Reports 10 (1): 1–10. https://doi.org/10.1038/s41598-020-63650-z.

Clark, W.G., Hare, S.R., Parma, A.M., Sullivan, P.J., and Trumble, R.J. 1999. Decadal changes in growth and recruitment of Pacific halibut (*Hippoglossus stenolepis*). Canadian Journal of Fisheries and Aquatic Sciences 56: 242-252.

Drinan, D.P., Galindo, H.M., Loher, T., and Hauser, L. 2016. Subtle genetic population structure in Pacific halibut Hippoglossus stenolepis. J Fish Biol 89(6): 2571-2594. doi:10.1111/jfb.13148.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124-1138.

Fry, Cristy. 2021. "Seawatch : Halibut Prices Soaring." Homer News, 2021. https://www.homernews.com/business/seawatch-halibut-prices-soaring/.

Lane, Daniel E, and R L Stephenson. 1995. "Fisheries Management Science: The Framework to Link Biological, Economic, and Social Objectives in Fisheries Management." Aquatic Living Resources 8 (3): 215–21.

Leaman, B.M., and Stewart, I.J. 2017. 2.12 Research basis for estimated discard mortality rates used for Pacific halibut in longline and trawl fisheries. IPHC Report of Assessment and Research Activities 2016: 133-172.

Lynch, Patrick D, Richard D Methot, and Jason S Link. 2018. "Implementing a Next Generation Stock Assessment Enterprise: Policymakers' Summary." NOAA Technical Memorandum NMFS-F/SPO-183.

OECD. 2020. "Fisheries, Aquaculture and COVID-19: Issues and Policy Responses." Tackling Coronavirus (Covid-19).

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.R., and Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society 78(6): 1069-1079.

Sadorus, Lauri L., Esther D. Goldstein, Raymond A. Webster, William T. Stockhausen, Josep V. Planas, and Janet T. Duffy-Anderson. 2020. "Multiple life-stage connectivity of Pacific halibut (*Hippoglossus stenolepis*) across the Bering Sea and Gulf of Alaska." Fisheries Oceanography 30 (2): 174-193. https://doi.org/10.1111/fog.12512

Stewart, I., and Webster, R. 2022. Overview of data sources for the Pacific halibut stock assessment, harvest policy, and related analyses. IPHC-2022-SA-02. 56 p.

Stewart, I., Hicks, A., Webster, R., and Wilson, D. 2021. Stock assessment: Summary of the data, stock assessment, and harvest decision table for Pacific halibut (Hippoglossus stenolepis) at the end of 2020. IPHC-2021-AM097-08. 19 p.

Webster, R.A., Clark, W.G., Leaman, B.M., and Forsberg, J.E. 2013. Pacific halibut on the move: a renewed understanding of adult migration from a coastwide tagging study. Canadian Journal of Fisheries and Aquatic Sciences 70(4): 642-653. doi:10.1139/cjfas-2012-0371.

ACKNOWLEDGEMENTS

This updated document was developed by Dr. David Wilson, Dr. Josep Planas, Dr. Ian Stewart, Dr. Allan Hicks, Dr. Ray Webster, and Dr. Basia Hutniczak, in collaboration with other current members of the IPHC Secretariat.

APPENDICES

Appendix I: Outcomes of the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21)

Appendix II: Biological research areas in the 5-Year Program of Integrated Research and Monitoring (2022-2026) and ranked relevance for stock assessment and management strategy evaluation

Appendix III: List of ranked research priorities for stock assessment

Appendix IV: List of ranked research priorities for management strategy evaluation

Appendix V: Proposed schedule of outputs

Appendix VI: Proposed schedule with funding and staffing indicators

APPENDIX I

Outcomes of the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21) (IPHC–2019–BESRP-5YP)

A. Outcomes by Research Area:

1. Migration and Distribution.

1.1. Larval and juvenile connectivity and early life history studies. Planned research outcomes: improved understanding of larval and juvenile distribution.

Main results:

- Larval connectivity between the Gulf of Alaska and the Bering Sea occurs through large island passes across the Aleutian Island chain.
- The degree of larval connectivity between the Gulf of Alaska and the Bering Sea is influenced by spawning location.
- Spawning locations in the western Gulf of Alaska significantly contribute Pacific halibut larvae to the Bering Sea.
- Pacific halibut juveniles counter-migrate from inshore settlement areas in the eastern Bering Sea into the Gulf of Alaska through Unimak Pass.
- Elemental signatures of otoliths from juvenile Pacific halibut vary geographically at a scale equivalent to IPHC regulatory areas.

Publications:

Sadorus, L.; Goldstein, E.; Webster, R.; Stockhausen, W.; Planas, J.V.; Duffy-Anderson, J. Multiple life-stage connectivity of Pacific halibut (*Hippoglossus stenolepis*) across the Bering Sea and Gulf of Alaska. *Fisheries Oceanography*. 2021. 30:174-193. doi: https://doi.org/10.1111/fog.12512.

Loher, T., Bath, G. E., Wischniowsky, S. The potential utility of otolith microchemistry as an indicator of nursery origins in Pacific halibut (*Hippoglossus stenolepis*) in the eastern Pacific: the importance of scale and geographic trending. *Fisheries Research*. 2021. 243: 106072. https://doi.org/10.1016/j.fishres.2021.106072.

Links to 5-Year Research Plan (2022-2026):

- Evaluate the level of genetic diversity among juvenile Pacific halibut in the Gulf of Alaska and the Bering sea due to admixture.
- Assignment of individual juvenile Pacific halibut to source populations.

Integration with Stock Assessment and MSE: The relevance of research outcomes from activities in this research area for stock assessment is in the improvement of estimates of productivity. Research outcomes will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region and represent one of the top three biological inputs into stock assessment. The relevance of these research outcomes for MSE is in the improvement of the parametrization of the Operating Model and represent the top ranked biological input into the MSE.

2. Reproduction.

2.1 Sex ratio of commercial landings. Planned research outcomes: sex ratio information.

Main results:

- Establishment of TaqMan-based genetic assays for genotyping Pacific halibut in the IPHC Biological Laboratory.
- Sex ratio information for the 2017-2020 commercial landings.
- Transfer of genotyping efforts for sex identification to IPHC monitoring program.

Links to 5-Year Research Plan (2022-2026):

- Monitoring effort.

2.2 Histological maturity assessment. Planned research outcomes: updated maturity schedule.

Main results:

- Oocyte developmental stages have been characterized and fully described in female Pacific halibut for the first time.
- Oocyte developmental stages have been used for the classification of female developmental stages and to be able to characterize female Pacific halibut as group synchronous with determinate fecundity.
- Female developmental stages have been used for the classification of female reproductive phases and to be able to characterize female Pacific halibut as following an annual reproductive cycle with spawning in January and February.
- Female developmental stages and reproductive phases of females collected in the central Gulf of Alaska have been used to identify the month of August as the time of the transition between the Vtg2 and Vtg3 developmental stages marking the beginning of the spawning capable reproductive phase.
- Future gonad collections for revising maturity schedules and estimating fecundity can be conducted in August during the FISS.

Publications:

Fish, T., Wolf, N., Harris, B.P., Planas, J.V. A comprehensive description of oocyte developmental stages in Pacific halibut, *Hippoglossus stenolepis*. *Journal of Fish Biology* 2020. 97: 1880-1885. doi: 10.1111/jfb.14551.

Fish, T., Wolf, N., Smeltz, T. S., Harris, B. P., and Planas, J. V. Reproductive Biology of Female Pacific Halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska. *Frontiers in Marine Science* 2022. 9:801759. doi: 10.3389/fmars.2022.801759.

Links to 5-Year Research Plan (2022-2026):

- Revision of maturity schedule by gonad collection during the FISS, as informed by previous studies on reproductive development.

- Estimation of fecundity by age and size, as informed by previous studies demonstrating determinate fecundity.

Integration with Stock Assessment and MSE: Research activities in this Research Area aim at providing information on key biological processes related to reproduction in Pacific halibut (maturity and fecundity) and to provide sex ratio information of Pacific halibut commercial landings. The relevance of research outcomes from these activities for stock assessment is in the scaling of Pacific halibut biomass and in the estimation of reference points and fishing intensity. These research outputs will result in a revision of current maturity schedules and will be included as inputs into the stock assessment and represent the most important biological inputs for stock assessment. The relevance of these research outcomes for MSE is in the improvement of the simulation of spawning biomass in the Operating Model.

3. **Growth**.

3.1 Identification of physiological growth markers and their application for growth pattern evaluation. Planned research outcomes: informative physiological growth markers.

Main results:

- Transcriptomic profiling by RNAseq of white skeletal muscle from juvenile Pacific halibut subjected to growth suppression and to growth stimulation resulted in the identification of a number of genes that change their expression levels in response to growth manipulations.
- Proteomic profiling by LC-MS/MS of white skeletal muscle from juvenile Pacific halibut subjected to growth suppression and to growth stimulation resulted in the identification of a number of proteins that change their abundance in response to growth manipulations.
- Genes and proteins that changed their expression levels in accordance to changes in the growth rate in juvenile Pacific halibut were selected as putative growth markers for future studies on growth pattern evaluation.

Publications:

Planas et al. 2022. In Preparation.

Links to 5-Year Research Plan (2022-2026):

- Application of identified growth markers in studies aiming at investigating environmental influences on growth patterns and at investigating dietary influences on growth patterns and physiological condition.

3.2 Environmental influences on growth patterns. Planned research outcomes: information on growth responses to temperature variation.

Main results:

- Laboratory experiments under controlled temperature conditions have shown that temperature affects the growth rate of juvenile Pacific halibut through changes in the expression of genes that regulate growth processes.

Publications:

Planas et al. 2022. In Preparation.

Links to 5-Year Research Plan (2022-2026):

- Identification of temperature-specific responses in skeletal muscle through comparison between transcriptomic responses to temperature-induced growth changes and to density- and stress-induced growth changes.
- Application of growth markers for additional studies investigating the link between environmental variability and growth patterns and the effects of diet (prey quality and abundance) on growth and physiological condition.

Integration with Stock Assessment and MSE: Research activities conducted in this Research Area aim at providing information on somatic growth processes driving size-at-age in Pacific halibut. The relevance of research outcomes from these activities for stock assessment resides, first, in their ability to inform yield-per-recruit and other spatial evaluations for productivity that support mortality limit-setting, and second, in that they may provide covariates for projecting short-term size-at-age and may help delineate between fishery and environmental effects, thereby informing appropriate management responses. The relevance of these research outcomes for MSE is in the improvement of the simulation of variability and to allow for scenarios investigating climate change.

4. Mortality and Survival Assessment.

4.1 Discard mortality rate estimation in the longline Pacific halibut fishery. Planned research outcomes: experimentally-derived DMR.

Main results:

- Different hook release methods used in the longline fishery result in specific injury profiles and viability classification.
- Plasma lactate levels are high in Pacific halibut with the lowest viability classification.
- Mortality of discarded fish with the highest viability classification is estimated to be between 4.2 and 8.4%.

Publications:

Kroska, A.C., Wolf, N., Planas, J.V., Baker, M.R., Smeltz, T.S., Harris, B.P. Controlled experiments to explore the use of a multi-tissue approach to characterizing stress in wild-caught Pacific halibut (*Hippoglossus stenolepis*). *Conservation Physiology* 2021. 9(1):coab001; doi:10.1093/conphys/coab001.

Loher, T., Dykstra, C.L., Hicks, A., Stewart, I.J., Wolf, N., Harris, B.P., Planas, J.V. Estimation of postrelease longline mortality in Pacific halibut using acceleration-logging tags. *North American Journal of Fisheries Management*. 2022. 42: 37-49. DOI: https://doi.org/10.1002/nafm.10711.

Links to 5-Year Research Plan (2022-2026):

- Integration of information on capture and handling conditions, injury and viability assessment and physiological condition will lead to establishing a set of best handling practices in the longline fishery.

4.2 Discard mortality rate estimation in the guided recreational Pacific halibut fishery. Planned research outcomes: experimentally-derived DMR.

Main results:

- Field experiments testing two different types of gear types (i.e. 12/0 and 16/0 circle hooks) resulted in the capture, sampling and tagging of 243 Pacific halibut in IPHC Regulatory Area 2C (Sitka, AK) and 118 in IPHC Regulatory Area 3A (Seward, AK).
- The distributions of fish lengths by regulatory area and by hook size were similar.

Links to 5-Year Research Plan (2022-2026):

- Estimation of discard mortality rate in the guided recreational fishery.
- Integration of information on capture and handling conditions, injury and viability assessment and physiological condition linked to survival.
- Establishment of a set of best handling practices in the guided recreational fishery.

Integration with Stock Assessment and MSE: The relevance of research outcomes from these activities for stock assessment resides in their ability to improve trends in unobserved mortality in order to improve estimates of stock productivity and represent the most important inputs in fishery yield for stock assessment. The relevance of these research outcomes for MSE is in fishery parametrization

5. Genetics and genomics.

5.1 Generation of genomic resources for Pacific halibut. Planned research outcomes: sequenced genome and reference transcriptome.

Main results:

- A first draft of the chromosome-level assembly of the Pacific halibut genome has been generated.
- The Pacific halibut genome has a size of 602 Mb and contains 24 chromosome-size scaffolds covering 99.8% of the complete assembly with a N50 scaffold length of 27 Mb at a coverage of 91x.
- The Pacific halibut genome has been annotated by NCBI and is available as NCBI Hippoglossus stenolepis Annotation Release 101 (https://www.ncbi.nlm.nih.gov/assembly/GCA_022539355.2/).
- Transcriptome (i.e. RNA) sequencing has been conducted in twelve tissues in Pacific halibut and the raw sequence data have been deposited in NCBI's Sequence Read Archive (SRA) under the bioproject number PRJNA634339 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA634339) and with SRA accession numbers SAMN14989915 - SAMN14989926.

Publications:

Jasonowicz, A.C., Simeon, A., Zahm, M., Cabau, C., Klopp, C., Roques, C., Iampietro, C., Lluch, J., Donnadieu, C., Parrinello, H., Drinan, D.P., Hauser, L., Guiguen, Y., Planas, J.V. Generation of a chromosome-level genome assembly for Pacific halibut (*Hippoglossus stenolepis*) and characterization of its sex-determining genomic region. *Molecular Ecology Resources. 2022. In Press*. doi: https://doi.org/10.1111/1755-0998.13641.

Jasonowicz et al. 2022. In Preparation.

Links to 5-Year Research Plan (2022-2026):

- Genome-wide analysis of stock structure and composition.

5.2 Determine the genetic structure of the Pacific halibut population in the Convention Area. Planned research outcomes: genetic population structure.

Main results:

- The collection of winter genetic samples in the Aleutian Islands completed the winter sample collection needed to conduct studies on the genetic population structure of Pacific halibut in the Convention Area.
- Initial results of low coverage whole genome resequencing of winter samples indicate that an average of 26.5 million raw sequencing reads per obtained per sample that provided average individual genomic coverages for quality filtered alignments of 3.2x.

Links to 5-Year Research Plan (2022-2026):

- Fine-scale delineation of population structure, with particular emphasis on IPHC Regulatory 4B structure.

Integration with Stock Assessment and MSE: The relevance of research outcomes from these activities for stock assessment resides in the introduction of possible changes in the structure of future stock assessments, as separate assessments may be constructed if functionally isolated components of the population are found (e.g. IPHC Regulatory Area 4B), and in the improvement of productivity estimates, as this information may be used to define management targets for minimum spawning biomass by Biological Region. These research outcomes provide the second and third top ranked biological inputs into stock assessment. Furthermore, the relevance of these research outcomes for MSE is in biological parametization and validation of movement estimates and of recruitment distribution.

B. List of ranked biological uncertainties and parameters for stock assessment (SA) and their links to research areas and activities contemplated in the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21)

SA Rank	Research outcomes	Relevance for stock assessment	Specific analysis input	Research Area	Research activities
1. Biological input	Updated maturity schedule	Scale biomass and reference point estimates	Will be included in the stock assessment, replacing the current schedule last updated in 2006	Reproduction	Histological maturity assessment
	Incidence of skip spawning		Will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment		Examination of potential skip spawning
	Fecundity-at-age and -size information		Will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points		Fecundity assessment
	Revised field maturity classification		Revised time-series of historical (and future) maturity for input to the stock assessment		Examination of accuracy of current field macroscopic maturity classification
2. Biological input	Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area	Altered structure of future stock assessments	If 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area	Genetics and Genomics	Population structure
3. Biological input	Assignment of individuals to source populations and assessment of distribution changes	Improve estimates of productivity	Will be used to define management targets for minimum spawning biomass by Biological Region		Distribution
	Improved understanding of larval and juvenile distribution		Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	Migration	Larval and juvenile connectivity studies
1. Assessment data collection and processing	Sex ratio-at-age	Scale biomass and fishing intensity	Annual sex-ratio at age for the commercial fishery fit by the stock assessment	Reproduction	Sex ratio of current commercial landings
	Historical sex ratio-at-age		Annual sex-ratio at age for the commercial fishery fit by the stock assessment		Historical sex ratios based on archived otolith DNA analyses
2. Assessment data collection and processing	New tools for fishery avoidance/deterence; improved estimation of depredation mortality	Improve mortality accounting	May reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude	Mortality and survival assessment	Whale depredation accounting and tools for avoidance
1. Fishery yield	Physiological and behavioral responses to fishing gear	Reduce incidental mortality	May increase yield available to directed fisheries	Mortality and survival assessment	Biological interactions with fishing gear
2. Fishery yield	Guidelines for reducing discard mortality	Improve estimates of unobserved mortality	May reduce discard mortality, thereby increasing available yield for directed fisheries	Mortality and survival assessment	Best handling practices: recreational fishery

C. List of ranked biological uncertainties and parameters for management strategy evaluation (MSE) and their links to research areas and activities contemplated in the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21)

MSE Rank	Research outcomes	Relevance for MSE	Research Area	Research activities
1. Biological parameterization and validation of movement estimates	Improved understanding of larval and juvenile distribution	Improve parametization of the Operating Model	Migration	Larval and juvenile connectivity studies
	Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area		Genetics and Genomics	Population structure
2. Biological parameterization and validation of recruitment variability and distribution	Assignment of individuals to source populations and assessment of distribution changes	Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model		Distribution
	Establishment of temporal and spatial maturity and spawning patterns	Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model	Reproduction	Recruitment strength and variability
3. Biological parameterization and validation for growth projections	Identification and application of markers for growth pattern evaluation	Improve simulation of variability and allow for scenarios investigating climate change	Growth	Evaluation of somatic growth variation as a driver for changes in size-at-age
	Environmental influences on growth patterns			
	Dietary influences on growth patterns and physiological condition			
1. Fishery parameterization	Experimentally-derived DMRs	Improve estimates of stock productivity	Mortality and survival assessment	Discard mortality rate estimate: recreational fishery

D. External funding received during the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21):

Project #	Grant agency	Project name	PI	Partners	IPHC Budget ($US)	Management implications	Grant period
1	**Saltonstall-Kennedy NOAA**	Improving discard mortality rate estimates in the Pacific halibut by integrating handling practices, physiological condition and post-release survival (NOAA Award No. NA17NMF4270240)	IPHC	Alaska Pacific University	$286,121	Bycatch estimates	September 2017 – August 2020
2	**North Pacific Research Board**	Somatic growth processes in the Pacific halibut (*Hippoglossus stenolepis*) and their response to temperature, density and stress manipulation effects (NPRB Award No. 1704)	IPHC	AFSC-NOAA-Newport, OR	$131,891	Changes in biomass/size-at-age	September 2017 – February 2020
3	**Bycatch Reduction Engineering Program - NOAA**	Adapting Towed Array Hydrophones to Support Information Sharing Networks to Reduce Interactions Between Sperm Whales and Longline Gear in Alaska	Alaska Longline Fishing Association	IPHC, University of Alaska Southeast, AFSC-NOAA	-	Whale Depredation	September 2018 – August 2019
4	**Bycatch Reduction Engineering Program - NOAA**	Use of LEDs to reduce Pacific halibut catches before trawl entrainment	Pacific States Marine Fisheries Commission	IPHC, NMFS	-	Bycatch reduction	September 2018 – August 2019
5	**National Fish & Wildlife Foundation**	Improving the characterization of discard mortality of Pacific halibut in the recreational fisheries (NFWF Award No. 61484)	IPHC	Alaska Pacific University, U of A Fairbanks, charter industry	$98,902	Bycatch estimates	April 2019 – November 2021
6	**North Pacific Research Board**	Pacific halibut discard mortality rates (NPRB Award No. 2009)	IPHC	Alaska Pacific University,	$210,502	Bycatch estimates	January 2021 –March 2022
7	**Bycatch Reduction Engineering Program - NOAA**	Gear-based approaches to catch protection as a means for minimizing whale depredation in longline fisheries (NA21NMF4720534)	IPHC	Deep Sea Fishermen's Union, Alaska Fisheries Science Center-NOAA, industry representatives	$99,700	Mortality estimations due to whale depredation	November 2021 – October 2022
8	**North Pacific Research Board**	Pacific halibut population genomics (NPRB Award No. 2110)	IPHC	Alaska Fisheries Science Center-NOAA	$193,685	Stock structure	December 2021-January 2024
				Total awarded ($)	**$1,020,801**		

E. Publications in the peer-reviewed literature resulting from the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21):

2020:

Fish, T., Wolf, N., Harris, B.P., Planas, J.V. A comprehensive description of oocyte developmental stages in Pacific halibut, *Hippoglossus stenolepis*. *Journal of Fish Biology*. 2020. 97: 1880-1885. https://doi: 10.1111/jfb.14551.

2021:

Carpi, P., Loher, T., Sadorus, L., Forsberg, J., Webster, R., Planas, J.V., Jasonowicz, A., Stewart, I. J., Hicks, A. C. Ontogenetic and spawning migration of Pacific halibut: a review. *Rev Fish Biol Fisheries*. 2021. https://doi.org/10.1007/s11160-021-09672-w.

Kroska, A.C., Wolf, N., Planas, J.V., Baker, M.R., Smeltz, T.S., Harris, B.P. Controlled experiments to explore the use of a multi-tissue approach to characterizing stress in wild-caught Pacific halibut (*Hippoglossus stenolepis*). *Conservation Physiology* 2021. 9(1):coab001. https://doi:10.1093/conphys/coab001.

Loher, T., Bath, G. E., Wischniowsky, S. The potential utility of otolith microchemistry as an indicator of nursery origins in Pacific halibut (*Hippoglossus stenolepis*) in the eastern Pacific: the importance of scale and geographic trending. *Fisheries Research*. 2021. 243: 106072. https://doi.org/10.1016/j.fishres.2021.106072.

Lomeli, M.J.M., Wakefield, W.W., Herrmann, B., Dykstra, C.L., Simeon, A., Rudy, D.M., Planas, J.V. Use of Artificial Illumination to Reduce Pacific Halibut Bycatch in a U.S. West Coast Groundfish Bottom Trawl. *Fisheries Research*. 2021. 233: 105737. doi: 10.1016/j.fishres.2020.105737.

Sadorus, L., Goldstein, E., Webster, R., Stockhausen, W., Planas, J.V., Duffy-Anderson, J. Multiple life-stage connectivity of Pacific halibut (*Hippoglossus stenolepis*) across the Bering Sea and Gulf of Alaska. *Fisheries Oceanography*. 2021. 30:174-193. doi: https://doi.org/10.1111/fog.12512.

2022:

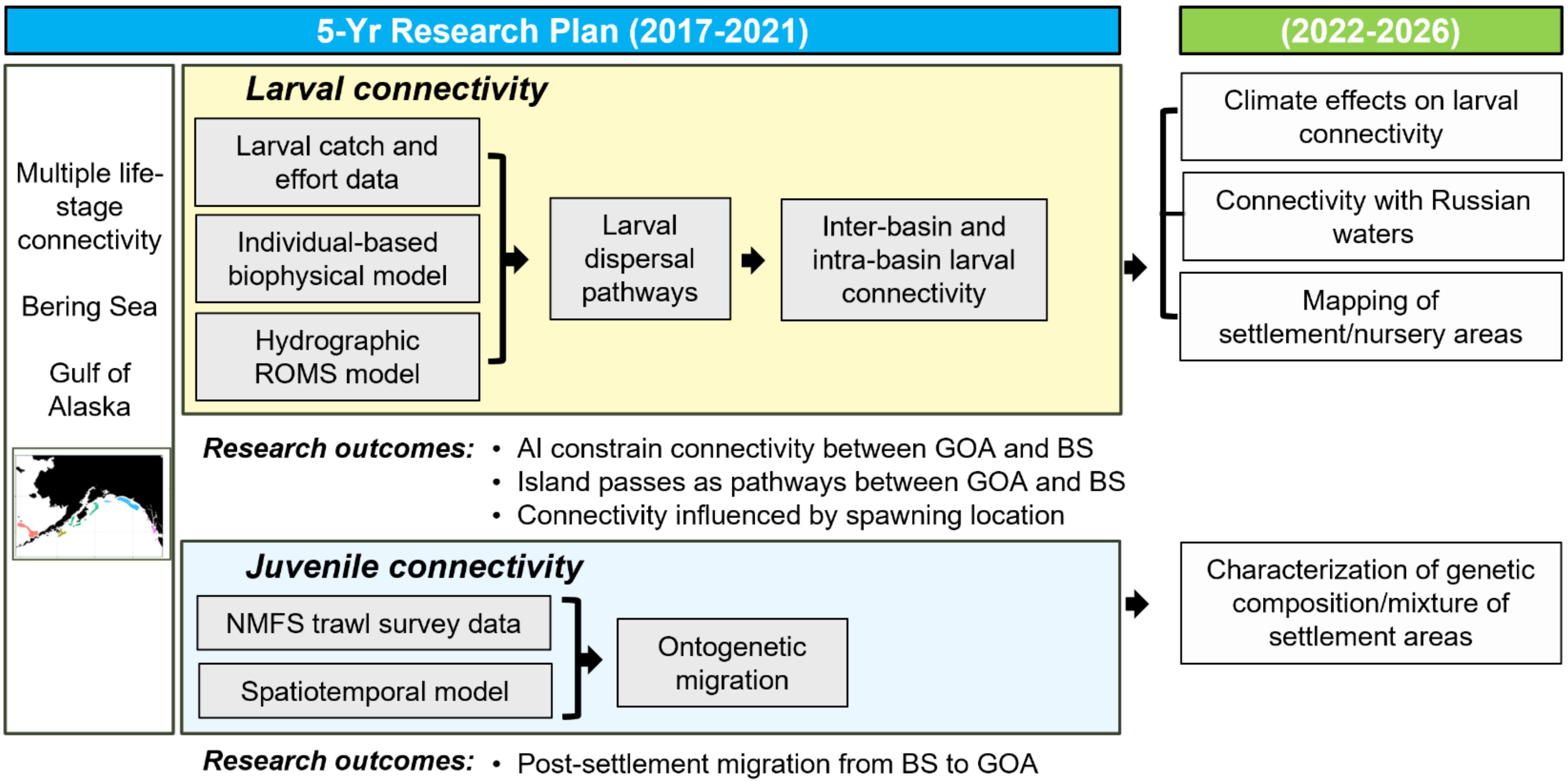
Fish, T., Wolf, N., Smeltz, T. S., Harris, B. P., and Planas, J. V. Reproductive Biology of Female Pacific Halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska. *Frontiers in Marine Science* 2022. 9:801759. doi: 10.3389/fmars.2022.801759.

Jasonowicz, A.C., Simeon, A., Zahm, M., Cabau, C., Klopp, C., Roques, C., Iampietro, C., Lluch, J., Donnadieu, C., Parrinello, H., Drinan, D.P., Hauser, L., Guiguen, Y., Planas, J.V. Generation of a chromosome-level genome assembly for Pacific halibut (*Hippoglossus stenolepis*) and characterization of its sex-determining genomic region. *Molecular Ecology Resources.* 2022. In Press. doi: https://doi.org/10.1111/1755-0998.13641.

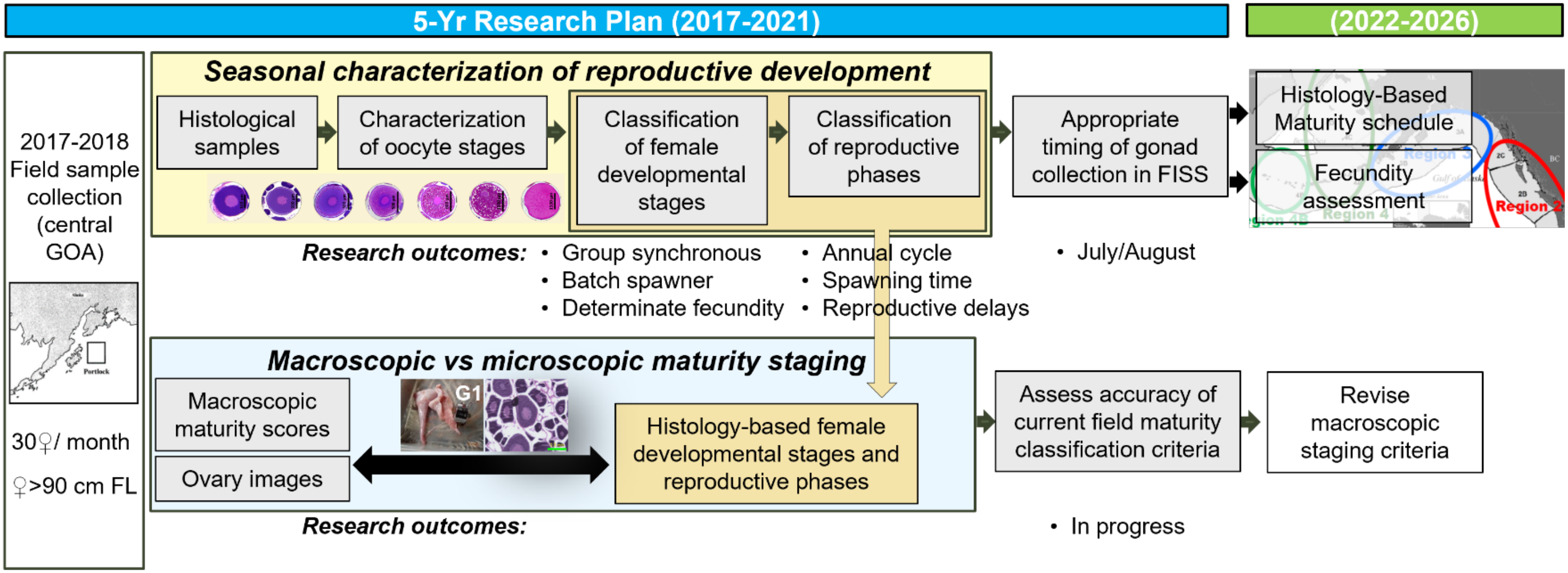
Loher, T., Dykstra, C.L., Hicks, A., Stewart, I.J., Wolf, N., Harris, B.P., Planas, J.V. Estimation of postrelease longline mortality in Pacific halibut using acceleration-logging tags. *North American Journal of Fisheries Management*. 2022. 42: 37-49. DOI: http://dx.doi.org/10.1002/nafm.10711.

F. Flow chart of progress resulting from the IPHC 5-Year Biological and Ecosystem Science Research Plan (2017-21) by research area leading to the IPHC 5-Year Program of Integrated Research and Monitoring (2022-2026)

1. Migration and Distribution

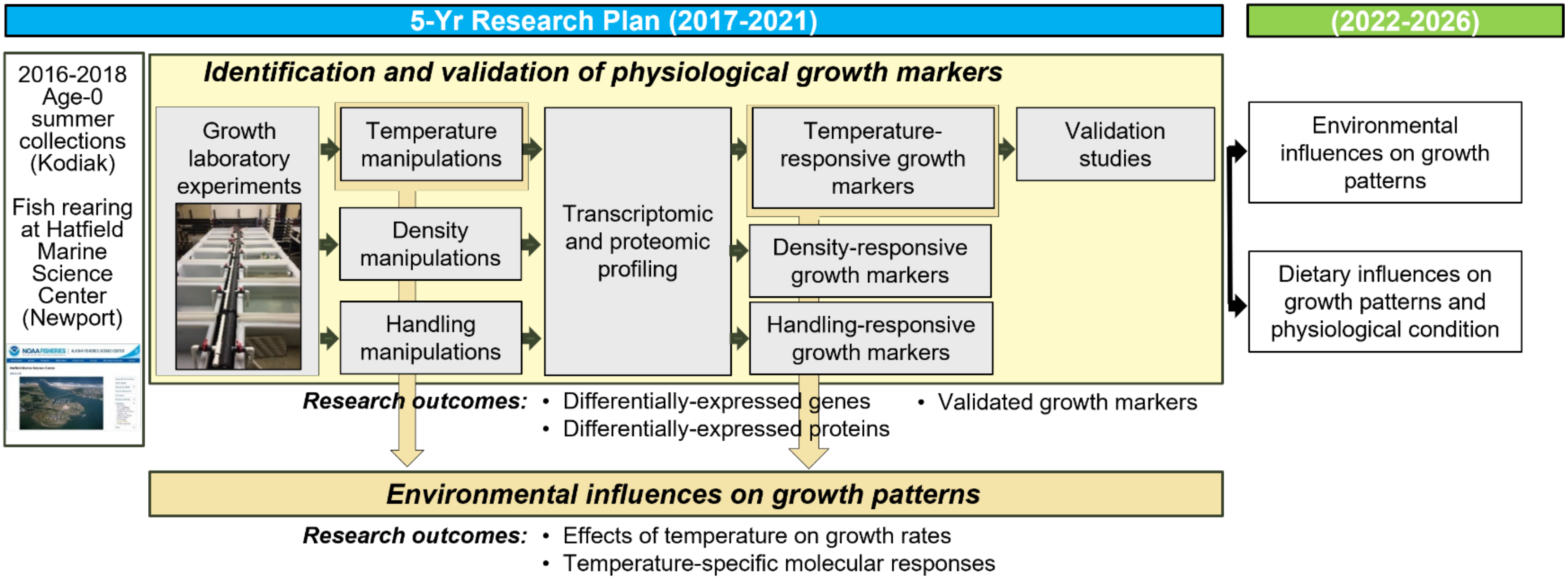


2. Reproduction



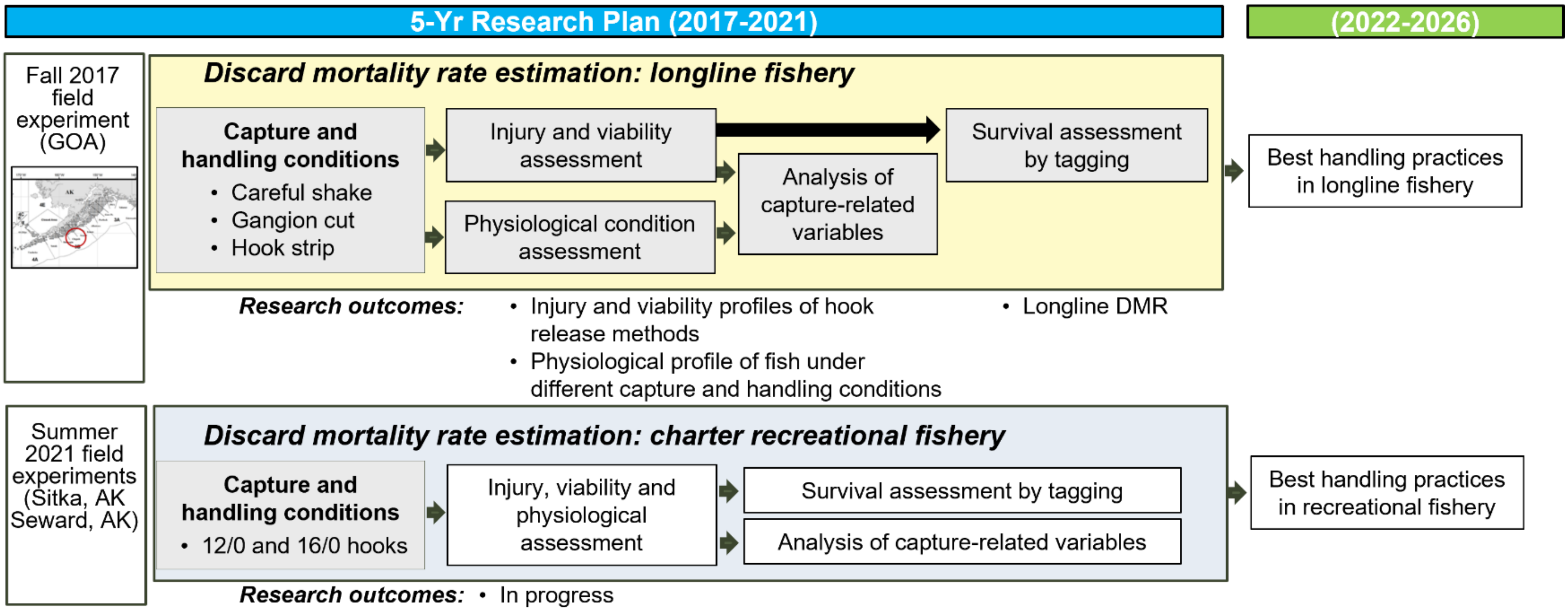
Staff involved: Teresa Fish, MSc APU (2018-2020), Crystal Simchick, Ian Stewart, Allan Hicks, Josep Planas
Funding: IPHC (2018-2020)
Publications (2): Fish et al. (2020) *J. Fish Biol*. **97**: 1880–1885 ; Fish et al. (2022) Front. Mar. Sci. 9:801759

3. Growth



Staff involved: Andy Jasonowicz, Crystal Simchick, Josep Planas
Funding: NPRB Grant#1704 (Sept. 2017-Feb. 2020)
Publications: Planas et al. (in preparation)

4. Mortality and Survival Assessment

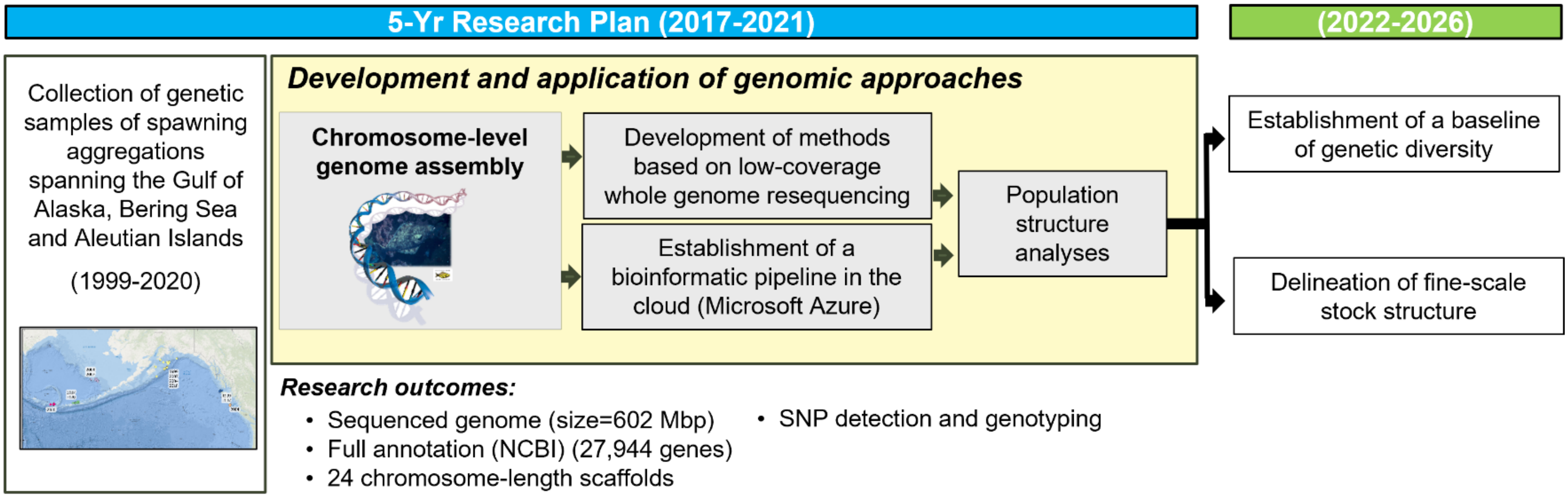


Staff involved: Claude Dykstra, Allan Hicks, Ian Stewart, Josep Planas

Funding (3): Saltonstall-Kennedy NOAA (Sept. 2017-Aug. 2020); NFWF (Apr. 2019-Nov. 2021); NPRB#2009 (Jan. 2021-Mar. 2022)

Publications (2): Kroska et al. (2021) *Conserv. Physiol.*; Loher et al. (2022) *North Amer. J. Fish. Manag. 42: 37-49*

5. Genetics and Genomics



Staff involved: Andy Jasonowicz, Josep Planas
Funding: IPHC, NPRB#2110
Publications: Jasonowicz et al. (2022) *Mol. Ecol. Resour. (In Review)*

APPENDIX II

Biological research areas in the 5-Year Program of Integrated Research and Monitoring (2022-2026) and ranked relevance for stock assessment and management strategy evaluation (MSE)

Research areas	Research activities	Research outcomes	Relevance for stock assessment	Relevance for MSE	Specific analysis input	SA Rank	MSE Rank	Research priorization
Migration and population dynamics	Population structure	Population structure in the Convention Area	Altered structure of future stock assessments	Improve parametization of the Operating Model	If 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area	2. Biological input	1. Biological parameterization and validation of movement estimates and recruitment distribution	2
	Distribution	Assignment of individuals to source populations and assessment of distribution changes	Improve estimates of productivity		Will be used to define management targets for minimum spawning biomass by Biological Region	3. Biological input		2
	Larval and juvenile connectivity studies	Improved understanding of larval and juvenile distribution	Improve estimates of productivity		Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	3. Biological input	1. Biological parameterization and validation of movement estimates	2
Reproduction	Histological maturity assessment	Updated maturity schedule	Scale biomass and reference point estimates	Improve simulation of spawning biomass in the Operating Model	Will be included in the stock assessment, replacing the current schedule last updated in 2006	1. Biological input		1
	Examination of potential skip spawning	Incidence of skip spawning			Will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment			1
	Fecundity assessment	Fecundity-at-age and -size information			Will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points			1
	Examination of accuracy of current field macroscopic maturity classification	Revised field maturity classification			Revised time-series of historical (and future) maturity for input to the stock assessment			1
Growth	Evaluation of somatic growth variation as a driver for changes in size-at-age	Identification and application of markers for growth pattern evaluation	Scale stock productivity and reference point estimates	Improve simulation of variability and allow for scenarios investigating climate change	May inform yield-per-recruit and other spatial evaluations of productivity that support mortality limit-setting		3. Biological parameterization and validation for growth projections	5
		Environmental influences on growth patterns			May provide covariates for projecting short-term size-at-age. May help to delineate between effects due to fishing and those due to environment, thereby informing appropriate management response			5
		Dietary influences on growth patterns and physiological condition			May provide covariates for projecting short-term size-at-age. May help to deleineate between effects due to fishing and those due to environment, thereby informing appropriate management response			5
Mortality and survival assessment	Discard mortality rate estimate: longline fishery	Experimentally-derived DMR	Improve trends in unobserved mortality	Improve estimates of stock productivity	Will improve estimates of discard mortality, reducing potential bias in stock assessment results and management of mortality limits	1. Fishery yield	1. Fishery parameterization	4
	Discard mortality rate estimate: recreational fishery				Will improve estimates of discard mortality, reducing potential bias in stock assessment results and management of mortality limits			4
	Best handling and release practices	Guidelines for reducing discard mortality			May reduce discard mortality, thereby increasing available yield for directed fisheries	2. Fishery yield		4
Fishing technology	Whale depredation accounting and tools for avoidance	New tools for fishery avoidance/deterence; improved estimation of depredation mortality	Improve mortality accounting	Improve estimates of stock productivity	May reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude	1. Assessment data collection and processing		3

APPENDIX III

List of ranked research priorities for stock assessment

SA Rank	Research outcomes	Relevance for stock assessment	Specific analysis input	Research Area	Research activities
1. Biological input	Updated maturity schedule	Scale biomass and reference point estimates	Will be included in the stock assessment, replacing the current schedule last updated in 2006	Reproduction	Histological maturity assessment
	Incidence of skip spawning		Will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment		Examination of potential skip spawning
	Fecundity-at-age and -size information		Will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points		Fecundity assessment
	Revised field maturity classification		Revised time-series of historical (and future) maturity for input to the stock assessment		Examination of accuracy of current field macroscopic maturity classification
2. Biological input	Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area	Altered structure of future stock assessments	If 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area	Migration and population dynamics	Population structure
3. Biological input	Assignment of individuals to source populations and assessment of distribution changes	Improve estimates of productivity	Will be used to define management targets for minimum spawning biomass by Biological Region		Distribution
	Improved understanding of larval and juvenile distribution		Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region		Larval and juvenile connectivity studies
1. Assessment data collection and processing	Sex ratio-at-age	Scale biomass and fishing intensity	Annual sex-ratio at age for the commercial fishery fit by the stock assessment	Reproduction	Sex ratio of current commercial landings
	Historical sex ratio-at-age		Annual sex-ratio at age for the commercial fishery fit by the stock assessment		Historical sex ratios based on archived otolith DNA analyses
2. Assessment data collection and processing	New tools for fishery avoidance/deterence; improved estimation of depredation mortality	Improve mortality accounting	May reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude	Fishing technology	Whale depredation accounting and tools for avoidance
1. Fishery yield	Physiological and behavioral responses to fishing gear	Reduce incidental mortality	May increase yield available to directed fisheries	Fishing technology	Biological interactions with fishing gear
2. Fishery yield	Guidelines for reducing discard mortality	Improve estimates of unobserved mortality	May reduce discard mortality, thereby increasing available yield for directed fisheries	Mortality and survival assessment	Best handling practices: recreational fishery

APPENDIX IV

List of ranked research priorities for management strategy evaluation (MSE)

MSE Rank	Research outcomes	Relevance for MSE	Research Area	Research activities
1. Biological parameterization and validation of movement estimates	Improved understanding of larval and juvenile distribution	Improve parametization of the Operating Model	Migration and population dynamics	Larval and juvenile connectivity studies
	Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area			Population structure
2. Biological parameterization and validation of recruitment variability and distribution	Assignment of individuals to source populations and assessment of distribution changes	Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model		Distribution
	Establishment of temporal and spatial maturity and spawning patterns	Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model	Reproduction	Recruitment strength and variability
3. Biological parameterization and validation for growth projections	Identification and application of markers for growth pattern evaluation	Improve simulation of variability and allow for scenarios investigating climate change	Growth	Evaluation of somatic growth variation as a driver for changes in size-at-age
	Environmental influences on growth patterns			
	Dietary influences on growth patterns and physiological condition			
1. Fishery parameterization	Experimentally-derived DMRs	Improve estimates of stock productivity	Mortality and survival assessment	Discard mortality rate estimate: recreational fishery

APPENDIX V

List of ongoing and planned research projects (Will be linked to the website)

Research Project #	Project Title	Abstract	Objectives	Deliverables	Progress report	5YPRIM Research area	Management implications	Specific inputs into management	Period of Performance	PI	Funding source	Budget	Research prioritization for SA/MSE
1	Leveraging multiple genomic approaches to investigate population structure and dynamics of Pacific halibut	The Pacific halibut (Hippoglossus stenolepis) is a key flatfish species in the North Pacific Ocean ecosystem that supports important commercial, recreational and subsistence fisheries and that is managed as a single stock by the International Pacific Halibut Commission. The overarching goal of the present study is to advance our understanding of Pacific halibut population structure and dynamics in a changing climate through the use of genomic approaches to inform fishery management. In particular, we seek to improve our current understanding of stock structure among spawning groups of Pacific halibut in the northeast Pacific Ocean by conducting low coverage whole genome resequencing, a method that allows the characterization of genomic variation at the highest resolution possible and with which we will establish a baseline of Pacific halibut genetic diversity. Subsequently, we will leverage the obtained genomic data to identify markers that display high differentiation among the different genetic baseline datasets. The results from this study will inform on the delimitation of management units and provide preliminary information on stock composition in the Pacific halibut fishery, as well as provide a tool to monitor changes in distribution associated with climate change.	1. Investigate fine scale Pacific halibut population structure in the northeast Pacific Ocean using low coverage whole genome resequencing: characterization of neutral and adaptive variation at very high resolution among spawning groups leading to the identification of millions of genome-derived genetic markers. 2. Develop a high-throughput genetic marker panel consisting of a selection of genome-derived, high resolution markers	1.Establishment of a baseline of Pacific halibut genetic diversity. The genomic data produced will represent a detailed baseline of Pacific halibut genetic structure and diversity at neutral and adaptive markers over a large geographical scale (Gulf of Alaska, Aleutian Islands and Bering Sea) and over a broad temporal scale (last 30 years). 2.Delineation of fine-scale Pacific halibut stock structure. 3. Assignment of individuals to source populations and assessment of distribution changes.	IPHC-2023-SRB022-09/NPRB Interim Report July 2023/IPHC-2023-WM2023-12	Migration and Population Dynamics	1. Altered structure of future stock assessments and MSE operating models. 2. Improve estimates of productivity. 3. Improve understanding of population distribution and the effects of distributing fishing effort.	If IPHC Regulatory Area 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area. Research outcomes will be used to define management targets for minimum spawning biomass by Biological Region.	12/01/2021-02/16/2024	Josep Planas	External (North Pacific Research Board; Project No. 2110)	$193,685	Priority Rank #2
2	Mapping of Pacific halibut juvenile habitat	The IPHC Secretariat recently completed a study to investigate the connectivity between spawning grounds and possible settlement areas based on a biophysical larval transport model (Sadorus et al., 2021; https://doi.org/10.1111/fog.12512). Although it is known that Pacific halibut, following the pelagic larval phase, begin their demersal stage as roughly 6-month-old juveniles, settling in shallow nursery (settlement) areas, near or outside the mouths of bays (Carpi et al., 2021; https://doi.org/10.1007/s11160-021-09672-w), very little information is available on the geographic location and physical characteristics of these areas. In order to fill this knowledge gap, the IPHC Secretariat has initiated studies to identify potential settlement areas for juvenile Pacific halibut throughout IPHC Convention Waters.	1. Collect data sources on juvenile Pacific halibut presence. 2. Create a map of suitable settlement habitat by combining available bathymetry information (e.g. benthic sediment composition and shoreline morphological data) and information on recorded presence of age-0, age-1 and age-2 Pacific halibut juveniles as well as absence of young Pacific halibut noted by various nursery habitat projects focused on other flatfish species.	Map of juvenile Pacific halibut habitat.	IPHC-2023-SRB022-09/IPHC-2023-WM2023-12	Migration and Population Dynamics	Improve estimates of productivity	Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	01/01/2023-12/31/2025	Josep Planas	Internal	$0	Priority Rank #2
3	Female reproductive assessment	In fisheries, understanding the reproductive biology of a species is important for estimating the reproductive potential and spawning biomass of the stock and, consequently, for optimizing management of the species. Recent sensitivity analyses have shown the importance of changes in spawning output in female Pacific halibut due to changes in maturity schedules, in fecundity estimations and/or in skip spawning for stock assessment (Stewart and Hicks, 2020). These results highlight the need for a better understanding of factors influencing reproductive biology and spawning success in Pacific halibut. In order to fill existing knowledge gaps related to the reproductive biology of female Pacific halibut, research efforts are being conducted to characterize female reproductive capacity in this species. Improved knowledge on key aspects of the reproductive physiology of Pacific halibut (e.g., maturity schedules, fecundity, etc.) will provide an updated and more comprehensive description of reproductive capacity and success in this important species.	1.Produce an accurate description of oocyte developmental stages in female Pacific halibut that can be used to classify female maturity stages. 2.Describe changes in female and male maturity stages throughout an entire annual reproductive cycle based on histological assessment and physiological parameters that will be used to revise current estimates of female and male age-at-maturity. 3.Compare macroscopic (based on field observations) and microscopic (based on histological assessment) female and male maturity stages and revise maturity criteria used in FISS. 4.Update maturity schedules based on histological classification of female maturity. 5. Conduct investigations on fecundity and on the incidence of skip-spawning in female Pacific halibut. 6.Conduct investigations on possible temporal and spatial changes in reproductive performance (maturity, fecundity, skip-spawning) in female Pacific halibut.	1. Updated maturity schedule coastwide. 2. Fecundity-at-age and -size estimates. 3. Revised field maturity classification. 4. Information on skip-spawning.	IPHC-2023-SRB022-09/IPHC-2023-WM2023-12	Reproduction	Scale biomass and reference point estimates. Improve estimates of spawning biomass in the stock assessment and improve simulations of spawning biomass in the MSE operating model. □ □ □	Research outcomes will be included in the stock assessment, replacing the current maturity schedule last updated in 2006. Research outcomes will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment. Research outcomes will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points. Research outcomes will result in revised time-series of historical (and future) maturity for input to the stock assessment.	01/01/2017-12/31/2026	Josep Planas	Internal	$51,834 (FY2024)	Priority Rank #1
4	Gear-based approaches to catch protection as a means for minimizing whale depredation in longline fisheries	In the north Pacific, both Killer (Orcinus orca) and Sperm (Physeter macrocephalus) whales are involved in depredation behavior in Pacific halibut (Hippoglossus stenolepis). In 2011 and 2012 fisheries observers estimated that 6.9% of Pacific halibut sets were affected by whale depredation in the Bering Sea. Reductions in catch per unit effort (CPUE) when whales were present ranged across geographic regions from 5 15-57% for Pacific halibut. These impacts also incur significant time, fuel, and personnel costs to fishing operations. From a fisheries management perspective, depredation creates an additional and highly uncertain source of mortality, loss of data (e.g. compromised survey activity), and reduces fishery efficiency. Stock assessments of both Pacific halibut (Stewart et al. 2020) and sablefish (Goethel et al. 2020) have adjusted their analysis of fishery independent data to account for the effects of whale depredation on catch rates. In the sablefish assessment, fishery limits are also adjusted downward to reflect expected depredation during the commercial fishery. Meanwhile, potential risks to the whales include physical injury due to being near vessels and gear, disruption of social structure , and developing an artificial reliance on food items that can be affected by fishery dynamics. Many efforts have been made over the years to mitigate this problem, with fishers generally limited to simple methods that can be constructed, deployed, or enacted without significantly disrupting normal fishing operations, or without violating gear regulations. Existing approaches include catch protection, physical and auditory deterrents, and spatial or temporal avoidance. These approaches have had variable degrees of success and ease of adoption in each fishery but none have provided a long-term solution. There are increasing data sources supporting the notion that technologies which reduce initial contact between gear and depredators will reduce the likelihood of foraging attempts around the gear, thereby sustaining levels of target catch while simultaneously reducing risk of depredator mortality and gear damage. Recent studies using physical catch protection methods include the development of underwater shuttles that unhook, and transport catch to the surface (Patagonian toothfish), light and expandable 'slinky' pots (sablefish), and flashers or mesh panels attached to the gear to obscure catch (tuna) (IPHC 2022). While slinky pots had quick uptake in the sablefish longline fishery, depredation occurring with this gear has been reported (Goethel et al. 2022), demonstrating the urgency of ongoing challenges to interrupting the reward cycle underpinning this problem.	1. Identify potential methods for protecting hook captured fish from whale depredation. 2. Develop and field-test several simple low-cost catch-protection designs that can be deployed effectively using current longline fishing techniques.	1. Cost effective prospective terminal gear modifications designed to protect longline catch from whale depredation. 2. Demonstration of the functionality of these proof-of-concept catch protection devices in field tests and provide direction for further modifications and larger scale experimental testing.	IPHC-2023-SRB022-09/IPHC-2023-WM2023-12/BREP Interim Report May 2023	Fishing technology	Improve mortality accounting. Improve estimates of stock productivity.	Research outcomes may reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude.	11/01/2021-10/30/2023	Claude Dykstra/Ian Stewart	External (Bycatch Reduction Engineering Program - NOAA: Project NA21NMF4720534)	$99,700	Priority Rank #3
5	Use of artificial intelligence (AI) for determining the age of Pacific halibut from images of collected otoliths	The IPHC Secretariat is looking at options for supplementing current Pacific halibut ageing protocol with automatized ageing that does not require extensive otolith-reader training. The IPHC is investigating the potential use of artificial intelligence (AI) for determining the age of Pacific halibut from images of collected otoliths. The Secretariat is in the process of initializing creation of a database of pictures with expert-provided labels, utilizing previously aged otoliths, and assessing the option for the development of a Convolutional Neural Network (CNN) model specifically designed for image classification to determine Pacific halibut age. The goal is to create an AI-based age determination system that complements traditional methods for reliable fish stock assessment and management advice.	1. Develop a labeled image database from previously aged otoliths 2.Train and validate a CNN model for automated ageing 3. Verify the accuracy of the CNN model against traditional ageing methods	1. Predictive CNN model for ageing Pacific halibut complementing traditional methods 2. A report comparing CNNI model performance to traditional ageing techniques	N/A	Age composition data (both fishery-dependent and fishery-independent)	Age data is a critical input for stock assessment.	AI-driven age determination offers a critical enhancement to stock assessment methodologies, aiding in the estimation of growth rates, maturity, and population structure of Pacific halibut.	09/2023-12/2024+	Barbara Hutniczak	Internal	$0	Priority Rank #1

Proposed

Research Project #	Project Title	Abstract	Objectives	Deliverables	Progress report	5YPRIM Research area	Management implications	Specific inputs into management	Requested period of performance	PI	Targeted funding source	Requested budget	Research prioritization for SA/MSE
1	Genomic analyses of Pacific halibut in Washington State waters to inform population structure and dynamics affecting coastal communities	Current studies at the IPHC, with funding from a grant from the North Pacific Research Board (Project #2110; 2022-2024), are devoted to the application of genome-based approaches (i.e. low coverage whole genome resequencing, lcWGR) to investigate stock structure among known spawning groups of Pacific halibut in the Gulf of Alaska (as far South as Haida Gwaii), Bering Sea and Aleutian Islands. By leveraging the recently sequenced and annotated reference Pacific halibut genome (Jasonowicz et al., 2022; GCF_022539355.2), the IPHC has conducted lcWGR for a total of 600 individual samples from the above-mentioned spawning groups at a coverage of 3X. This effort has so far resulted in the identification of 11.5 million autosomal single nucleotide polymorphisms (SNPs), of which 4 million SNPs have a minor allele frequency higher than 0.05. Considerable progress is currently being made towards using genome approaches to establish a genetic baseline for the available spawning groups, and towards the development of genomic tools aimed at addressing important ecological, environmental, and management-related issues with respect to Pacific halibut in the Gulf of Alaska, Bering Sea and Aleutian Islands. However, the lack of genetic samples from spawning groups off the WA coast limits the application of the above-mentioned genomic tools to advance our understanding of population structure, movement, connectivity, adaptive characteristics, and environmental responses of Pacific halibut in Convention waters. Although no major spawning ground has been mapped south of Cape St. James in the southern tip of Haida Gwaii (St. Pierre, 1984), archeological records along with traditional and ecological knowledge from Indian Tribes (e.g., Makah tribe, etc.) that fished Pacific halibut in the winter off the WA coast indicate that Pacific halibut, at least historically, spawned in what is now IPHC Regulatory Area 2A (Salmen-Hartley, 2018). Additionally, contemporary reports of spawning Pacific halibut south of Cape Flattery and the existence of suitable spawning habitat for Pacific halibut (i.e., deep areas off the continental slope, 200-600 m) are strongly indicative of the presence of spawning grounds for Pacific halibut off the WA coast. Therefore, the identification of potential winter spawning groups of Pacific halibut in WA waters and their biological (i.e., genetic and reproductive) characterization are important for addressing key issues related to Pacific halibut that impact coastal communities within Convention Waters. The overarching goal of this proposal is to characterize the genetic composition of Pacific halibut found off the WA coast using state-of-the-art genomic approaches. The results of this proposal will improve our understanding	1. To identify winter spawning groups of Pacific halibut off the WA coast with the use of traditional and ecological knowledge and collect biological samples. 2. To characterize the reproductive condition of female and male Pacific halibut off the WA coast during the winter spawning season. 3. To generate and incorporate genomic data from winter spawning groups off the WA coast to existing data from winter spawning groups in other geographic areas in the northeastern Pacific Ocean to establish an expanded baseline of Pacific halibut genetic diversity.	1. Information on Pacific halibut spawning groups off the WA coast: location information, spawning time and collection of biological (genetic and reproductive) samples. 2. Extended baseline of Pacific halibut genetic diversity and delineation of fine-scale Pacific halibut stock structure in WA waters and coastwide.	N/A	Migration and Population Dynamics	Altered structure of future stock assessments and MSE operating models. Improved estimates of productivity coastwide.	Information of stock structure of the Pacific halibut population in Convention waters will inform management actions by validating management units. Research outcomes will be used to define management targets for minimum spawning biomass by Biological Region.	02/01/2024-1/31/2026	Josep Planas	External (Washington Sea Grant). Full proposal submitted in May 2023. Proposal not selected for funding.	$288,652	Priority Rank #2
2	Full scale testing of devices to minimize whale depredation in longline fisheries	In the north Pacific, both Killer (Orcinus orca) and Sperm (Physeter macrocephalus) whales are involved in depredation behavior in Pacific halibut (Hippoglossus stenolepis). In 2011 and 2012 fisheries observers estimated that 6.9% of Pacific halibut sets were affected by whale depredation in the Bering Sea (Peterson et al. 2014). Reductions in catch per unit effort (CPUE) when whales were present ranged across geographic regions from 5 15-57% for Pacific halibut (Peterson et al. 2014). These impacts also incur significant time, fuel, and personnel costs to fishing operations. From a fisheries management perspective, depredation creates an additional and highly uncertain source of mortality, loss of data (e.g. compromised survey activity), and reduces fishery efficiency. Stock assessments of both Pacific halibut (Stewart et al. 2020) and sablefish (Goethel et al. 2020) have adjusted their analysis of fishery independent data to account for the effects of whale depredation on catch rates. In the sablefish assessment, fishery limits are also adjusted downward to reflect expected depredation during the commercial fishery. Meanwhile, potential risks to the whales include physical injury due to being near vessels and gear, disruption of social structure (e.g., Chilvers and Corkeron 2001), and developing an artificial reliance on food items that can be affected by fishery dynamics. Many efforts have been made over the years to mitigate this problem, with fishers generally limited to simple methods that can be constructed, deployed, or enacted without significantly disrupting normal fishing operations, or without violating gear regulations. Existing approaches include catch protection, physical and auditory deterrents, and spatial or temporal avoidance. These approaches have had variable degrees of success and ease of adoption in each fishery (Werner et al. 2015) but none have provided a long-term solution. There are increasing data sources supporting the notion that technologies which reduce initial contact between gear and depredators will reduce the likelihood of foraging attempts around the gear, thereby sustaining levels of target catch while simultaneously reducing risk of depredator mortality and gear damage. Recent studies using physical catch protection methods include the development of underwater shuttles that unhook, and transport catch to the surface (Patagonian toothfish), light and expandable 'slinky' pots (sablefish), and flashers or mesh panels attached to the gear to obscure catch (tuna) (IPHC 2022). While slinky pots had quick uptake in the sablefish longline fishery, depredation occurring with this gear has been reported (Goethel et al. 2022), demonstrating the	1. Assess the performance of catch protection devices to effectively reduce depredation of longline captured fish in the presence of toothed whales. 2. Assess the performance metrics of catch protection devices on the size, number, and condition of fish successfully entrained in the devices	1. Further define and develop previously identified high priority work that can break the reward cycle of depredation behavior and thereby suppress its prevalence. 2. Build on strategies to protect already captured fish in cost effective manners that are compatible with currently employed hook and line fishing practices in the North Pacific halibut fishery.	N/A	Fishing technology	Improved accuracy of mortality estimates. Improve estimates of productivity	Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	11/1/2023-04/30/2025	Claude Dykstra/Ian Stewart	External (Bycatch Reduction Engineering Program - NOAA). Full proposal submitted in March 2023. Awarded.	$199,870	Priority Rank #3
3	Development of a non-lethal genetic-based method for aging Pacific halibut	Robust methods to estimate the ages of commercially exploited fish species are critical for stock assessment. Furthermore, when combined with data on other biological characteristics; such as length/weight, maturity, movement, and distribution; the age distribution or age structure of a fish population provides essential information on population dynamics related to age, predicted reproductive status, life history stage, etc. For Pacific halibut; an ecologically, economically and culturally important fish species in Alaska; age estimations are critical to our understanding of the composition of the stock for sustainable management, of historical changes in size-at-age, maturity-at-age, year class strength, mortality, etc., as well as of the response of the Pacific halibut stock to current and future climate variability. For many managed groundfish species, such as Pacific halibut, age has been traditionally estimated by manually counting the number of annuli or concentric lamellae present in sagittal otoliths (i.e. calcified structures located in the head that are used for balance and hearing) under a compound microscope. The International Pacific Halibut Commission (IPHC) has used sagittal otoliths for aging Pacific halibut since 1914, employing a method referred to as "surface aging" until 2002 and switching to a methodological variation known as "break-and-burn" thereafter (Forsberg, 2001). However, for various reasons, alternative methods to traditional otolith age estimations are being explored, developed and applied in fisheries. One of these is a genetic method for aging based on the known observation that the methylation patterns on genomic DNA change predictably with age. DNA methylation (DNAm) is an epigenetic modification of the DNA that consists in the covalent modification of cytosine, one of the four nucleobases found in DNA, and that regulates the expression of genes. Therefore, age-associated DNA methylation patterns can be modelled to generate molecular (i.e., epigenetic) age predictors capable of estimating chronological age with high accuracy. These are referred to as "epigenetic clocks" and can be developed from DNA isolated from any tissue, including non-lethal biological samples, such as a fin clip. Epigenetic clocks have been developed for many vertebrate species, including fish, with high accuracy (r between 0.84 and 0.99) and an average MAE of 0.87 years, that corresponds to 3.5% of the total lifespan of the species examined. Since DNA sequencing for measuring methylation levels is becoming cost effective and is a high throughput technique with little or no inherent human error or bias, epigenetic clocks have moved to the forefront among the alternative methods for aging that are currently available for fish species. The	1. To identify DNA methylation signals in Pacific halibut fin tissue. 2. To develop an age prediction model based on DNA methylation patterns: an epigenetic clock for Pacific halibut. 3. To develop a targeted DNA methylation assay for larger scale age estimations.	1. Reduced representation genome-wide map of DNA methylation at single base-pair resolution for Pacific halibut fin tissue. 2. Age predicting model for Pacific halibut using fin tissue.	N/A	Migration and Population Dynamics/Female Reproductive Assessment/Growth	Age is a critical input for stock assessment.	Age is a key biological input into stock assessment as it is used for estimating fish growth, fish maturity and fecundity-at-age, and mortality rates as well as population structure. Age distribution of Pacific halibut captured in the different fisheries and surveys is used in stock assessment.	02/01/2024-1/31/2026	Josep Planas	External (Alaska Sea Grant). Full proposal submitted in May 2023. Decision expected January 2024.	$60,374	Priority Rank #1

APPENDIX VI

Proposed schedule of outputs

	2022		2023		2024		2025		2026	
Biology and Ecology										
Migration and population dynamics										
Reproduction										
Growth										
Mortality and survival assessment										
Fishing technology										
Stock Assessment										
Management Strategy Evaluation										
Monitoring										

APPENDIX VII

Proposed schedule of funding and staffing indicators: Biology and Ecology

Research areas	Research activities	Required FTEs/Year	IPHC FTEs/Year	2022	2023	2024	2025	2026	IPHC Funds	Grant Funds
Migration and Population Dynamics	Larval and juvenile connectivity and early life history studies	0.45	0.45		RB1	RB2			Yes	NPRB #2100
	Population structure	0.4	0.8		RB1				No	NPRB #2110
	Adult migration and distribution	0.4							No	NPRB #2110
	Close-kin mark-recapture studies	1	0						No	Planned
	Seascape genomics	1	0						No	Planned
	Genome-wide association analyses	1	0						No	Planned
	Genomic-based aging methods	1	1		RS 1				Yes	No
Reproduction	Maturity-at-age estimations	0.75	0						Yes	No
	Fecundity assessment	0.5	0.25			RB4	RS 2		Yes	No
	Examination of accuracy of current field macroscopic maturity classification	0.25							Yes	No
	Sex ratio of current commercial landings	0.5	0.75	LT					Yes	No
	Recruitment strength and variability	0.5	0				RS 2		Yes	Planned
Growth	Environmental influences on growth patterns	0.5	0.5		MSc student				No	Planned
	Dietary influences on growth patterns and physiological condition	0.5	0.2		RB3				No	Planned
Mortality and survival assessment	Discard mortality rate estimate: recreational fishery	0.5	1						No	NPRB #2009
	Best handling practices: recreational fishery	0.5		RB 3					No	NPRB #2009
	Whale depredation accounting and tools for avoidance	0.5							No	BREP
	Biological interactions with fishing gear	0.5							No	BREP

IPHC staff (Planned):

RS1: Research Scientist 1(PhD; Life History Modeler I). Full time temporary position (100% research;
RS2: Research Scientist 1(PhD; Life History Modeler II). Full time temporary position (100% research;
RB1: Research Biologist 1 (Geneticist; MSc). Full time temporary position (until April 2022; 1 FTE). 55% of salary covered by Grant NPRB#2110.
RB2: Research Biologist 2 (Early Life History; MSc). Full time permanent position (40% research; 0.4 FTE)
RB3: Research Biologist 3 (DMR; MSc). Full time permanent position (100% research; 1 FTE)
RB4: Research Biologist 4 (Maturity and Fecundity; MSc). Full time permanent position (100% research; 1 FTE)
LT: Laboratory Technician (MSc). Full time temporary position (100% research; 1 FTE)

Report on Current and Future Biological and Ecosystem Science Research Activities

PREPARED BY: IPHC SECRETARIAT (J. PLANAS, C. DYKSTRA, A. JASONOWICZ, C. JONES, 07 MAY 2025)

PURPOSE

To provide the Scientific Review Board with a description of progress towards research activities described in the IPHC's five-year Program of Integrated Research and Monitoring (2022-2026).

BACKGROUND

The primary biological and ecological research activities at the IPHC that follow Commission objectives are identified and described in the IPHC Five-Year Program of Integrated Research and Monitoring (2022-2026). These activities are integrated with stock assessment (SA) and the management strategy evaluation (MSE) processes (Appendix I) and are summarized in five main areas, as follows:

1) Migration and Population Dynamics. Studies are aimed at improving current knowledge of Pacific halibut migration and population dynamics throughout all life stages in order to achieve a complete understanding of stock structure and distribution across the entire distribution range of Pacific halibut in the North Pacific Ocean and the biotic and abiotic factors that influence it.
2) Reproduction. Studies are aimed at providing information on the sex ratio of the commercial catch and to improve current estimates of maturity and fecundity.
3) Growth. Studies are aimed at describing the role of factors responsible for the observed changes in size-at-age and at evaluating growth and physiological condition in Pacific halibut.
4) Mortality and Survival Assessment. Studies are aimed at providing updated estimates of discard mortality rates in the guided recreational fisheries and at evaluating methods for reducing mortality of Pacific halibut.
5) Fishing Technology. Studies are aimed at developing methods that involve modifications of fishing gear with the purpose of reducing Pacific halibut mortality due to depredation and bycatch.

A ranked list of biological uncertainties and parameters for SA (Appendix II) and the MSE process (Appendix III) and their links to research activities and outcomes derived from the five-year research plan are provided.

SRB RECOMMENDATIONS AND REQUESTS

The SRB issued several recommendations and requests in their report of SRB024 (IPHC-2024-SRB024-R) in relation to presentation IPHC-2024-SRB024-09:

SRB025–Rec.09 (para. 35) The SRB ***RECOMMENDED*** *that when incorporating the new maturity ogive derived from the use of generalised additive models into the stock assessment, that the Secretariat consider using annual calculation of a regionally weighted ogive for years where FISS regional abundance estimates are available rather than one weighted by the 2023 FISS relative abundances by biological region.*

*SRB025–Rec.010 (para. 36) The SRB **NOTED** a decrease in the coastwide A50, driven largely by changes in Biological Region 2 from 2022 to 2023 and **RECOMMENDED**:*

a) not to pool years to inspect potential decreasing trends in the age at maturity;

b) investigating separately the maturity ogives and the age at the first maturity by determining, where possible, whether an individual has spawned previously.

The IPHC Secretariat has addressed this recommendation in Section 2 of this report and will present results at the SRB026 meeting.

*SRB025–Req.04 (para. 37) The SRB **REQUESTED** a preliminary evaluation of the feasibility for using information on the genetic differentiation of Pacific halibut parasites as a possible stock structure marker.*

The IPHC Secretariat has conducted literature searches on the types and prevalence of parasites in Pacific halibut and their outcomes will be discussed at the SRB026 meeting.

UPDATE ON PROGRESS ON THE MAIN RESEARCH ACTIVITIES

1. Migration and Population Dynamics.

   The IPHC Secretariat is currently focusing on studies that incorporate genomics approaches in order to produce useful information on population structure, distribution and connectivity of Pacific halibut. The relevance of research outcomes from these activities for stock assessment (SA) resides (1) in the introduction of possible changes in the structure of future stock assessments, as separate assessments may be constructed if functionally isolated components of the population are found (e.g. IPHC Regulatory Area 4B), and (2) in the improvement of productivity estimates, as this information may be used to define management targets for minimum spawning biomass by Biological Region. These research outcomes provide the second and third top ranked biological inputs into SA (Appendix II). Furthermore, the relevance of these research outcomes for the MSE process is in biological parameterization and validation of movement estimates, on one hand, and of recruitment distribution, on the other hand (Appendix III).

1.1. Population genomics. The primary objective of these studies is to investigate the genetic structure of the Pacific halibut population and to conduct genetic analyses to inform on Pacific halibut population dynamics and distribution within the Convention Area

Details on sample collection, sequencing, bioinformatic processing and proposed analyses utilizing low-coverage whole genome sequencing (lcWGR) to investigate Pacific halibut population structure were provided in documents IPHC-2021-SRB018-08, IPHC-2022-SRB021-09, IPHC-2023-SRB022-09 and IPHC-2024-SRB024-09.

1.1.1. Methods. We have conducted additional sequencing in order to balance the sample sizes for the sample collections that comprise our genetic baseline (i.e. samples collected in the winter during the spawning season) (Figure 1) and to increase the total number of samples available for analysis. We also included samples collected during the summer foraging season that will provide additional information over a larger geographic area that includes the latitudinal extremes of the range of Pacific halibut (Figure 1). In total, 384 additional samples were sequenced on a single Illumina NovaSeqX 25B lane by Novogene (Sacramento CA).

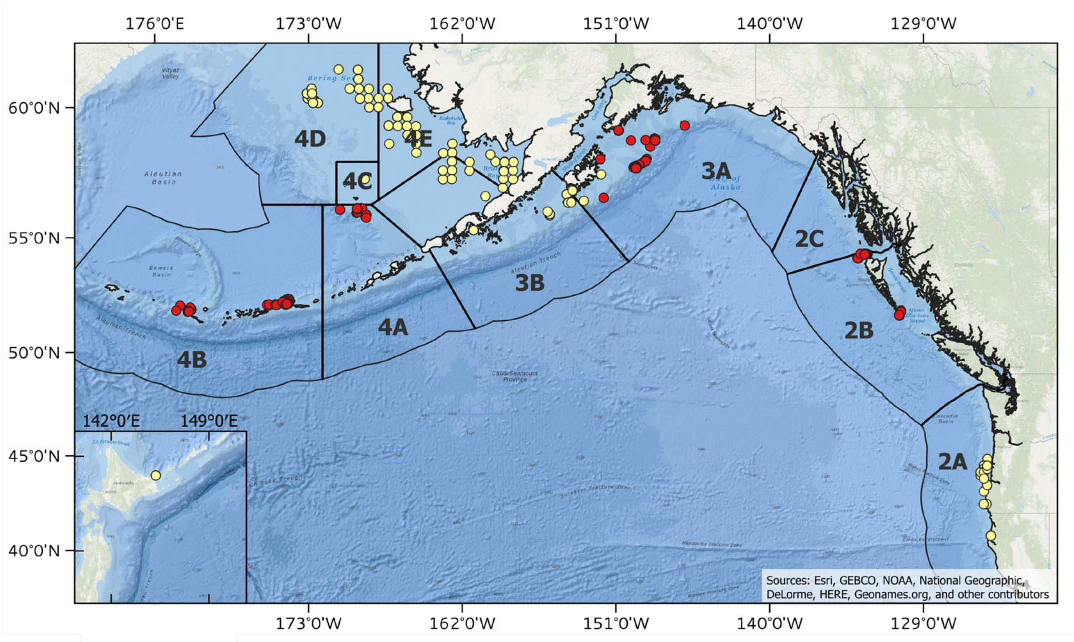


Figure 1. *Map showing sampling locations of Pacific halibut used for low-coverage whole genome resequencing. Red points indicate the locations of winter collected samples in the baseline set of samples and yellow points indicate the locations of samples collected during the summer.*

<u>Bioinformatic Processing</u>

The procedure used for the bioinformatic processing of the raw sequence reads and genotype likelihood estimation are detailed in IPHC-2023-SRB022-09 with the following modifications: individual samples were removed from the dataset if the average sequencing depth was < 1x, and we also applied the extended base alignment quality (baq) (Li 2011) adjustment during genotype likelihood estimation. We used ngsParalog (v1.3.3) (Linderoth 2018) to identify regions of the genome that may be problematic for mapping sequence reads. We first used ngsParalog to calculate a likelihood ratio that mis-mapped reads are covering individual SNPs. We then used average read depth observed at each SNP and the likelihood ratios in a Hidden Markov Model to identify the start and stop coordinates of genomic regions problematic for read mapping. SNPs contained within these regions were removed prior to any downstream analysis.

Population structure

To visualize patterns of population structure among spawning groups of Pacific halibut, we conduced Principal Component Analysis (PCA) using PCAngsd (v1.36.10) (Meisner and Albrechtsen 2018) to estimate a covariance matrix from the set of baseline samples. Eigendecomposition was performed in R (v4.3.3) (R Core Team 2022) using the eigen function. The percent variance explained for each principal component (PC) was calculated by dividing each eigenvalue associated with each PC by the sum of all eigenvalues. To determine an appropriate number of PCs to retain for downstream analyses, a scree plot of the first 20 eigenvalues was visually inspected and Cattell's rule (Cattell 1966) was used for this purpose. We also removed individuals identified as outliers in the PCA using a technique similar to that of Patterson et al. (2006). If any sample was > $|6\sigma|$ along one of the top 3 PCs we consider it an outlier, the sample was removed from the covariance matrix and performed eigendecomposition again. This procedure was conducted for up to 10 iterations or until no outliers remained. We then ran again PCAngsd one final time with the outliers removed to estimate the final covariance matrix to be analyzed. K-means clustering was then performed on the retained PCs using the kmeans function in R. To determine the optimal number of clusters (K) present in the data, we tested a range of K values (1 to 20) and used total within-cluster sum of squares (WSS) and Bayesian information criterion (BIC) to compare the K values tested and identify the best supported number of clusters.

Assignment testing

We also conducted assignment testing using the same procedure that is detailed in IPHC-2024-SRB024-09. With the increased samples sizes afforded by the additional baseline samples, we are able to potentially increase the accuracy of the population specific allele frequencies required for conducting individual assignment tests. In addition to the 50-50 train/test split, the increased sample sizes enabled us to construct a training set consisting of 45 individuals randomly selected from each geographic area and still have a reasonable number of samples (at least 16) left in each geographic area to be used in the test set.

1.1.2. Results

The initial bioinformatic processing of the raw sequence reads from the most recent sequencing run, yielded 16.9 million sequence reads per individual in average, resulting in an average coverage per sample of 2.4x. At a minimum coverage threshold of 1x, we were able to add 161 samples to our baseline dataset. Therefore, the final collection of genetic samples representing the complete baseline dataset to finalize our population genomic studies consists of 731 separate individuals (Figure 1, Table 1). Additionally, 136 summer collected samples were sequenced on the most recent sequencing run that, together with the summer samples already sequenced in previous sequencing runs, amount to a total of 327 summer collected samples (Figure 1, Table 2) that will be used to examine patterns of structure over a larger geographic scale.

We identified 8,460,466 SNPs in fully assembled autosomal regions of the Pacific halibut genome. Following the removal of 751,285 SNPs in regions of the genome identified as problematic for read mapping and removing SNPs with a global minor allele frequency (MAF) < 0.05, we retained 3,676,428 SNPs for further analysis.

	Winter Collections (baseline samples)				
	1999	2004	2007	2018	2020
British Columbia (winter)	59	63	61		
GOA (winter)	61	61	61	60	
Bering Sea (winter)		61	61		
Central AI (winter)			61		61
Western AI (winter)					61

Table 1. *Final sample sizes for each area in the baseline dataset by year of sample collection after a minimum sequencing depth threshold of 1x is applied.*

	Summer Collections				
	2013	2016	2019	2022	2024
Northern CA (summer)	46				
Southern OR (FISS)				45	
GOA (NMFS)			41		
Bering Sea (NMFS)		20	75	48	
Nothern Bering Sea (FISS)				48	
Rausu (summer)					4

Table 2. *Sample sizes for each area sampled during the summer by year of sample collection after a minimum sequencing depth threshold of 1x is applied.*

Population structure

The PCA outlier removal procedure identified 22 outliers in the baseline dataset resulting in 709 samples being retained for PCA and K-means clustering. Visual inspection of the top two PCs revealed a single cluster of samples with a large degree of overlap among the geographic areas (Figure 2). We retained the top 3 PCs for K-means clustering (Figure 3). Inspecting model selection measures of total within-clusters sum of squares and BIC, we observe a constant and continual decay as larger K-values are tested (Figure 4). Following the guidance of Jombart et al. (2010) on the use of BIC for selecting the best value of K, we were unable to confidently select an optimal value for K, the true number of clusters in the dataset. This is consistent with the lack of discrete genetic groups observed in Figure 2.

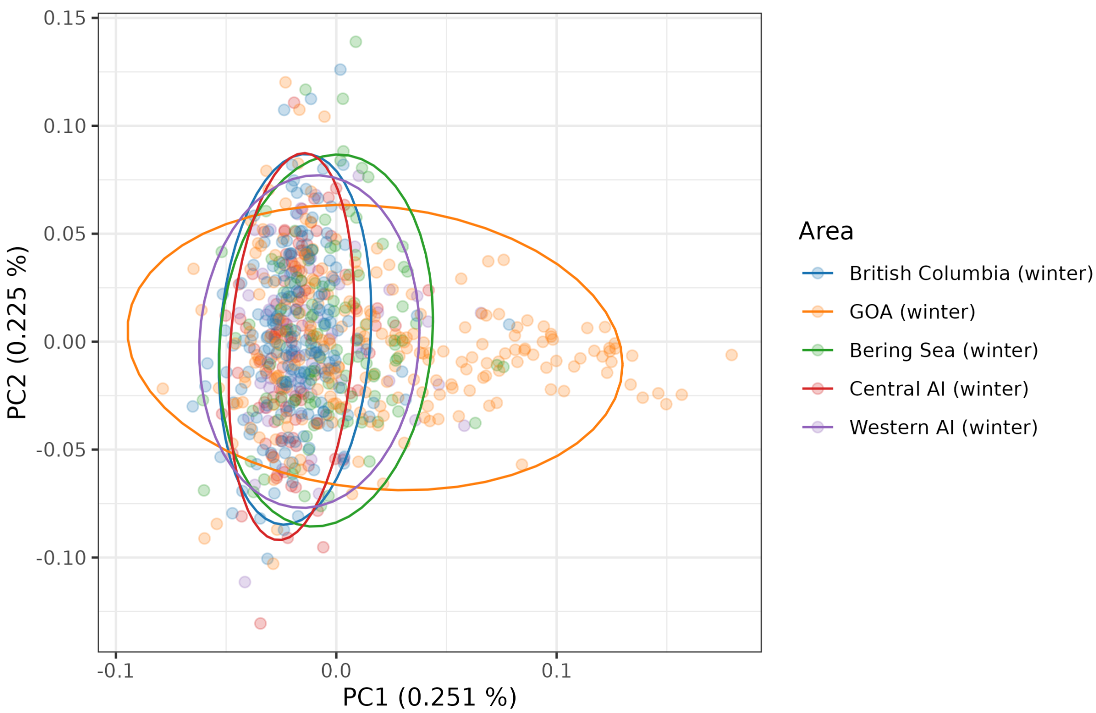


Figure 2. *PCA biplot of the first two PC axes for 709 Pacific halibut collected during the spawning season (winter) in IPHC Convention Waters. Individuals are colored by geographic area in all panels with 95% confidence ellipses drawn for each geographic area.*

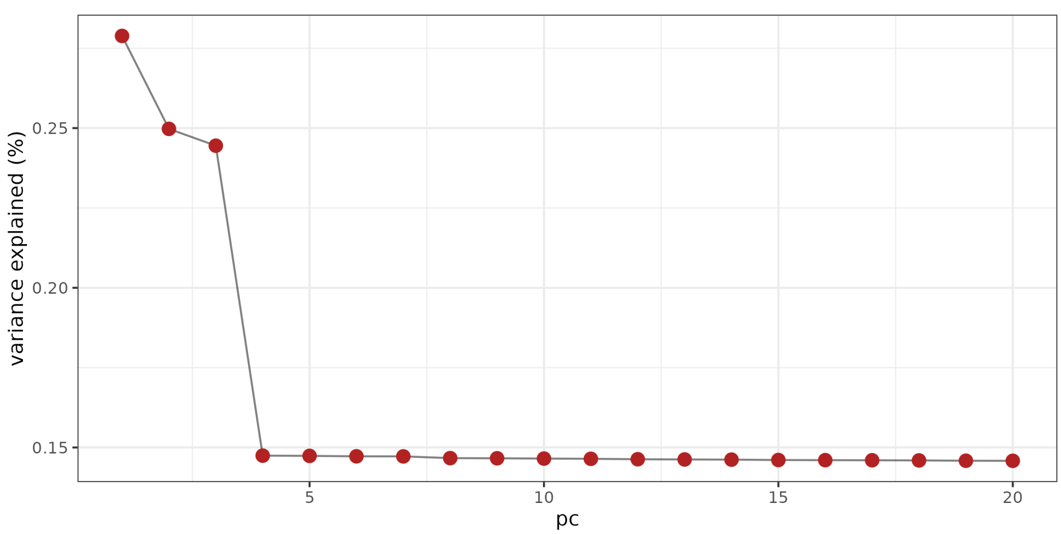


Figure 3. *Scree plot of the percent variance explained for the first 20 principal components (PCs) for the baseline set of samples.*

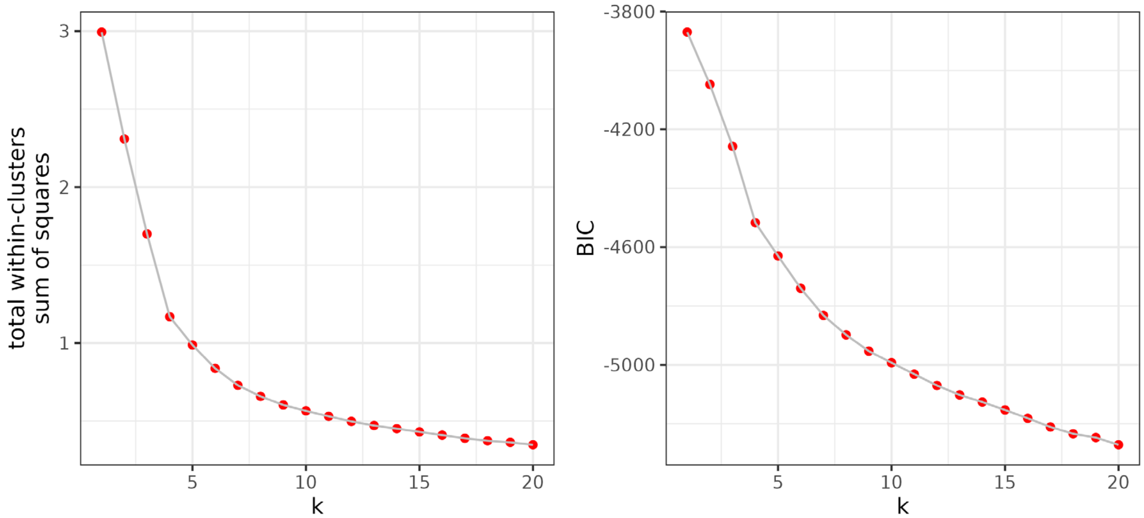


Figure 4. *Plots of total within-clusters sum of squares (A) and Bayesian information criterion (BIC) (B) for each value of K tested (1-20) for clustering analysis conducted using 709 Pacific halibut collected during the spawning season (winter) in IPHC Convention Waters.*

Assignment testing

After combining the top 1,000 SNPs from each pairwise comparison and removing any duplicates, the resulting set of SNPs consisted of 8,535 SNPs for the 50-50 train/test split, and 9,078 SNPs when the training set was constructed with equal number of samples (n=45) from each area. Results for the assignment testing for the 50-50 train/test split are nearly identical to those reported in IPHC-2024-SRB024-09, prior to the addition of the new samples (Table 1). We observed that all of the individuals in the test set were assigned back to the Gulf of Alaska with high confidence (> 95%), resulting in a relatively low overall assignment accuracy of 33.14% for the 50-50 train/test split (Figure 5A). Evaluation of the training set (50-50 split) using leave-one-out cross-validation yielded a 100% self-assignment rate of with all of the samples assigning back to the geographic area in which they were collected (Figure 5B). While increased samples sizes enabled us to construct a test set consisting of an equal number of samples per area and produce allele frequency estimates with similar precision for all areas in the baseline, we actually observed reduced overall assignment accuracy of 27.27% with 8.06% of the individuals being classified as unassigned (Figure 6A). Similar to the 50-50 train/test split, we observed 100% self-assignment when this training set was evaluated using leave-one-out cross-validation (Figure 6B). The assignment testing conducted here is very sensitive to how the training and test sets are constructed and, therefore, our interpretation of the assignment testing results is that we are capturing noise in the training set due to the lack of genetic structure.

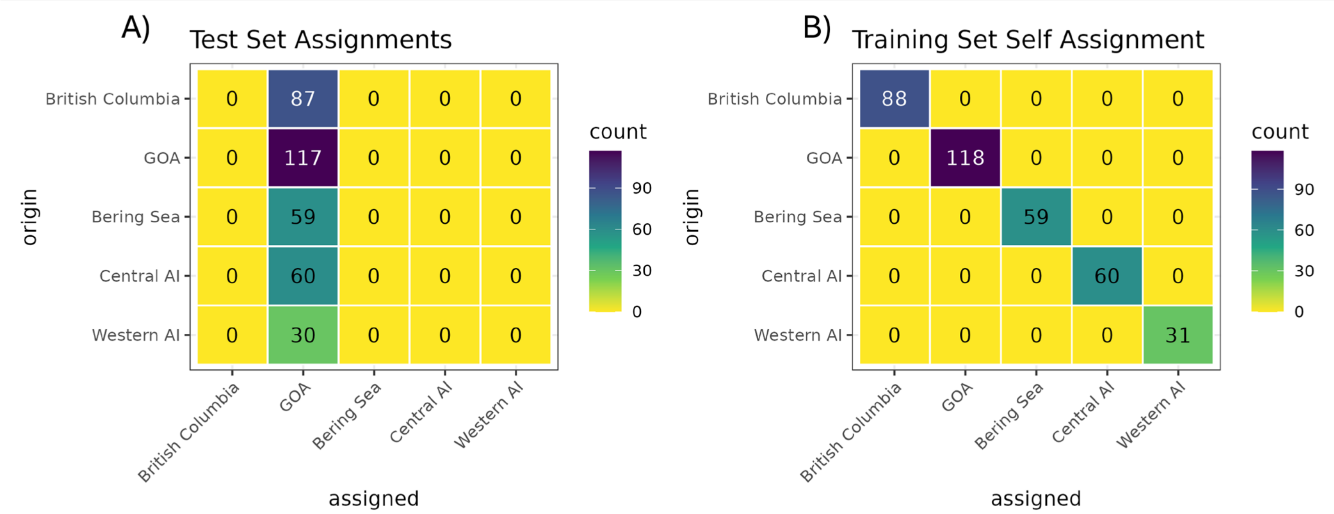


Figure 5. *Confusion matrices for individual population assignments when using a 50-50 train/test split and a set of 8,535 SNPs, requiring a minimum assignment probability of 95% for an individual to be assigned to a reference population. Geographic area of origin and assigned population are respectively shown on the x and y axes. A) Count of individuals in the validation set with assignments to the reference populations established using the training set. B) Assignment counts of individuals in the training set that self-assign to the reference populations, established using leave-one-out cross-validation.*

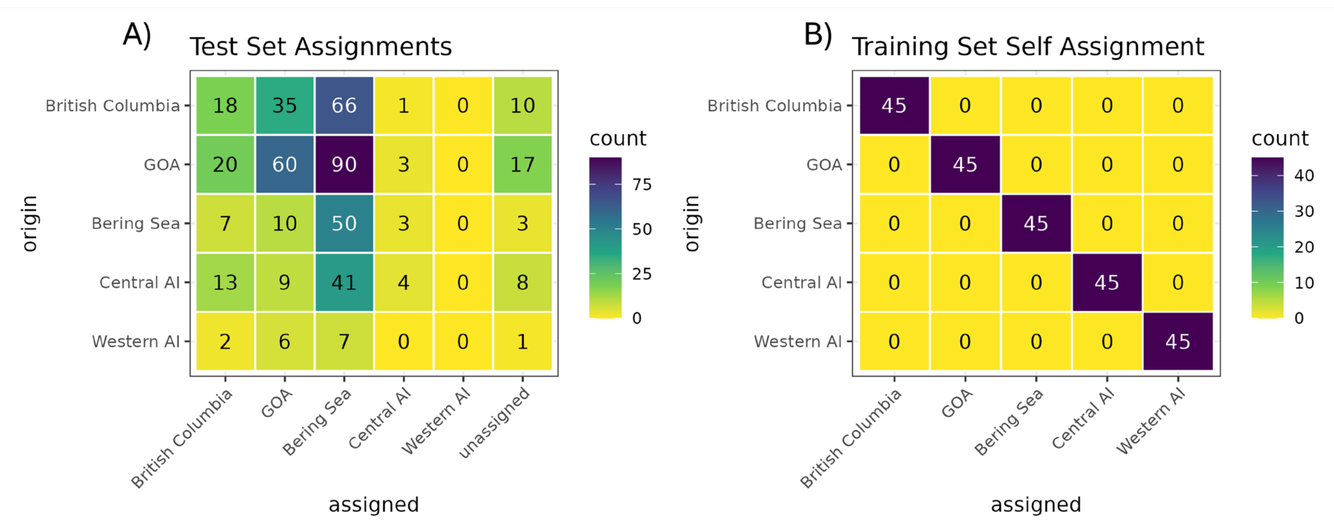


Figure 6. *Confusion matrices for individual population assignments when an equal number of samples per geographic area (n=45) were used for the training set and a set of 9,078 SNPs, requiring a minimum assignment probability of 95% for an individual to be assigned to a reference population. Geographic area of origin and assigned population are respectively shown on the x and y axes. A) Count of individuals in the validation set with assignments to the reference populations established using the training set. B) Assignment counts of individuals in the training set that self-assign to the reference populations, established using leave-one-out cross-validation.*

1.1.3. Conclusions. The inclusion of additional samples enabled us to improve the quality of our dataset by balancing the samples sizes for each sample collection in our baseline dataset and by increasing the total number of samples available for analysis. The results observed after adding additional baseline samples are very similar to those reported in IPHC-2024-SRB024-09. Overall, these results continue to support the notion that a single genetic group of Pacific halibut is present in IPHC Convention Waters. We are unable to confidently identify the presence of discrete genetic groups using unsupervised clustering analyses in the set of baseline samples despite the extensive set of samples being collected over broad geographic and temporal scales. The lack of population structure limits our ability to assign samples back to the location from which they were sampled from despite the increased sample sizes afforded by this final sequencing run.

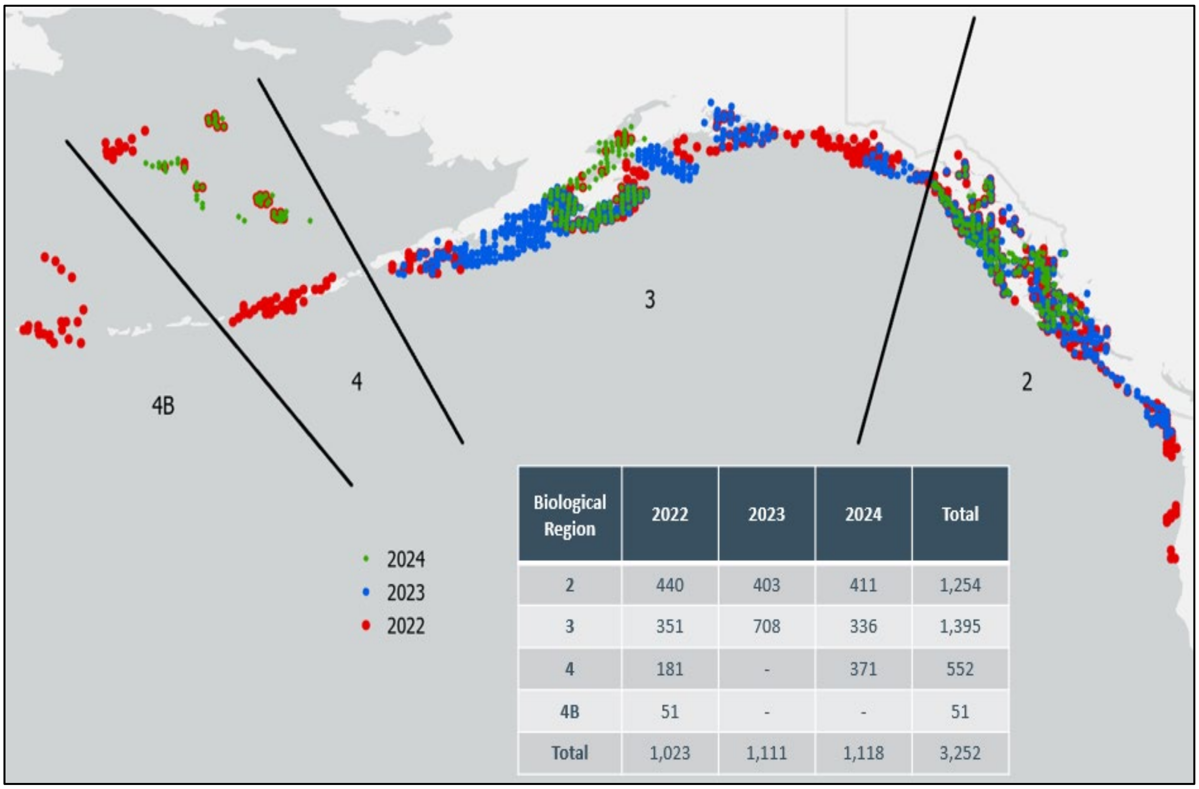
2. Reproduction.

Research activities in this Research Area aim at providing information on key biological processes related to reproduction in Pacific halibut (maturity and fecundity) and to provide sex ratio information of Pacific halibut commercial landings. The relevance of research outcomes from these activities for stock assessment (SA) is in the scaling of Pacific halibut biomass and in the estimation of reference points and fishing intensity. These research outputs will result in a revision of current maturity schedules and will be included as inputs into the SA (Appendix II), and represent some of the most important biological inputs for stock assessment (please see document IPHC-2021-SRB018-06). The relevance of these research outcomes for the management and strategy evaluation (MSE) process is in the improvement of the simulation of spawning biomass in the Operating Model (Appendix III).

2.1. Sex ratio of the commercial landings. The IPHC Secretariat is finalizing the processing of genetic samples from the 2024 aged commercial landings.

2.2. Reproductive assessment. Recent sensitivity analyses have shown the importance of changes in spawning output due to changes in maturity schedules and/or skip spawning and fecundity for SA (Stewart and Hicks, 2018). Information on these key reproductive parameters provides direct input to the SA. For example, information on fecundity-at-age and -size could be used to replace spawning biomass with egg output as the metric of reproductive capability in the SA and management reference points. This information highlights the need for a better understanding of factors influencing reproductive biology and success of Pacific halibut. To fill existing knowledge gaps related to the reproductive biology of female Pacific halibut, research efforts are devoted to characterizing female reproduction in this species. Specific objectives of current studies include: 1) update of maturity schedules based on histological-based data; and 2) calibration of historical visual maturity schedules using histological-based data.

2.2.1. Update of maturity schedules based on histological-based data. The IPHC Secretariat is undertaking studies to revise maturity schedules in all four IPHC Biological Regions through histological (i.e. microscopic) characterization of maturity, as reported previously. The coastwide maturity schedule (i.e. the proportion of mature females by age) that is currently used in SA was based on visual (i.e. macroscopic) maturity classification in the field (Fishery-independent Setline Survey (FISS)). To revise currently used maturity schedules, the IPHC Secretariat has collected ovarian samples for histology during the 2022, 2023 and 2024 FISS. The 2022 FISS sampling resulted in a total of 1,023 ovarian samples collected. Due to a reduced FISS design in 2023, sampling only occurred in Biological Regions 2 and 3 and resulted in a total of 1,111 ovarian samples collected. In 2024, 411, 336 and 371 ovarian samples were collected in Biological Regions 2, 3 and 4, respectively. In total, 3,252 ovarian samples have been collected for histology between 2022 and 2024 (Figure 7).



Biological Region	2022	2023	2024	Total
2	440	403	411	1,254
3	351	708	336	1,395
4	181	-	371	552
4B	51	-	-	51
Total	1,023	1,111	1,118	3,252

Figure 7. Map of 2022, 2023 and 2024 maturity samples for histology collected on FISS. Red dots (2022), blue dots (2023) and green dots (2024) indicate a distinct FISS station in which a sample was collected.

The IPHC Secretariat will continue to collect ovarian samples in the 2025 FISS. Targets for 2025 are to collect 400 samples in Biological Regions 2 and 3, 188 in Biological Region 4, and 414 in Biological Region 4B. These samples will allow us to further investigate both spatial and temporal differences in histological-based female Pacific halibut maturity.

Ovarian samples from 2022 to 2024 were processed for histology and we finalized scoring samples for maturity using histological maturity classifications, as previously described in Fish et al. (2020, 2022). Following this maturity classification criteria, all sampled Pacific halibut females were assigned to either the mature or immature categories. Mature female Pacific halibut are deemed to have at least reached the early vitellogenesis (Vtg1) stage of oocyte development.

Maturity ogives (i.e., the relationships between the probability of maturity determined by histological assessments and variables including IPHC Biological Region, age, and year) were estimated by fitting generalized additive models (GAM) with logit link (i.e., logistic regression). For example, if p_i is the probability that the *i*th sampled fish is mature, then a simple model with one explanatory variable x_i (e.g., age, log(age), length) would be:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{m=1}^{M} \beta_m x_{m,i}$$

where *f( )* is a smoothing function (thin plate regression spline) and β_0 is an intercept term. Our models also include variables for Biological Region and year to allow for spatial and temporal differences in the maturity relationships.

Alternative models were compared using the Akaike information criterion (AIC, Akaike 1973), with smaller AIC values indicating better fitting model. The models fitted with log(age) provided a better fit, with the estimated curves better matching the steep rise in the proportion of mature females from ages 6 to 8, and subsequent slower increase for older fish. For logistic GAM, models were fitted using function gam from the mgcv library (Wood 2006) in R 4.4.3.

We first ran again the best-fit logistic GAM models using log(Age), Biological Region, and year for the 2022-2024 samples. By examining the 2024 output for the logistic GAM (Figure 8), Biological Region 2 once again shows older maturity-at-age (indicated by the dashed lines for A50 and A95) and lower maturity-at-age from ages 10-20 than Biological Region 3. Biological Region 3 once again in 2024 shows a steep increase in maturity-at-age when compared to all other Biological Regions, with over 80% of mature females by age 9. Biological Region 4 shows a delayed start to maturation with only 5% of mature females at age 9 but maturation rapidly increases to ~90% mature females at age 15.

To examine temporal changes across all Biological Regions, we overlayed all three years of histological data by region (Figure 9). Overall, there is a significant shift to the left in maturity ogives from 2022 to 2024 in the three Biological Regions (2, 3, and 4) that have multiple years of data, indicating younger maturing females in 2024 than in 2022 and 2023. This could be indicative of a particular year class maturing through the population; however, this is difficult to discern with only three years of data. Biological Region 2 had a significant change from 2022 to 2023. With more individuals classified as mature between the ages of 8-20 in 2023 than in 2022, the rate of maturation in Biological Region 2 increased at younger ages causing the steepness of the curve to rapidly increase. There did not appear to be a difference between 2023 and 2024 for Biological Region 2. For Biological Region 3, there is a similar trend in that the maturity ogive has progressively shifted slightly to the left from 2022 to 2024. This indicates that a higher proportion of females at any given age are mature in 2024 compared to the previous two years. Biological Region 4 also showed a shift to the left from 2022 to 2024 (no data in 2023). It will be important to continue to monitor temporal trends in histological-based maturity ogives to determine if the observed shifts in maturity ogives continue.

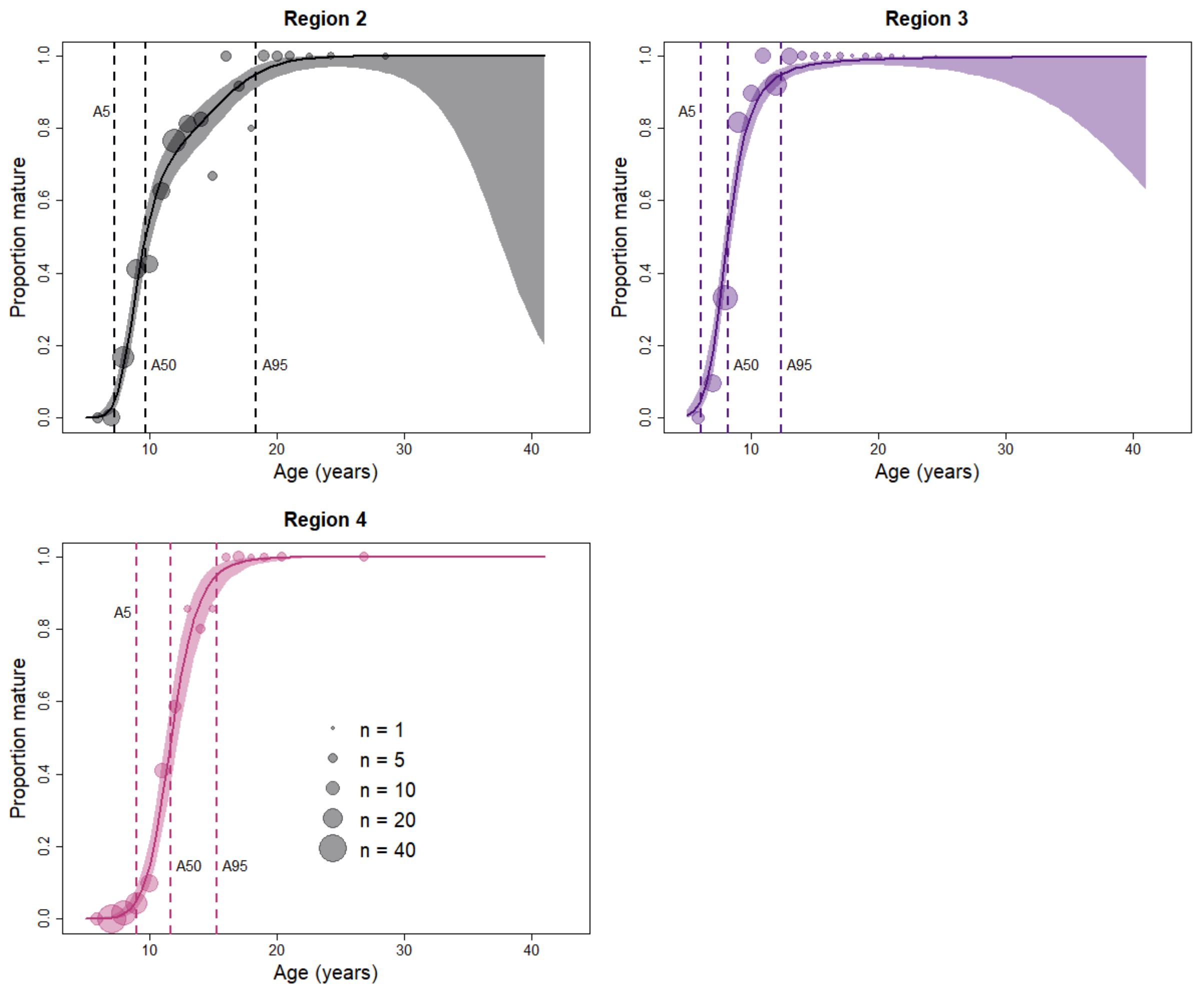


Figure 8. Female Pacific halibut age at maturity by IPHC Biological Region in 2024 using best-fit logistic GAM, with color shading indicating 95% CI for each IPHC Biological Region. Vertical dashed lines indicate proportion mature at 5% (A5), 50% (A50), and 95% (A95).

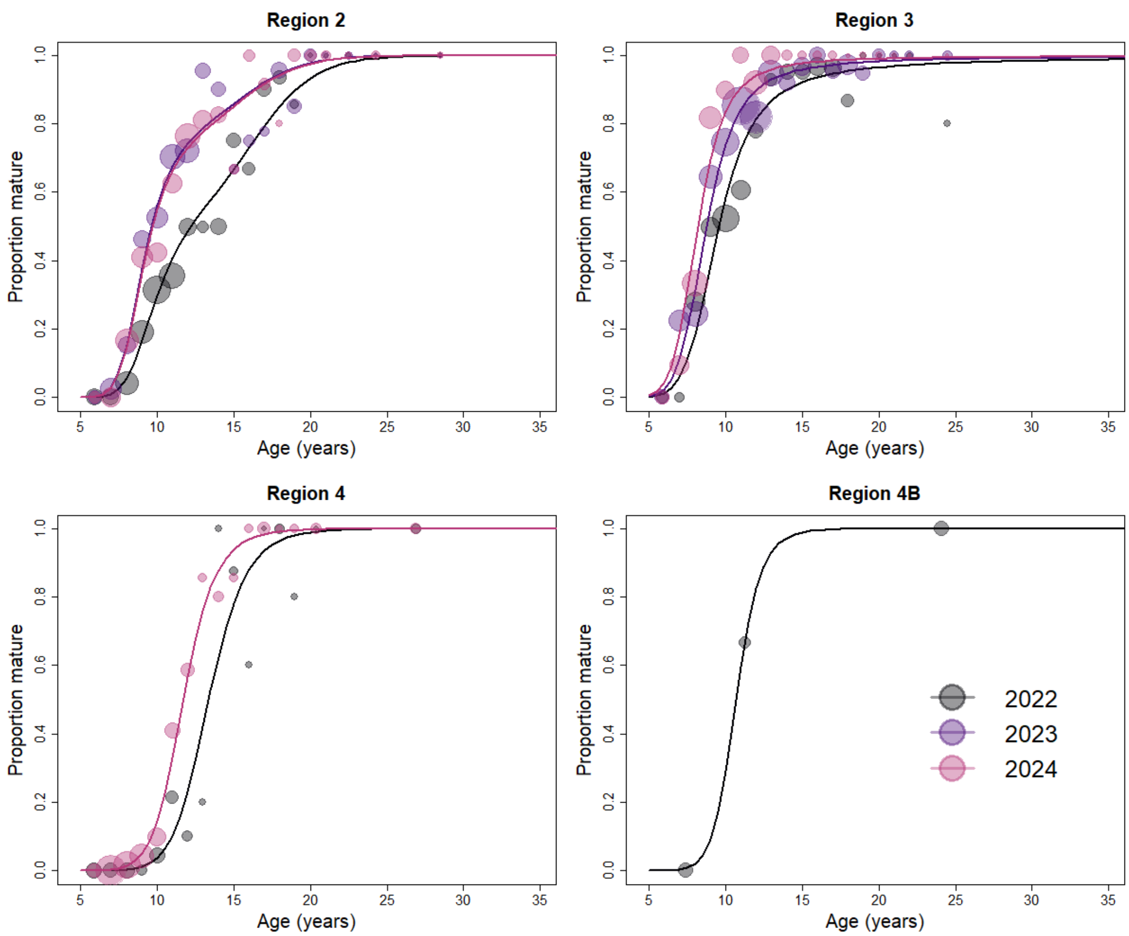


Figure 9. Female Pacific halibut age at maturity by IPHC Biological Region and year using best-fit logistic GAM.

To estimate a coastwide ogive with the 2022-2024 histology-based maturity information, we removed the year effect from the logistic GAM model and pooled all years by Biological Region. The logistic GAM estimated maturity curves for each IPHC Biological Region (Figure 10). Noting that sample size was not proportional to population size for each region, we used the average estimated regional abundance proportions from 2022-2024 from IPHC's space-time modeling of FISS numbers per unit effort (NPUE) data as weights in estimating a coastwide maturity ogive (Figure 11). The value of the coastwide ogive at each age is calculated as the abundance proportion at age times the proportion mature at age summed across regions. For example, for age, let q_j be the estimate of the abundance proportion for Biological Region *j*, and p_j(age) be the probability of maturity at age *a* estimated from fitting the model including both region and age as explanatory variables. Then, the coastwide maturity probability at age is estimated by

$$p_{CW}(a) = \sum_{j=1}^{4} q_j p_j(a)$$

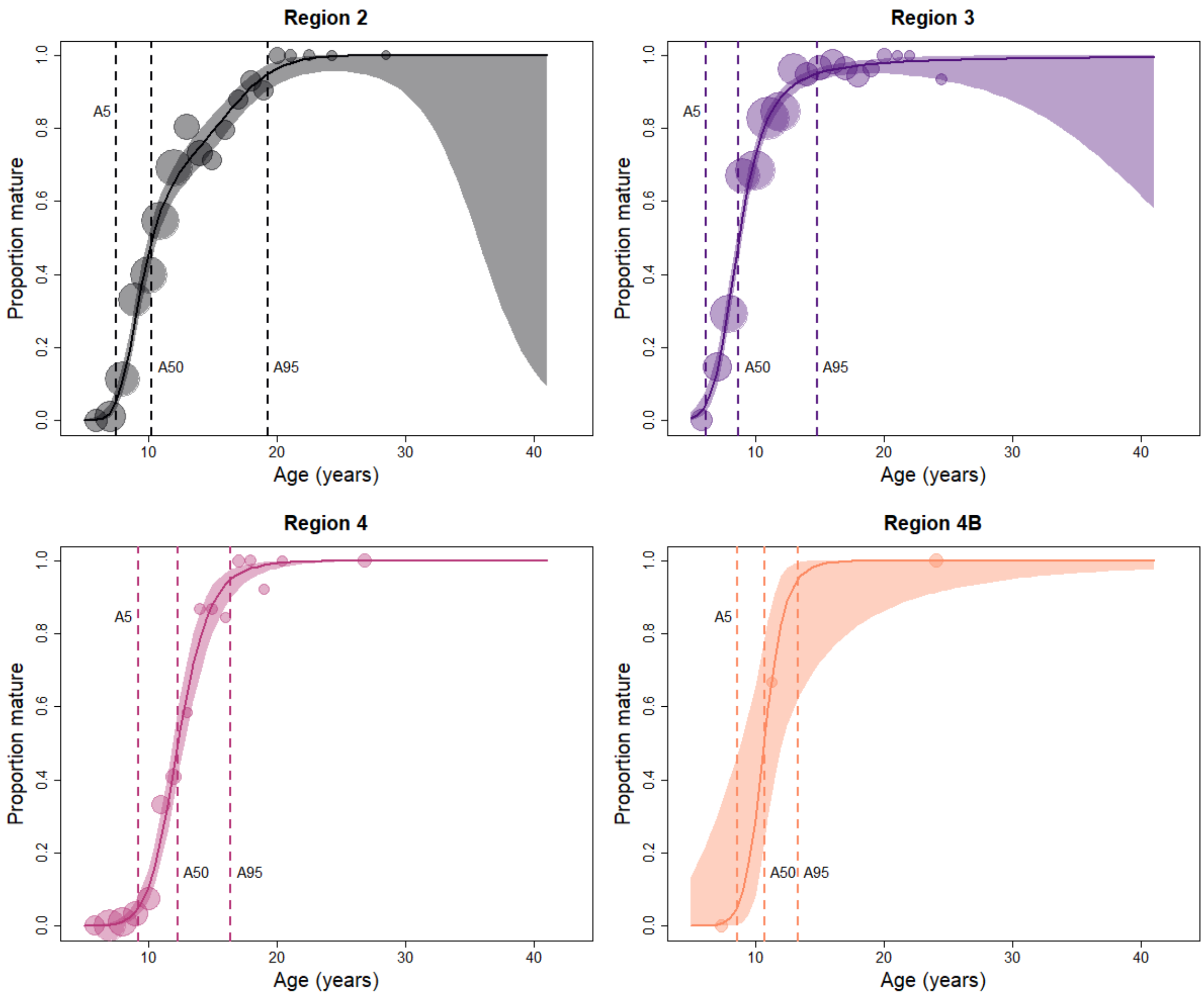


Figure 10. Female Pacific halibut age at maturity by IPHC Biological Region using best-fit logistic GAM, with color shading indicating 95% CI for each IPHC Biological Region. Vertical dashed lines indicate proportion mature at 5% (A5), 50% (A50), and 95% (A95).

The modeled coastwide ogive for maturity-at-age falls between the maturity ogives for Biological Regions 2 and 3 (Figure 11). This outcome was expected as these two Biological Regions currently have the highest estimated abundance. Maturity is used to assign the numbers of fish at each age in the SA model to either a reproductive or non-reproductive state. The total reproductive output of these fish in the SA is then estimated by multiplying the number of reproductive fish at each age by their average somatic weight and then by the fecundity per age or body weight (currently assumed to be 1 for all body weights and ages). Therefore, defining our coastwide maturity ogive in terms of numbers of fish is consistent with its use in the SA. Conversely, defining it in terms of biomass would require converting back to maturity in numbers for use in the SA. Age at 50% maturity (A50) was estimated from the coastwide ogive using an optimizing routine in R 4.4.3 (function optim) and was calculated to be 9.8

years, an almost two-year shift to younger maturing females when compared to our current maturity estimates from visual (field) data of 11.6 years.

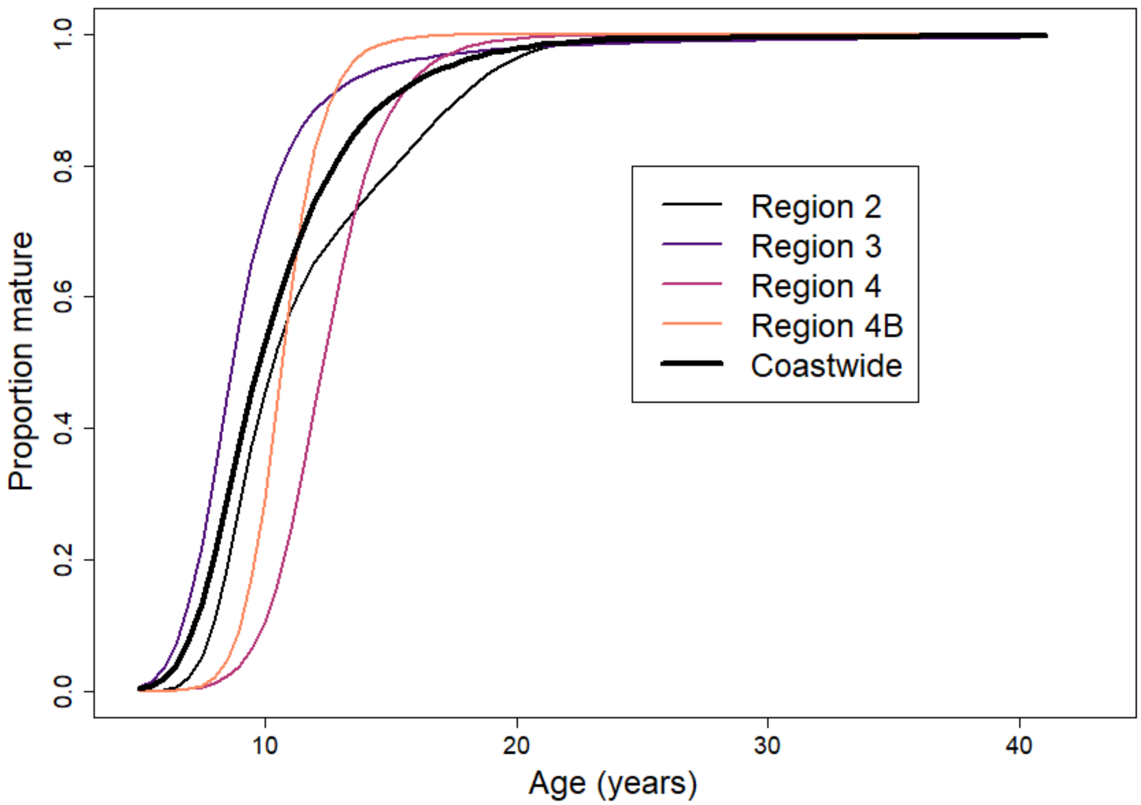


Figure 11. Coastwide maturity ogive generated from 2022-2024 average estimated regional abundance proportions (thick black line) and individual Biological Region ogives. Ogives shown without CI to better visualize differences between the coastwide and Biological Region ogives.

With the inclusion of histological data from 2024, we plotted the progression of female developmental stages by month of sampling and by Biological Region during the 2022-2024 period (Figure 12). Females in Biological Region 2 show a clear increase in the proportion of mature individuals from May (30%) until August (70%), with females advancing from Vtg1 to Vtg3 during this period. In contrast, the proportion of mature females in Biological Region 3 was already high in May (75%) and stayed elevated through September, with mature females progressively advancing through and nearing completion of vitellogenesis by that time. Biological Region 4 had a smaller proportion of mature females from June (25%) to August (40%), but about 50% of mature females were already at the Vtg3 stage in August. With only samples collected in June, mature females in Biological Region 4B appeared to undergo earlier ovarian development than females in other Biological Regions with ~50% of individuals at Vtg2 or more advanced stages, showing even signs of completion of vitellogenesis. With three years (2022-2024) of ovarian samples, the temporal analysis of ovarian development in mature females is consistent across Biological Regions and years, and provides useful insights into the existence of differences related to the timing of ovarian development in mature females throughout Convention waters.

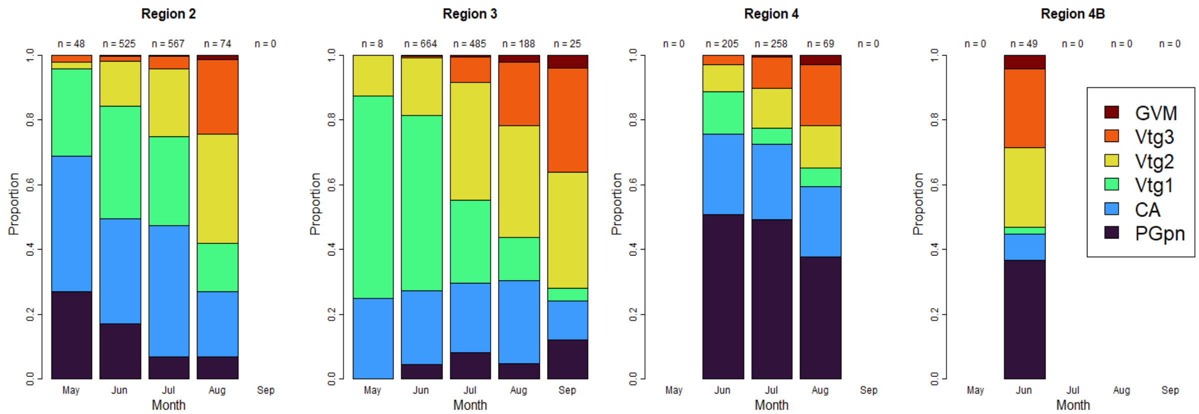


Figure 12. Reproductive development of female Pacific halibut by month sampled and IPHC Biological Region from 2022 to 2024. Number of samples (n) collected by month shown at the top of each stacked bar.

2.2.2. <u>Calibration of historical visual maturity schedules using histology-based data.</u> After creating a new coastwide maturity ogive using histology-based maturity estimates from 2022 to 2024, we investigated how visual maturity estimates have changed over the same timeframe. All females that we obtained a histology sample from also received a visual maturity estimate in the field. Using the same logistical GAM and methods used to create a coastwide ogive from the histology-based maturity data, we created a new coastwide visual maturity ogive (Figure 13, blue line).

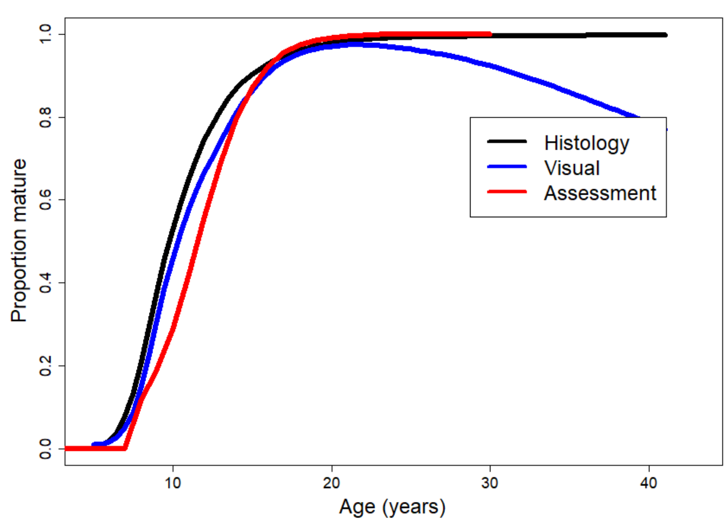


Figure 13. Coastwide maturity ogive generated from 2022-2024 average estimated regional abundance proportions using histological (black) and visual (blue) maturity estimation methods. Current coastwide ogive (red) used in stock assessment shown for reference.

The A50 value of the 2022-2024 coastwide visual maturity ogive was calculated to be 10.3 years. When comparing the new coastwide visual ogive to the current SA ogive (Figure 13, red line), a shift to the left is observed, with a higher proportion of mature females observed between the ages of 8 to 13 years. The drop in the proportion of mature females for older individuals in the new visual maturity ogive was caused by two older females (25-30 years old) that were visually classified as immature in the field.

The IPHC Secretariat has visual maturity assessment data from the FISS going back to 2002 with ages determined using the current break-and-burn ageing method. In order to create a time series consistent with the more accurate histological assessments, we first developed a calibration between histological and visual maturity curves from the 2022-2024 data. Let $p_h(a)$ and $p_v(a)$ be, respectively, the histological and visual maturity values at age. We can think of the histological curve as an adjusted version of the visual curve:

$$\operatorname{logit}\left\{p_h\left(a\right)\right\} = \operatorname{logit}\left\{p_v\left(a\right)\right\} + \delta\left(a\right)$$

Here $\delta(a)$ is the calibration factor at age a. Note that working on the logit scale ensures that all probabilities remain between 0 and 1 for the resulting calibrated curve. Rearranging, and using the probabilities estimated by fitting GAM models, we get estimates of the calibration factors, $\hat{\delta}(a)$:

$$\hat{\delta}(a) = \log\left\{\frac{\hat{p}_h(a)}{\hat{p}_v(a)} . \frac{1-\hat{p}_v(a)}{1-\hat{p}_h(a)}\right\}$$

These values were estimated based on the 2022-2024 maturity data and then used to create calibrated maturity curves from estimated visual curves for any given year, y:

$$\hat{p}_h(a,y) = \operatorname{logit}^{-1}\left\{\operatorname{logit}\left\{\hat{p}_v(a,y)\right\} + \hat{\delta}(a)\right\}$$

Just as maturity curves are estimated for each Biological Region, we estimated separate calibration factors for each region. The estimated values are shown in Figure 14. Positive values indicate that the curves based on visual assessments are shifted upwards for a given age, while negative values indicate that the calibration shifts the curves downwards. It is possible that differences between visual and histological assessments vary with time, due to observer differences and to other factors. This is something we can examine as we collect histological data over a greater number of years, although our ability to account for such factors when calibrating historical curves could be limited.

Figure 15 shows coastwide maturity curves by year estimated from visual maturity assessment data. Each curve was estimated using three-year rolling data windows, e.g., the 2003 curve is estimated from 2002-2004 data. Three years is the minimum timeframe that ensures that there are data in all Biological Regions within each rolling window. For the ends of the visual assessment time series, i.e. 2002 and 2024, where the three-year data window includes years with no observations (2001 or 2025), we expanded the window to ensure that three years of data were included in the analysis. This indicates that the logistic GAM models for 2002 and 2003 use the same data (from 2002-2004), as is the case for 2023 and 2024 (data from 2022-2024). Corresponding calibrated curves are shown in Figure 16. To obtain a final coastwide calibrated visual maturity ogive for the 2002-2024 time series, we averaged across all three-year rolling data windows (i.e. 2002-2004, 2003-2005, 2004-2006, etc.). This is depicted with the mean calibrated visual ogive shown in Figure 17.

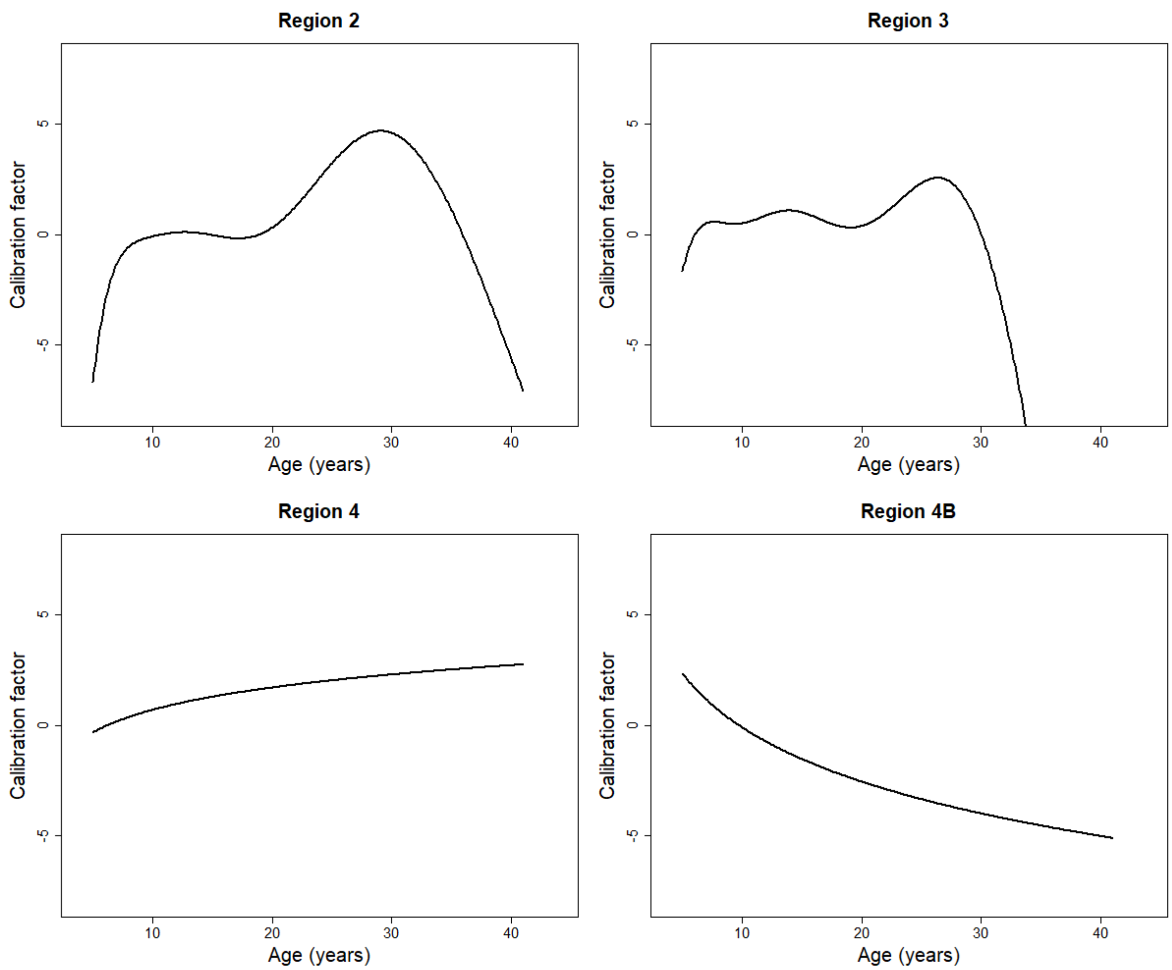


Figure 14. Calibration factors estimated from 2022-2024 visual and histological maturity ogives by Biological Region.

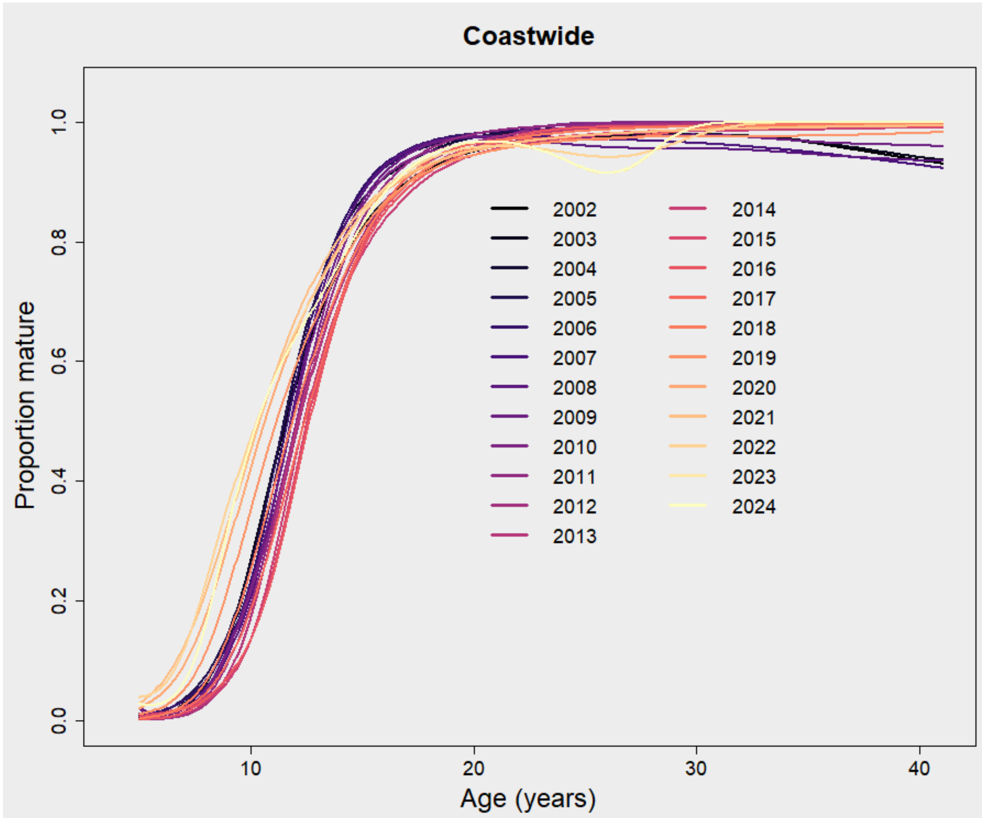


Figure 15. Estimated maturity ogives as a function of age based on visual maturity assessment data from rolling three-year data windows from 2002-2024.

When comparing the new coastwide calibrated visual maturity ogive to the current ogive used in the SA, the curve shifted slightly to the left from ages 8-15 (Figure 18, overlapping black and green lines). The calibrated visual ogive has a calculated A50 of 11.0 years, lower than the A50 value of 11.6 from the current SA ogive (red line), and evidences a slight decrease in the proportion of mature females from ages 15-20 years. These shifts in the maturity curves are to be expected as the histology-based data provide a better indicator of younger maturing females, but also older immature females. It is important to note is that these maturity ogives do not offer a direct comparison, given that the current SA ogive is based on visual estimates exclusively from IPHC Regulatory Areas 2B and 3A, whereas the new calibrated ogive incorporates data from all four Biological Regions. For input into the SA, we truncated the new calibrated ogive at age 7 years (Figure 18, green line) as histology-based maturity estimations did not find females < 7 years old that were mature. Previous maturity ogives using visual estimates truncated the curve at age 8 years. The impacts and sensitivity analyses of the new calibrated and truncated visual maturity ogive for female SSB in the SA are shown in document IPHC-2025-SRB026-07.

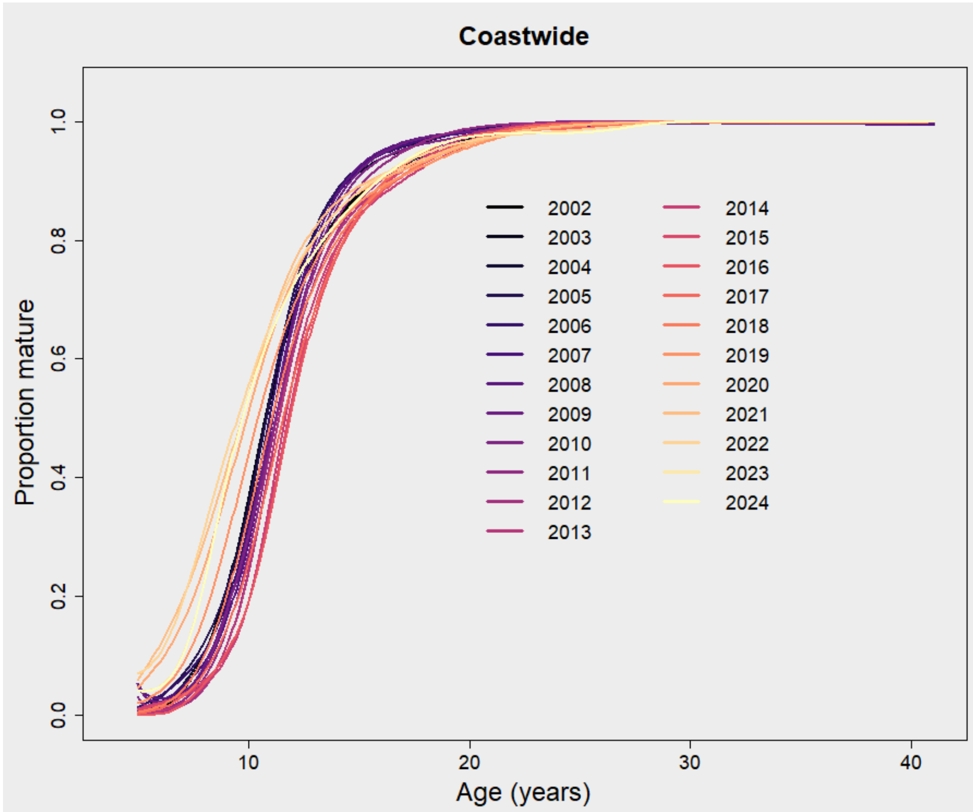


Figure 16. Estimated maturity ogives as a function of age calculated by applying the estimated calibration factors to the curves estimated from visual maturity assessment data from Figure 15.

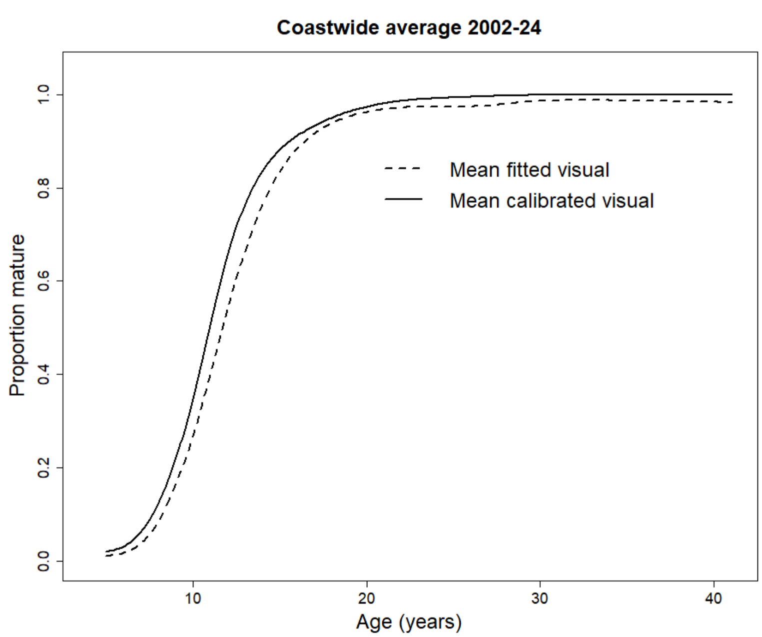


Figure 17. Mean of estimated coastwide visual and calibrated visual maturity ogives from Figures 15 and 16.

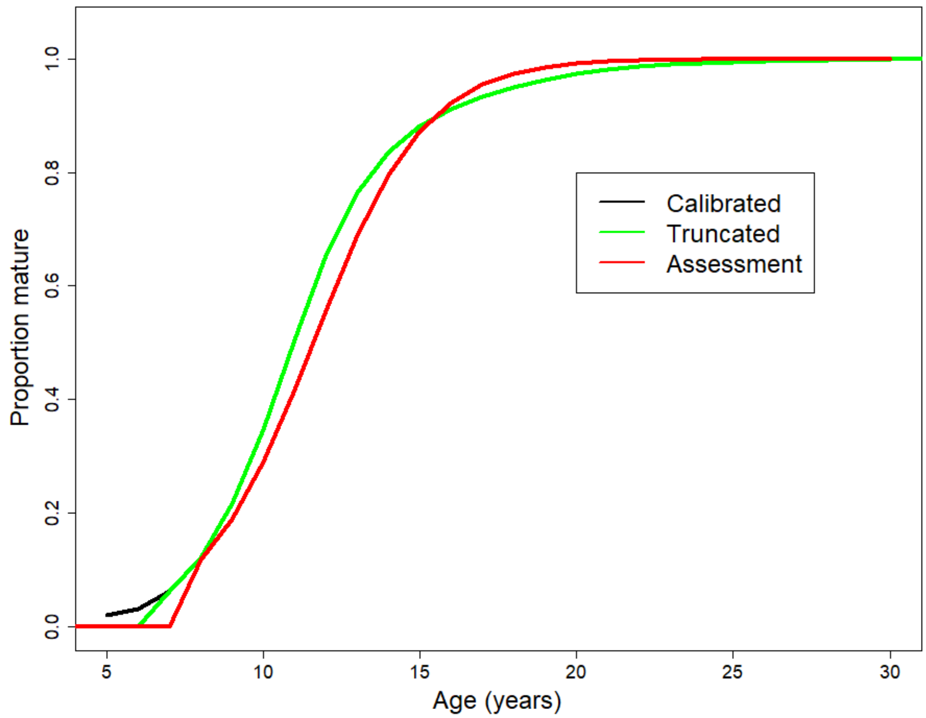


Figure 18. Estimated mean calibrated visual maturity ogive (black) with same ogive overlayed but truncated to zero at age 7 (green). Current coastwide ogive (red) used in stock assessment shown for reference.

3. Growth.

Research activities conducted in this Research Area aim at providing information on somatic growth processes driving size-at-age in Pacific halibut. The relevance of research outcomes from these activities for stock assessment (SA) resides, first, in their ability to inform yield-per-recruit and other spatial evaluations for productivity that support mortality limit-setting, and, second, in that they may provide covariates for projecting short-term size-at-age and may help delineate between fishery and environmental effects, thereby informing appropriate management responses (Appendix II). The relevance of these research outcomes for the management and strategy evaluation (MSE) process is in the improvement of the simulation of variability and to allow for scenarios investigating climate change (Appendix III).

The IPHC Secretariat has conducted studies aimed at elucidating the drivers of somatic growth leading to the decline in SAA by investigating the physiological mechanisms that contribute to growth changes in the Pacific halibut. The two main objectives of these studies have been: 1) the identification and validation of physiological markers for somatic growth; and 2) the application of molecular growth markers for evaluating growth patterns in the Pacific halibut population.

A manuscript describing the results of these studies is being prepared for submission to a peer-reviewed journal.

4. Mortality and Survival Assessment.

Information on all Pacific halibut removals is integrated by the IPHC Secretariat, providing annual estimates of total mortality from all sources for its stock assessment. Bycatch and wastage of Pacific halibut, as defined by the incidental catch of fish in non-target fisheries and by the mortality that occurs in the directed fishery (i.e. fish discarded for sublegal size or regulatory reasons), respectively, represent important sources of mortality that can result in significant reductions in exploitable yield in the directed fishery. Given that the incidental mortality from the commercial Pacific halibut fisheries and bycatch fisheries is included as part of the total removals that are accounted for in stock assessment, changes in the estimates of incidental mortality will influence the output of the stock assessment and, consequently, the catch levels of the directed fishery. Research activities conducted in this Research Area aim at providing information on discard mortality rates and producing guidelines for reducing discard mortality in Pacific halibut in the longline and recreational fisheries. The relevance of research outcomes from these activities for stock assessment (SA) resides in their ability to improve trends in unobserved mortality to improve estimates of stock productivity and represent the most important inputs in fishery yield for stock assessment (Appendix II). The relevance of these research outcomes for the management and strategy evaluation (MSE) process is in fishery parametrization (Appendix III).

4.1. Estimation of discard mortality rates in the charter recreational sector. Results from a recently completed study investigating discard mortality rates and characteristics of fish captured and released using guided recreational fishery practices are currently being prepared for publication in a peer-reviewed journal.

5. Fishing technology.

The IPHC Secretariat has determined that research to provide the Pacific halibut fishery with tools to reduce whale depredation is considered a high priority (Appendix I). This research is now contemplated as one of the research areas of high priority within the 5-year Program of Integrated Research and Monitoring (2022-2026). Towards this goal, the IPHC secretariat has been investigating gear-based approaches to catch protection as a means for minimizing whale depredation in the Pacific halibut and other longline fisheries with funding from NOAA's Bycatch Research and Engineering Program (BREP) (NOAA Awards NA21NMF4720534 and NA23NMF4720414; Appendix IV). The results and outcomes of the initial pilot phase of this project were reported in the documentation provided for the previous SRB meetings: IPHC-2022-SRB020-08 and IPHC-2024-SRB024-09.

The second phase of this project will involve refining effective methods for protecting longline captured fish from depredation, and conducting tests in the presence of toothed whales in known depredation hotspots to demonstrate the efficacy and safety of the gear. Plans are underway to conduct the second phase of this project in IPHC Regulatory Area 4A during the second half of May 2025.

RECOMMENDATION/S

That the SRB:

a) **NOTE** paper IPHC-2025-SRB026-06 which provides a response to Recommendations and Requests from SRB025, and a report on current biological research activities contemplated within the IPHC's five-year Program of Integrated Research and Monitoring (2022-26).

REFERENCES

Akaike, H. 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika, 60(2), 255-265.

Cattell, R.B. 1966. The Scree Test For The Number Of Factors. Multivariate Behavioral Research 1(2): 245--276. doi:10.1207/s15327906mbr0102_10.

Fish, T., Wolf, N., Harris, B.P., Planas, J.V. 2020. A comprehensive description of oocyte developmental stages in Pacific halibut, *Hippoglossus stenolepis*. Journal of Fish Biology. 97: 1880-1885. doi: https://doi.org/10.1111/jfb.14551.

Fish, T., Wolf, N., Smeltz, T.S., Harris, B.P., Planas, J.V. 2022. Reproductive biology of female Pacific halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska. Frontiers in Marine Science. 9: 801759. doi: https://doi.org/10.3389/fmars.2022.801759.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet **11**(1): 94. doi:10.1186/1471-2156-11-94Li, H. 2011. Improving SNP discovery by base alignment quality. Bioinformatics **27**(8): 1157-1158. doi:10.1093/bioinformatics/btr076.

Linderoth, T.P. 2018. Identifying Population Histories, Adaptive Genes, and Genetic Duplication from Population- Scale Next Generation Sequencing. University of California, Berkeley.

Meisner, J., and Albrechtsen, A. 2018. Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. Genetics 210(2): 719--731. doi:10.1534/genetics.118.301336.

Patterson, N., Price, A.L., and Reich, D. 2006. Population structure and eigenanalysis. PLoS Genet **2**(12): e190. doi:10.1371/journal.pgen.0020190.

R CoreTeam. 2024. R: A language and environment for statistical computing (v4.4.3).

Stewart, I., and Hicks, A. 2018. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2017. Int. Pac. Halibut Comm. Annual Meeting Report: IPHC-2018-AM094-10.

Wood, S. 2006. Generalized additive models: an introduction with R. Chapman and Hall/CRC, New York, New York, USA.

APPENDIX I

Integration of biological research, stock assessment (SA) and management strategy evaluation (MSE): rationale for biological research prioritization

Research areas	Research activities	Research outcomes	Relevance for stock assessment	Relevance for MSE	Specific analysis input	SA Rank	MSE Rank	Research priorization
Migration and population dynamics	Population structure	Population structure in the Convention Area	Altered structure of future stock assessments	Improve parametization of the Operating Model	If 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area	2. Biological input	1. Biological parameterization and validation of movement estimates and recruitment distribution	2
	Distribution	Assignment of individuals to source populations and assessment of distribution changes	Improve estimates of productivity		Will be used to define management targets for minimum spawning biomass by Biological Region	3. Biological input		2
	Larval and juvenile connectivity studies	Improved understanding of larval and juvenile distribution	Improve estimates of productivity		Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	3. Biological input	1. Biological parameterization and validation of movement estimates	2
Reproduction	Histological maturity assessment	Updated maturity schedule	Scale biomass and reference point estimates	Improve simulation of spawning biomass in the Operating Model	Will be included in the stock assessment, replacing the current schedule last updated in 2006	1. Biological input		1
	Examination of potential skip spawning	Incidence of skip spawning			Will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment			1
	Fecundity assessment	Fecundity-at-age and -size information			Will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points			1
	Examination of accuracy of current field macroscopic maturity classification	Revised field maturity classification			Revised time-series of historical (and future) maturity for input to the stock assessment			1
Growth	Evaluation of somatic growth variation as a driver for changes in size-at-age	Identification and application of markers for growth pattern evaluation	Scale stock productivity and reference point estimates	Improve simulation of variability and allow for scenarios investigating climate change	May inform yield-per-recruit and other spatial evaluations of productivity that support mortality limit-setting		3. Biological parameterization and validation for growth projections	5
		Environmental influences on growth patterns			May provide covariates for projecting short-term size-at-age. May help to delineate between effects due to fishing and those due to environment, thereby informing appropriate management response			5
		Dietary influences on growth patterns and physiological condition			May provide covariates for projecting short-term size-at-age. May help to deleineate between effects due to fishing and those due to environment, thereby informing appropriate management response			5
Mortality and survival assessment	Discard mortality rate estimate: longline fishery	Experimentally-derived DMR	Improve trends in unobserved mortality	Improve estimates of stock productivity	Will improve estimates of discard mortality, reducing potential bias in stock assessment results and management of mortality limits	1. Fishery yield	1. Fishery parameterization	4
	Discard mortality rate estimate: recreational fishery				Will improve estimates of discard mortality, reducing potential bias in stock assessment results and management of mortality limits			4
	Best handling and release practices	Guidelines for reducing discard mortality			May reduce discard mortality, thereby increasing available yield for directed fisheries	2. Fishery yield		4
Fishing technology	Whale depredation accounting and tools for avoidance	New tools for fishery avoidance/deterence; improved estimation of depredation mortality	Improve mortality accounting	Improve estimates of stock productivity	May reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude	1. Assessment data collection and processing		3

APPENDIX II

List of ranked biological uncertainties and parameters for stock assessment (SA) and their links to biological research areas and research activities

SA Rank	Research outcomes	Relevance for stock assessment	Specific analysis input	Research Area	Research activities
1. Biological input	Updated maturity schedule	Scale biomass and reference point estimates	Will be included in the stock assessment, replacing the current schedule last updated in 2006	Reproduction	Histological maturity assessment
	Incidence of skip spawning		Will be used to adjust the asymptote of the maturity schedule, if/when a time-series is available this will be used as a direct input to the stock assessment		Examination of potential skip spawning
	Fecundity-at-age and -size information		Will be used to move from spawning biomass to egg-output as the metric of reproductive capability in the stock assessment and management reference points		Fecundity assessment
	Revised field maturity classification		Revised time-series of historical (and future) maturity for input to the stock assessment		Examination of accuracy of current field macroscopic maturity classification
2. Biological input	Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area	Altered structure of future stock assessments	If 4B is found to be functionally isolated, a separate assessment may be constructed for that IPHC Regulatory Area	Genetics and Genomics	Population structure
3. Biological input	Assignment of individuals to source populations and assessment of distribution changes	Improve estimates of productivity	Will be used to define management targets for minimum spawning biomass by Biological Region		Distribution
	Improved understanding of larval and juvenile distribution		Will be used to generate potential recruitment covariates and to inform minimum spawning biomass targets by Biological Region	Migration	Larval and juvenile connectivity studies
1. Assessment data collection and processing	Sex ratio-at-age	Scale biomass and fishing intensity	Annual sex-ratio at age for the commercial fishery fit by the stock assessment	Reproduction	Sex ratio of current commercial landings
	Historical sex ratio-at-age		Annual sex-ratio at age for the commercial fishery fit by the stock assessment		Historical sex ratios based on archived otolith DNA analyses
2. Assessment data collection and processing	New tools for fishery avoidance/deterence; improved estimation of depredation mortality	Improve mortality accounting	May reduce depredation mortality, thereby increasing available yield for directed fisheries. May also be included as another explicit source of mortality in the stock assessment and mortality limit setting process depending on the estimated magnitude	Mortality and survival assessment	Whale depredation accounting and tools for avoidance
1. Fishery yield	Physiological and behavioral responses to fishing gear	Reduce incidental mortality	May increase yield available to directed fisheries	Mortality and survival assessment	Biological interactions with fishing gear
2. Fishery yield	Guidelines for reducing discard mortality	Improve estimates of unobserved mortality	May reduce discard mortality, thereby increasing available yield for directed fisheries	Mortality and survival assessment	Best handling practices: recreational fishery

APPENDIX III

List of ranked biological uncertainties and parameters for management strategy evaluation (MSE) and their links to biological research areas and research activities

<table>
<tr><th>MSE Rank</th><th>Research outcomes</th><th>Relevance for MSE</th><th>Research Area</th><th>Research activities</th></tr>
<tr><td rowspan="2">1. Biological parameterization and validation of movement estimates</td><td>Improved understanding of larval and juvenile distribution</td><td rowspan="2">Improve parametization of the Operating Model</td><td>Migration</td><td>Larval and juvenile connectivity studies</td></tr>
<tr><td>Stock structure of IPHC Regulatory Area 4B relative to the rest of the Convention Area</td><td rowspan="2">Genetics and Genomics</td><td>Population structure</td></tr>
<tr><td rowspan="2">2. Biological parameterization and validation of recruitment variability and distribution</td><td>Assignment of individuals to source populations and assessment of distribution changes</td><td>Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model</td><td>Distribution</td></tr>
<tr><td>Establishment of temporal and spatial maturity and spawning patterns</td><td>Improve simulation of recruitment variability and parametization of recruitment distribution in the Operating Model</td><td>Reproduction</td><td>Recruitment strength and variability</td></tr>
<tr><td rowspan="3">3. Biological parameterization and validation for growth projections</td><td>Identification and application of markers for growth pattern evaluation</td><td rowspan="3">Improve simulation of variability and allow for scenarios investigating climate change</td><td rowspan="3">Growth</td><td rowspan="3">Evaluation of somatic growth variation as a driver for changes in size-at-age</td></tr>
<tr><td>Environmental influences on growth patterns</td></tr>
<tr><td>Dietary influences on growth patterns and physiological condition</td></tr>
<tr><td>1. Fishery parameterization</td><td>Experimentally-derived DMRs</td><td>Improve estimates of stock productivity</td><td>Mortality and survival assessment</td><td>Discard mortality rate estimate: recreational fishery</td></tr>
</table>

APPENDIX IV
Summary of current external research grants

Project #	Grant agency	Project name	PI	Partners	IPHC Budget ($US)	Management implications	Grant period
1	**Bycatch Reduction Engineering Program - NOAA**	Full scale testing of devices to minimize whale depredation in longline fisheries (NA23NMF4720414)	IPHC	NOAA Fisheries - Alaska Fisheries Science Center (Seattle)	$199,870	Mortality estimations due to whale depredation	November 2023 – April 2026
2	**Alaska Sea Grant (pending award)**	Development of a non-lethal genetic-based method for aging Pacific halibut (R/2024-05)	IPHC, Alaska Pacific Univ. (APU)	Alaska Fisheries Science Center-NOAA (Juneau)	$60,374	Stock structure	December 2024-December 2026
Total awarded ($)					**$260,244**		

IPHC-2025-SRB026-07

Development of the 2025 Pacific halibut (*Hippoglossus stenolepis*) stock assessment

PREPARED BY: IPHC SECRETARIAT (I. STEWART AND A. HICKS; 9 MAY 2025)

Table of Contents

Summary

This document reports preliminary analyses in development of the 2025 Pacific halibut (*Hippoglossus stenolepis*) stock assessment. It follows the previous full stock assessments conducted in 2022 (Stewart and Hicks 2022) and 2019, including the independent peer review in 2019 (Stewart and Hicks 2019b; Stewart and Hicks 2020; Stokes 2019). Since the 2022 full stock assessment, two updates have been completed in 2023 (Stewart and Hicks 2024) and 2024 (Stewart and Hicks 2025) which included little change to the data or methods. Following the review of this document in June 2025 (SRB026), requested revisions will be considered and presented for additional review in September 2025 (SRB027), and the final 2025 assessment will be produced for the IPHC's Interim (IM101) and Annual (AM102) meetings. Updated data sources, including the results of the 2025 Fishery-Independent Setline Survey (FISS), logbook and biological data from the 2025 commercial fishery, and sex-ratio information from the 2024 commercial landings-at-age will be included for the final 2025 analysis.

Creating robust, stable, and well-performing stock assessment models for the Pacific halibut stock has historically proven to be challenging due to the highly dynamic nature of the biology, distribution, and fisheries (Stewart and Martell 2014). The stock assessment for Pacific halibut has evolved through many different modeling approaches over the last 30 years (Clark 2003). These changes have reflected improvements in fisheries analysis methods, changes in model assumptions, and responses to recurrent retrospective biases and other lack-of-fit metrics (Stewart and Martell 2014). The use of multiple models provides a solution to the endless search for a better stock assessment model and allows for structural as well as estimation uncertainty to be better captured. The IPHC adopted the ensemble approach for its 2012 stock assessment (Stewart et al. 2013a) and has continued to develop and refine the set of models used to provide tactical management information each year. The ensemble approach integrates the results of multiple hypotheses with the uncertainty associated with parameter estimation (Stewart and Martell 2015). This reduces potential for abrupt changes in management quantities as improvements and additional data are added to individual models (Stewart and Hicks 2018), and provides a more realistic perception of uncertainty than any single model, and therefore a stronger basis for probabilistic risk assessment.

Development of the current ensemble of stock assessment models began in 2012 with a single model using three alternative fixed values of natural mortality (Stewart et al. 2013a). In subsequent years, ensemble development included exploration of highly varied model approaches, including a Virtual Population Analysis (VPA) and a simple biomass production model (Cox et al. 2014) and a spatially explicit model including migration rates and recruitment distribution (Cox et al. 2017). The treatment of the historical data through long and short modelled time-series', and the treatment of spatial patterns via coastwide aggregation of data and an Areas-As-Fleets (AAF) approach have emerged as two critically important axes over which to describe the uncertainty in both the scale and trends of the Pacific halibut stock and population dynamics. Therefore, recent ensembles have included four equally weighted models representing a two-way cross of time-series length (short and long) and data aggregation (coastwide and by Biological Region).

Starting with the final 2024 stock assessment data, models and results (Stewart and Webster 2025; Stewart and Hicks 2025), this analysis provides a sequentially updated ‘bridge’ of the changes made thus far toward a preliminary assessment for 2025. This bridging analysis included a series of steps for which intermediate results and comparisons are provided. These steps included:

1) Extending the time series to include projected mortality based on limits adopted for 2025 (IPHC 2025),
2) updating to the newest stock synthesis software version (3.30.23.1; Methot Jr 2024),
3) updating the time-series information for the Pacific Decadal Oscillation, used as a covariate to the stock-recruitment relationship,
4) retuning the constraint on the scale of male time-varying fishery selectivity (the sex-ratio of the commercial fishery) and extending this variability into the forecast,
5) improving the bootstrapping approach to pre-model calculation of maximum effective sample sizes to include ageing imprecision (Hulson and Williams 2024),
6) re-tuning the process and observation error components of these models to achieve internal consistency within each,
7) and updating the maturity ogive to reflect the recent histology-based estimates produced by the IPHC’s Biological and Ecosystem Sciences Branch.

Briefly, extending the time-series, updating the software version, and updating the treatment of the fishery sex-ratio all had no of very little effect on the model results. Moving to the new PDO covariate increased the estimated spawning biomass in the coastwide long model and little effect on the AAF long model. Adding the new bootstrapping results and retuning the sample sizes and process error variance terms for internal model consistency generally increased the estimated spawning biomass, except for the coastwide long model near the end of the time-series where it was nearly unchanged. Finally, updating the maturity ogive resulted in a larger spawning biomass across all four models and especially in the historical period. Convergence, sensitivity and retrospective analyses were performed on all models contributing to the ensemble. The coastwide long model was most sensitive to the fixed value of steepness, with lower values corresponding to higher spawning biomass, while the other three models showed little difference at higher or lower values. The spawning biomass estimated by the coastwide short model scaled nearly linearly with the fixed value of *M*; higher *M* corresponding to larger estimates of spawning biomass. Excluding the PDO relationship resulted in a larger estimated spawning biomass in the coastwide long model across the entire time-series but had little effect on the recent years in the AAF long model. Evaluation of potential increased whale depredation in recent years was unable to explain the reduced recruitment observed since 2005. Retrospective analyses showed generally downward trends as data were added for all but the coastwide short model. Jitter analyses indicate that the AAF models were the least robust to a wide range of initial parameter estimates; however, there was no evidence that convergence was not achieved for the results provided here.

In aggregate, the results of the preliminary ensemble indicate that the uncertainty in stock dynamics remains similar to previous assessments and high relative to that frequently reported for many single-model or simple stock assessment analyses. This uncertainty will continue to be captured via the annual decision table (Stewart and Hicks 2025), reporting the trade-offs between yield and various stock and fishery risks. Given the challenges and uncertainties of the Pacific halibut population dynamics and stock assessment it is unlikely that future assessment models will provide substantially more precise and stable results, even as data time-series grow longer. In light of the uncertainty and variability within which the Pacific halibut management occurs, a robust management procedure, tested via the IPHC's Management Strategy Evaluation (MSE) process (Hicks and Stewart 2025) may provide a stronger basis for future management success and stability than annual decisions based on stock assessment results.

Data sources

The Pacific halibut data sources are collected with sampling designs created to produce results first for each IPHC Regulatory Area, and then to be aggregated to Biological Regions and to the entire range of the species in U.S. and Canadian waters (Figure 1). This section provides a brief overview of the key types of data available for analysis. A more in-depth summary can be found in the annual overview of data sources created each year and most recently for the 2024 stock assessment (Stewart and Webster 2025).

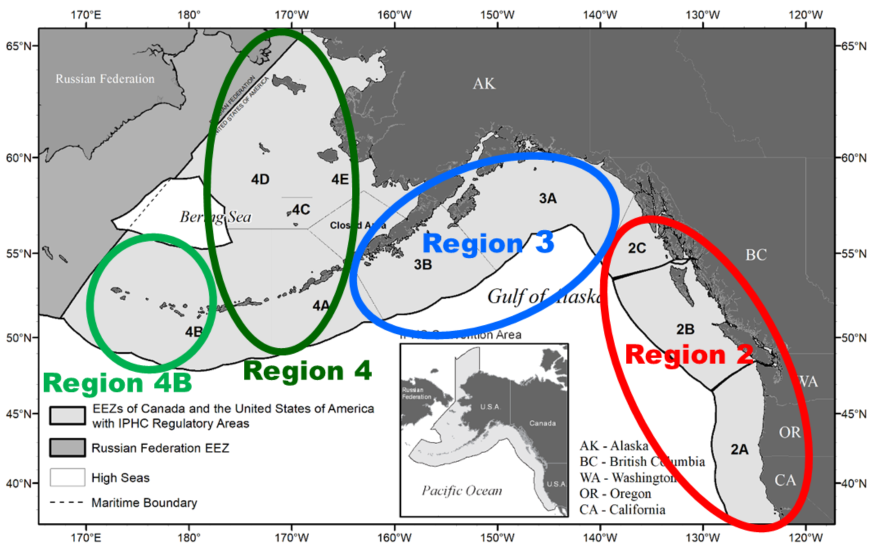


Figure 1. IPHC Regulatory Areas, Biological Regions, and the Pacific halibut geographical range within the territorial waters of Canada and the United States of America.

Overview of existing data

The time-series' of Pacific halibut data (described and plotted in much more detail in Stewart and Webster 2025) provide a rich historical record including mortality estimates, abundance indices (Catch-Per-Unit-Effort; CPUE) and age-composition data that extend back to the late 1800s and early 1900s (Figure 2). The IPHC's Fishery Independent Setline Survey (Ualesi et al. 2025; Webster 2025) provides the primary index of abundance and the most rich source of demographic information via individual weight, length and age data. The FISS includes Pacific halibut as young as 4-5 years old, which are below the IPHC's 32 inch (82 cm) minimum size limit (Stewart et al. 2021). Thus, these fish are observed several years prior to entry into the retained fishery landings which are sampled at the point of landing (Kong et al. 2022) and do not contain biological or catch-rate information on younger fish. Annual mortality estimates are provided to the IPHC from a variety of sources (Hutniczak et al. 2025) including the directed halibut fisheries (commercial, recreational and subsistence) as well as incidental mortality associated with discards in directed fisheries and discard mortality in non-directed fisheries ('bycatch') that are not allowed to legally retain Pacific halibut. Each of these sources have differing levels of precision and likely accuracy associated with the estimates used for stock assessment.

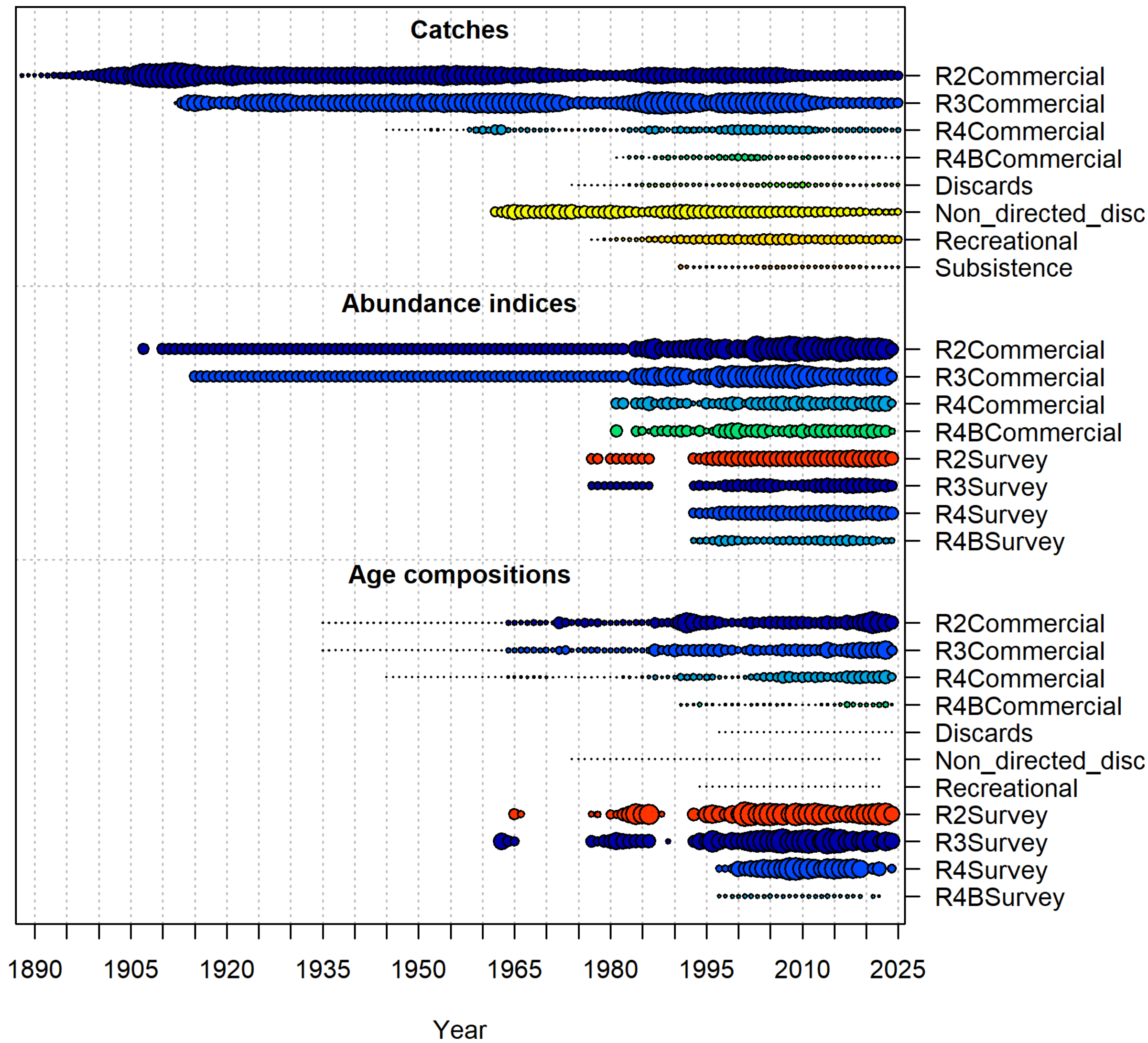


Figure 2. Data used in the stock assessment. Circle size is proportional to the magnitude of mortality (catches), inversely proportional to the variance (abundance indices) or proportional to the input sample size based on bootstrapping and prior to tuning for internal consistency (age-composition data).

Mortality

The industrial Pacific halibut fishery developed first off the west coast of the United States and Canada and sequentially moved to the north (Stewart and Webster 2025), only reaching full exploitation across all spatial areas in the last several decades. Mortality from non-directed discards increased rapidly with the arrival of foreign fleets into U.S. and Canadian waters in the 1960s. Recreational mortality has also increased over the time-series, although somewhat more gradually, since its initiation in the 1970s (Figure 3).

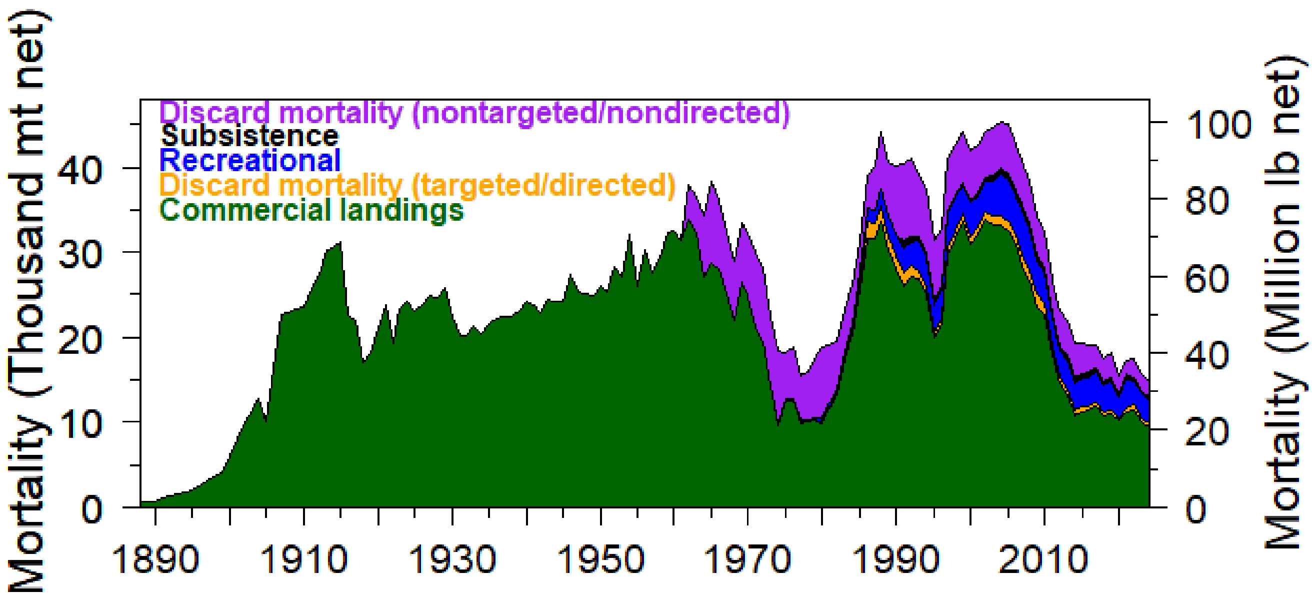


Figure 3. Time-series of mortality estimates by source.

Index data

The IPHC's FISS comprises the primary index of recent abundance and source of biological data for use in the stock assessment. Index values (Table 1) are used in this assessment in numbers of halibut captured per unit effort (NPUE). The recent time-series (1993-2024) is based on the output of the IPHC's space-time model (Webster 2025; Webster et al. 2020) which estimates the degree of spatial and temporal correlation among survey stations in order to predict trends in biomass and abundance across the entire range of Pacific halibut within the IPHC Convention Area. This index provides precise trend information by IPHC Regulatory Area, even when annual sampling is reduced, which are weighted by the relative spatial bottom area and combined to Biological Regions and a coastwide index. The variances are summed, accounting for the square of the weights, and converted to log(SE) for use in the assessment model assuming log-normal error. There were geographically limited surveys conducting during 1963-1989, with summarized catch rates, but no variance estimates available from 1977 (Table 1). For the period prior to 1993 where there are no variance estimates, twice the recent average value is used, and for the coastwide series where spatial coverage is incomplete values are doubled again.

Commercial fishery CPUE (generally referred to as Weight-Per-Unit-Effort or WPUE as landings are recorded in weight) is reported through mandatory logbooks (voluntary only for vessels under 26 feet, 7.9 m, in length), collected by IPHC port samplers, or returned directly to the IPHC by mail. Commercial CPUE is available as far back as the early 1900s (Stewart and Webster 2025) providing a valuable historical record, but spanning a period of continuous fishery development and change, including an important transition to circle hooks in 1984 that substantially increased average catchability (Table 2-4).

Table 1. Modelled survey Numbers-Per-Unit-Effort (NPUE) and log(SE) 1993-2024, raw average observed NPUE 1977-1986; assumed values in italics.

	Region 2		Region 3		Region 4		Region 4B		Coastwide	
Year	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)
1977	0.60	*0.107*	2.00	0.176	--	--	--	--	1.47	*0.227*
1978	0.80	*0.107*	1.30	0.176	--	--	--	--	1.11	*0.227*
1979	--	--	1.90	0.176	--	--	--	--	--	--
1980	1.20	*0.107*	2.50	0.176	--	--	--	--	2.01	*0.227*
1981	0.80	*0.107*	3.80	0.176	--	--	--	--	2.67	*0.227*
1982	1.84	*0.107*	3.80	0.176	--	--	--	--	2.87	*0.227*
1983	2.30	*0.107*	3.40	0.176	--	--	--	--	2.88	*0.227*
1984	6.74	*0.107*	11.60	0.176	--	--	--	--	9.30	*0.227*
1985	5.65	*0.107*	11.90	0.176	--	--	--	--	8.94	*0.227*
1986	4.54	*0.107*	7.80	0.176	--	--	--	--	6.26	*0.227*
1993	6.20	0.101	23.79	0.141	1.82	0.129	9.87	0.291	7.26	0.097
1994	7.41	0.104	23.42	0.120	2.10	0.111	10.12	0.263	7.56	0.080
1995	8.76	0.072	25.21	0.135	2.10	0.108	10.39	0.231	8.14	0.087
1996	7.76	0.059	26.40	0.173	2.32	0.094	10.59	0.182	8.36	0.111
1997	7.18	0.055	28.51	0.167	2.55	0.062	10.76	0.110	8.83	0.108
1998	6.20	0.054	24.60	0.084	2.65	0.063	11.10	0.111	7.98	0.054
1999	5.03	0.052	23.50	0.087	2.32	0.066	9.51	0.127	7.30	0.058
2000	5.58	0.055	25.26	0.080	2.47	0.062	8.43	0.151	7.77	0.053
2001	6.47	0.050	22.32	0.096	2.34	0.061	6.38	0.175	7.15	0.061
2002	6.40	0.051	24.42	0.069	2.22	0.059	4.66	0.205	7.40	0.047
2003	5.51	0.054	24.13	0.068	2.13	0.062	4.00	0.230	7.12	0.048
2004	5.02	0.052	27.55	0.067	2.13	0.059	3.73	0.214	7.70	0.049
2005	5.52	0.052	22.87	0.059	2.19	0.056	3.68	0.202	6.89	0.041
2006	5.42	0.052	21.79	0.087	2.25	0.049	4.21	0.197	6.72	0.057
2007	6.09	0.051	23.67	0.116	2.21	0.055	5.28	0.188	7.23	0.076
2008	6.14	0.050	21.34	0.127	2.49	0.061	5.30	0.167	6.94	0.079
2009	6.31	0.052	20.05	0.126	2.46	0.058	4.54	0.180	6.66	0.077
2010	6.11	0.050	20.17	0.087	2.32	0.052	4.28	0.179	6.55	0.055
2011	6.09	0.048	20.42	0.095	2.19	0.053	4.24	0.162	6.52	0.060
2012	7.26	0.047	21.10	0.060	2.14	0.050	3.82	0.154	6.79	0.039
2013	7.13	0.047	16.03	0.057	1.93	0.047	5.13	0.128	5.70	0.035
2014	7.30	0.046	19.06	0.056	1.97	0.044	4.47	0.135	6.32	0.036
2015	8.04	0.048	19.11	0.057	1.99	0.045	4.50	0.137	6.46	0.036
2016	8.09	0.046	19.44	0.064	1.88	0.049	5.14	0.122	6.51	0.040
2017	5.90	0.044	13.91	0.050	1.74	0.052	4.04	0.092	4.92	0.032
2018	5.23	0.041	12.69	0.050	1.66	0.052	4.05	0.128	4.53	0.031
2019	5.36	0.042	11.36	0.055	1.62	0.055	4.02	0.146	4.26	0.033
2020	5.05	0.042	11.76	0.054	1.57	0.070	3.95	0.176	4.26	0.036
2021	5.73	0.044	15.40	0.062	1.50	0.053	3.81	0.139	5.04	0.040
2022	5.81	0.045	12.74	0.080	1.45	0.049	3.58	0.204	4.48	0.048
2023	6.00	0.045	12.04	0.074	1.42	0.059	3.74	0.253	4.37	0.044
2024	6.64	0.066	12.20	0.111	1.38	0.074	3.90	0.290	4.49	0.066

Table 2. Commercial fishery Weight-Per-Unit-Effort (WPUE) 1907-1949 and log(SE); assumed values in italics.

	Region 2		Region 3		Region 4		Region 4B		Coastwide	
Year	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)
1907	280.00	*0.100*	--	--	--	--	--	--	280.00	*0.100*
1910	271.00	*0.100*	--	--	--	--	--	--	271.00	*0.100*
1911	237.00	*0.100*	--	--	--	--	--	--	237.00	*0.100*
1912	176.00	*0.100*	--	--	--	--	--	--	176.00	*0.100*
1913	128.94	*0.100*	--	--	--	--	--	--	129.00	*0.100*
1914	124.13	*0.100*	--	--	--	--	--	--	124.00	*0.100*
1915	118.02	*0.100*	266.10	*0.100*	--	--	--	--	118.00	*0.100*
1916	114.60	*0.100*	202.80	*0.100*	--	--	--	--	137.00	*0.100*
1917	81.80	*0.100*	157.90	*0.100*	--	--	--	--	98.00	*0.100*
1918	87.50	*0.100*	125.40	*0.100*	--	--	--	--	96.00	*0.100*
1919	82.30	*0.100*	129.90	*0.100*	--	--	--	--	93.00	*0.100*
1920	84.10	*0.100*	147.90	*0.100*	--	--	--	--	96.00	*0.100*
1921	76.46	*0.100*	141.17	*0.100*	--	--	--	--	88.00	*0.100*
1922	62.44	*0.100*	133.79	*0.100*	--	--	--	--	73.00	*0.100*
1923	56.68	*0.100*	149.97	*0.100*	--	--	--	--	78.00	*0.100*
1924	55.39	*0.100*	109.13	*0.100*	--	--	--	--	74.00	*0.100*
1925	51.21	*0.100*	94.63	*0.100*	--	--	--	--	68.00	*0.100*
1926	51.67	*0.100*	93.73	*0.100*	--	--	--	--	67.00	*0.100*
1927	48.83	*0.100*	86.32	*0.100*	--	--	--	--	65.00	*0.100*
1928	47.27	*0.100*	72.34	*0.100*	--	--	--	--	58.00	*0.100*
1929	38.55	*0.100*	70.79	*0.100*	--	--	--	--	51.00	*0.100*
1930	34.44	*0.100*	65.91	*0.100*	--	--	--	--	46.00	*0.100*
1931	38.48	*0.100*	76.17	*0.100*	--	--	--	--	50.00	*0.100*
1932	47.50	*0.100*	83.49	*0.100*	--	--	--	--	60.00	*0.100*
1933	50.16	*0.100*	83.99	*0.100*	--	--	--	--	63.00	*0.100*
1934	54.07	*0.100*	74.97	*0.100*	--	--	--	--	62.00	*0.100*
1935	61.77	*0.100*	97.57	*0.100*	--	--	--	--	76.00	*0.100*
1936	54.66	*0.100*	96.70	*0.100*	--	--	--	--	71.00	*0.100*
1937	61.48	*0.100*	109.99	*0.100*	--	--	--	--	80.00	*0.100*
1938	70.33	*0.100*	114.29	*0.100*	--	--	--	--	88.00	*0.100*
1939	61.90	*0.100*	112.21	*0.100*	--	--	--	--	80.00	*0.100*
1940	61.71	*0.100*	116.38	*0.100*	--	--	--	--	81.00	*0.100*
1941	62.54	*0.100*	122.26	*0.100*	--	--	--	--	85.00	*0.100*
1942	65.43	*0.100*	132.54	*0.100*	--	--	--	--	90.00	*0.100*
1943	72.24	*0.100*	131.27	*0.100*	--	--	--	--	95.00	*0.100*
1944	86.84	*0.100*	149.23	*0.100*	--	--	--	--	110.00	*0.100*
1945	79.69	*0.100*	130.86	*0.100*	--	--	--	--	102.00	*0.100*
1946	83.78	*0.100*	123.82	*0.100*	--	--	--	--	101.00	*0.100*
1947	86.30	*0.100*	114.56	*0.100*	--	--	--	--	99.00	*0.100*
1948	88.61	*0.100*	112.20	*0.100*	--	--	--	--	99.00	*0.100*
1949	85.01	*0.100*	105.89	*0.100*	--	--	--	--	95.00	*0.100*

Table 3. Commercial fishery Weight-Per-Unit-Effort (WPUE) 1950-1991 and log(SE); assumed values in italics.

	Region 2		Region 3		Region 4		Region 4B		Coastwide	
Year	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)
1950	87.66	*0.100*	103.60	*0.100*	--	--	--	--	95.00	*0.100*
1951	87.63	*0.100*	108.93	*0.100*	--	--	--	--	96.00	*0.100*
1952	95.58	*0.100*	128.86	*0.100*	--	--	--	--	110.00	*0.100*
1953	128.65	*0.100*	134.32	*0.100*	--	--	--	--	131.00	*0.100*
1954	137.97	*0.100*	127.43	*0.100*	--	--	--	--	133.00	*0.100*
1955	122.20	*0.100*	116.32	*0.100*	--	--	--	--	119.00	*0.100*
1956	132.02	*0.100*	126.05	*0.100*	--	--	--	--	129.00	*0.100*
1957	100.95	*0.100*	119.84	*0.100*	--	--	--	--	110.00	*0.100*
1958	101.96	*0.100*	139.96	*0.100*	--	--	--	--	121.00	*0.100*
1959	98.67	*0.100*	160.62	*0.100*	--	--	--	--	129.00	*0.100*
1960	105.02	*0.100*	156.08	*0.100*	--	--	--	--	132.00	*0.100*
1961	96.00	*0.100*	159.79	*0.100*	--	--	--	--	127.00	*0.100*
1962	84.76	*0.100*	136.89	*0.100*	--	--	--	--	115.00	*0.100*
1963	77.73	*0.100*	123.89	*0.100*	--	--	--	--	105.00	*0.100*
1964	75.27	*0.100*	120.10	*0.100*	--	--	--	--	100.00	*0.100*
1965	86.47	*0.100*	107.07	*0.100*	--	--	--	--	99.00	*0.100*
1966	82.59	*0.100*	112.72	*0.100*	--	--	--	--	100.00	*0.100*
1967	81.44	*0.100*	113.00	*0.100*	--	--	--	--	101.00	*0.100*
1968	86.58	*0.100*	111.62	*0.100*	--	--	--	--	103.00	*0.100*
1969	81.53	*0.100*	105.07	*0.100*	--	--	--	--	95.00	*0.100*
1970	73.62	*0.100*	103.67	*0.100*	--	--	--	--	91.00	*0.100*
1971	76.05	*0.100*	96.31	*0.100*	--	--	--	--	89.00	*0.100*
1972	69.47	*0.100*	82.87	*0.100*	--	--	--	--	78.00	*0.100*
1973	64.41	*0.100*	62.13	*0.100*	--	--	--	--	63.00	*0.100*
1974	60.89	*0.100*	61.95	*0.100*	--	--	--	--	61.00	*0.100*
1975	61.87	*0.100*	66.76	*0.100*	--	--	--	--	61.00	*0.100*
1976	44.39	*0.100*	61.91	*0.100*	--	--	--	--	55.00	*0.100*
1977	64.17	*0.100*	65.57	*0.100*	--	--	--	--	63.00	*0.100*
1978	54.06	*0.100*	68.47	*0.100*	--	--	--	--	71.00	*0.100*
1979	55.80	*0.100*	67.33	*0.100*	--	--	--	--	75.00	*0.100*
1980	59.54	*0.100*	116.09	*0.100*	--	--	--	--	94.00	*0.100*
1981	73.84	*0.100*	148.86	*0.100*	136.84	*0.100*	99.00	0.078	111.00	*0.100*
1982	71.85	*0.100*	181.34	*0.100*	98.68	*0.100*	--	--	127.00	*0.100*
1984	151.95	0.045	491.33	0.046	386.90	*0.100*	161.00	0.103	316.00	0.035
1985	161.59	0.051	535.06	0.039	456.18	0.099	234.00	0.160	352.00	0.034
1986	137.26	0.035	506.00	0.042	308.70	0.062	238.00	0.372	315.00	0.041
1987	135.53	0.027	490.38	0.036	360.93	0.159	220.00	0.111	316.00	0.038
1988	168.40	0.054	560.55	0.042	405.68	0.105	224.00	0.122	363.00	0.036
1989	154.92	0.042	507.69	0.031	387.41	0.078	268.00	0.094	353.00	0.025
1990	194.64	0.043	403.54	0.036	370.26	0.095	209.00	0.103	315.00	0.029
1991	170.62	0.039	375.02	0.041	367.06	0.157	329.00	0.085	314.00	0.038

Table 4. Commercial fishery Weight-Per-Unit-Effort (WPUE) 1992-2024 and log(SE); assumed values in italics.

	Region 2		Region 3		Region 4		Region 4B		Coastwide	
Year	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)	Index	log(SE)
1992	167.66	0.040	413.39	0.048	324.01	0.117	280.00	0.095	315.00	0.035
1993	200.04	0.031	439.11	0.096	399.87	0.448	218.00	0.220	369.00	0.100
1994	175.74	0.027	362.77	0.049	343.14	0.333	197.00	0.101	302.00	0.069
1995	190.73	0.025	439.48	0.043	330.22	0.100	189.00	0.336	326.00	0.037
1996	208.81	0.042	505.01	0.046	427.58	0.138	269.00	0.185	387.00	0.039
1997	237.52	0.035	498.02	0.026	417.44	0.107	275.00	0.064	400.00	0.025
1998	221.23	0.029	512.59	0.036	411.86	0.089	287.00	0.058	402.00	0.025
1999	249.48	0.079	475.49	0.024	385.64	0.061	310.00	0.045	390.00	0.023
2000	227.94	0.036	492.21	0.025	403.74	0.082	318.00	0.046	396.00	0.020
2001	202.84	0.039	454.52	0.029	363.00	0.213	270.00	0.076	358.00	0.042
2002	214.81	0.032	466.46	0.025	296.56	0.082	245.00	0.081	356.00	0.020
2003	208.95	0.018	439.27	0.024	251.12	0.072	196.00	0.068	325.00	0.018
2004	192.88	0.028	425.79	0.026	235.23	0.072	202.00	0.061	315.00	0.019
2005	178.98	0.024	387.69	0.023	219.59	0.063	238.00	0.093	293.00	0.017
2006	180.22	0.024	360.70	0.022	178.26	0.064	218.00	0.111	268.00	0.019
2007	155.80	0.022	338.41	0.023	154.12	0.054	231.00	0.109	246.00	0.016
2008	135.02	0.018	314.08	0.022	162.55	0.071	193.00	0.069	227.00	0.018
2009	152.95	0.020	277.22	0.020	174.43	0.054	189.00	0.100	220.00	0.018
2010	185.68	0.034	242.32	0.024	143.97	0.079	143.00	0.062	203.00	0.020
2011	180.42	0.019	226.65	0.025	143.25	0.056	165.00	0.103	196.00	0.015
2012	193.96	0.020	214.08	0.032	137.37	0.074	149.00	0.067	193.00	0.021
2013	192.78	0.026	189.98	0.033	122.70	0.072	127.00	0.064	178.00	0.017
2014	210.44	0.026	182.93	0.039	116.04	0.092	146.00	0.070	183.00	0.022
2015	217.37	0.024	224.46	0.045	136.04	0.065	149.00	0.076	202.00	0.025
2016	212.66	0.019	216.22	0.044	128.30	0.066	123.00	0.083	196.00	0.020
2017	212.49	0.020	218.98	0.037	129.11	0.077	119.00	0.076	202.00	0.020
2018	195.67	0.027	189.88	0.055	115.12	0.058	134.00	0.071	177.00	0.028
2019	184.34	0.027	213.12	0.037	101.45	0.100	115.00	0.085	179.00	0.022
2020	175.50	0.024	215.84	0.040	100.92	0.081	105.00	0.059	178.00	0.022
2021	178.43	0.025	194.57	0.041	127.08	0.049	88.00	0.057	168.00	0.024
2022	154.56	0.024	133.57	0.038	122.22	0.049	80.00	0.085	134.00	0.022
2023	138.94	0.029	114.76	0.036	85.35	0.052	84.00	0.090	114.00	0.018
2024	135.08	*0.069*	105.23	*0.100*	80.98	*0.157*	115.00	*0.278*	112.00	*0.062*

Age data

At each FISS station, otoliths are sampled randomly at rates selected to generate approximately 1500 per IPHC Regulatory Area per year. The number of stations contributing to the annual age information varies considerably over the time-series, with Biological Region 3 the most heavily sampled, followed by Region 2, Region 4 and far fewer samples collected in Region 4B (Table 5). There are also a small number of geographically limited surveys from the period 1963-1966 for which there are age samples, but no corresponding index. Otoliths from the commercial fishery landings are also sampled in proportion to the weight of the catch with different rates by IPHC Regulatory Area (Hutniczak et al. 2025). This has led to a relatively larger number of commercial trips sampled in Biological Region 2 over most of the historical period, with Region 3, Region 4, and Region 4B each contributing fewer samples (Table 6-7).

Table 5. Number of stations contributing to FISS age data (1963-2024).

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
1963	--	236	--	--	236
1964	--	305	--	--	305
1965	121	146	--	--	267
1966	66	--	--	--	66
1977	58	100	--	--	158
1978	62	98	--	--	160
1979	--	104	--	--	104
1980	80	101	--	--	181
1981	72	102	--	--	174
1982	154	148	--	--	302
1983	192	101	--	--	293
1984	241	198	--	--	439
1985	166	103	--	--	269
1986	178	97	--	--	275
1988	72	--	--	--	72
1989	--	33	--	--	33
1993	66	70	--	--	136
1994	14	147	--	--	161
1995	103	120	--	--	223
1996	198	424	--	--	622
1997	211	424	220	74	929
1998	228	507	100	42	877
1999	332	554	61	82	1,029
2000	239	548	149	83	1,019
2001	330	520	146	83	1,079
2002	313	555	154	82	1,104
2003	323	516	153	82	1,074
2004	327	523	145	70	1,065
2005	340	507	144	81	1,072
2006	317	526	240	84	1,167
2007	330	538	176	73	1,117
2008	338	549	166	76	1,129
2009	333	537	171	84	1,125
2010	333	521	172	76	1,102
2011	358	549	166	79	1,152
2012	354	522	168	71	1,115
2013	364	528	167	78	1,137
2014	381	556	227	76	1,240
2015	352	529	239	81	1,201
2016	350	538	220	72	1,180
2017	371	521	166	118	1,176
2018	466	537	167	77	1,247
2019	482	560	167	81	1,290
2020	370	494	--	--	864
2021	393	550	77	37	1,057
2022	321	266	117	27	731
2023	378	411	--	--	789
2024	175	157	55	--	387

Table 6. Number of commercial fishing trips contributing to fishery age data (1935-1982); historical values in italics are assumed.

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
1935	*50*	*50*	--	--	*100*
1936	*50*	*50*	--	--	*100*
1937	*50*	*50*	--	--	*100*
1938	*50*	*50*	--	--	*100*
1939	*50*	*50*	--	--	*100*
1940	*50*	*50*	--	--	*100*
1941	*50*	*50*	--	--	*100*
1942	*50*	*50*	--	--	*100*
1943	*50*	*50*	--	--	*100*
1944	*50*	*50*	--	--	*100*
1945	*50*	*50*	*5*	--	*100*
1946	*50*	*50*	*5*	--	*100*
1947	*50*	*50*	*5*	--	*100*
1948	*50*	*50*	*5*	--	*100*
1949	*50*	*50*	*5*	--	*100*
1950	*50*	*50*	*5*	--	*100*
1951	*50*	*50*	*5*	--	*100*
1952	*50*	*50*	*5*	--	*100*
1953	*50*	*50*	*5*	--	*100*
1954	*50*	*50*	*5*	--	*100*
1955	*50*	*50*	*5*	--	*100*
1956	*50*	*50*	*5*	--	*100*
1957	*50*	*50*	*5*	--	*100*
1958	*50*	*50*	*5*	--	*100*
1959	*50*	*50*	*5*	--	*100*
1960	*50*	*50*	*5*	--	*100*
1961	*50*	*50*	*5*	--	*100*
1962	*50*	*50*	*5*	--	*100*
1963	*50*	*50*	*5*	--	*100*
1964	116	100	14	--	230
1965	118	106	12	--	238
1966	102	113	12	--	228
1967	125	133	20	--	278
1968	135	132	14	--	282
1969	113	102	12	--	227
1970	97	125	18	--	241
1971	82	77	9	--	168
1972	552	196	3	--	752
1973	311	262	5	--	578
1974	153	68	3	--	226
1975	234	76	7	--	320
1976	332	135	7	--	476
1977	247	138	7	--	401
1978	241	120	4	--	377
1979	125	101	6	--	244
1980	140	113	1	--	262
1981	146	90	7	--	248
1982	168	137	11	--	316

Table 7. Number of commercial fishing trips contributing to fishery age data (1983-2024).

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
1983	133	106	23	6	268
1984	170	90	9	13	282
1985	171	99	14	2	286
1986	158	152	34	1	345
1987	531	498	76	12	1,117
1988	278	258	19	16	571
1989	318	371	39	24	752
1990	491	560	50	3	1,104
1991	718	496	62	12	1,288
1992	1,027	478	61	20	1,586
1993	959	471	65	11	1,506
1994	896	474	89	31	1,490
1995	887	468	72	37	1,464
1996	859	437	76	27	1,399
1997	676	429	183	58	1,346
1998	515	277	127	47	966
1999	454	303	118	24	899
2000	512	358	119	27	1,016
2001	505	233	117	13	868
2002	561	284	163	53	1,061
2003	545	266	118	49	978
2004	491	200	75	9	775
2005	461	193	125	13	792
2006	483	256	81	22	842
2007	429	218	95	12	754
2008	385	221	98	11	715
2009	432	240	68	14	754
2010	354	260	97	25	736
2011	383	224	83	14	704
2012	421	217	81	13	732
2013	455	196	73	14	738
2014	426	221	64	8	719
2015	476	192	119	15	802
2016	466	164	112	15	757
2017	410	175	106	17	708
2018	337	178	105	17	637
2019	409	199	116	10	734
2020	406	176	47	12	641
2021	379	160	43	11	593
2022	467	190	60	11	728
2023	495	220	64	9	788
2024	511	216	161	14	902

As has been the case since the 2015 stock assessment (Stewart and Martell 2016), all age data used in the stock assessment is aggregated into bins of ages from age-2 to age-25, with age 2 representing a ‘minus’ group including all fish of age 2 and younger, and age 25 representing a ‘plus’ group including all fish age 25 and older. For years prior to 2002 (except the survey ages from 1998 which were re-aged in 2013), surface ages were the standard ageing method, replaced by break-and-bake in recent years. Because surface ages are known to be biased at older ages (Forsberg and Stewart 2015), the age data are aggregated at a lower 'plus' group, age 20+, for all years where this was the primary method.

Beginning with the 2019 stock assessment, sex-specific fishery age data has been available via the collection of fin clips and subsequent genetic assay based on sampling begun in 2017. The processing of these samples lags one-year, thus for the preliminary 2025 stock assessment there were seven years of sex-specific fishery age compositions used (2017-2023). They are compiled in an identical manner to the standard fishery age data, but delineating males and females through the weighting and aggregation up to Biological Regions and coastwide. The sex-specific fishery age compositions for 2024 will be available for the final 2025 stock assessment later this year, along with re-aged FISS data from 2023 and 2024. The re-ageing of the 2023 and 2024 samples was prompted by increasingly difficult ring identification in recent years leading to some patterns that appeared inconsistent with previous year's data (tracking of year-classes).

Other biological and fishery information

There are several other sources of information contributing to the stock assessment models. These include:

1) the time-series of the Pacific Decadal Oscillation (PDO) index
2) the maturity ogive
3) priors on natural mortality (*M*)
4) fecundity information
5) estimated weight-at-age
6) the length-weight relationship
7) ageing error (bias and imprecision)
8) data based 'priors' on bycatch, discard, and recreational selectivity

The PDO index and the maturity schedule were the subjects of specific development and sensitivity analyses reported in this stock assessment. There have been no significant changes to the treatment of other sources of information since the 2015 stock assessment (Stewart and Martell 2016), except for the length-weight relationship which was updated as part of the 2022 stock assessment based on an analysis conducted in 2021 (Webster and Stewart 2022). Directly measured weights have been collected during the FISS (since 2019) and the commercial sampling (since 2015) and used directly in the stock assessment data preparation. Therefore, the length-weight relationship is primarily used for estimation by domestic agencies of mortality in weight from piece counts (this is relevant to non-directed discard mortality, recreational mortality and subsistence mortality).

All other sources of information are updated (where appropriate) and described each year in the annual overview of data sources (Stewart and Webster 2022). For convenience, the treatment of each is briefly summarized in Table 8.

Table 8. Summary of other information sources contributing directly to stock assessment input files (Stewart and Webster 2025).

Input	Summary	Key assumptions
Pacific Decadal Oscillation index[1]	Monthly values averaged and compiled into a binary index for each year based on assignment to 'positive' and 'negative' phases. **Updated as part of the bridging analysis for 2025.**	Used as a binary indicator rather than annually varying values.
Maturity	Trimmed Generalized Additive Model based on calibrated visual estimates from 2002-2024; 50% female maturity at 11.0 years old. **Updated as part of the bridging analysis for 2025.**	Based on histological assessments, treated as age-based and time-invariant.
Priors on natural mortality (*M*)	Prior on age 3+ *M* based on longevity (Hamel and Cope 2022) and elevated values for ages 0-2 based on life history theory and analyses of other flatfish.	Age 55 is a reasonable proxy for longevity based on multiple observations of male and female halibut age 50 and greater.
Fecundity	Assumed to be proportional to body weight.	Temporal variability is included via changes in weight-at-age.
Weight-at-age	Reconstructed from survey and fishery information by Biological Region.	Historical variability has been similar for female and male Pacific halibut.
Length-weight relationship	Not used directly in the assessment, most of the historical data relies on a constant average length-weight relationship.	Measured weights are used preferentially where available.
Ageing error	Pacific halibut are relatively easy to age accurately and with a high degree of precision using the break-and-bake method (Clark 2004a, 2004b; Clark and Hare 2006; Piner and Wischnioski 2004). Surface ages are biased and less precise (Stewart 2014).	Multi-decadal comparison suggest that accuracy and precision have not changed appreciably over the entire historical record (Forsberg and Stewart 2015).
Bycatch selectivity prior	Age-distributions are created from weighted and aggregated length frequencies from a variety of sources and age-length keys from trawl surveys.	Due to incomplete sampling, poor data quality in many years, and other uncertainties, data are considered unreliable for estimation of recruitment.
Discard selectivity prior	Age-distributions of sub-legal (<32 inch) Pacific Halibut captured by the FISS are used as a proxy for poorly sampled directed commercial fishery discards.	Survey data may not be representative of commercial fishing behavior but are currently the only source of information on the age range of discarded fish.
Recreational selectivity prior	Weighted age-frequency data from the IPHC Regulatory Area 3A recreational fishery are the only comprehensive source available.	These data may not be representative of all recreational mortality but provide the best information currently available.

[1] Data can be accessed at: https://www.ncei.noaa.gov/access/monitoring/pdo/

External information on M

The approach taken to natural mortality remains unchanged from that in the 2022 stock assessment (Stewart and Hicks 2022). That analysis was thoroughly documented and is briefly summarized here. It is based on two primary concepts stemming from the 2021 CAPAM workshop (Hamel et al. 2023):

1) Although results are varied, simulations experiments have generally indicated that estimation of *M* is preferable to the use of fixed values, where this is possible. The use of informative priors is frequently necessary, with the most common prior based on longevity.
2) Elevated *M* at the youngest ages/smallest sizes should be expected due to increased size-dependent predation mortality.

For the 2022 Pacific halibut assessment both of these concepts were included in the four stock assessment models. An age independent prior on M was developed based on published meta-analyses (Hamel 2014; Hamel and Cope 2022), which uses the prediction interval based on a meta-analysis of the maximum observed age for a wide range of species. Both male and female Pacific halibut have been observed to age-55 (with multiple fish of both sexes exceeding age-50 indicating that this is likely to be an accurate estimate of longevity, and not an artifact of a single case of ageing imprecision). The prior median is given by:

$$M = \frac{5.4}{Age_{max}}$$

which results in a value of 0.0982, and a log(SD) of 0.438. With such a large variance, this prior is only weakly informative (Figure 4), but still may provide some stability for estimation of *M*.

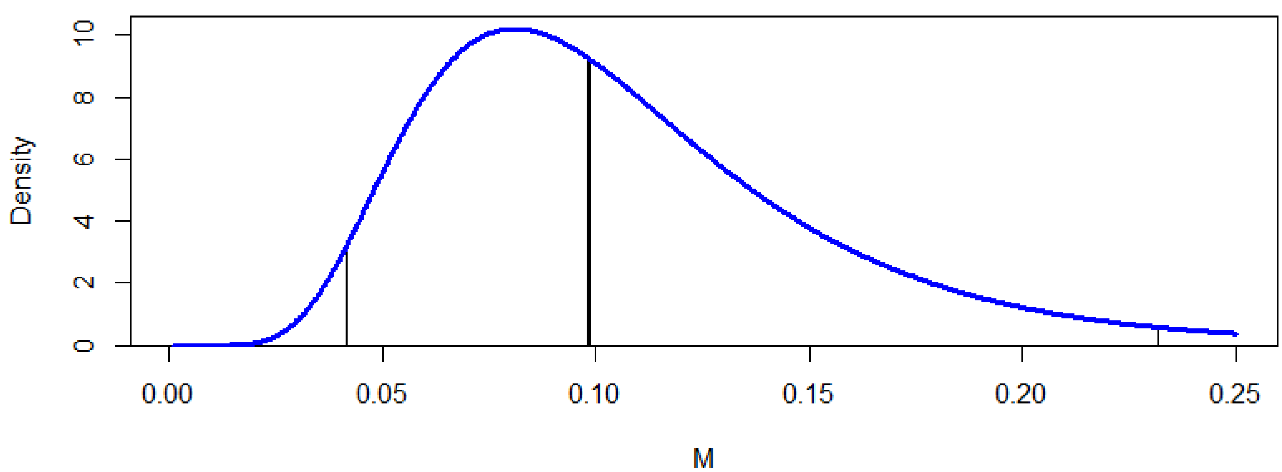


Figure 4. Informative prior for *M*. Thick vertical line denotes the median, thin lines the 2.5 and 97.5 percentiles of the distribution.

Pacific halibut were compared to other flatfish species via a summary of sex specific asymptotic size (either L_{inf} or L_{old}, depending on the parameterization) and *M* from all available Northeast

Pacific flatfish stock assessments (Stewart and Hicks 2022). Complete data were available for 26 stocks spread over four geographical regions comprising the U.S. West Coast, British Columbia, Gulf of Alaska, and the Bering Sea. A key result of this comparison was that flatfish with maximum sizes of >35 cm had natural mortality values both above and below those used for Pacific halibut. As Pacific halibut reach this size by age-3, this suggests that strongly elevated natural mortality due to predation common across flatfish species is likely to occur primarily below this age. Ecosystem models that include predator-prey dynamics generally suggest much higher *M* for the youngest age classes of NE Pacific groundfish (Adams et al. 2022). Where this information has been applied in other assessments used for management advice somewhat arbitrary scalars such as 1.5 x *M* for age 2, 2 x *M* for age-1 and *3* x *M* for age-0 are generally consistent with ecosystem models (e.g., Ianelli et al. 2021). Applying this general approach to Pacific halibut allows for size-dependent *M* that is consistent with theoretical concepts but does not appreciably change the natural mortality used for ages represented in observed fishery and survey data (exclusively age 2+). With little to no data at these youngest ages, any effect is likely to 'scale out' in the absolute estimates of recruitment deviations; however, when an index of recruitment is evaluated (i.e., the PDO in this assessment; see sensitivity analyses below) it may be important to include elevated *M* at these ages.

Bootstrapping input sample sizes for age compositions

Data weighting in the Pacific halibut stock assessment historically relied on the number of sampled FISS stations and number of sampled commercial fishery trips as a starting point for all models. Investigation of alternative tuning procedures and likelihoods was necessarily conditioned on those starting values, yet those starting values had not been evaluated specifically until the 2022 stock assessment. That analysis followed the method developed in Stewart and Hamel (2014) for estimating the maximum effective sample size based on the actual distribution and weighting of both the samples and the fish within samples. The effective sample size calculated in this manner is analogous to a minimum variance estimate – the actual effective sample size may be lower than calculated if not all strata are fully sampled (measurement error), of the source of the data differs from that assumed in the assessment (structural or process error). However, the effective sample size cannot be larger than the bootstrapped value simply due to the among and within sample variability and the sample sizes achieved. Thus, although time-consuming to produce, the approach provides an objective starting point for data weighting, and a logical upper bound on sample sizes used in the stock assessment models. Since the development of the 2022 stock assessment an improved method was developed that included ageing imprecision as part of the inherent variability in the observations (Hulson and Williams 2024). This new method not only resamples hauls (or trips) and fish within those hauls, but also individual ages from an empirical matrix of multiple age estimates from the same otoliths. Specifically, for a given observed age, a random age is drawn from all fish that had the original age assigned by one reader and a different age by a second read. In this way, the precision of each realization of the observed age composition information is reduced based on how likely that specific age would be given multiple reads. For Pacific halibut there are two ageing methods, break-and-bake (Table 9) and surface (Table 10), each with a differing degree of imprecision.

Table 9. Distribution of multiple reads of the same otoliths using the break-and-bake method.

	Second read																							
First read	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25+
2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	79	9	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	26	411	44	5	1	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
5	-	1	216	931	63	11	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	-	-	8	450	1,122	166	18	6	2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	1	17	1,004	1,849	332	48	10	2	1	-	2	-	-	-	-	-	-	-	-	-	-	-
8	-	-	-	4	38	735	2,572	451	86	20	10	1	1	-	1	-	-	-	-	-	-	-	-	1
9	-	-	-	-	7	74	680	2,753	496	115	25	8	3	1	3	-	-	1	-	-	-	-	-	-
10	-	-	-	-	2	7	112	749	2,961	513	123	38	9	5	2	1	-	1	1	-	-	-	-	-
11	-	-	-	-	-	2	11	119	683	2,678	595	104	33	11	3	2	2	1	1	-	-	-	-	1
12	-	-	-	-	2	-	6	26	134	709	2,527	540	132	63	13	4	2	3	2	-	-	-	-	-
13	-	-	-	1	-	1	1	7	34	154	676	1,872	417	132	22	9	4	2	1	-	-	-	-	-
14	-	-	-	-	-	-	-	4	13	42	146	483	1,377	436	112	32	8	5	1	-	1	1	-	1
15	-	-	-	-	-	-	1	13	16	67	153	262	623	1,414	356	141	51	10	5	1	3	1	-	-
16	-	-	-	-	-	-	2	1	4	4	13	40	110	349	822	230	68	18	13	-	2	-	1	-
17	-	-	-	-	-	-	1	1	2	7	20	26	43	119	295	646	173	62	16	4	-	2	1	1
18	-	-	-	-	-	-	2	3	2	4	20	26	34	42	118	226	510	159	55	27	8	2	1	2
19	-	-	-	-	-	-	1	-	-	1	-	3	7	13	19	72	176	336	107	33	9	3	1	-
20	-	-	-	-	-	-	-	-	-	2	10	13	23	25	30	67	135	189	420	121	24	19	7	5
21	-	-	-	-	-	-	-	-	-	1	-	1	1	2	2	7	12	34	86	183	66	13	7	8
22	-	-	-	-	-	-	1	-	-	-	-	1	-	1	-	-	8	14	38	91	128	38	22	17
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	9	14	22	59	113	39	35
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	1	2	4	8	26	42	88	50
25+	-	-	-	-	-	-	-	-	-	2	1	1	1	-	3	5	17	19	17	28	30	48	82	801

Table 10. Distribution of multiple reads of the same otoliths using the surface method.

		Second read																		
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+
First read	2	439	72	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	3	22	1,111	108	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	4	-	54	1,123	74	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	5	-	1	45	420	79	9	1	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	1	6	49	803	95	16	2	2	-	-	-	-	-	-	-	-	-	-
	7	-	-	-	15	140	1,197	196	24	7	3	1	-	-	-	-	-	-	-	-
	8	-	-	-	2	22	245	1,802	343	65	12	2	-	2	1	-	-	-	-	-
	9	-	-	-	2	2	45	378	2,460	434	84	22	3	3	-	1	-	-	-	-
	10	-	-	-	-	3	13	77	470	2,348	411	97	31	16	3	3	-	-	-	-
	11	-	-	-	-	1	3	13	102	479	2,068	489	120	34	13	2	1	-	1	-
	12	-	-	-	-	2	2	8	26	88	438	1,963	492	113	49	16	6	1	-	-
	13	-	-	-	-	2	-	2	7	18	100	442	1,468	373	95	15	6	3	1	2
	14	-	-	-	-	-	-	1	1	2	25	86	302	1,031	250	59	24	12	7	4
	15	-	-	-	-	-	1	-	1	3	5	29	76	200	577	156	42	20	7	1
	16	-	-	-	-	-	-	-	-	2	-	8	20	46	142	335	91	25	15	11
	17	-	-	-	-	-	-	-	-	-	-	8	4	19	33	91	209	69	20	21
	18	-	-	-	-	-	-	-	-	-	1	1	-	5	18	24	58	131	42	23
	19	-	-	-	-	-	-	-	-	-	-	-	1	2	2	6	15	44	79	43
	20+	-	-	-	-	-	-	-	-	-	1	-	-	-	1	2	8	21	35	222

The results of this updated bootstrapping analysis indicated that the effective sample size across all composition data was approximately 1.5 times the raw number of samples collected, more variable and much lower than the 2022 analysis (Figure 7). There were differences between the fishery data and the FISS data and among geographical aggregations (Table 11) present with and without accounting for ageing imprecision, but the largest effect of the updated bootstrapping approach was on the historical data due to the lower precision of the surface method. On average, the effective sample size for the FISS data decreased by about 40%, the sexed fishery data by about 60% and the unsexed fishery data (including the historical surface ages) by about 80%.

Because the early fishery data are unavailable in current IPHC data bases, age compositions prior to 1991 were unable to be bootstrapped. Instead, the average relationship between the number of samples and the bootstrapped effective sample size was used to approximate effective sample sizes for use as starting values in the assessment models. Bootstrapped FISS (Table 12) and fishery (Table 13-14) maximum effective sample sizes are provided below.

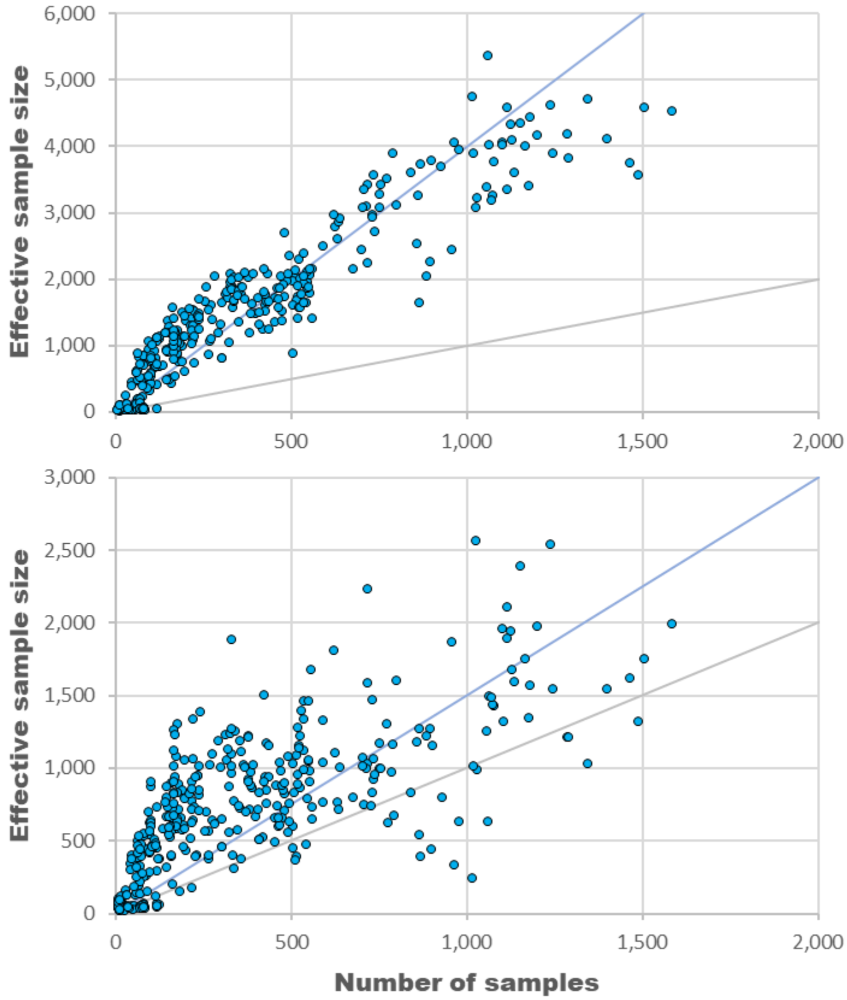


Figure 5. Number of samples vs. bootstrapped effective sample size for all FISS and fishery age compositions data. Upper panel indicates the relationship estimated for the age data used 2022 stock assessment, lower panel indicates the relationship estimated for the age data used in this stock assessment including ageing imprecision. Grey line indicates a 1:1 relationship, blue line indicates a 4:1 relationship (upper panel), or 1.5:1 relationship (lower panel). Note that the y-axes differ between the two panels.

Table 11. Summary of bootstrapping results by data type and spatial aggregation.

Data type	Aggregation	Mean effective N per sample
FISS	Coastwide	2.3
FISS	Region 2	3.6
FISS	Region 3	3.1
FISS	Region 4	4.5
FISS	Region 4B	0.6
All fishery	Coastwide	1.2
All fishery	Region 2	1.6
All fishery	Region 3	2.9
All fishery	Region 4	4.6
All fishery	Region 4B	3.2
Sexed fishery	Coastwide	1.6
Sexed fishery	Region 2	1.9
Sexed fishery	Region 3	3.8
Sexed fishery	Region 4	6.2
Sexed fishery	Region 4B	5.9

Table 12. Bootstrapped effective sample size for FISS age data (1963-2024).

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
1963	--	943	--	--	943
1964	--	456	--	--	456
1965	371	316	--	--	371
1966	187	--	--	--	187
1977	126	419	--	--	521
1978	139	285	--	--	395
1979	--	453	--	--	453
1980	249	642	--	--	884
1981	212	901	--	--	1,076
1982	536	752	--	--	1,004
1983	849	622	--	--	1,184
1984	1,383	578	--	--	1,151
1985	1,121	597	--	--	1,022
1986	1,303	561	--	--	1,096
1988	129	--	--	--	129
1989	--	124	--	--	124
1993	527	436	--	--	584
1994	88	871	--	--	897
1995	871	729	--	--	1,063
1996	1,053	1,503	--	--	1,807
1997	818	907	175	48	796
1998	393	446	152	34	535
1999	962	905	252	58	989
2000	701	788	729	56	1,008
2001	1,881	956	663	70	1,429
2002	1,227	982	586	59	1,320
2003	868	1,088	661	43	1,435
2004	1,233	1,155	792	50	1,495
2005	1,249	1,024	765	59	1,483
2006	1,055	1,217	847	38	1,747
2007	1,095	1,463	832	43	2,108
2008	772	1,047	1,258	32	1,671
2009	1,266	1,123	1,230	33	1,943
2010	1,012	1,278	937	39	1,956
2011	1,111	1,456	1,011	40	2,387
2012	1,183	1,139	828	40	1,888
2013	1,021	866	673	42	1,592
2014	904	1,678	940	52	2,536
2015	734	1,396	1,014	38	1,970
2016	571	1,335	855	35	1,570
2017	693	888	823	48	1,344
2018	655	987	906	36	1,538
2019	897	731	751	42	1,212
2020	1,000	1,078	--	--	1,272
2021	1,020	1,049	275	44	1,250
2022	1,126	399	461	16	1,467
2023	1,211	831	--	--	1,164
2024	601	594	184	--	966

Table 13. Bootstrapped effective sample size for commercial fishery age data (1964-2011).

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
1964	166	195	40	--	182
1965	169	207	34	--	189
1966	146	221	34	--	181
1967	179	260	57	--	220
1968	193	258	40	--	224
1969	162	199	34	--	180
1970	139	244	52	--	191
1971	117	150	26	--	133
1972	790	383	9	--	596
1973	445	512	14	--	458
1974	219	133	9	--	179
1975	335	148	20	--	254
1976	475	264	20	--	377
1977	354	270	20	--	318
1978	345	234	11	--	299
1979	179	197	17	--	193
1980	200	221	3	--	208
1981	209	176	20	--	197
1982	240	268	32	--	251
1983	190	207	66	--	213
1984	243	176	26	--	224
1985	245	193	40	--	227
1986	226	297	98	--	274
1987	760	973	218	--	886
1988	398	504	55	--	453
1989	455	725	112	--	596
1990	703	1,094	143	--	875
1991	1,586	754	355	73	1,210
1992	2,565	849	353	76	1,991
1993	1,865	977	382	44	1,746
1994	1,270	882	261	156	1,322
1995	1,215	1,040	324	52	1,617
1996	1,180	936	268	51	1,542
1997	799	1,170	152	31	1,031
1998	393	638	60	52	330
1999	491	644	68	45	439
2000	368	372	42	38	238
2001	600	395	115	22	386
2002	649	613	199	47	633
2003	476	560	291	28	629
2004	559	432	387	36	621
2005	736	460	373	32	671
2006	707	562	449	42	827
2007	753	767	698	29	1,167
2008	882	642	624	37	997
2009	846	816	464	20	999
2010	720	700	541	17	917
2011	717	731	528	21	1,010

Table 14. Bootstrapped effective sample size for commercial fishery age data (2012-2024). 2017-2023 represent bootstrapping of the sex-specific age data.

Year	Region 2	Region 3	Region 4	Region 4B	Coastwide
2012	522	623	544	20	832
2013	655	607	474	42	952
2014	937	1,338	415	32	2,231
2015	654	822	449	45	1,596
2016	601	724	508	83	995
2017	553	549	417	104	766
2018	399	611	466	61	760
2019	664	711	474	43	1,072
2020	860	628	394	54	1,102
2021	1,216	654	342	53	1,324
2022	878	655	378	67	1,024
2023	823	905	495	97	1,301
2024	833	510	387	25	1,149

Pacific Decadal Oscillation (PDO) Indices

The PDO (Mantua et al. 1997) was been identified and used as a covariate to the scale of Pacific halibut recruitment for decades (Clark and Hare 2002; Clark et al. 1999). Monthly values were averaged to generate an annual deviation which was then assigned a binary 'regime' (Figure 6). The previous approach[2] included 1900 through the present, but there were differences in how the most recent years had been calculated when compared with earlier years. In 2023, the methods for generating the time-series changed and were applied to a longer period. There is now a longer time-series available[3], including the period from 1854 to the present, using consistent methods across the entire period. However, in recalculating the index the period of years over which it was standardized was also changed, so not only do the individual years differ but the transition between regimes also differs, most importantly at the end of the time-series where all values after 1997 would be considered part of a negative regime (i.e., no more than three consecutive positive values; Figure 7). The transition between these two data sets is described below as part of the bridging analysis for 2025.

[2] Data from 1950 available here: https://oceanview.pfeg.noaa.gov/erddap/tabledap/cciea_OC_PDO.htmlTable?time,PDO
[3] Data available here: https://www.ncei.noaa.gov/access/monitoring/pdo/

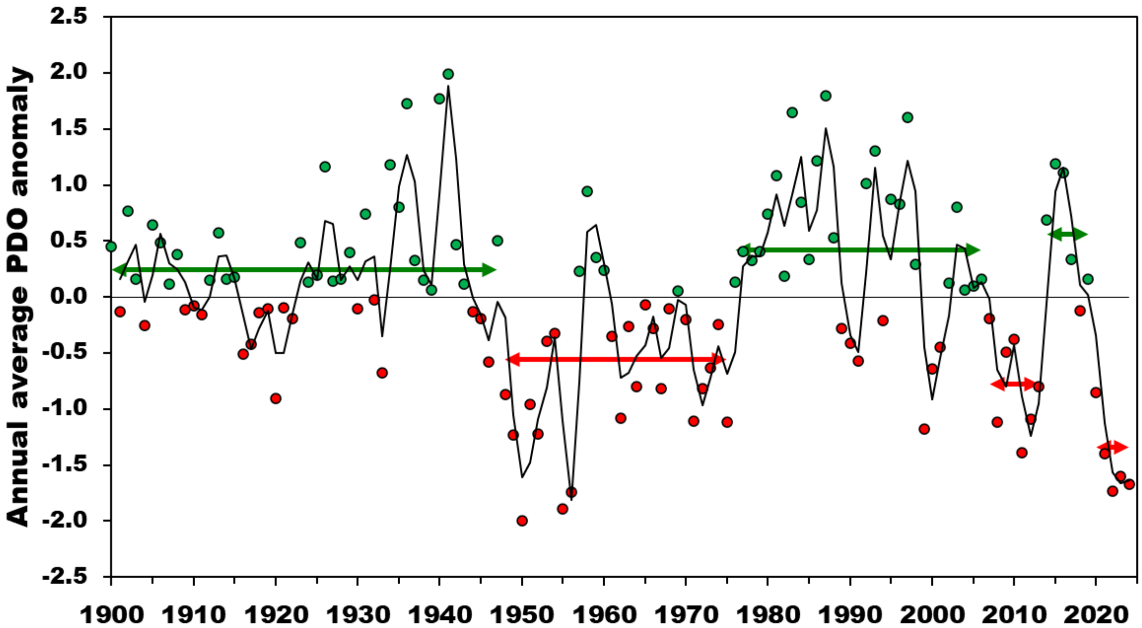


Figure 6. Time-series of PDO anomalies used in previous stock assessments. Lines with arrows indicate positive (green) and negative (red) regimes.

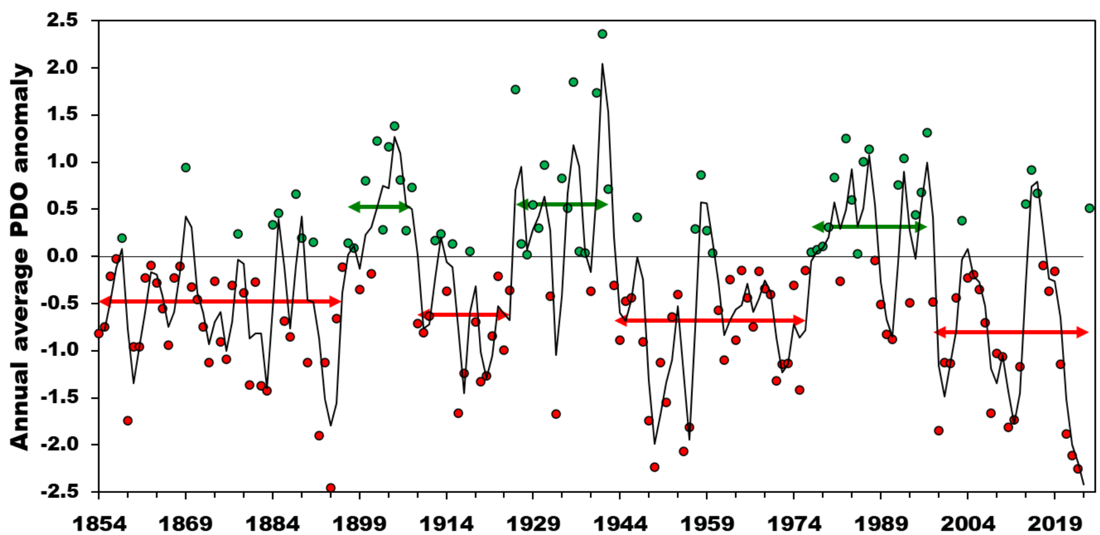


Figure 7. Updated time-series of PDO anomalies used in this stock assessment. Lines with arrows indicate positive (green) and negative (red) regimes.

Maturity

The maturity ogive used in Pacific halibut stock assessments, based on the work of Clark and Hare (2006), has remained unchanged since 2006. That analysis used visually estimated maturity from a subset of the IPHC Regulatory Areas collected from 2002-2004 to generate a logistic relationship, which was then truncated to be zero below the first age for which a fish had been observed to be mature (age 8). In 2025, an extensive analysis of histological maturity was completed by the IPHC (*IPHC-2025-SRB026-06*). This is described in detail in the referenced document. Briefly, Histological maturity ogives were estimated with Generalized Additive Models (GAMs) fit by Biological Region to data collected between 2022 and 2024. This information was then used to calibrate the corresponding visual estimates of maturity, and those extending back to 2002. The reason for truncating this analysis in 2002 is the change in ageing method from break-and-bake (2002+) to biased surface methods (<=2001), which would create a bias in the maturity ogive for older fish. The curves for each Biological Region over the time-series were combined into coastwide ogives based on the relative abundance of Pacific halibut in each Region. Finally, a time-series average was calculated with and without truncating the youngest ages for which no Pacific halibut have been observed to be mature (Ages 5-6; Figure 8). The transition between the older curve and the newer curve is described below as part of the bridging analysis for 2025 and the effect on model results of truncating the youngest ages was explored as part of the sensitivity analyses.

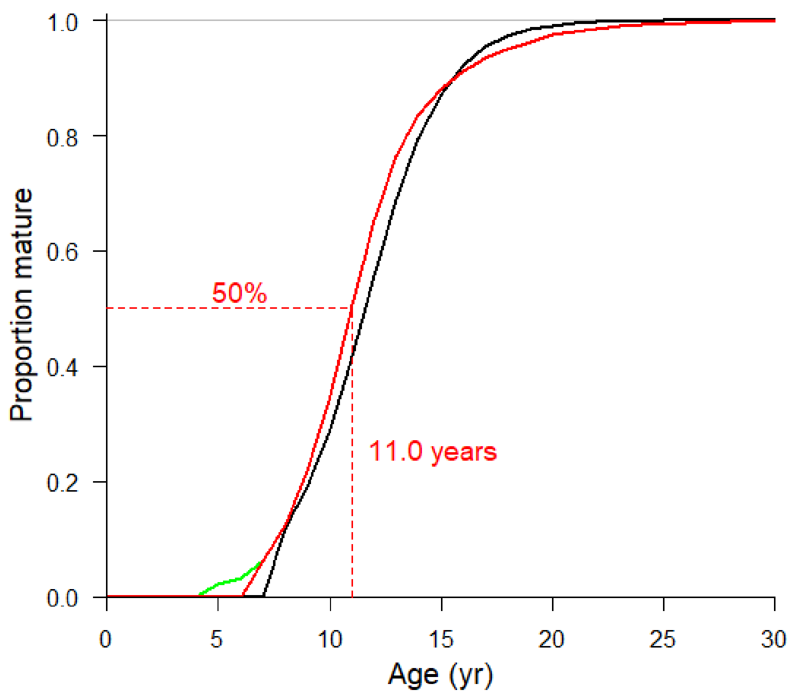


Figure 8. Historical maturity ogive from Clark and Hare (2006; black line), updated curve (green line) and updated curve truncated to be zero below age 7 (red line).

Model development

Multimodel approach

Creating robust, stable, and well-performing stock assessment models for the Pacific halibut stock has proven extremely challenging due to the highly dynamic nature of the biology, distribution, and fisheries (Stewart and Martell 2014). The stock assessment for Pacific halibut has evolved through many different modeling approaches over the last 30 years (Clark 2003; Clark and Hare 2006). These changes have reflected improvements in fisheries analysis methods, changes in model assumptions, and responses to recurrent retrospective biases and other evidence of model mis-specification and concurrent degradation of model performance (Stewart and Martell 2014). Perhaps the most influential of these changes was the transition from separate IPHC Regulatory Area-specific assessment models to a coastwide model in 2006, as the understanding of adult movement among areas was substantially updated by the results of the IPHC's extensive PIT-tagging experiment in 2003-2009 (Clark and Hare 2006; Webster et al. 2013). Some simulation studies have found that dividing a migratory population into several discrete assessment units tends to overestimate the total biomass (e.g., Li et al. 2014; McGilliard et al. 2014).

Although recent modelling efforts have created some new alternatives, no single model satisfactorily approximates all aspects of the available data and scientific understanding. Building on simpler approaches in 2012 and 2013, in 2014, the current ensemble of four stock assessment models, representing a two-way cross of short vs. long time series', and aggregated coastwide vs. AAF models was developed for the full assessment analysis and review in 2015 (Stewart and Martell 2016). The models were further improved in 2019 to accommodate sex-specific age composition data from the commercial fishery (Stewart and Hicks 2019b) and again in 2022 to improve the input sample sizes for age composition data and to better inform estimates of natural mortality (Stewart and Hicks 2022).

AAF models are commonly applied when biological or sampling differences among geographical areas make coastwide summary of data sources problematic (Waterhouse et al. 2014). AAF models continue to treat the population dynamics as a single aggregate stock, but fit to each of the spatial datasets individually, allowing for differences in selectivity and catchability of the fishery and survey among regions. In addition, AAF models more easily accommodate temporal and spatial trends in where and how data have been collected, and fishery catches have occurred. This is achieved through explicitly accounting for missing information in some years, rather than making assumptions to expand incomplete observations to the aggregate coastwide level. Both aggregating the data into a single series and approximating spatial dynamics via AAF approaches may be useful under some circumstances; however, there is no clear best-performing configuration under all conditions. Not surprisingly, models that most closely match the biology, which is only known under simulated conditions, tend to perform the best (Punt et al. 2015).

To capture the structural uncertainty inherent among the Pacific halibut stock assessment models, it is necessary to use multi-model inference, here referred to as an 'ensemble' of models (e.g., Ianelli et al. 2016; Karp et al. 2018; Stewart and Martell 2015). The ensemble approach, applied in many fields in addition to fisheries (Du 2014; Hamill et al. 2012), recognizes that there is no "perfect" or "true" assessment model, and that a robust risk assessment can be best achieved via the inclusion of multiple models in the estimation of management quantities and the uncertainty about these quantities (Stewart and Martell 2015). This stock assessment is based on the approximate probability distributions derived from an ensemble of models, thereby incorporating the uncertainty within each model as well as the uncertainty among models. This approach reduces potential for abrupt changes in management quantities as improvements and additional data are added to individual models (Stewart and Hicks 2018), and provides a more realistic perception of uncertainty than any single model, and therefore a stronger basis for risk assessment.

The current ensemble explicitly captures two critically important dimensions of uncertainty: how the time-series data are used via short and long models, and how the spatial information is treated in the models via data aggregation to the level of Biological Regions treated as separate fleets (AAF) or to the coastwide level. Inclusion of these sources of structural uncertainty results in wider confidence intervals than are commonly seen in single-model stock assessments (Stewart and Hicks 2019a). More detail on how the models are weighted and integrated can be found in the Ensemble section below.

Structural rationale

Consistent with analyses since 2015, this stock assessment is implemented using the generalized software stock synthesis (Methot and Wetzel 2013b), a widely used modeling platform developed at the National Marine Fisheries Service. This platform allows for a wide range of structural choices with regard to biology and growth, catchability, selectivity, spatial processes, stock-recruitment dynamics as well as error distributions and integrated projections. A benefit of using this code is that it is well documented, and the inputs and output formats are standardized (Methot Jr et al. 2021b), regardless of model configuration, allowing easy interpretation of model files and rapid evaluation of the results without re-running the fitting algorithm using the r4ss package (Taylor et al. 2021) implemented in the R programming language (https://cran.r-project.org/).

A primary structural stock assessment model choice is whether or not to model growth explicitly (and often parametrically) or empirically. Many stock assessments assert/estimate a growth function of some type and rely on this growth function to translate between numbers and biomass for model calculations. This approach has the benefits of allowing direct fitting to observed length observations, interpolating and/or extrapolating predictions for years where direct observations may be missing, as well as direct inclusion of the potential effects of selectivity at length on the observed data. The cost of such an approach is that growth can be an extremely complex process, varying over time, space and by cohort (via density dependence). When there is appreciable growth variability, a great deal of complexity may be required to adequately model this population process, even before sampling and selectivity issues have been addressed. Failure to account for this type of variability can lead to poor fits to composition data, potentially biasing the assessment results (Maunder et al. 2015, and subsequent special issue papers).

Pacific halibut show a very high degree of growth variability, with a 20 year-old female potentially as small as 60 cm or as large as 240 cm (Figure 9), males do not reach the same maximum sizes but also show a high degree of variability. Both sexes show nearly linear growth over their entire lifetime making the use of common parametric growth relationships (e.g., the von Bertalanffy curve) unreasonable. Previous efforts to fit to length data in the assessment model have shown little information content and a very high computational overhead.

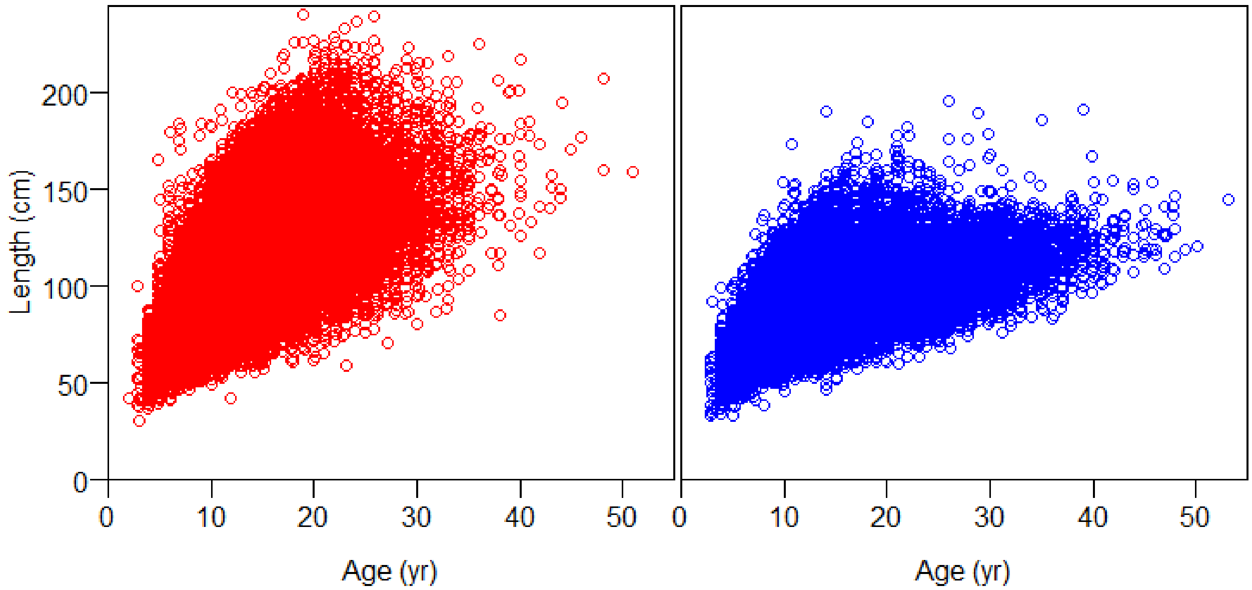


Figure 9. Observed length at age for female (left panel) and male (right panel) Pacific halibut collected by the FISS through 2023.

The Pacific halibut stock assessment models, like many other stock assessments with relatively complete age and size information, take a simpler approach to growth by using empirically derived weights-at-age. The empirical weight-at-age approach has the benefit of reducing complexity with regard to growth modelling but has several costs in other modelling areas. These include the need for more complexity in modelling selectivity, particularly where some of the selectivity process may be a function of size rather than age alone. This is the case for Pacific halibut, where the interaction of changes in size-at-age, gear selectivity that is likely at least partially a function of fish size, and a minimum size limit thus requires the treatment of selectivity-at-age as a time-varying process (Stewart and Martell 2014). However, the treatment of selectivity as time-varying appears to be a necessity for Pacific halibut even if treated as a function of size; static selectivity for a spatially aggregated model in the face of changes in availability was identified as a primary contributor to severe historical retrospective patterns (Stewart and Martell 2014).

There are relatively few examples of stock assessments used for management purposes that are explicitly spatial: modelling movement among areas, distributing recruitment events, and tracking spatial variability in biological characteristics (e.g., McGilliard and Palsson 2021; Stewart et al. 2009). Most such cases rely on low rates of movement to allow for estimation of recruitment distribution among areas. More frequently assessments either aggregate the available data across spatial heterogeneity (preferably weighting appropriately such that the aggregate information reflects the underlying distribution), or retain separate data series representing spatial areas, but fit to them in the context of a single instantaneously mixing population model (the AAF approach). These methods for dealing implicitly with spatial dynamics are by necessity gross approximations, with performance properties specific to a particular application that are unknown, and almost certainly depend on the true underlying processes. Some simulation studies have shown that fisheries operating in different areas with differing selectivity schedules can be reasonably approximated by an AAF approach (e.g., Waterhouse 2014). Other studies have found acceptable performance of AAFs when simulating actual spatial variability (e.g., Hurtado et al. 2014, McGilliard et al. 2014); however additional studies have found that combining spatial data into weighted aggregates also performs acceptably and may be more stable than more complex AAF approaches (Punt et al. 2015, Li et al. 2015). A primary conclusion from simulation-based studies is that if the true underlying process is well-represented, then models reflecting these dynamics tend to perform well (Bosley et al. 2021; Goethel and Berger 2017). Unfortunately, for Pacific halibut it is not clear whether aggregated or AAF models might be the best choice as neither approach accurately represents the complex spatial dynamics.

The choice of how long a time-series to model generally represents a compromise among: data availability, data quality, model complexity, and technical convenience (e.g., data preparation and model convergence times). As assessment model time series' are extended to include more historical data, the quality of those data generally becomes increasingly lower as standardization of sampling programs has a greater likelihood of having changed appreciably. In the case of Pacific halibut, fishery-independent survey information has been reasonably comprehensive since approximately 1997, and sufficient to support the IPHC's geostatistical model since 1993 (Webster 2018; Webster et al. 2020). Current fishery sampling approaches have also not changed dramatically over the same period. The completeness of this time period with regard to data availability was one of the primary incentives for stock assessment models used by the IPHC since 2006 to begin the modelled period in 1996. Notable differences prior to that period included the transition in the survey and fishery from "J" to circle hooks (1984), variable and much less comprehensive survey coverage (<1997), lack of access to raw historical fishery data (ages, catch rates, etc.; <1991), and many others. The costs of using only a relatively short time-series include: a lack of integration between harvest strategy calculations derived from the full historical period, a lack of perspective on recent trends, the need for careful treatment of initial model conditions, inability to estimate some model parameters with only a shorter time-series, and increased sensitivity to additional data, as each year represents a greater fraction of the total information available in the model. These trade-offs prompted the development of the first long time-series model in 2013, with the recognition that neither the short or long time-series approach was clearly superior, and that differences in the results reflected a meaningful source of uncertainty in the assessment results.

All of the halibut models considered here treat male and female halibut separately. Like many broadcast spawning fishes, there is a basic assumption that spawning is likely to be limited primarily by female spawning output and not by male abundance (at least over a reasonable range of sex-ratios; this is generally not a concern except for cases such as some crab stocks where fishery mortality may operate primarily on males). If the sex-ratio could be expected to be stable over time, it might be reasonable to structure assessment models without regard to sex and/or just assume half of the mature biomass represented females. This is a common approach for species where there is little dimorphic growth. However, for Pacific halibut, highly dimorphic growth interacting with gear selectivity for larger fish, and a fishery in which there are strong incentives to target the larger females (due to the minimum size limit and graduated price structure) results in sex-ratios of the catch and of the landings skewed largely toward females. Historical modelling suggested that the potential for a static assumption regarding sex-ratio could lead to a highly biased interpretation of stock status and that females and males are best modelled separately.

In aggregate, these considerations led to the choice of four stock assessment models during the 2014 assessment process: a two-way cross of: coastwide vs. AAF data structuring, and long vs. short time-series. Each of these models explicitly treated male and female halibut separately and employed empirical weight-at-age rather than an explicit growth function. All models fit to both fishery and survey index trends and age compositions and allowed for temporal variability in selectivity and catchability. Additional alternative modelling approaches were considered, including a simple surplus production model and a Virtual Population Analysis model. Both of these approaches suggested that recent removals and stock trends were on a similar scale to the four models included in that assessment (Stewart and Martell 2015) but presented sufficiently substantial issues in interpretation or application to the management process that they were not formally included in that stock assessment.

General model configuration

There are a number of basic technical settings and features that are common to all four stock assessment models described here. This section provides an overview, which is supplemented by a description of specific individual model details below.

The stock synthesis software separates inputs into several files read in prior to model estimation including the primary data file, the primary control file (including parameter setup and estimation switches), the weight-at-age file, the forecast file (including settings for reference point calculations), and the starter file (including some general estimation and reporting switches and settings). Each of these input files for each of the four stock assessment models described here are included in the background documents, along with the primary report file of estimated and derived quantities and the directory of summary and diagnostic figures created using the R package r4ss (Taylor et al. 2021). A full summary of supplemental material is provided in Appendix A. Note that not all automatically created diagnostic material, nor all of the model output is relevant to the model configurations employed here.

These models were configured to make use of relatively standard stock assessment practice in the population structuring. There were no seasonal dynamics, and catches were assumed to be removed halfway through the year via Pope's approximation. This approach does not require iterative estimation of fleet- and year-specific fishing mortality rate parameters (often reducing model run times) and should reasonably approximate the dynamics unless fishing mortality rates are extremely high or within year growth increments very large. Catches were input in thousands of pounds (net weight; head-off and gutted, approximately 75% of round weight), so that the mean weight-at-age inputs were in net pounds and the numbers-at-age are tracked in thousands of individuals. Population dynamics contain ages 0-30, and female and male halibut are modelled separately in the underlying dynamics.

The input data were partitioned via a fleet structure of: the directed fishery (by area in the AAF models), discard mortality from the directed fishery, non-directed discard mortality ('bycatch'), recreational, subsistence, and survey (FISS; by area in the AAF models). Table 15 summarizes the data and key features of each model (note that changes from the 2024 model are described in greater detail below). Age data were partitioned by sex (the vectors for each year contain females, then males, such that the sex-ratio is inherently included in the age compositions), where this information was available and assigned the appropriate ageing method in the data file (see section above). Where few fish contribute to the 'tails' of the age distributions for each fleet and year combination, the model was set to automatically aggregate observations and predictions at each of the low and high ages with proportions less than 0.1%. This choice avoids large vectors of zeroes in the multinomial calculations. The model was also set up to add a very small constant (0.0001) to all age proportions in order to stabilize the computation.

All model growth specifications were bypassed in order to use the empirical weight-at-age approach; therefore, the settings in the control file and the results included in model outputs related to these settings are not meaningful (this includes length-at-age, weight-at-length, and maturity-at-length; these are all integrated directly in the weight-at-age inputs). The weight-at-age file also included a matrix of spawning output-at-age representing the product of annual weight-at-age (a matrix) and the vector of maturity-at-age (Stewart and Webster 2025).

For most estimated parameters, uniform priors were implemented, with bounds sufficiently wide to avoid maximum likelihood estimates falling on or very near a bound, unless the bound was structurally logical. Exceptions included process-error deviations, which are constrained by variance parameters, and the longevity-based prior on natural mortality, as described above. Table 16 summarizes the counts of estimated parameters in each model. Natural mortality was allowed to differ by sex, with the value for male halibut estimated in all four models, and the value for females in all but the short coastwide model. Treatment of both the stock-recruitment relationship and the initial conditions at the start of the modelled time-series differed among the four models and are described below.

The double-normal selectivity parameterization is used in all four models, as it represents a flexible, but still parametric approach that can easily be made time-varying via just one or two parameters with annual deviations. There are more flexible nonparametric selectivity options, but these generally require all the parameters to vary over time, creating a substantial increase in complexity. The double-normal selectivity can be easily configured to be either asymptotic or dome-shaped, by adjusting the width of the peak and/or descending slope and final selectivity parameters. It also includes an option for male selectivity to be offset from female selectivity, based directly on the parameters of the selectivity curve, such that time-varying selectivity for one sex can be mapped into temporal variability for both sexes without estimating a second set of deviation parameters. The double-normal was implemented for all model fleets, with at least the ascending limb of selectivity (ascending width and peak parameters) allowed to vary over time for all four models (described further below).

Table 15. Comparison of structural assumptions among models.

	Model			
	Coastwide Short	Coastwide Long	AAF Short	AAF Long
Modelled period[1]	1992+	1888+	1992+	1888+
Data partitions	N/A	N/A	Regions 2, 3, 4, 4B	Regions 2, 3, 4, 4B
Directed Fishery fleets	1	1	4	4
Other fishing fleets	4	4	4	4
Survey fleets	1	1	4	4
Fishery CPUE (weight)	1992+	1907+	1992+	1907+, 1915+, 1981+, 1981+
Fishery age data years	1992+	1935+	1992+	1935+, 1935+, 1945+, 1991+
Survey CPUE (numbers)	1993+	1977+	1993+, 1993+, 1997+, 1997+	1977+, 1977+, 1997+, 1997+
Survey age data years	1993+	1963+	1993+, 1993+, 1997+, 1997+	1965+, 1963+, 1997+, 1997+
Weight-at-age	Aggregate	Aggregate	Areas 2, 3, 4	Areas 2, 3, 4
Female *M*	Fixed at 0.15	Estimated	Estimated	Estimated
Male *M*	Estimated	Estimated	Estimated	Estimated
Stock-recruit relationship	B-H	B-H	B-H	B-H
Initial conditions estimated	R_{init} *N*-at-age: 1-19	R_0, *N*-at-age: 1-29	R_{init}, *N*-at-age: 1-19	R_0, *N*-at-age: 1-29
Environmental regime effects on recruitment	No	Estimated	No	Estimated
Steepness (*h*)	0.75	0.75	0.75	0.75
$\sigma_{\text{recruitment deviations}}$	1.0	0.54	0.72	0.50
Survey selectivity	Asymptotic, by sex	Asymptotic, by sex	Domed, by sex (R2, R3) Asymptotic, by sex (R4, R4B)	Domed (R2, R3), Asymptotic (R4, R4B)
Fishery selectivity	Asymptotic, by sex	Asymptotic, by sex	Domed, by sex (R2, R3) Asymptotic, by sex (R4, R4B)	Domed, by sex (R2, R3) Asymptotic, by sex (R4, R4B)
Scale of male fishery selectivity	Estimated, time-varying	Estimated, time-varying	Estimated, time-varying	Estimated, time-varying
Non-directed discard selectivity	Domed	Asymptotic	Domed	Domed
Recreational selectivity	Asymptotic	Domed	Domed	Domed
Discard selectivity	Domed, by sex	Domed, by sex	Domed, by sex	Domed, by sex
Subsistence selectivity	Mirrored to recreational	Mirrored to recreational	Mirrored to recreational	Mirrored to recreational

[1]Mortality estimates for 2025 were projected based on adopted IPHC limits.

As has been the case in all recent halibut models, the catch-per-unit-effort index derived from the directed halibut fishery is included in each of the models, but the catchability is allowed to vary over time. In principle, there are many factors which can create changes in the proportionality of the catch-rate in a fishery with the underlying population. The most obvious of these are abrupt changes in fishing methods, such as the change from “J” to circle-hooks in 1984. This type of change was accommodated (in the long time-series models) via an unconstrained deviation on catchability in that year (effectively a separate *q* for the two parts of the time series). Beyond abrupt changes, there are many factors that can ‘drift’ over time but may not be so obvious, including technological improvements, changes in spatial areas or times of year being fished, targeting of areas with large vs. small fish, etc. This type of change suggests a random walk in catchability, which was the approach taken in all four models here. To implement this, a catchability parameter was estimated for the first year for which index data were available, and then a deviation (from the previous year’s value, not the mean) was estimated for each subsequent year of the time-series. The annual catchability deviations were constrained by a single σ for each fleet. The iterative tuning algorithm for identifying the internally consistent values for each σ is described below.

In all models, fit to the age data used a multinomial likelihood with initial input sample sizes based on the revised bootstrap results described above, subsequently adjusted downward via a multiplicative scalar for each fleet in the control file (more discussion below). Indices of abundance from both the FISS and commercial fishery (by area in the AAF models) were fit using a log-normal likelihood and input log(*SE*)s based on the space-time modelling (FISS) or the between trip variability (fishery). Survey indices were fit in numbers of fish to avoid converting catch in numbers to weights in the data and then weights back to numbers in the model predictions (as informally recommended by the Scientific Review Board in 2014). Weight-per-unit-effort is the native scale for the fishery indices based on logbook records.

Using the method first developed for the 2015 assessment, discard mortality, bycatch and recreational selectivity are estimated, but the age composition data are down-weighted to avoid imparting any significant information on recruitment strengths from these uncertain and potentially non-representative data sets. In this way, the data that are available serve as an informative ‘prior’ on the selectivity for each of these fleets, and therefore propagate some uncertainty associated with selectivity estimation (vs. simply specifying selectivity as fixed parameters), but do not strongly inform other model parameters and population dynamics estimates.

Discards in the directed commercial fishery are treated as a separate fleet in each model. This approach was taken for several reasons: discard rates may be a function of spatial fishing effort and not simply contact selectivity as is often assumed in stock assessments - there has been little relationship between the magnitude of discards and the magnitude of commercial landings when this has been evaluated for previous reviews. Further, modelling discards with a retention curve in the empirical weight-at-age approach within SS does not allow for separate mean weight-at-age vectors to be applied to landings and discards (which may differ significantly for younger ages due to the size limit). Sex-specific selectivity curves were estimated in each model informed by the observations from the sublegal fish captured by the setline survey. The selectivity was configured to be a double normal, with female halibut offset from male halibut to account for the dimorphic growth (the opposite of all other fleets), and the relative scale of females to males estimated directly. Both sexes were allowed to be dome-shaped, with differing descending limbs. Because the sublegal survey age data were already included in the likelihood as part of the survey age compositions, it would be a misrepresentation of the uncertainty to naively fit them again equally as part of the discard data set. Instead, previous analyses showed that down-weighting these data such that they had a very small input sample size had little appreciable effect on the model results but still allowed for the direct estimation of selectivity. This approach lends itself to direct inclusion of observer data on discarded halibut when/if sampling expansion methods that are representative of the entire fleet become available.

Table 16. Comparison of estimated parameter counts among models.

	Model			
	Coastwide Short	Coastwide Long	AAF Short	AAF Long
Static				
Female M	--	1	1	1
Male M	1	1	1	1
Log(R_0)	1	1	1	1
Initial R_0 offset	1	--	1	--
Environmental link coefficient	--	1	--	1
Fishery catchability	1	2	4	7
Survey catchability	1	4	--[1]	4
Fishery selectivity	5	5	21	20
Discard selectivity	8	7	5	5
Non-directed discard selectivity	4	2	3	3
Recreational selectivity	5	6	5	6
Survey selectivity	5	5	21	18
Total static	32	35	63	67
Time-varying[2]				
Recruitment deviations[3]	57	171	57	171
Fishery catchability deviations	37	118	148	322
Fishery selectivity deviations	85	200	345	668
Survey selectivity deviations	108	159	270	324
Total deviations	287	648	820	1,485
Total	319	683	883	1,552

[1]The analytic solution is used for this catchability parameter.
[2]Includes five uninformed forecast years, in order to propagate uncertainty.
[3]Includes deviations representing the initial age structure at the beginning of the modelled time-period.

Bycatch and recreational selectivity curves were also allowed to be dome-shaped given the relative frequency of younger halibut in the observed distributions. Where descending limb parameters were estimated to be at the upper bounds, these parameters were fixed (making the curves asymptotic) to avoid any negative behavior during minimization and approximation of the variance in model quantities via the Hessian matrix. Since the 2019 assessment, sex-specific age composition data for the recreational fishery has become available (Stewart and Webster 2025), and so additional offset parameters were added to allow for sex-specific selectivity as in the treatment of the discards. Because of the down-weighting of the data for these series, and the unknown or potentially poorly spatially representative nature of the data themselves, no attempt was made to allow these selectivity curves to vary over time.

The presence of both observation error (in the indices and age composition data) and process error (in fishery catchability and selectivity for the survey and fishery) creates a challenge for standard weighting and tuning practices employed in many assessment models. Specifically, if process error is not modelled (and/or a fixed value is asserted), the input sample sizes (and sometimes index variances) can be relatively easily iteratively tuned or estimated (Maunder 2011). This approach is useful for reducing the potential effects of outliers, lack-of-fit, or model misspecification with regard to composition data (Francis 2011). At the other extreme, if the observation error is assumed to be known (and assigned a fixed value), then the degree process error can be estimated via random effects, or iteratively tuned using a maximum likelihood-based approximation (the 'Thompson and Lauth method'; Annex 2.1.1 in Thompson and Lauth 2012). When data are sufficient, both components can be iteratively, or by more statistically rigorous means, estimated simultaneously (Thorson 2019; Thorson et al. 2016).

The general goal for the treatment of process error in selectivity and catchability and observation error in the data is to first reduce clear signs of bias to the degree possible and then to achieve internal consistency among error distributions and sample sizes/variances. In all four models developed here, the initial input sample sizes, derived from the revised bootstrapping analysis described above were considerably larger than commonly applied weighting for stock assessment models would suggest (Table 12-14). These values were iteratively reduced based on evaluation of three considerations: the relative magnitude of the standardized Pearson residuals, comparison of the input value for each fleet with the harmonic mean effective sample size which is an unbiased estimator for a set of independent multinomial samples (Stewart and Hamel 2014), and the scaling suggested by the Francis (2011) method (as implemented in the r4ss package). For almost all fleets and all models, this approach led to a substantial reduction from initial sample sizes. In no cases were the input values increased from the maximum values derived via bootstrapping.

Starting from a small value for the input σ for each fleet and parameter combination where temporal variability was allowed, process error was increased until the tuned value was consistent with the degree of variability observed among the deviations (SE^2_{devs}) and the average uncertainty of the deviations themselves $\bar{\sigma}_{dev}{}^2$. This approach is very close to that outlined by Thompson and Lauth (2012) and is consistent with the preferred method for tuning this and other types of process error (such as recruitment deviations) in stock synthesis (Methot and Taylor 2011; Methot et al. 2019):

$$\sigma_{tuned} \sim \sqrt{SE_{devs}^2 + \bar{\sigma}_{dev}{}^2}$$

In addition to providing internal consistency, this approach makes intuitive sense: under perfect information the average variance of the deviations will be zero and the variability among the deviations will exactly match the process error, conversely, under no information the variance of the deviations will be the input constraint. After initial process error tuning, the input sample sizes were adjusted downward until the weights suggested by the fit to the mean age over the time series were approximately equivalent to the input values (the "Francis method'; Francis 2011). There were only minor changes to the tuned σ values required after iteration of the input sample sizes, suggesting the two processes were relatively separable and stable; further there were only minor changes in the process error variances in this assessment relative to the 2019 and 2022 assessment despite the revised input sample sizes.

As a final model-building step, models were regularized via adjusting parameterizations through removing and/or fixing selectivity parameters that consistently remain stuck to bounds or are not contributing to the likelihood in a meaningful way (<1% correlation with other model parameters). This regularization does not include forecast recruitment deviations, which are expected to be uncorrelated with other model parameters (and the objective function) but are 'estimated' in order to appropriately propagate the uncertainty in recent recruitments into forecasts.

The tuning approach for the stock-recruitment relationship was very similar, ensuring that the input σ governing recruitment variability was consistent with the observed variability and variance estimates; the calculation for this tuning is automated in the r4ss package, and the output was used as a guide for the scale of the bias correction, including ramps to and from the peak value consistent with the information content of the data and variability in the deviations observed in the output. This step is important for recruitment variability as it also provides for a better approximation for the bias correction in recruitment deviations (Methot and Taylor 2011) in the 'main' or best informed period of the time-series of recruitments. Again here, after initial tuning, little change was observed across alternative models or from previous results.

In the end, this tuning process provides a model that is internally consistent: the error distributions are commensurate with the fit to the data and the degree of process error is consistent with the signal (information content) in the data. Importantly, accounting for process error in selectivity was the primary solution for historically observed retrospective patterns in the Pacific halibut stock assessment models (Stewart and Martell 2014). Tuning diagnostics and results specific to each model are provided below.

Coastwide short

The initial conditions for a model starting after an extensive historical fishery and appreciable recruitment variability must be structured to avoid simple assumptions that may have strong effects on the subsequent time-series. For the coastwide short model, the initial conditions included estimating the population numbers at age 1-19 in the first year of the model (1992 after extension of the time-series; see below). Since the age data available for the initial year were aggregated at age-20 (due to the historical use of the surface ageing method), there was no specific information on additional individual year-classes. To accommodate a non-equilibrium value in the plus group, an offset to initial equilibrium recruitment (via a single time 'block') was also estimated. The effect of these two approaches was to essentially decouple the numbers-at-age at the beginning of the time-series from any equilibrium assumptions.

As in previous assessments, the coastwide short model employed a Beverton-Holt stock recruitment relationship with estimated equilibrium recruitment level (R_0) setting the scale of the stock-recruit relationship. Steepness (*h*) was fixed at a value of 0.75 for this and all other models, an assumption that has been explored extensively in previous assessments. Fixing steepness, but iteratively solving for the internally consistent level of recruitment variability generally does not have a large effect on year-class strengths where data are informative, but does have very strong effects on direct estimates of Maximum Sustainable Yield (Mangel et al. 2013); however, this quantity is not of specific interest for the Pacific halibut assessment. A summary of the number of estimated parameters contributing to each aspect of the model is provided in Table 16.

Age-based selectivity for female halibut for both the FISS and commercial fishery was estimated using the double normal, forced to be asymptotic once it reached peak selectivity. This required two parameters: the ascending width of the curve and the age at which the peak selectivity is reached. Both parameters are allowed to vary over time with a random walk of annual deviations. These deviations were initiated in the first year for which age composition data were available, and extend into the forecast period (three years) to propagate the variance associated with potential future changes in selectivity. Male selectivity for the survey was estimated via offsets to the female ascending width and peak parameters, and a third parameter defining the scale of male selectivity relative to that for females. Male selectivity offset parameters for the fishery allow for the time-varying process to apply to both males and females with only two additional parameters. The scale of male selectivity for both the survey and fishery were allowed to vary over time as a random walk. For the fishery, these deviations are estimated beginning in 2018, since the sex-specific age composition derived from genetic analyses begins in 2017. In aggregate, there were five estimated base parameters each for the survey and fishery and annual deviations on the ascending limb parameters (Table 16).

Coastwide long

Initial conditions for the coastwide long time-series model include the initial age structure and a long period of uninformed recruitments with the model period beginning in 1888 and the first age data available for 1935 (Table 15); therefore, there was a substantial 'burn in' for recruitment variability prior to any data. The treatment of the stock-recruitment function in the coastwide long model was substantially different from that of the coastwide short model. Consistent with historical IPHC analyses (Clark and Hare 2002, 2006) and previous stock assessments, the coastwide long model allowed for the possibility that the scale of the stock-recruitment function is correlated with the regimes of the Pacific Decadal Oscillation (PDO; Mantua et al. 1997). To implement this approach, a Beverton-Holt relationship was used, parameterized with an estimated value for the equilibrium recruitment level (R_0) parameter, and a fixed value of steepness (*h*) of 0.75. The annual average of the PDO index (see description above for updates to this index) was converted to a binary indicator (PDO_{regime}) where productive regimes (e.g., 1977-1997) were assigned a value of 1.0, and poor regimes (e.g., 1943-1976) a value of 0.0. These regimes were linked to the scale of the stock-recruit function via an adjusted equilibrium level of recruits (R_0') based on an estimated coefficient (β) creating an offset to the unadjusted value:

$$R_0{}' = R_0 * e^{\beta * PDO_{regime}}$$

The adjusted equilibrium recruitment value was then used in the stock-recruit function with bias-corrected annual deviations:

$$R_y = f\left(SB_y, R_0{}', SB_0, h\right) * e^{r_y - \frac{\sigma^2}{2}}$$

This parameterization has the desirable property that if there is no correlation between the putative environmental index and underlying mean recruitment, the β parameter will be estimated at a value of 0.0 and the recruitment estimates will be unaffected. In that case R_0' is simply equal to R_0. As was the case for the coastwide short time-series model, fixing steepness precludes the naïve use of *MSY* estimates.

The approach to selectivity in the coastwide long model was identical to that in the coastwide short model. Selectivity deviations on the ascending limb parameters of the fishery and survey series were initiated in the first year for which age composition data were available for both the fishery (1935) and the survey (1963).

Natural mortality (*M*) is estimated separately for males and females in the coastwide long model using the informative prior described above.

AAF short

The AAF short model was configured very similarly to the coastwide short model. The most notable difference was in the treatment of selectivity for the survey and fishery in Biological Regions 2 and 3: these were allowed to be dome-shaped relative to the coastwide population dynamics. Implementing dome-shaped selectivity for these four model fleets requires the addition of a third selectivity parameter defining the width of the descending limb. This additional parameter was not allowed to vary over time. Similar to the coastwide long model, the three parameters defining the annual male offset to female selectivity for the commercial fishery in each area were only estimable beginning with the 2017 sex-ratio data. Temporal variability in selectivity parameters occurred over a slightly longer range of years in the AAF short model, as there were Region-specific survey data available for the entire time-series from Biological Regions 2 and 3. Beginning with the 2022 assessment, the AAF short model estimates female and male M.

AAF long

The only structural differences between the AAF long and AAF short models were the years over which deviations in recruitment, selectivity and catchability are estimated. The AAF long model treated the stock-recruitment function in the same manner as the coastwide long model, including the PDO as an estimated covariate to equilibrium recruitment.

Changes from 2024

In the intervening period between the last full stock assessment analysis in 2022 and this preliminary analysis for 2025, the length and information content of the data sets has grown, and new information, such as the revised bootstrapping results (described above) has become available. Changes to specific data sets have been documented in the recent assessments and their effects evaluated individually in each year (Stewart and Hicks 2024, 2025). Key changes for 2025 included:

1) Extending the time series to include projected mortality based on limits adopted for 2025 (IPHC 2025),
2) updating to the newest stock synthesis software version (3.30.23.1; Methot Jr 2024),
3) updating the time-series information for the PDO, used as a covariate to the stock-recruitment relationship,
4) retuning the constraint on the scale of male time-varying fishery selectivity (the sex-ratio of the commercial fishery) and extending this variability into the forecast,
5) improving the bootstrapping approach to pre-model calculation of maximum effective sample sizes to include ageing imprecision (Hulson and Williams 2024),
6) re-tuning the process and observation error components of these models to achieve internal consistency within each,
7) and updating the maturity ogive to reflect the recent histology-based estimates produced by the IPHC's Biological and Ecosystem Sciences Branch.

The sequential effects on the model results of each of these changes are described below as a 'bridging' analysis from the 2024 stock assessment.

Extending the time-series

In order to provide for transparent comparisons from this preliminary stock assessment through the final results for 2025, the initial step in this analysis was to extend the modelled time-series to 2025, using the projected mortality associated with the limits set by the IPHC (IPHC 2025). Weight-at-age was assumed to remain constant from 2024 to 2025; however, it will be updated prior to the final 2025 sock assessment when the new data become available. No other information was needed for this single year projection and all model results and parameter estimates remained unchanged relative to the final 2024 stock assessment.

Software version update

The Pacific halibut stock assessment has updated to newer versions of the stock synthesis software (Methot and Wetzel 2013a; Methot and Wetzel 2013b) as new features have been added, and in order to avoid major changes as input/output changes have evolved over time. The 2024 stock assessment was implemented in version 3.30.22.1 (Methot et al. 2024), which was updated to 3.30.23.1 (Methot Jr 2024) for the 2025 stock assessment. The results were unaffected as there were no changes made that were related to any of the features used for the analyses of Pacific halibut; therefore, for simplicity, this step has been omitted from the bridging figures below.

PDO index

As described above, a revised PDO index using consistent methods for an extended time-series (1854-present) is now available. In order to compare how this new series explained the historical recruitment both the effect size (% difference in average recruitment over positive and negative regimes) and the Standard Deviation (SD) of the recruitment deviations were summarized for the previous index and the updated index. A lower SD implies less residual variability in recruitment and conversely more of the process explained by the underlying stock-recruitment curve and the environmental effect.

The effects of updating the PDO index differed between the CW long and AAF long models. For the coastwide long model, the updated index resulted in a larger regime effect (62% vs 59% higher average recruitment during a positive regime) and a slightly lower SD of the estimate recruitment deviations (0.364 vs 0.375; Table 17). For the AAF long model the updated index resulted in a slightly lower effect size (50% vs 53%) and no change in the SD of estimated recruitment deviations (Table 17). The estimated historical time-series of spawning biomass was adjusted to better align with the revised regime definitions in both models (Figure 10-11). Over the most recent portion of the time-series the CW long estimated spawning biomass was scaled downward, while the AAF long estimated spawning biomass was virtually unchanged (Figure 11-12). In aggregate, there was no strong support for remaining with the previous index and therefore the index was updated for the preliminary 2025 stock assessment. Further evaluation of the treatment PDO is provided as part of the sensitivity analyses described below.

Table 17. Comparison of effect size and SD of the recruitment deviations for the CW and AAF long models for the previous and updated PDO indices.

	Positive regime effect (increase)		SD of recruitment deviations	
Model	AAF long	CW long	AAF long	CW long
Previous index	53%	59%	0.322	0.375
Updated index	50%	62%	0.322	0.364

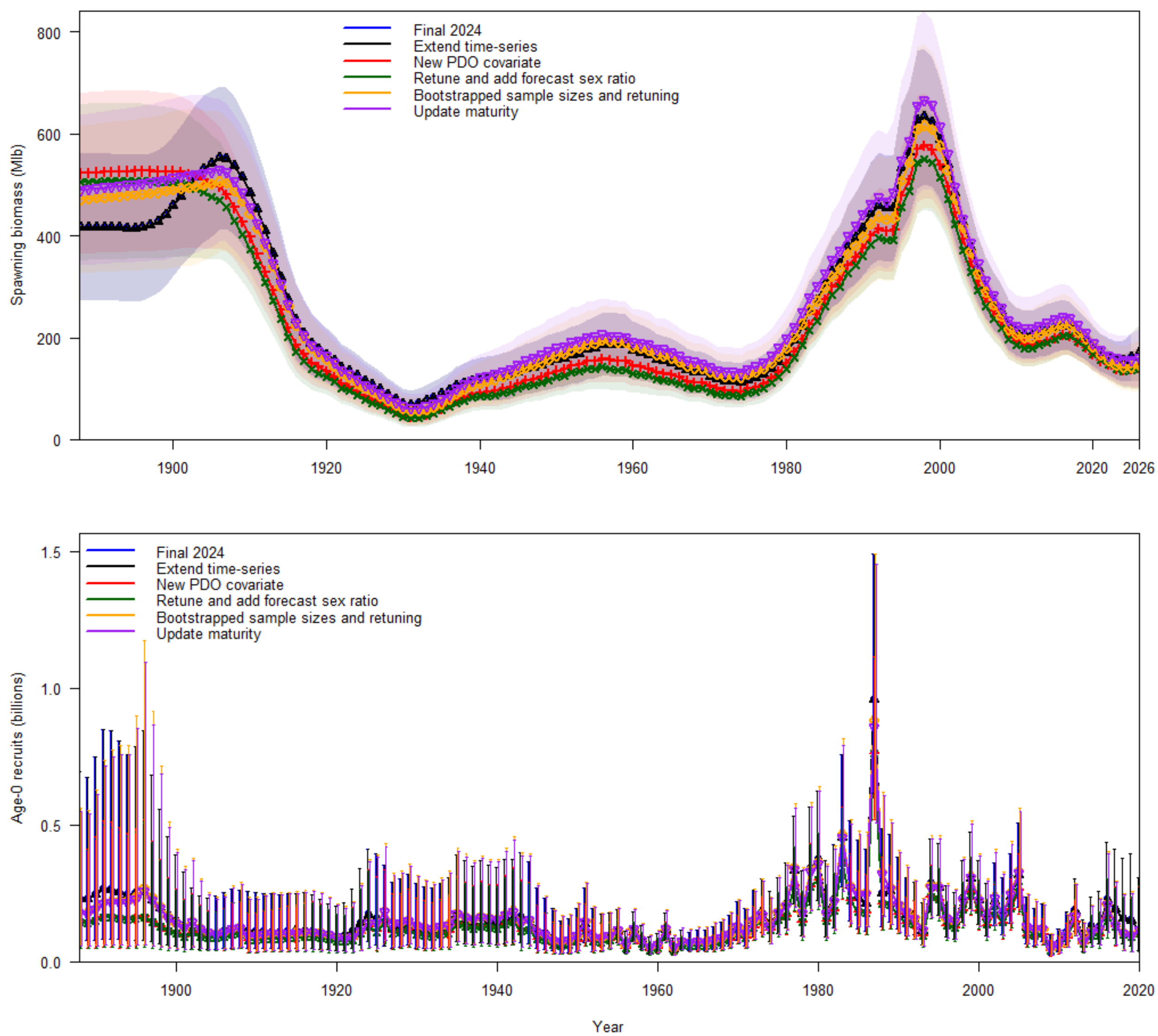


Figure 10. Comparison of estimated spawning biomass (upper panel) and recruitment time series (lower panel) over sequential changes from the 2024 to preliminary 2025 coastwide long models.

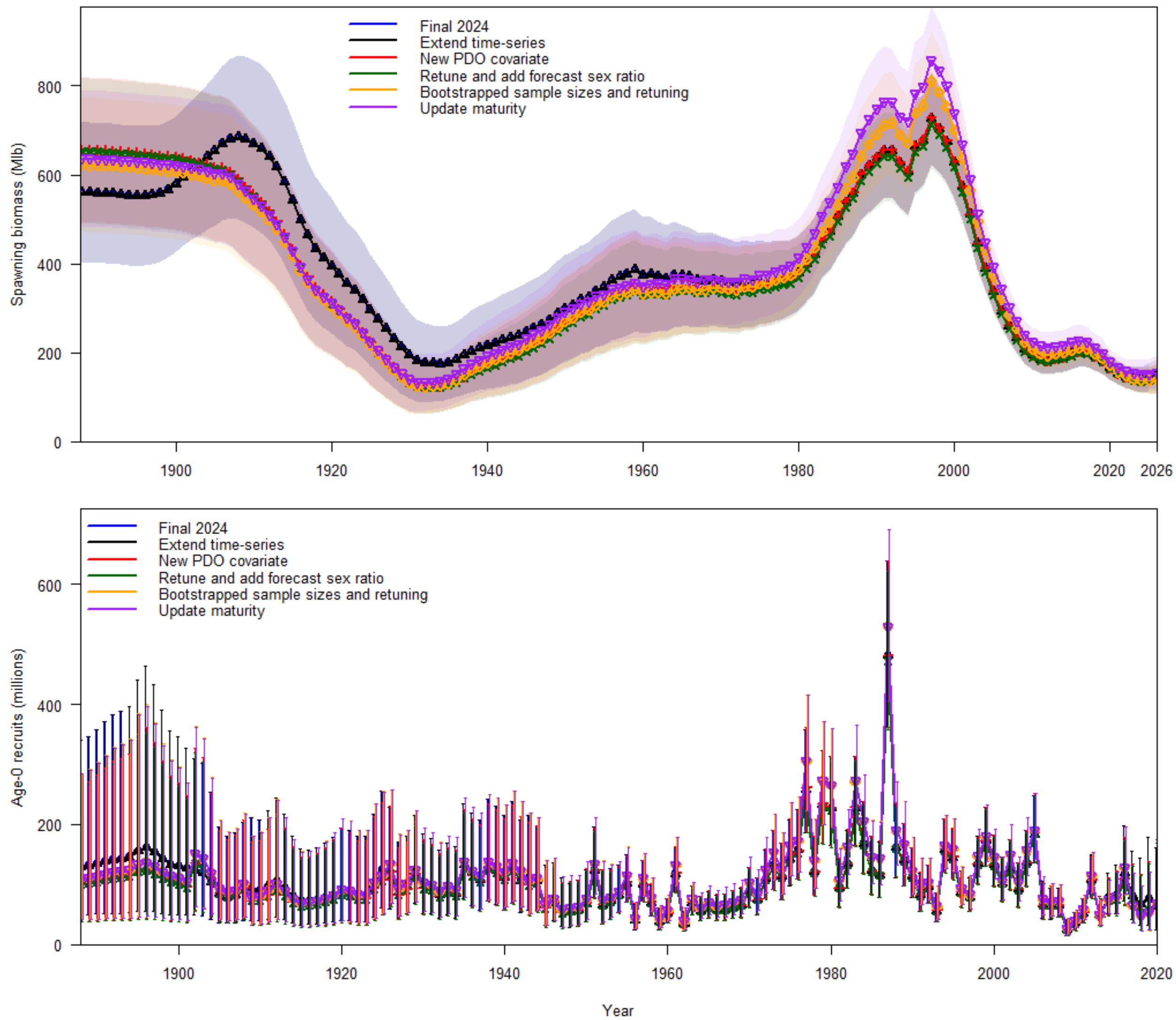


Figure 11. Comparison of estimated spawning biomass (upper panel) and recruitment time series (lower panel) over sequential changes from the 2024 to preliminary 2025 AAF long models.

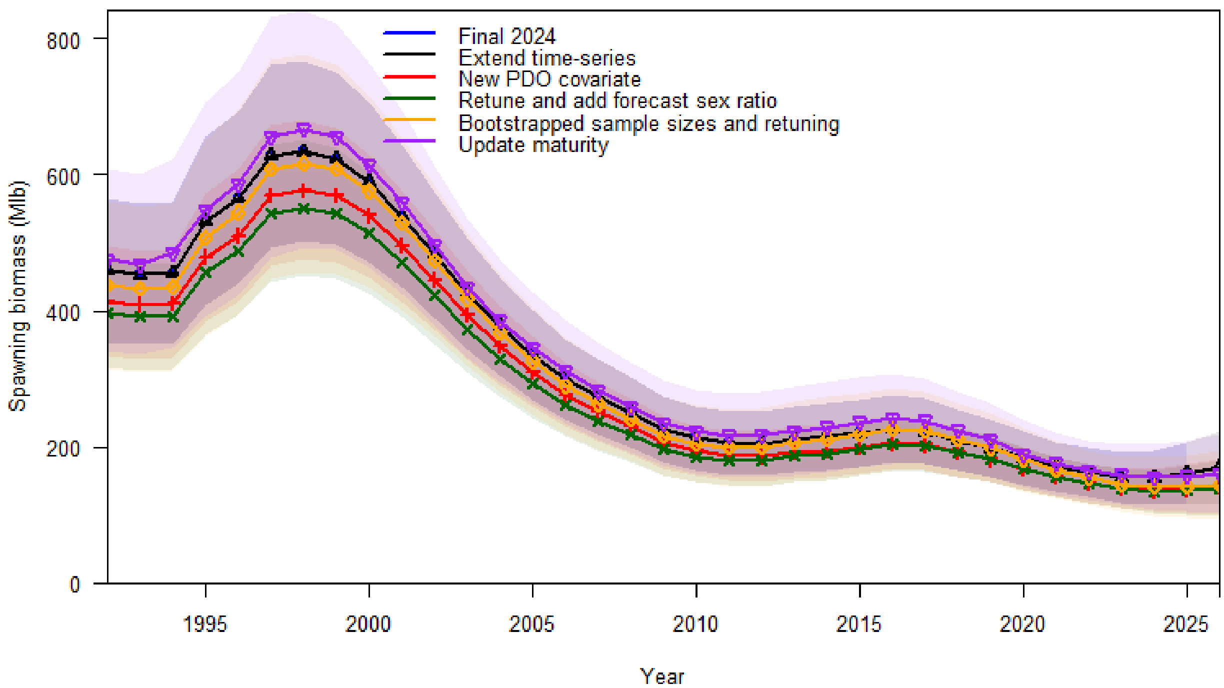


Figure 12. Comparison of recent estimated spawning biomass (1992-2026) over sequential changes from the 2024 to preliminary 2025 coastwide long models.

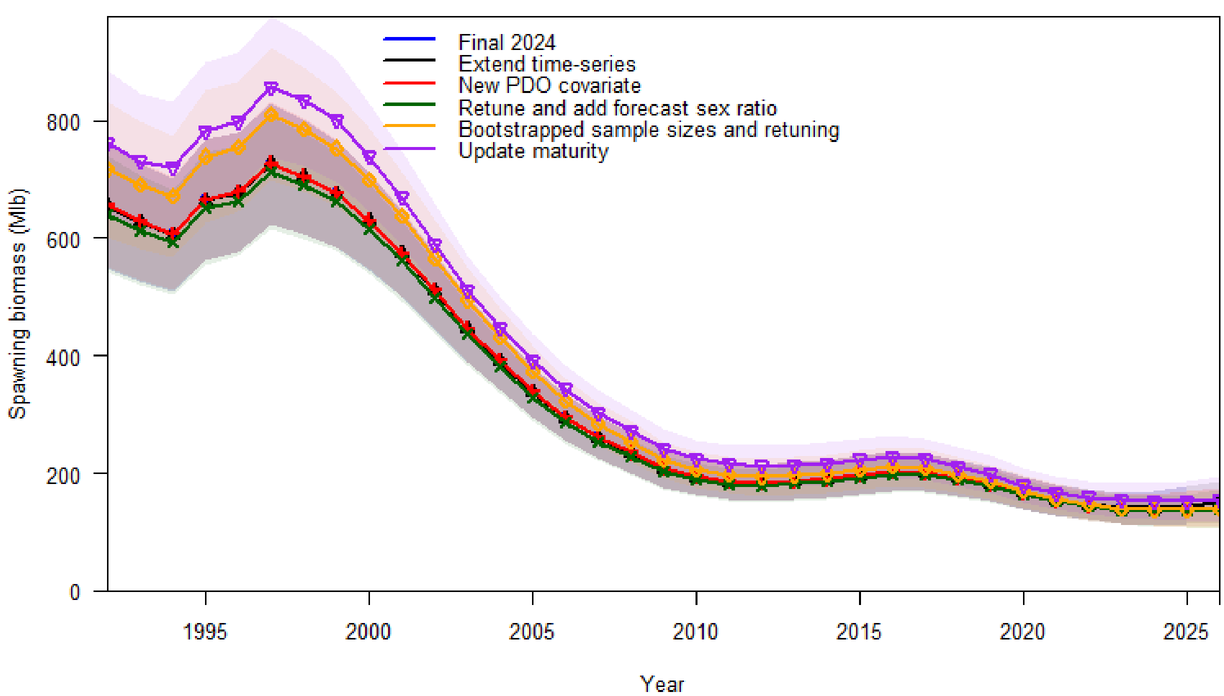


Figure 13. Comparison of recent estimated spawning biomass (1992-2026) over sequential changes from the 2024 to preliminary 2025 AAF long models.

Treatment of male selectivity

The next step in the bridging analysis was to revisit the treatment of time-varying male selectivity scale relative to female selectivity. In the offset approach used to define male selectivity, the scale (asymptote) of male selectivity is parameterized as a random walk beginning in 2018. This approach was first implemented in the 2019 stock assessment, and extended in the 2022 stock assessment as the sex-specific age composition data grew from two years to four years. However, the sigma constraining the random walk was still poorly informed in 2022 and the initial value used in 2019 (0.02, for a parameter that can logically vary from 0.0 to 1.0) had not been iteratively tuned. Further, the time-varying deviations in fishery selectivity were extended into the forecast period as part of the 2024 stock assessment (Stewart and Hicks 2025), in order to propagate the variance associated with unknown future selectivity but the male scale parameters were not included in this extension. For the preliminary 2025 stock assessment the deviations were extended to the full forecast period for consistency with all other modelled deviations (recruitment, fishery catchability, selectivity) and the sigmas iteratively retuned to be consistent with the estimated variability. The results of this change were negligible for all four models (Figure 10-15).

Updated sample sizes and data weighting

The next step in the bridging analysis was to replace the previously-used bootstrapped sample sizes with the updated bootstrapped maximum effective sample sizes (including ageing imprecision) described above. The effective sample sizes and process deviations were then iteratively retuned (as described above) to regain model internal consistency. Due to the substantial downweighting of the historical age composition information relative to the more recent data (as described above) this bridging step had a relatively large effect on model estimates of spawning biomass over the historical period for all four models (Figure 10-15). For the two long time-series models there was little change over the most recent spawning biomass estimates (Figure 12-13); however, the two short time-series models both estimated a larger recent spawning biomass for this bridging step, apparently largely as a result of increased estimates of the 2012 and 2016 year-classes (Figure 14-15).

Updated maturity ogive

The final step in the bridging analysis was to replace the historical maturity ogive with the newly estimated relationship described above and in *IPHC-2025-SRB026-06*. As is expected, shifting the maturity ogive toward younger fish results in a larger estimate of spawning biomass across all four models and across the entire time-series (Figure 10-15). Despite the upward scaling of this new information there was little effect on the trends. This is because the updated maturity is not treated as a time-varying process (a single estimate is applied to the entire time-series) and the only feedback to modelled dynamics is through the stock-recruitment function which estimated quite variable recruitment deviations.

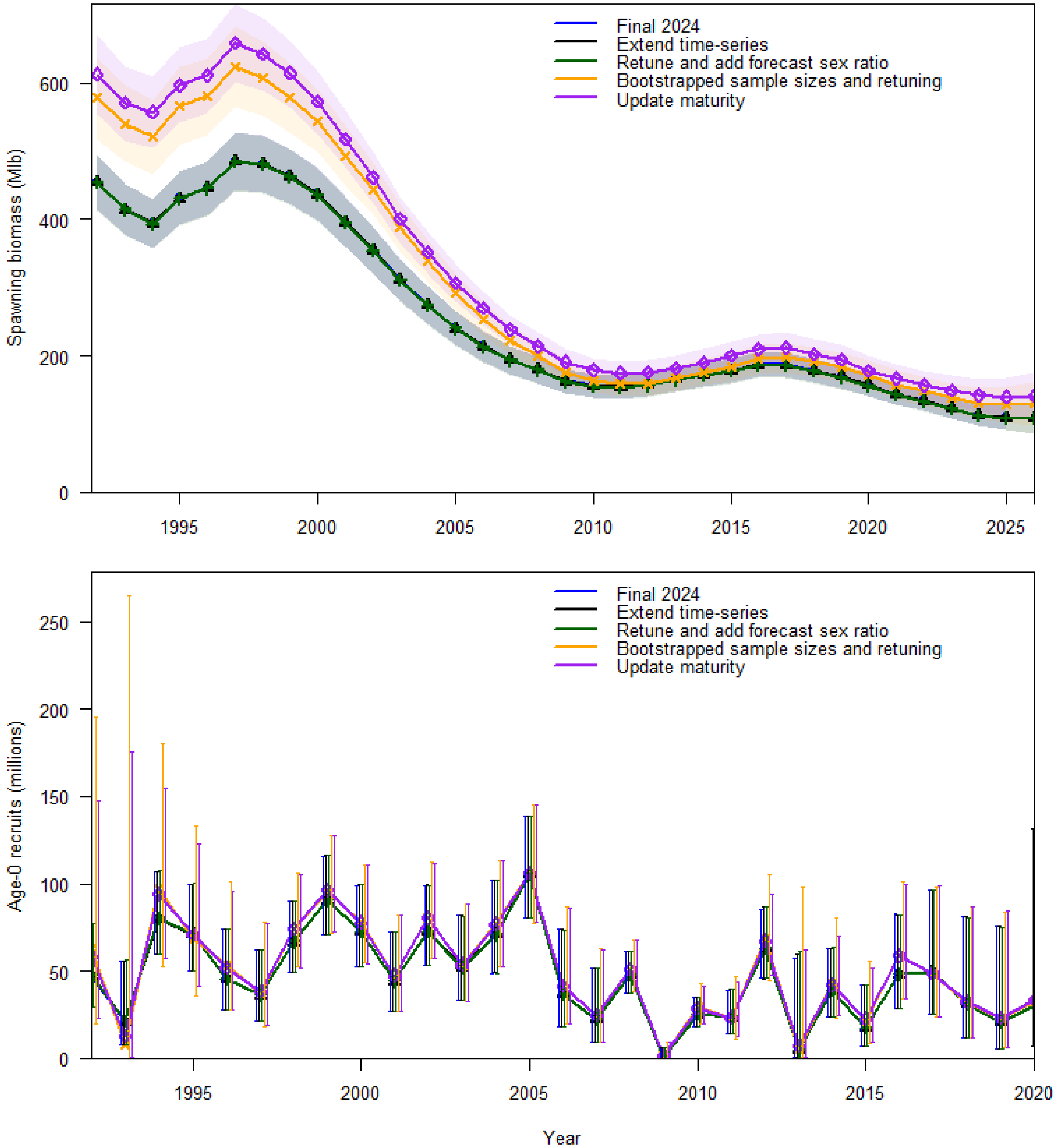


Figure 14. Comparison of estimated spawning biomass (upper panel) and recruitment time series (lower panel) over sequential changes from the 2024 to preliminary 2025 coastwide short models.

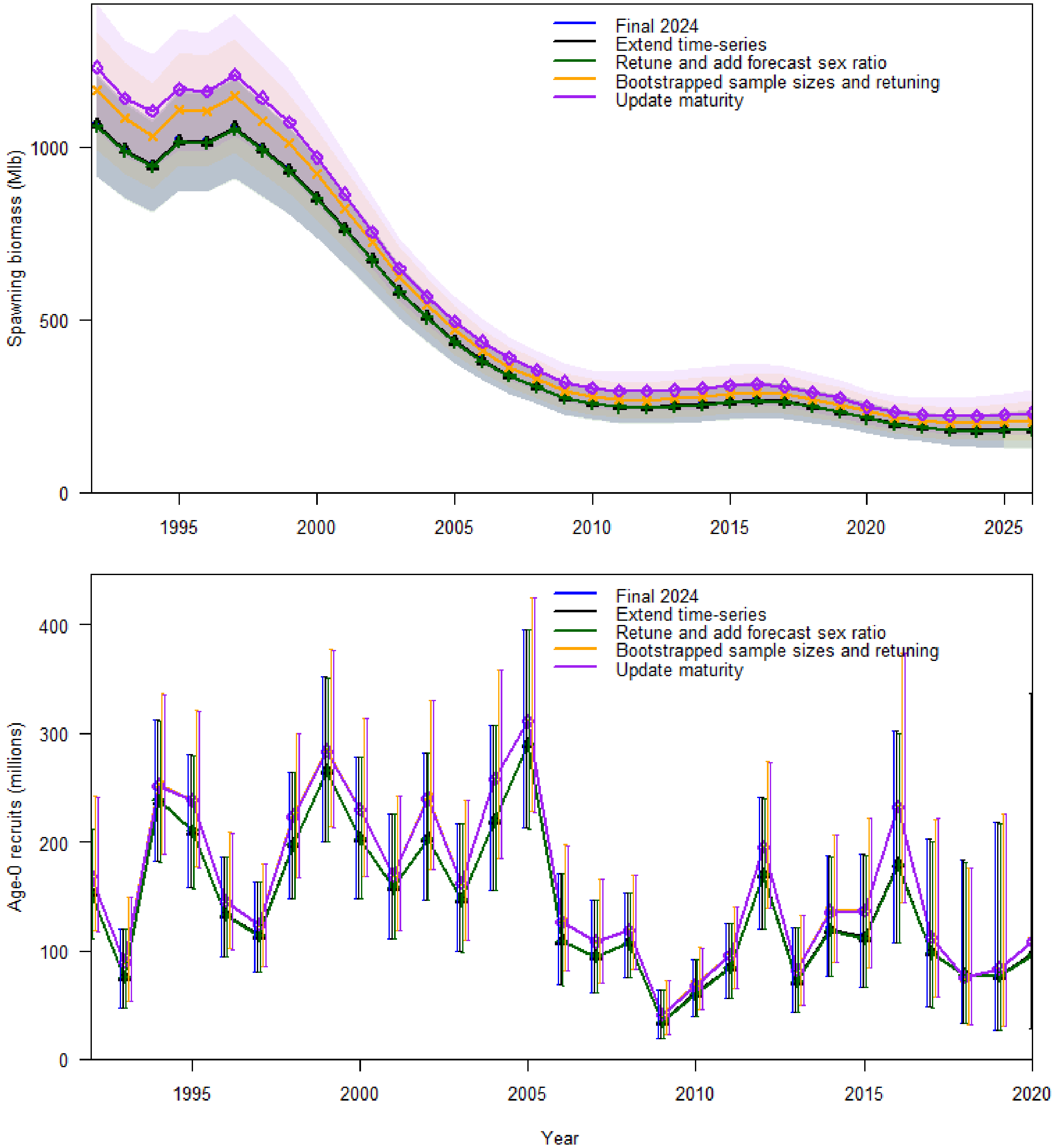


Figure 15. Comparison of estimated spawning biomass (upper panel) and recruitment time series (lower panel) over sequential changes from the 2024 to preliminary 2025 AAF short models.

Convergence criteria

Standard tools for monitoring convergence criteria include assessing the maximum gradient component, sensitivity to alternative phasing and initial values, use of overdispersed starting points or 'jitter analyses', as well as likelihood profiles, and Bayesian integration.

Wherever parameters were hitting bounds either the bounds were adjusted (if biologically plausible) or the parameters were fixed. For example, the descending limb of the 4B commercial fishery in the AAF models was estimated to be at the bound of 1.0 (as has been the case for all recent assessments) and so was fixed at this value. This approach reduces the likelihood that variances calculations will be (undesirably) effected by parameters stuck to bounds but does require periodic revisitation to ensure that the signal for parameters hitting bounds remains, and that fixing those parameters does not have an appreciable effect on the maximum likelihood solution.

For this preliminary 2025 assessment, all individual models all had a maximum gradient component < 0.002. A series of preliminary and intermediate runs did not indicate any signs that the estimates reported here represented local minima, nor did the models have difficulty producing a positive definite Hessian matrix under the range of alternative and sensitivity analyses (some presented in this document, but many used only for development). Both the AAF models did have trouble resolving the historical deviation parameters under some starting and phasing (the order in which the parameters were added to the minimization) configurations. In the stock synthesis framework random walk deviations cannot be estimated prior to the base selectivity parameters (and must start from a value of 0.0), therefore it is difficult to establish a general pattern of estimating the scaling parameters first and then adding less influential parameters later in the estimation phases. For this reason, the setup for these two models utilizes a parameter file, starting estimation for subsequent models at or near the solution from a previous run. Whenever a parameter file is used, it is important to periodically (and especially for the final model) rerun the model from dispersed starting points.

Convergence was explored for all four models specifically through a 'jitter' analysis perturbing all parameter values simultaneously and repeating minimization. A strong test using this method provides over-dispersed starting values such that the model is traversing a broad range of parameter space to ensure that the Maximum Likelihood Estimate (MLE) does not represent a local minimum that might be a poorer solution than another point in the likelihood space. For each of the four models 100 sets of dispersed starting points were used to initiate minimization for this analysis. Convergence to the MLE occurred for 46/100 for the CW long model, 26/100 for the CW short model, 65/100 for the AAF short model, and 50/100 for the AAF long model. None of the solutions resulted in a minimum that was better than the MLE. Although true convergence to a global minimum can never be proven, all convergence criteria indicate that the results of the preliminary 2025 assessment provide a robust solution.

Individual model diagnostics and results

This section provides more detail on the specific diagnostics and results of each of the four assessment models. It is not intended to provide the fit and residuals to every data component, but to summarize the basic performance of the model and specifically highlight areas of potential deficiency. Figures showing comprehensive diagnostics and results and the full report files, as output directly from stock synthesis, are provided electronically as described in Appendix A. Each model section finishes with a brief summary of the relative strengths and weaknesses of that model.

Coastwide short

Predictions of both the fishery and survey indices of abundance fit the observed data very well in the coastwide short model (Figure 16). Prior to the 2019 assessment, a small amount of process error was allowed on fishery catchability. In the 2019 analysis this process error was effectively zero and was turned off for that and the 2021-2024 stock assessments. Re-evaluation as part of this analysis via iterative tuning of the annual catchability deviations indicated a small value was again consistent with model fit and was therefore included. The predicted aggregate age distributions also matched the observed distributions well, for both the fishery and survey indicating that the selectivity parameterization was generally capturing differences in both the age-structure and the sex-ratio (Figure 17).

The coastwide short model tuning resulted in a higher weight on the coastwide FISS ages than for the commercial fishery age data (Table 18). The discard, non-directed discard and recreational age data were all intentionally heavily down-weighted (as described above) and so input sample sizes were not iterated to larger values, despite fits to the data that implied a higher weight. Fit to the annual FISS age compositions were generally good (Figure 18), although some patterning was visible in the standardized residuals (Figure 19). Specifically, there was a clear pattern of negative residuals in the plus group for male halibut; however, this was almost imperceptible in the fits themselves due to the very small observed and predicted values in this age bin. The fits to the annual fishery data were also acceptable, noting some patterning associated with the 1987 cohort and ages 15 and 20 in the most recent decade (Figure 20-21). The implied fit to the sex ratio information for the commercial fishery (Figure 22) was similar to that for the FISS (Figure 23); both show year-to-year variability in the scale and patterns. Additional diagnostics and diagnostic figures (such as fits to the down-weighted annual compositions for the discard, bycatch, and recreational fleets) are included in the background materials.

Neither the FISS nor the fishery female selectivity was estimated to have a highly variable ascending limb over the short time-series (Figure 24). The estimated fishery selectivity showed a small increase in the selection of males at the end of the time-series, somewhat the opposite of that estimated for the FISS (Figure 25), perhaps a function of the catch distribution shifting toward the Eastern side of the stock where fast-growing males are much more common. For the discard fleet, estimated selectivity included fewer and younger females than males (Figure 26). Estimated selectivity for the non-directed discards fleet showed a peak at ages 4-5 and a slightly domed relationship. Recreational/subsistence selectivity was shifted to the left of the commercial fishery discards (and therefore the FISS).

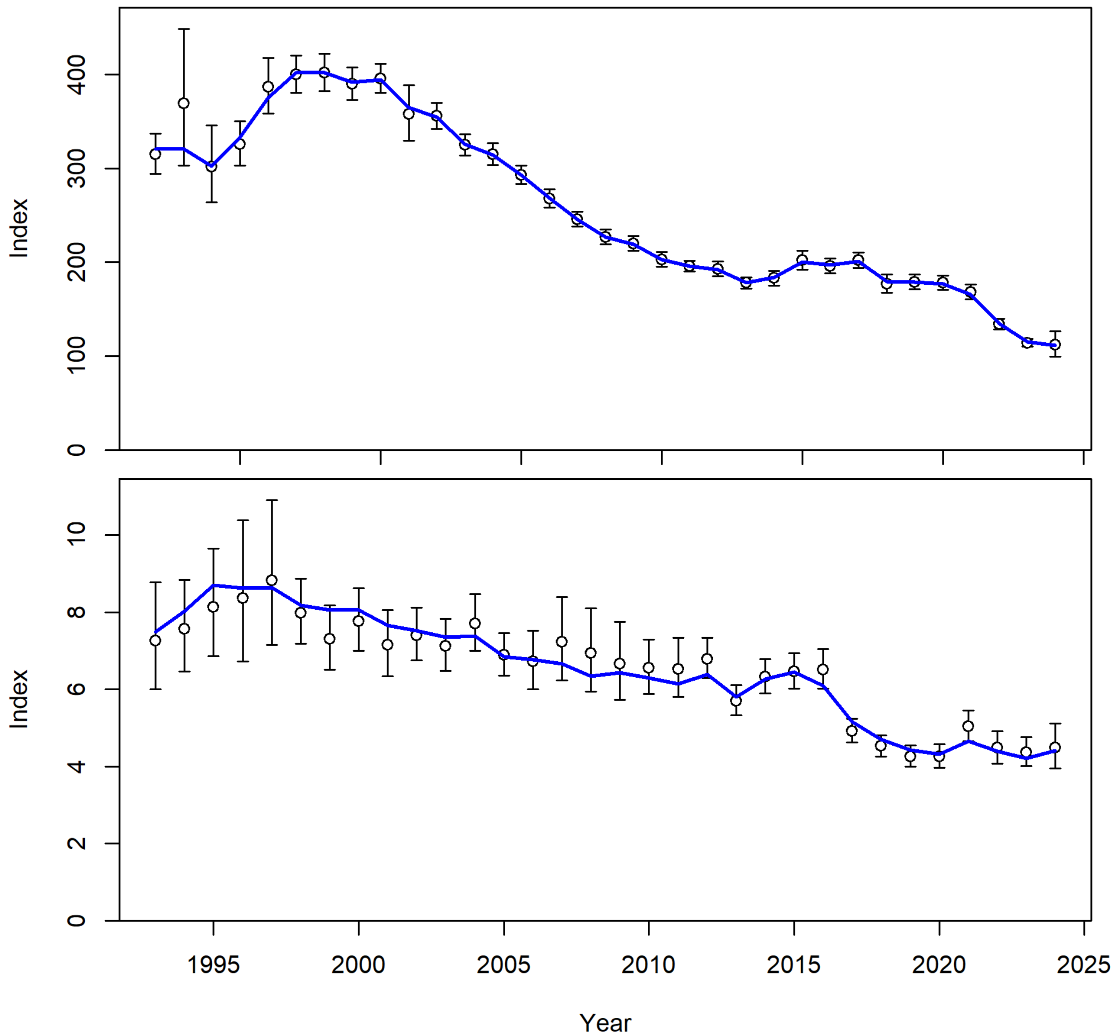


Figure 16. Fit to fishery (upper panel) and FISS (lower panel) indices of abundance in the coastwide short model; note that the scale of the y-axes differ as do the units (the fishery index is in weight and the FISS in numbers).

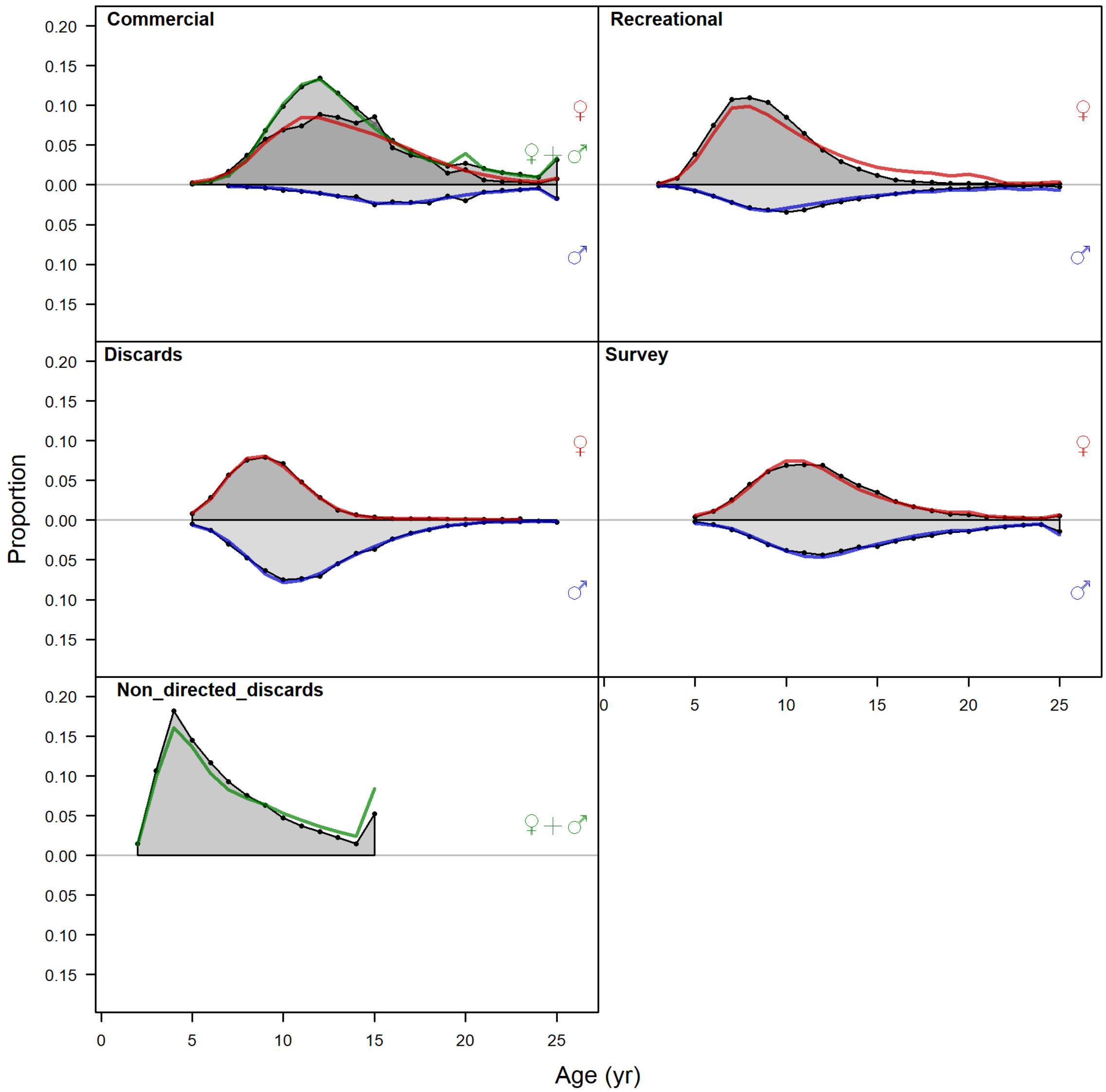


Figure 17. Aggregate fit to all age data by model fleet in the coastwide short model; sex-specific distributions for the commercial fishery represent only 2017-2023 and are plotted on top of sexes-aggregated distributions spanning 1992-2016 + 2024.

Table 18. Post-iteration sample size diagnostics for age-composition data by model and fleet. Average iterated input denotes the value used for model runs reported here, after iterating the bootstrapped starting points.

	Average iterated input	Harmonic mean effective	Francis weight effective	Maximum Pearson residual
Coastwide short				
Fishery	159	423	174	2.08
Discards[1]	6	221	116	0.67
Non-directed discards[1]	3	51	56	1.61
Recreational[1]	3	104	24	0.63
FISS	164	810	163	2.78
Coastwide long				
Fishery	144	318	148	3.01
Discards[1]	6	213	100	0.65
Non-directed discards[1]	3	38	7	1.30
Recreational[1]	3	131	20	0.61
FISS	97	208	97	4.18
AAF short				
Region 2 fishery	456	600	825	4.91
Region 3 fishery	599	609	733	4.10
Region 4 fishery	58	85	61	2.08
Region 4B fishery[2]	49	130	65	2.43
Discards[1]	6	198	80	0.66
Non-directed discards[1]	3	49	28	0.84
Recreational[1]	3	128	21	0.56
Region 2 FISS	7	77	5	1.06
Region 3 FISS	27	317	28	1.33
Region 4 FISS	86	156	90	2.68
Region 4B FISS[2]	41	147	38	2.56
AAF long				
Region 2 fishery	256	293	563	4.87
Region 3 fishery	319	286	468	3.93
Region 4 fishery	47	65	49	2.50
Region 4B fishery[2]	49	122	58	2.34
Discards[1]	6	157	85	1.22
Non-directed discards[1]	3	39	8	1.30
Recreational[1]	3	103	20	0.65
Region 2 FISS	5	63	5	1.35
Region 3 FISS	23	145	21	1.44
Region 4 FISS	114	158	121	2.80
Region 4B FISS[2]	34	147	35	2.00

[1]Inputs intentionally down-weighted – see text.
[2]Iterated sample size equal to maximum (bootstrapped input).

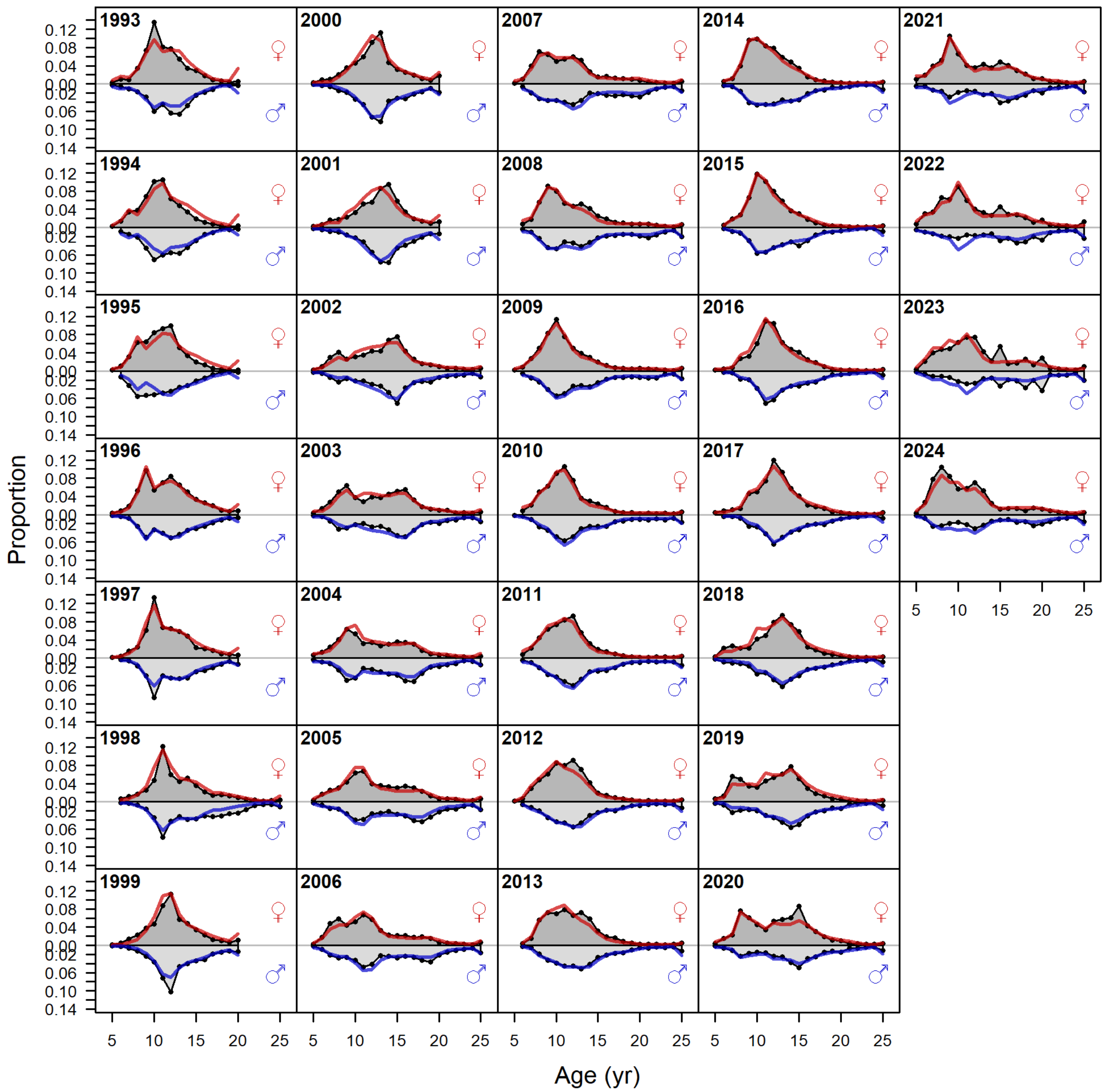


Figure 18. Fit to annual age data from the FISS survey in the coastwide short model.

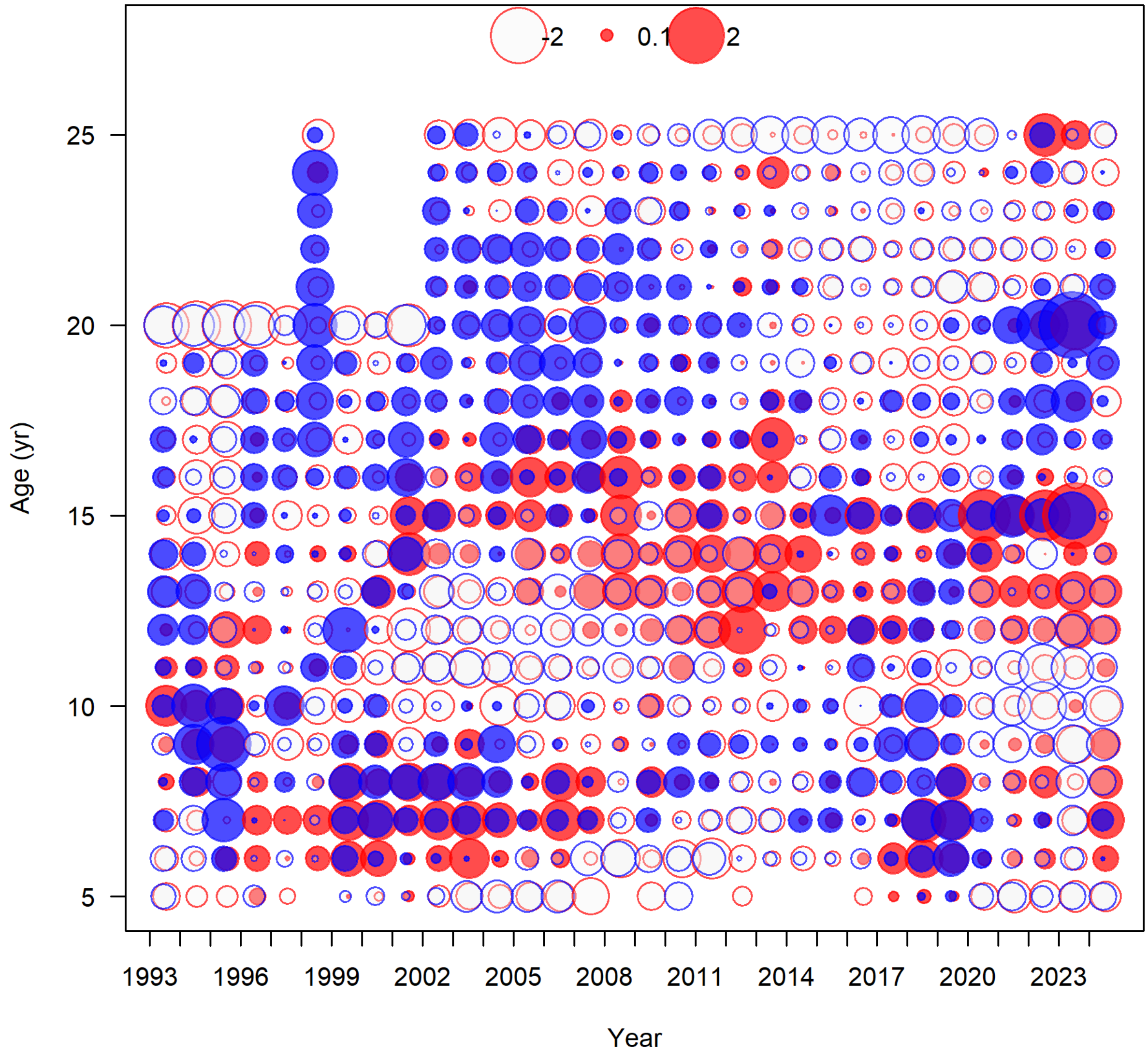


Figure 19. Pearson residuals for fit to annual age data from the FISS survey in the coastwide short model; red circles denote female residuals, and blue circles denote male residuals.

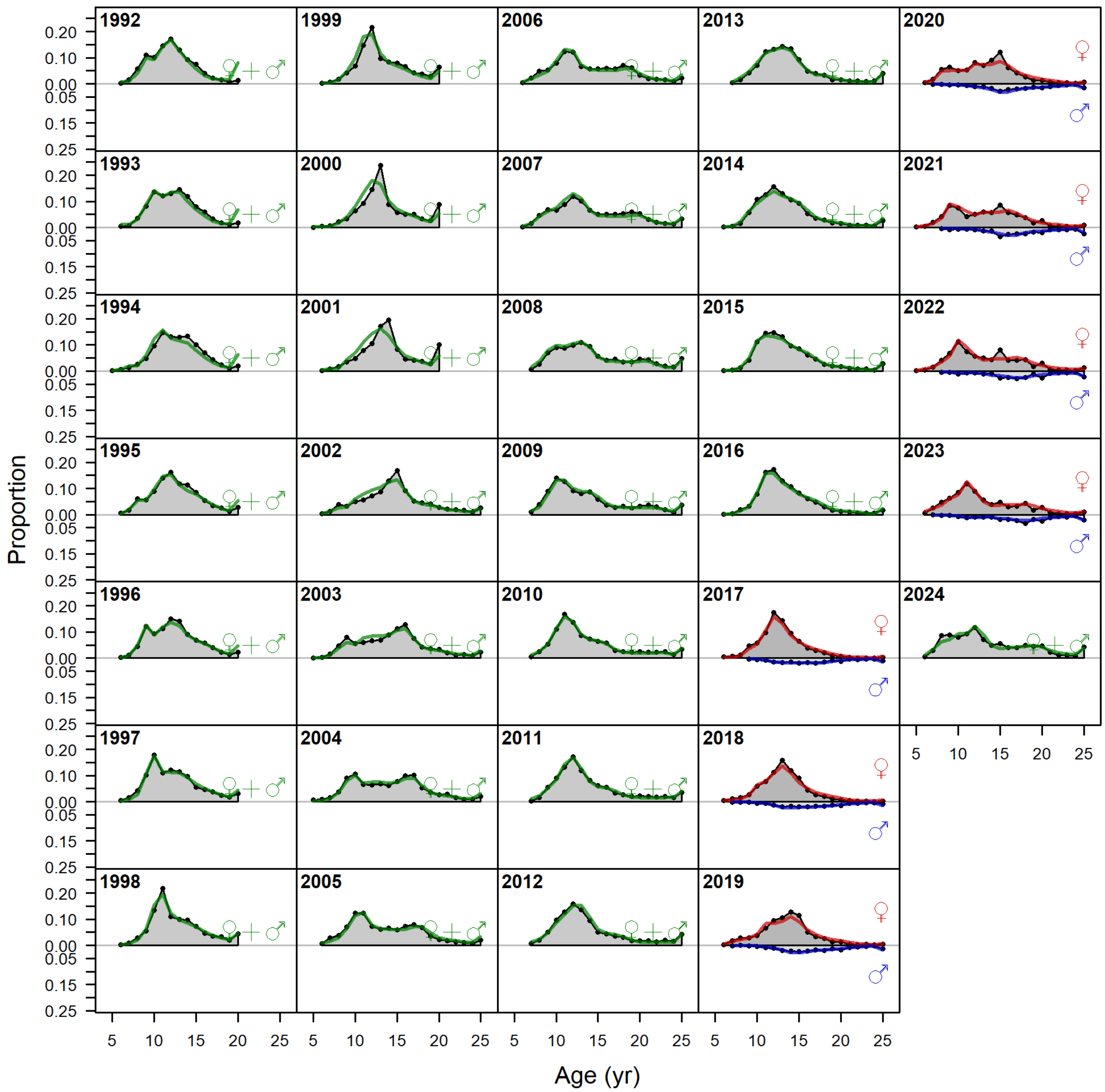


Figure 20. Fit to annual age data from the commercial fishery landings in the coastwide short model.

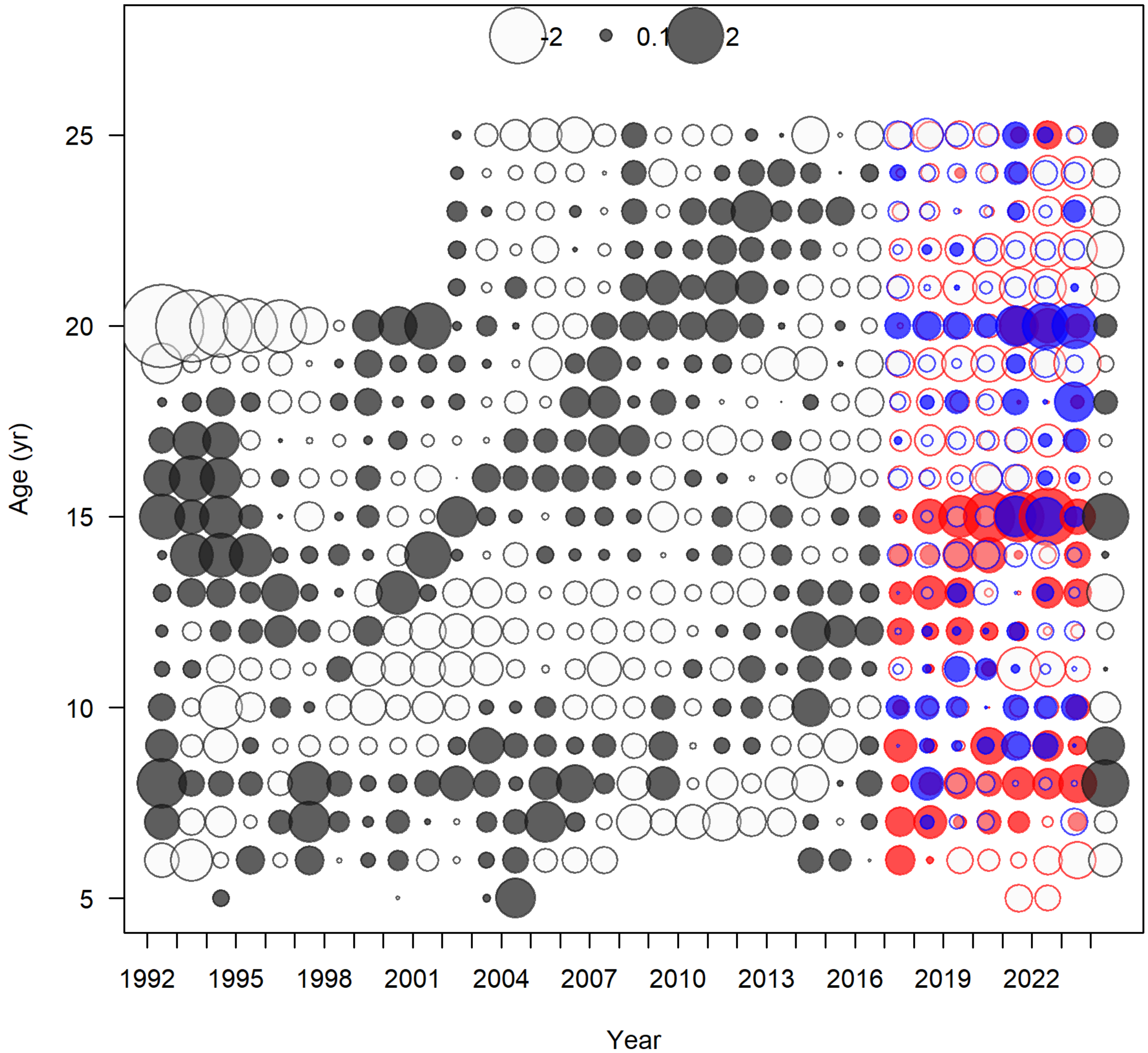


Figure 21. Pearson residuals for the fit to annual age data from the commercial fishery landings in the coastwide short model; grey circles denote unsexed residuals, red circles denote female residuals, and blue circles denote male residuals.

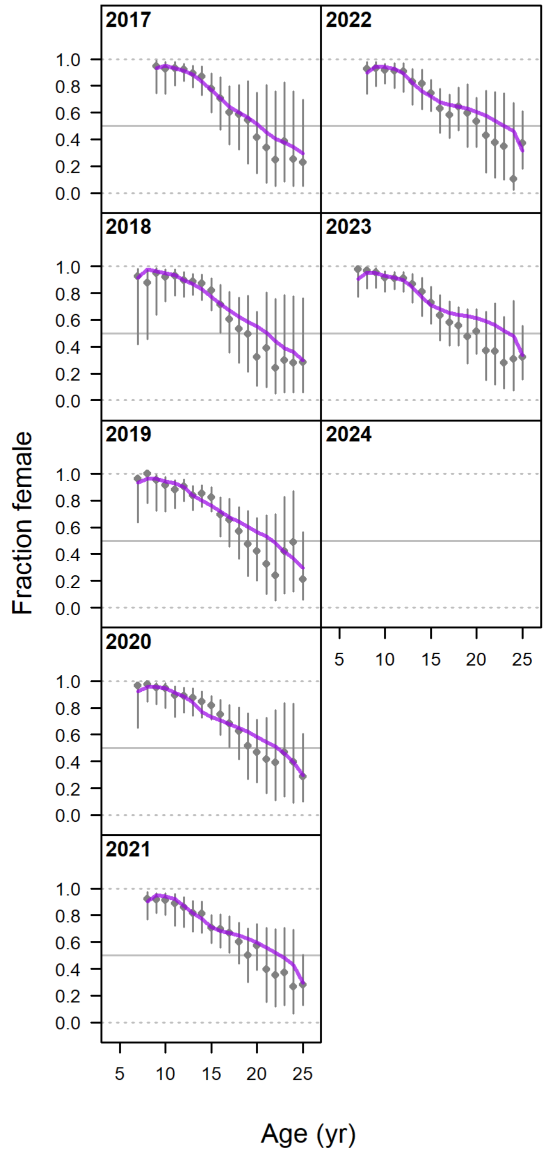


Figure 22. Observed and predicted sex-ratio in the commercial fishery landings from the coastwide short model for years with sex-specific age composition data (2017-2023).

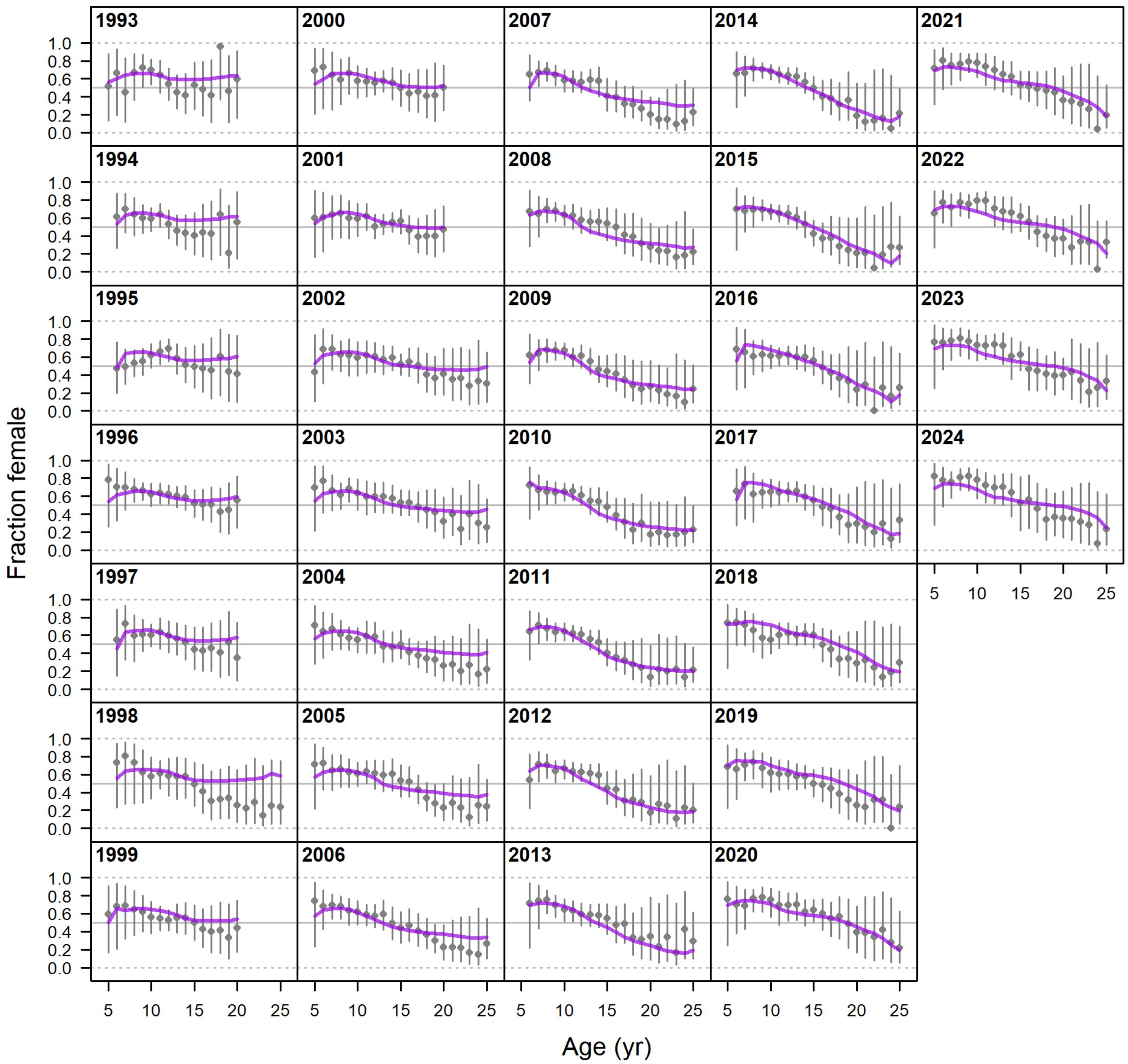


Figure 23. Observed and predicted sex-ratio in the FISS from the coastwide short model.

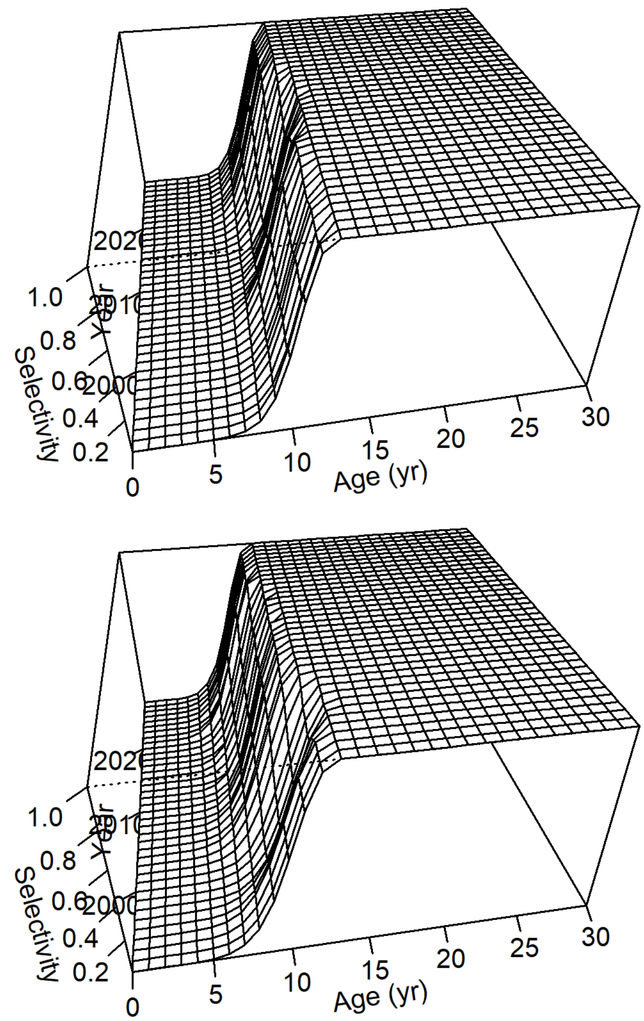


Figure 24. Estimated time-varying female selectivity curves for the commercial fishery landings (upper panel) and the FISS (lower panel).

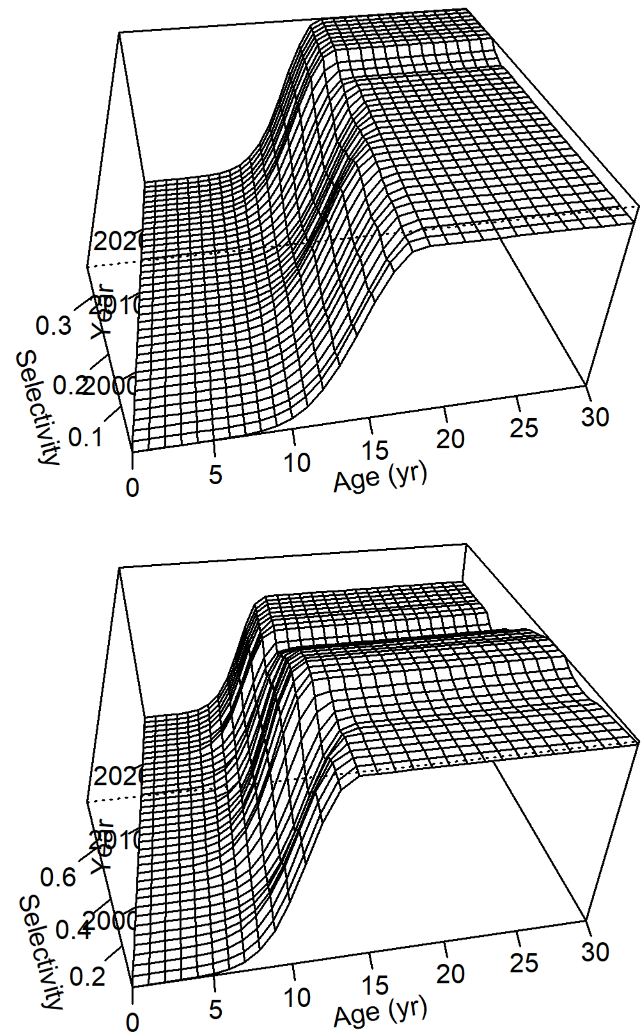


Figure 25. Estimated time-varying male selectivity curves for the commercial fishery landings (upper panel) and the FISS (lower panel).

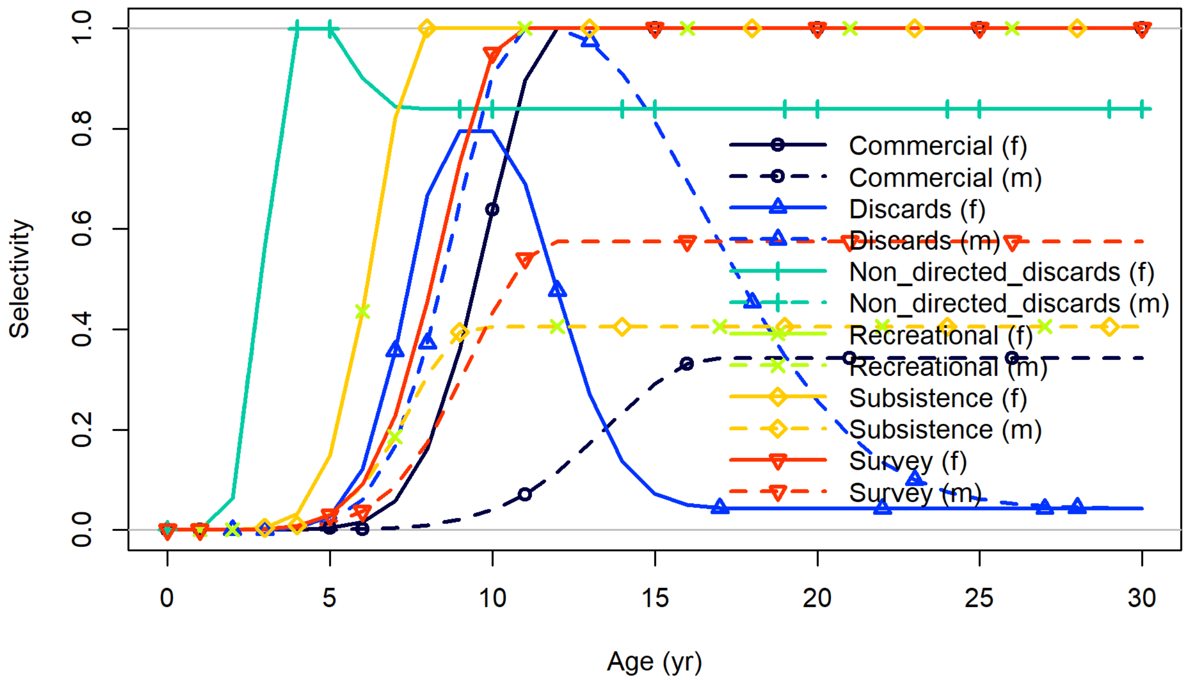


Figure 26. Estimated ending year selectivity curves by sex for the commercial fishery, discard, non-directed discard, recreational and FISS fleets in the coastwide short model.

Male *M* was estimated to be slightly higher (0.164) than the fixed value assumed for females of 0.15 (Table 19); this represented a slight increase from the value estimated in the 2022 and earlier assessments. The large negative estimated initial recruitment offset is consistent with the start year occurring after a very long time-series of fishing. The lower *M* fixed in the coastwide short model corresponded to lower recruitment and female spawning biomass estimates (Table 19) than the other three models, as has been the case for all recent assessments.

Summary of strengths and weaknesses for the coastwide short model:

Strengths:

- Lowest technical overhead (complexity) of the four models in the ensemble
- Fit the fishery and FISS indices very well
- Fit the survey age data (males and females) relatively well
- Parameter estimates are derived from the most recent time-period
- Internally consistent data weighting
- Similar weighting of commercial fishery and FISS age composition data

Weaknesses:

- Basis for fixed female *M* is unclear
- Does not include uncertainty in female *M* (see sensitivity analyses below)
- Does not include extensive historical data

- May lose Region-specific trends and biological patterns due to aggregation
- Does not use environmental information to inform recruitment

Table 19. Select parameter estimates (maximum likelihood value and approximate 95% confidence interval) and important recent population estimates by model and Biological Region (where applicable).

	Model			
	Coastwide Short	Coastwide Long	AAF Short	AAF Long
Biological				
Female M	0.150 (*Fixed*)	0.221 (0.185-0.257)	0.220 (0.204-0.236)	0.186 (0.169-0.204)
Male M	0.164 (0.155-0.172)	0.198 (0.181-0.216)	0.179 (0.169-0.189)	0.163 (0.154-0.171)
Log(R_0)	11.43 (11.19-11.67)	11.91 (11.51-12.32)	12.30 (12.06-12.54)	11.56 (11.32-11.79)
Initial log(R_0) offset	-1.512 (-1.746--1.278)	NA	-0.193 (-0.411-0.019)	NA
Environmental Link (β)	NA	0.456 (0.238-0.675)	NA	0.430 (0.225-0.636)
Survey Log(q) Δ1984 (transition to circle hooks)	NA	0.933 (0.485-1.381)	NA	R2: 1.344 (0.756-1.513) R3: 1.876 (1.631-2.120)
Fishery Log(q) Δ1984	NA	0.823 (0.647-0.999)	NA	R2: 0.562 (0.373-751) R3: 0.942 (0.751-1.133) R4: 0.850 (0.645-1.055) R4B: 0.381 (0.187-0.575)
2012 Age-0 recruitment (Millions)	67 (48-94)	164 (96-282)	195 (139-273)	115 (86-153)
2025 SB (Million lb)	139 (111-167)	156 (105-208)	226 (165-287)	153 (119-187)

Coastwide long

Both the fishery and FISS indices of abundance were fit well (Figure 27), with breaks in catchability to accommodate the change from “J” to circle hooks (1984) which were very large in both series (Table 19). In aggregate, the predicted age compositions matched the observed data well (Figure 28); however, there were notable differences among years within the time-series. Fits to the FISS were quite poor in the early portion of the time series when the spatial coverage was very limited (Figure 29), but improved where the data became more spatially comprehensive in the mid-1990s, and quite good in the most recent years (Figure 30). Fishery data fit reasonably well for the entire time-series (Figure 31-32), with patterns in the residuals corresponding to relatively small differences with observed distributions. The small contribution of males to the fishery landed catch is quite clear from the seven years that have sex-specific information (Figure 32). Harmonic mean effective sample sizes were much larger than adjusted inputs when Francis weights were close to 1.0; commercial fishery data were weighted slightly more heavily than FISS data, largely reflecting the spatial coverage of the early FISS years (Table 18).

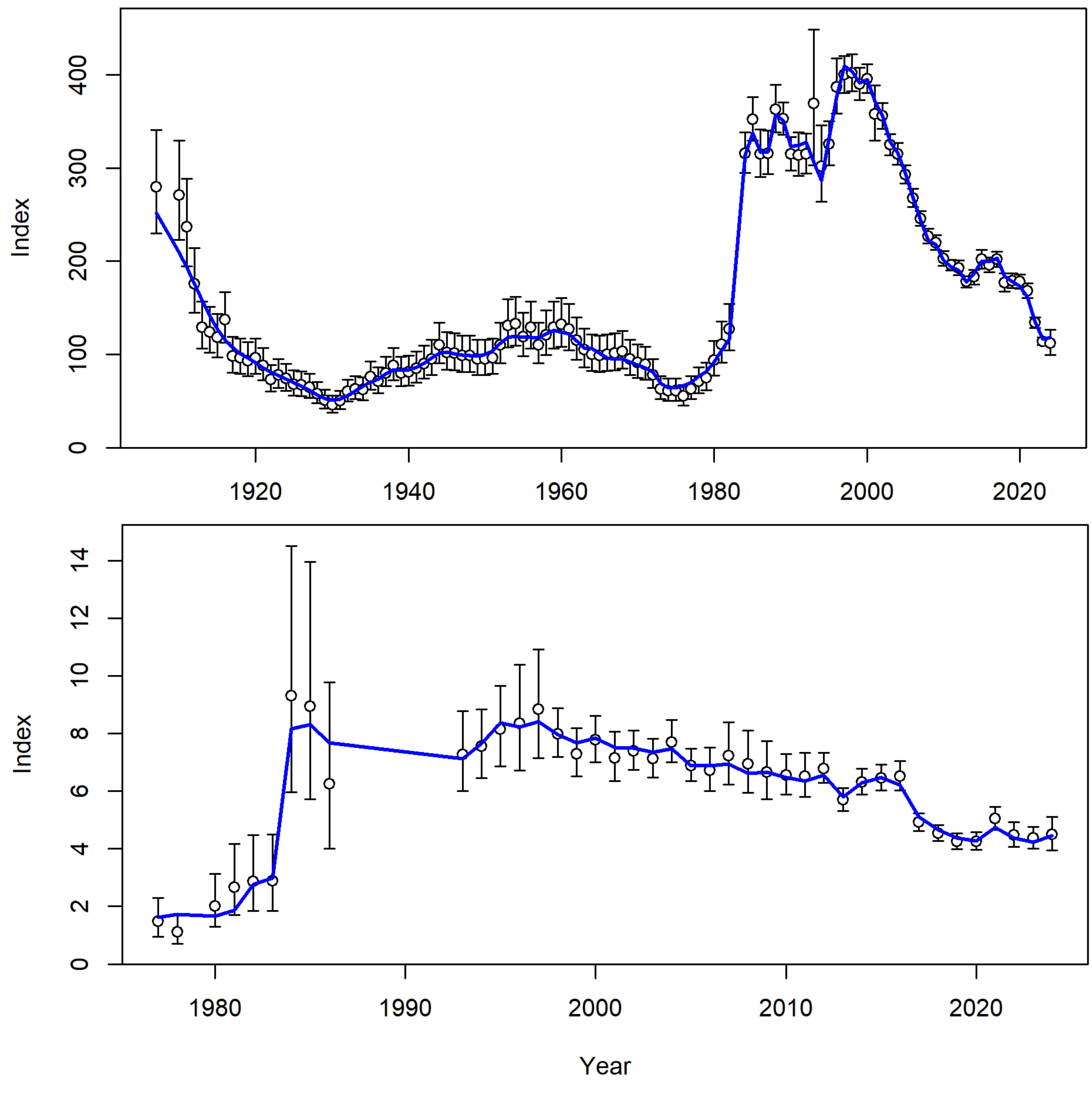


Figure 27. Fit to fishery (upper panel) and FISS (lower panel) indices in the coastwide long model.

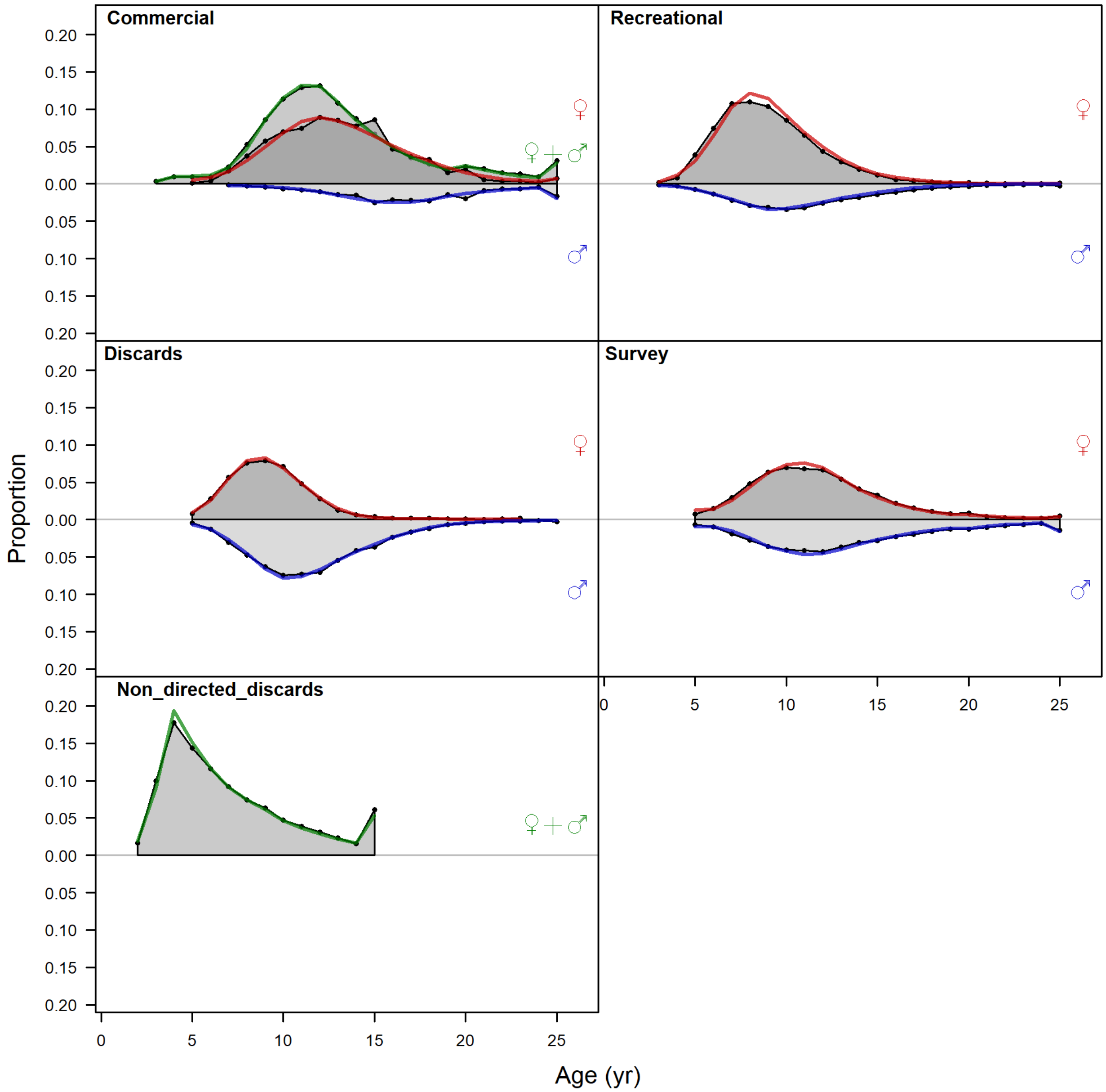


Figure 28. Aggregate fit to all age data by model fleet in the coastwide long model..

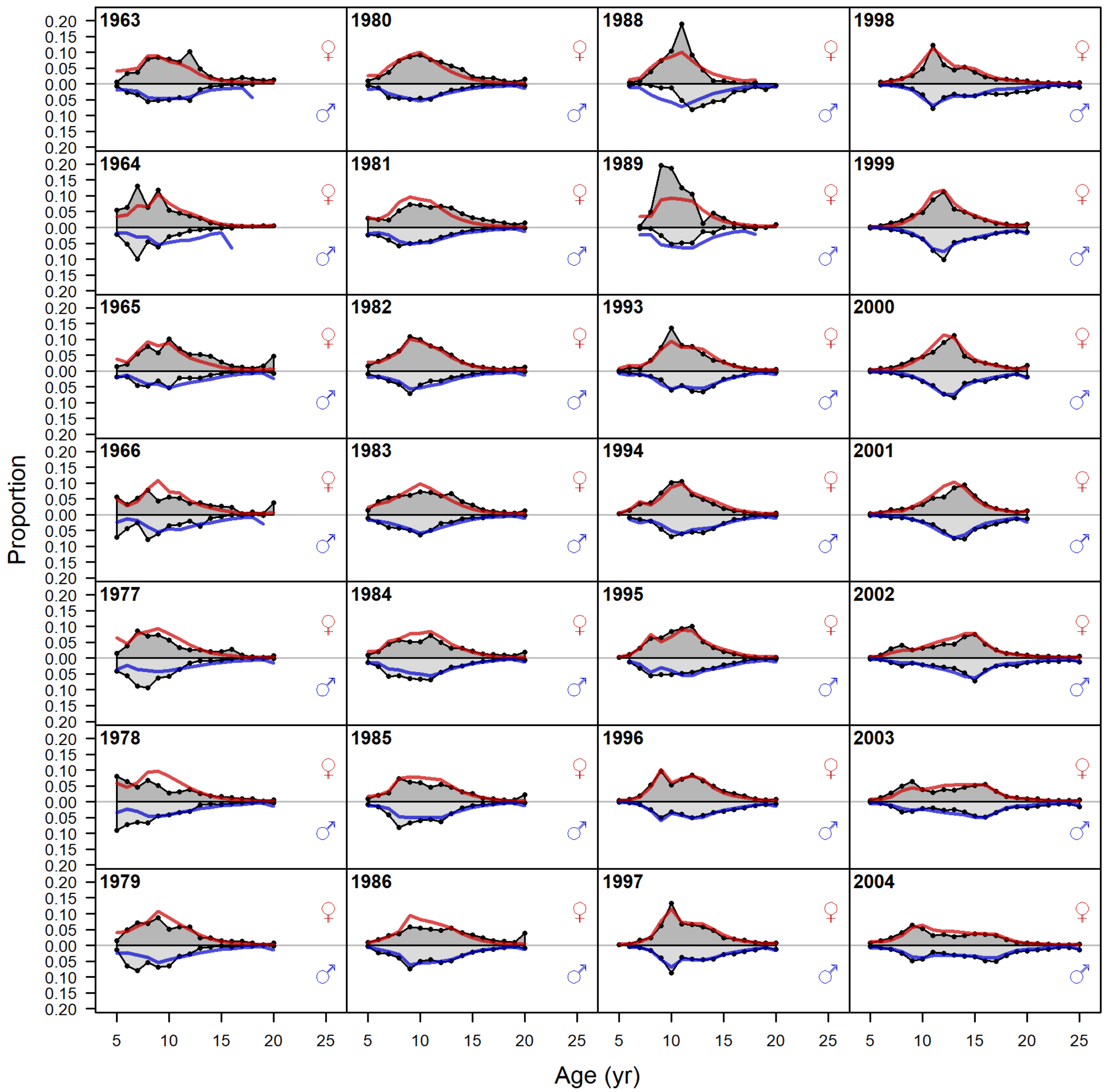


Figure 29. Fit to early years of FISS age data in the coastwide long model.

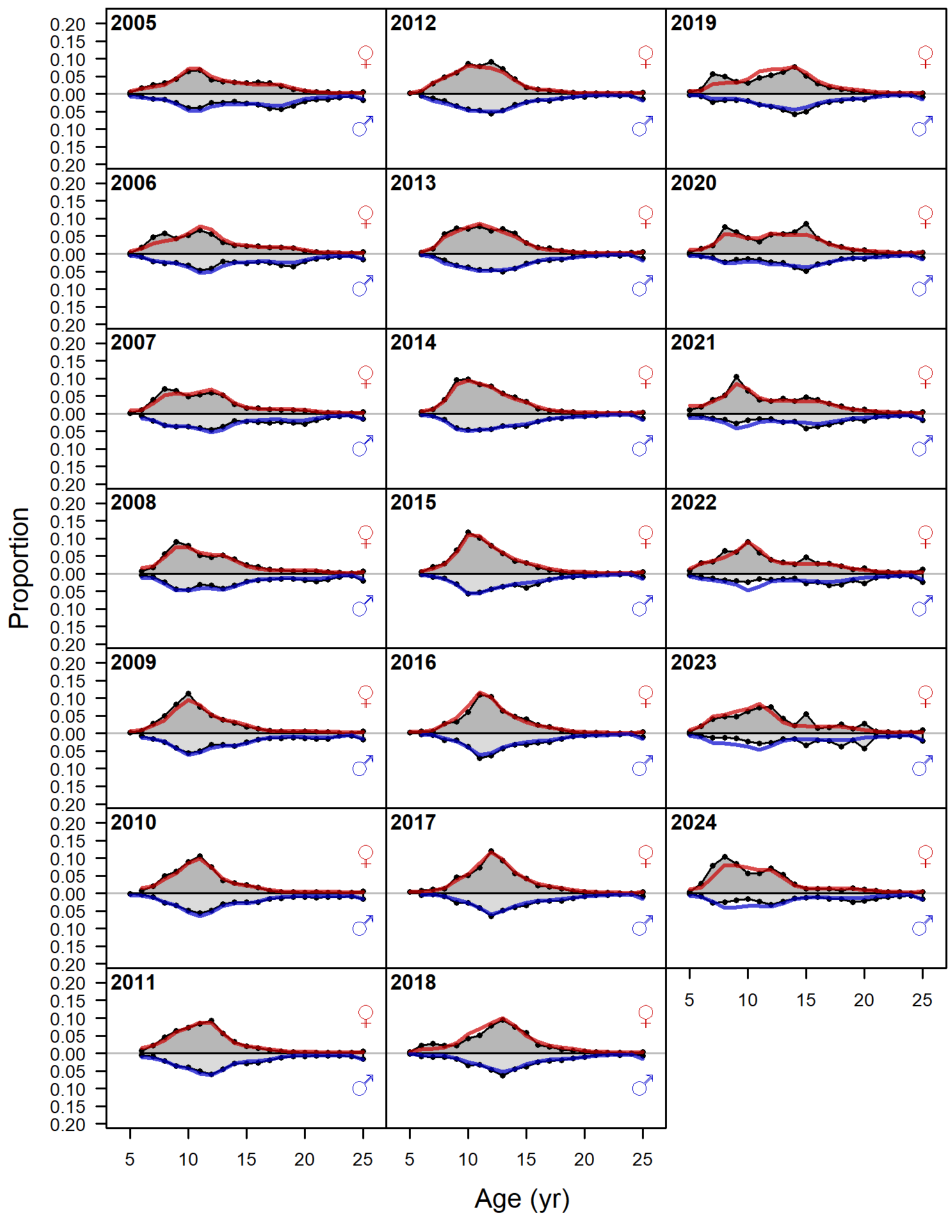


Figure 30. Fit to later years of FISS age data in the coastwide long model.

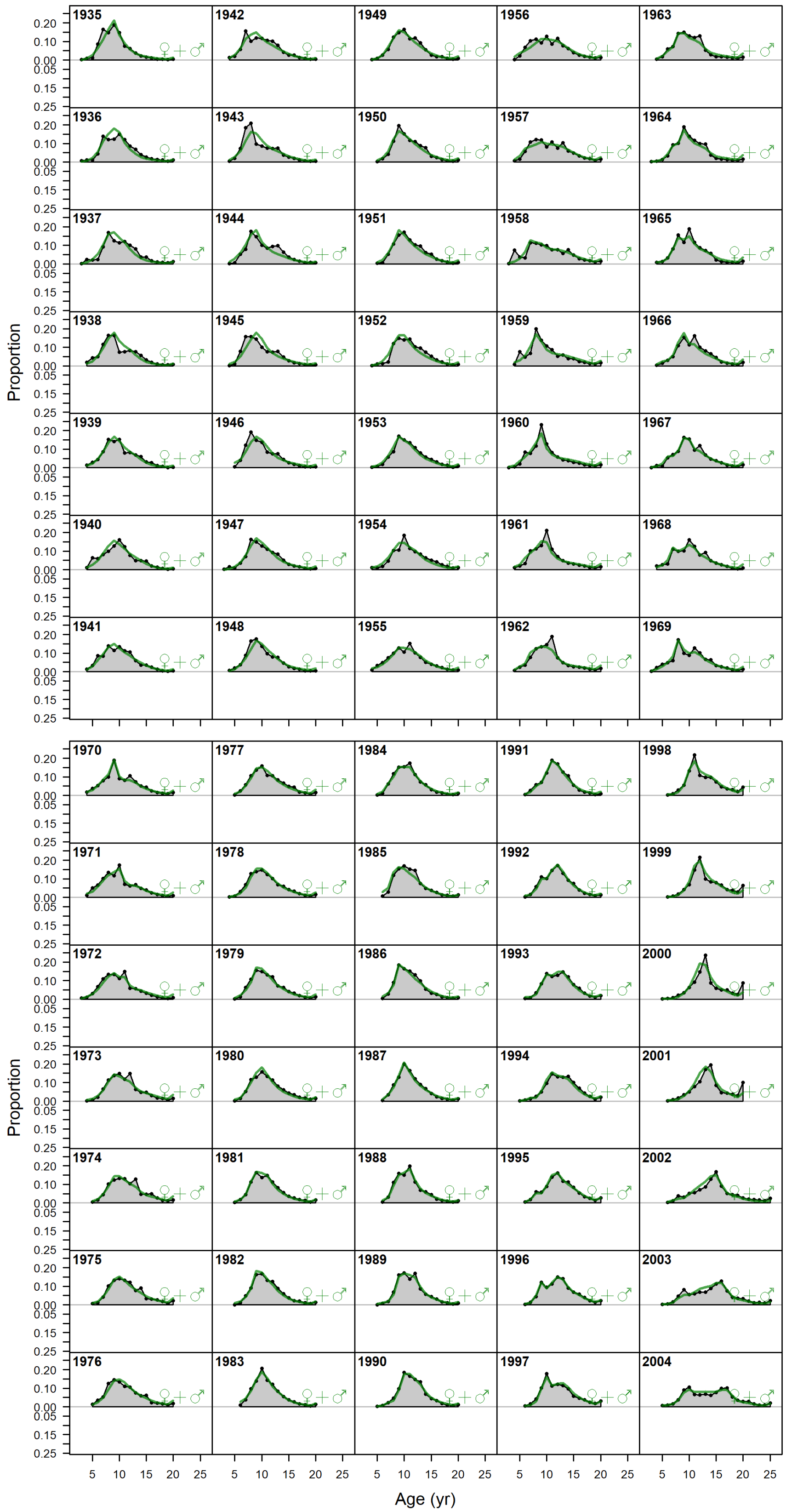


Figure 31. Fit to early years of fishery age data in the coastwide long model.

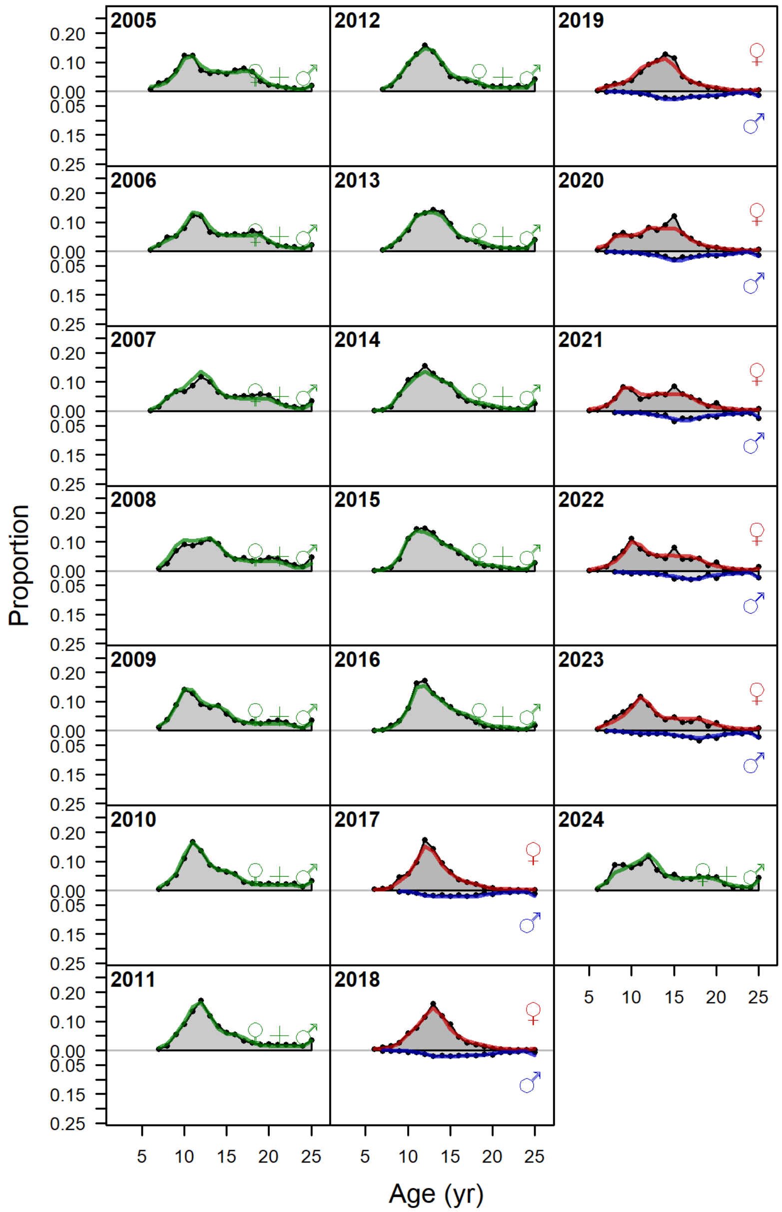


Figure 32. Fit to later years of fishery age data in the coastwide long model.

Fishery selectivity generally showed a pattern toward selecting fewer younger fish in the latter half of the time series, and a similar trend was estimated for the FISS (Figure 33). The apparent deviation toward lower selectivity of males around 2020 for the FISS may reflect the abrupt change in spatial coverage in that year due to logistical challenges and a reduced design. The overall shift toward lower selectivity for younger fish may be consistent with changes in both the age-structure of the stock, the trends in size-at-age interacting with age-based selectivity and the spatial distribution creating changes in availability. Fishery catchability was estimated to have a large (unconstrained) increase associated with the change from “J” to circle hooks (Table 19, Figure 34). Older halibut were more represented in the non-directed fishery discards age data prior to 1992, and therefore the estimated selectivity was nearly asymptotic. Recreational and discard selectivity estimates were relatively similar to those from the coastwide short model.

Female natural mortality in the coastwide long model was estimated to be higher (0.221) than for males (0.198) although the 95% intervals overlap (Table 19). The environmental link parameter (β) was estimated to be positive (0.456), with no density below a value of 0.0, thus suggesting a strong and significant relationship between average recruitment and the phase of the PDO (based on the updated PDO index described above; Table 19). However, the time series of estimated recruitments (Figure 35) and deviates from the PDO-informed stock-recruitment relationship (Figure 36) still show some temporal patterns, suggesting the potential for unmodelled effects on the stock-recruitment relationship might still be present. Specifically, the poor PDO period from the 1940s to the 1970s and the positive phase from the 1970s to the early 2000s generally correspond to negative and positive deviations even with the relationship included (Figure 36).

Summary of strengths and weaknesses for the coastwide long model:

Strengths:

- Includes uncertainty in female natural mortality
- Includes extensive historical data
- Uses environmental information to inform recruitment
- Modest technical overhead (complexity)
- Fits the fishery and survey indices well
- Fits both the survey and fishery age data well
- Internally consistent data weighting

Weaknesses:

- May lose Region-specific trends and biological patterns due to aggregation
- Relies heavily on only fishery trends over the historical period
- Implicitly assumes stationarity in some processes (e.g., the stock-recruitment function after accounting for the PDO, *M*) over the long historical period
- Implicitly assumes that availability to the fishery did not change over the historical period, despite known patterns in geographical expansion prior to the 1960s

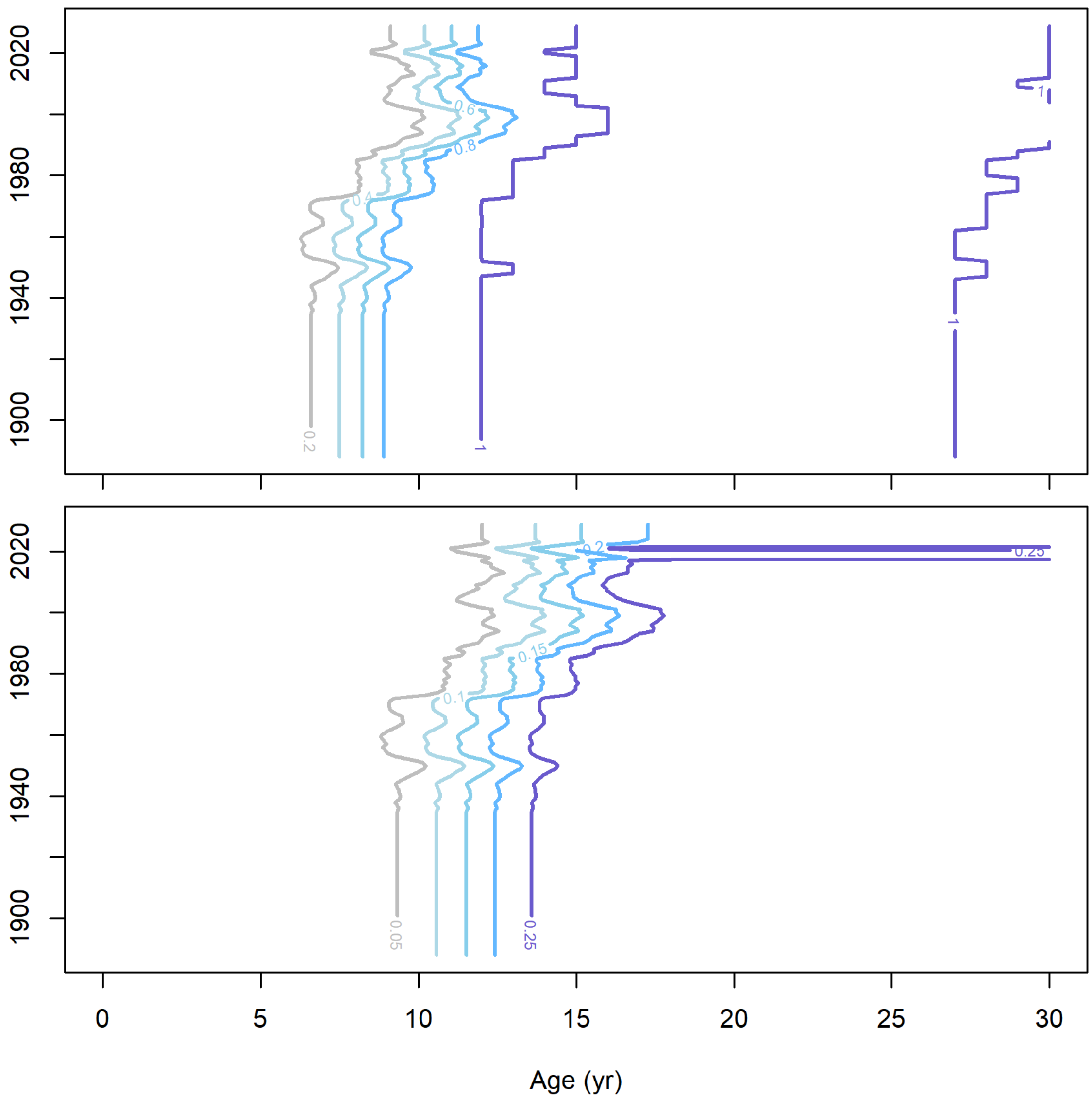


Figure 33. Estimated selectivity for females in the commercial fishery landings (upper panel) and survey (lower panel) in the coastwide long model.

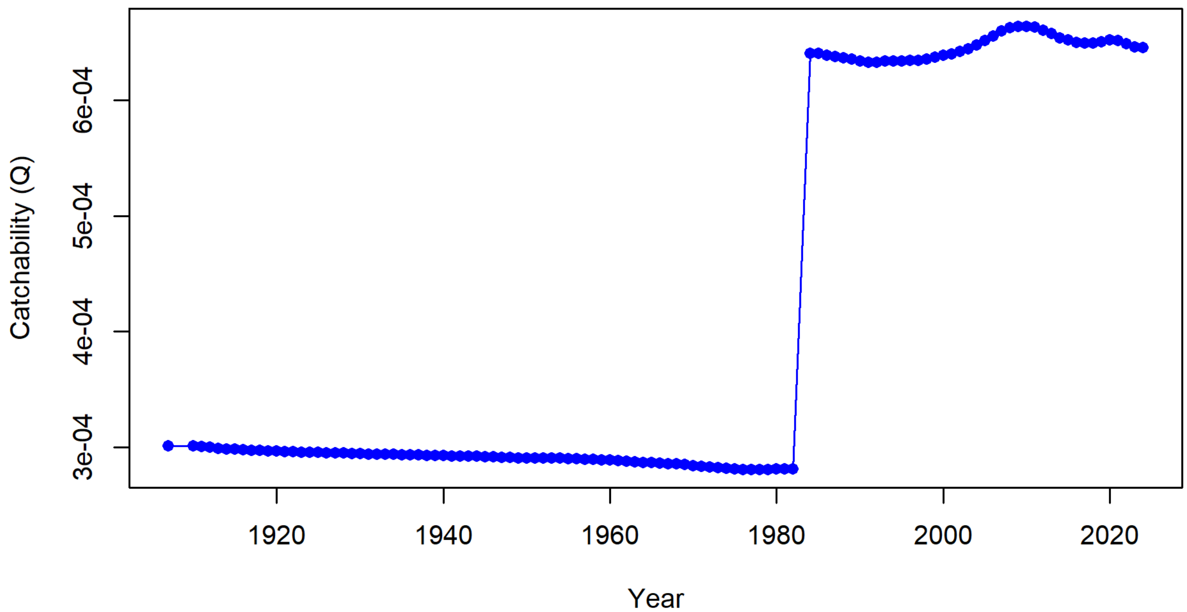


Figure 34. Time-varying fishery catchability in the coastwide long model. The change corresponding to the transition to circle hooks in 1984 is unconstrained.

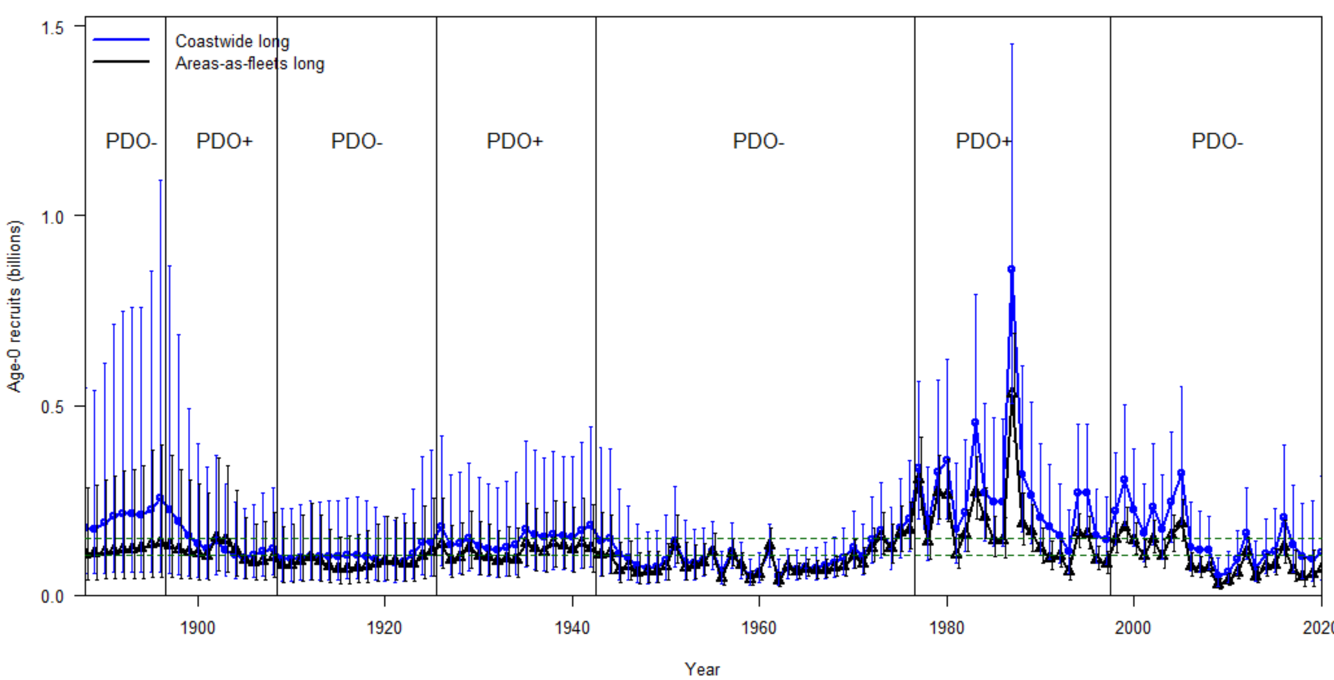


Figure 35. Estimated recruitments and assumed PDO regimes from the coastwide long and AAF long models (right panel); horizontal lines indicate equilibrium values in the absence of the PDO.

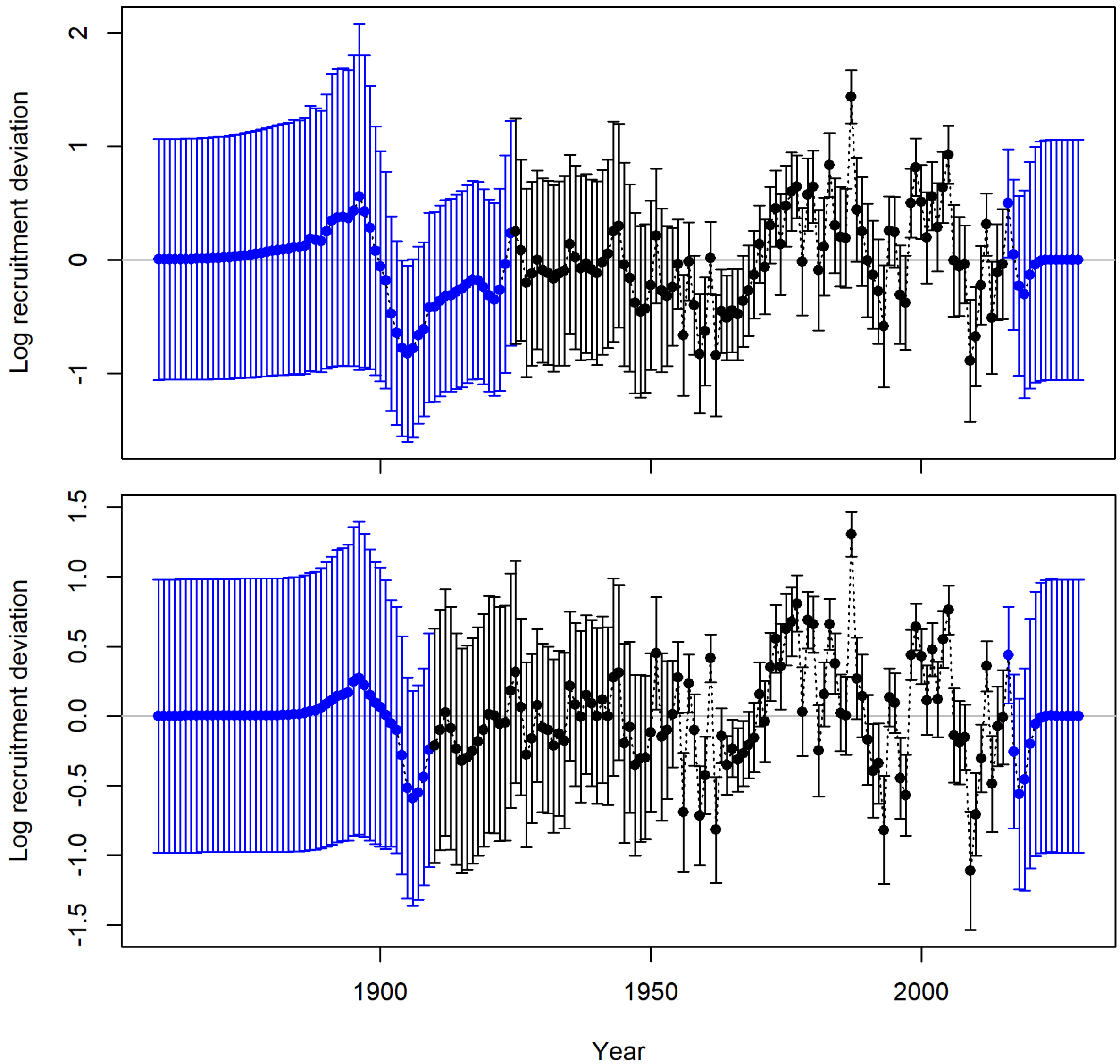


Figure 36. Estimated recruitment deviations in the coastwide long (upper panel) and AAF long (lower panel) models; horizontal lines indicate expected values based on the stock-recruitment functions as modified by the estimated PDO relationships.

AAF short

The AAF short model fit the observed trends in all fishery and FISS indices relatively well (Figure 37-38). Fit to the aggregate age data for each fleet clearly illustrated the differences in age structure among the data from each biological region and among fishery sectors (Figure 39). The biggest differences between the age of female and male halibut observed from the FISS occurred in Region 3, and generally Regions 4 and 4B were predicted (and observed) to have the greatest fraction of older halibut, a majority of which were males. The fit to the annual FISS age data generally captured these patterns, with the worst fit occurring for the data from Region 2 (Figure 40); the model weighting suggested a low effective sample size for the Region 2 FISS data consistent with these patterns in lack of fit (Table 18). Considerable exploration was made toward improving the fit to the Region 2 FISS data and addressing the clear residual patterns (see sensitivity analyses below; Figure 40); however, so satisfactory replacement approach was identified. The fit to the age composition data from Region 4 clearly shows the very small proportion of males in the landings (Figure 41), with a much greater proportion observed in Region 4B (Figure 42). Although showing a reasonably good aggregate fit, predicted annual commercial fishery landings in Biological Regions 4 and 4B did not capture the strong peaks created by the 1987 year-class in the late 1990s and early 2000s, suggesting that this large year class may have moved toward Regions 2 and 3 as those fish grew older and therefore the fit represents a compromise between fitting the younger and older observations from that cohort. This type of spatial dynamic is not fully approximated by an Areas-As-Fleets approach and would require a fully spatial model to model more accurately. Both of these Regions were weighted similarly after iterative tuning (Table 18).

The estimate of female natural mortality in the AAF short model (0.220) was slightly lower than in the coastwide long model and males were estimated to have a much lower value (0.179; Table 19). The lack of overlap on the 95% intervals indicates the clearly different explanation in this model for the observed sex-ratios, albeit restricted to the most recent portion of the time-series. This result likely indicates the trade-off between the assumption of asymptotic selectivity in the coastwide model and domed selectivity for most Regions in the AAF models. The AAF short model estimated a negative but somewhat smaller initial offset to recruitment as the coastwide short model. Due to the higher estimated *M*, the AAF short model estimated a higher absolute level of recent recruitment and spawning biomass than the coastwide short model (Table 19).

Summary of strengths and weaknesses for the AAF short model:

Strengths:

- Parameter estimates are derived from the most recent time-period
- Avoids aggregating data over Biological Regions with differing trends and biological patterns
- Fits the Regional fishery and FISS indices well
- Fits Regions 2 and 3 fishery age data well
- Internally consistent data weighting
- Propagates uncertainty in female and male *M* estimates

Weaknesses:

- Does not include environmental information to inform recruitment
- Increased technical overhead (complexity)
- Residual patterns in Region 4 and 4B fishery and survey age data
- Fits Region 2 FISS age data poorly
- Does not include extensive historical data

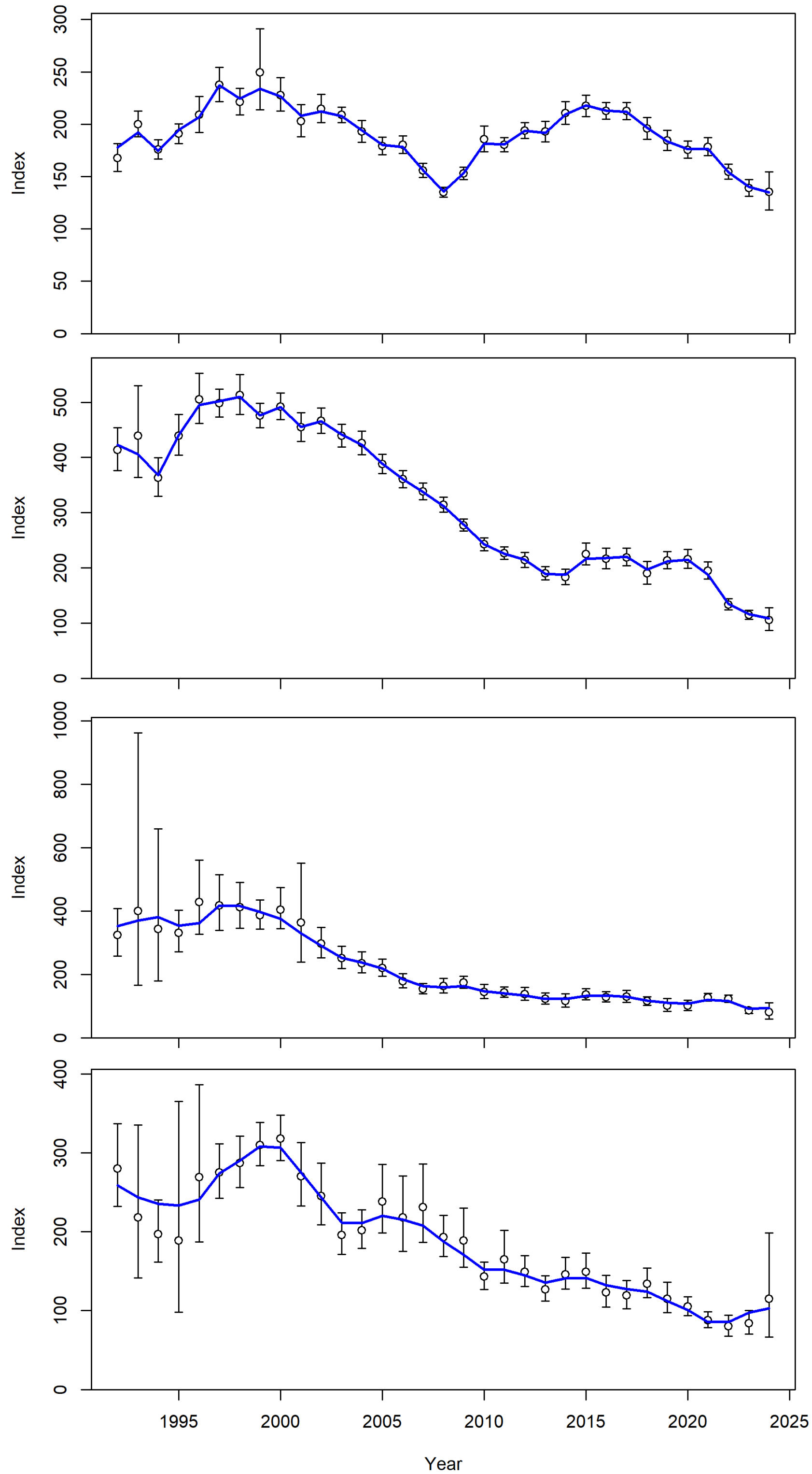


Figure 37. Fit to fishery trends in Biological Regions 2, 3, 4, and 4B (top to bottom) in the AAF short model.

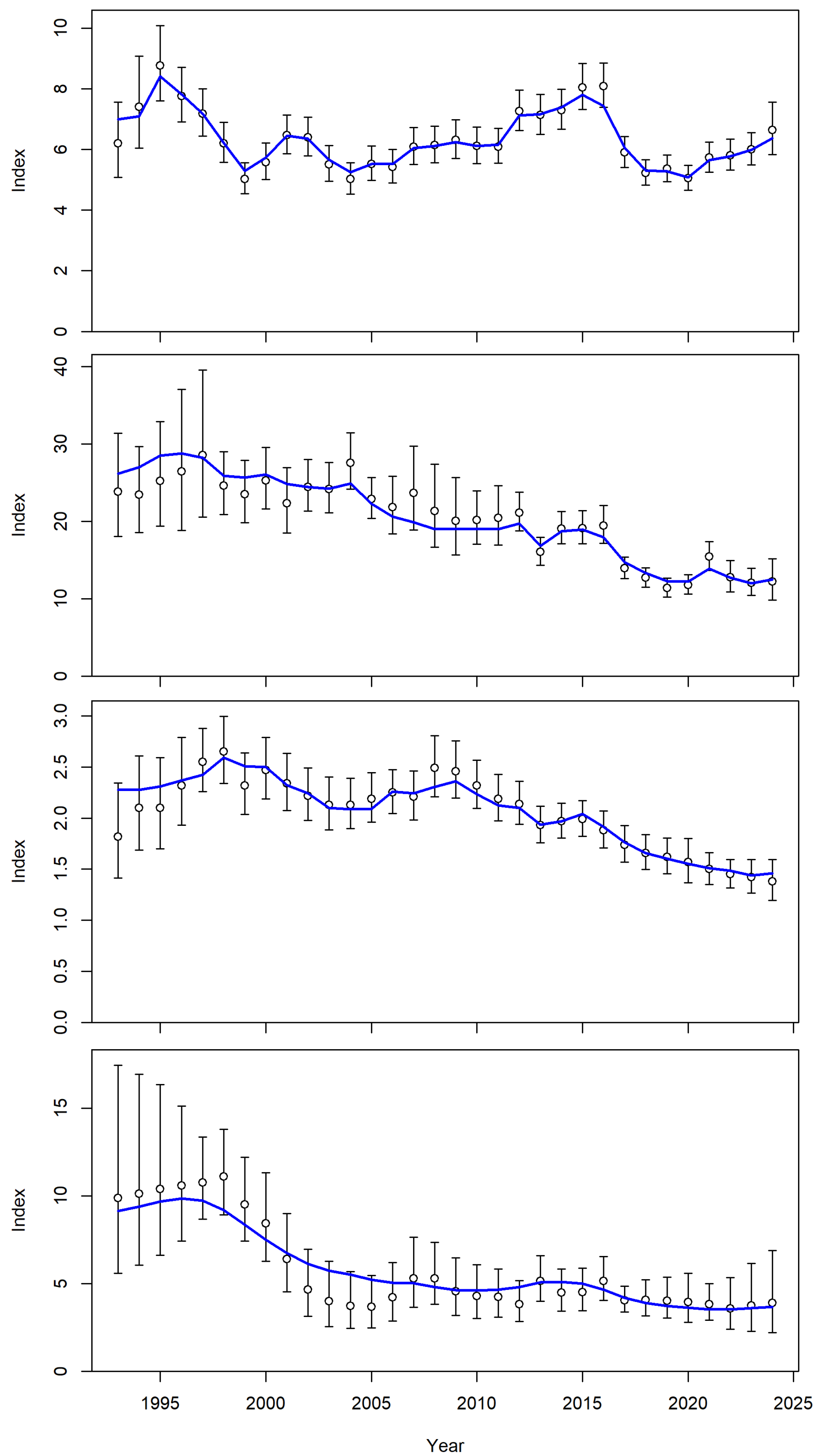


Figure 38. Fit to survey trends in Biological Regions 2, 3, 4, and 4B (top to bottom) in the AAF short model.

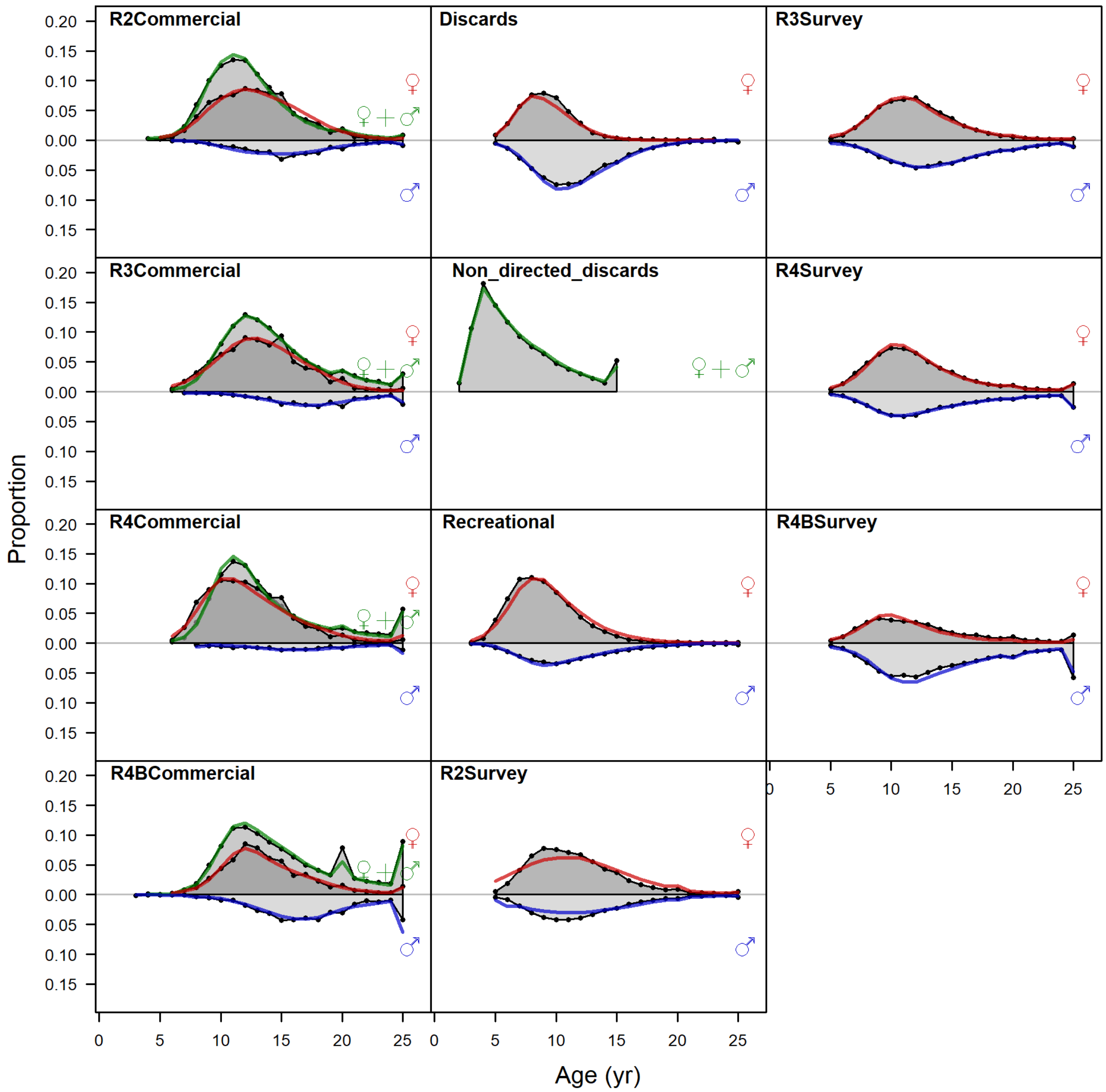


Figure 39. Aggregate fit to age data for each model fleet in the AAF short model.

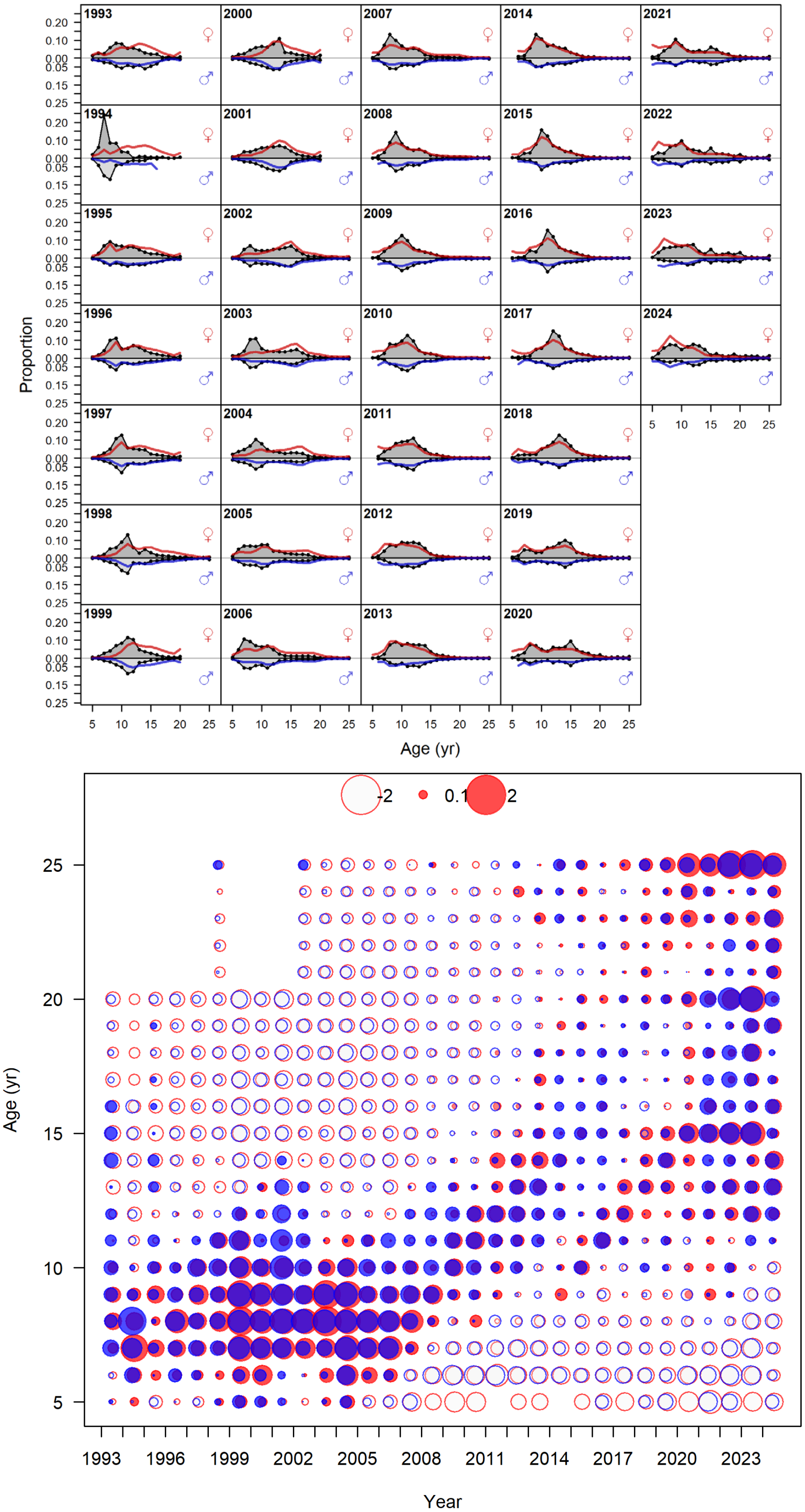


Figure 40. Fit to age data (upper panel) and Pearson residuals (lower panel) from the Region 2 FISS in the AAF short model; red circles denote female residuals, and blue circles denote male residuals.

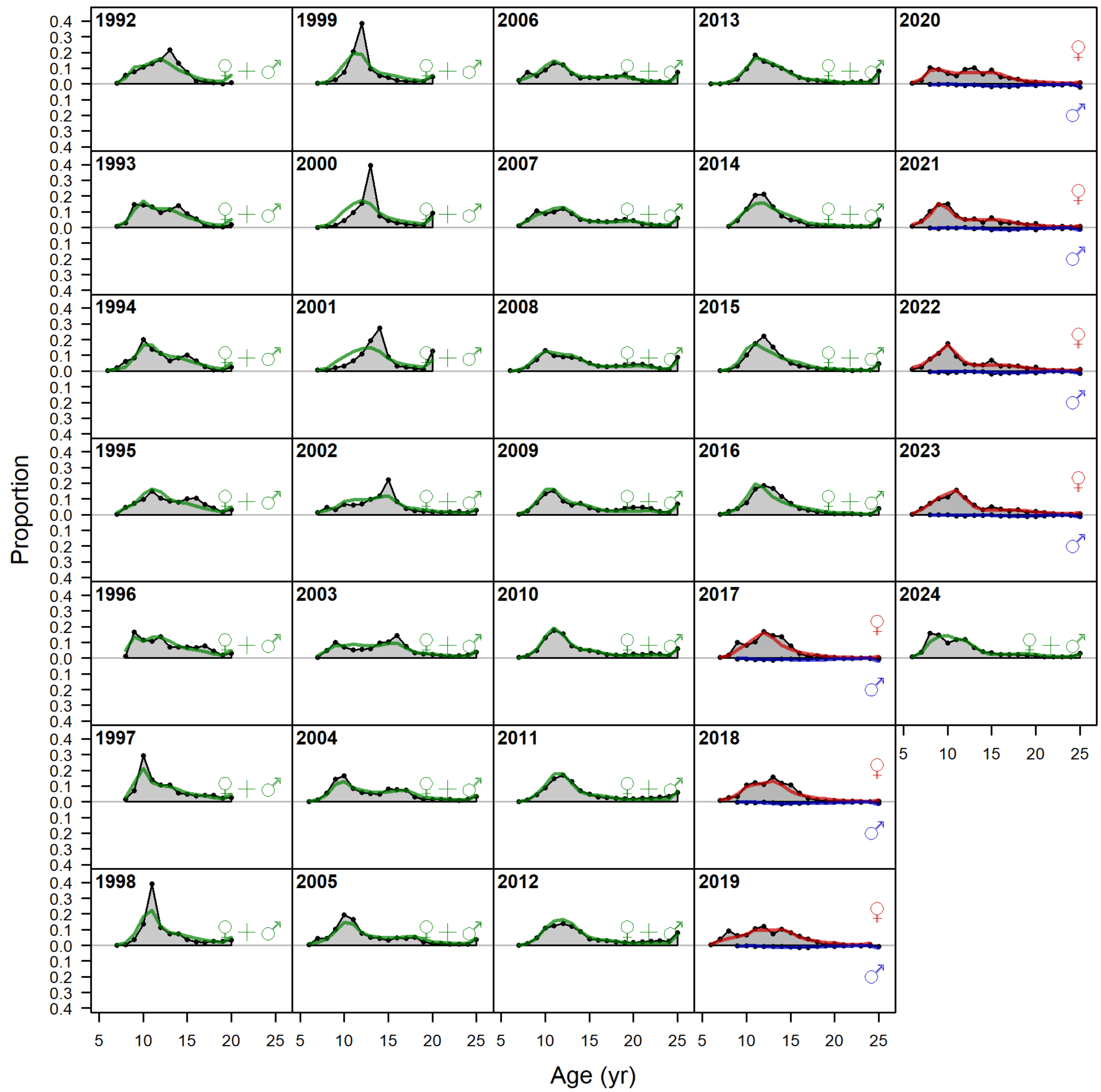


Figure 41. Fit to age data from the Region 4 commercial fishery landings in the AAF short model.

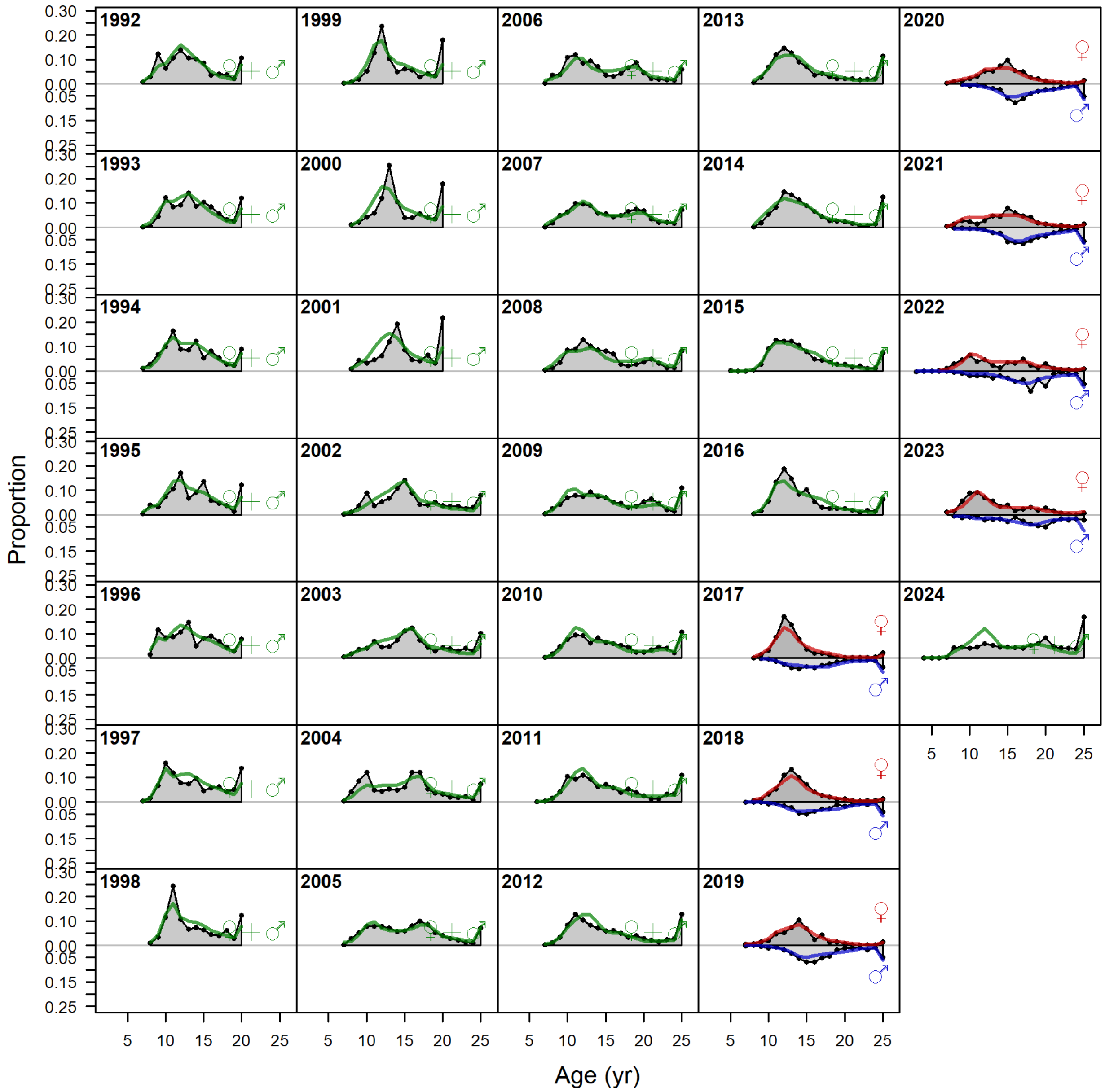


Figure 42. Fit to age data from the Region 4B commercial fishery landings in the AAF short model.

AAF long

Like the AAF short model, the AAF long model fit both the fishery and FISS trends well (Figure 43-44). Aggregate fits to the FISS age composition data showed similar patterns to those observed in the AAF short model (Figure 45). The fit to the FISS age data improved over the time series, but the Region 2 and 3 FISS age data was heavily down-weighted in order to achieve internally consistent weighting (Table 18). This corresponded to poor fits to the Region 2 age data over much of the time series (Figure 46-47). Lack of fit to the Region 3 FISS data occurred primarily in the early part of the time-series (Figure 48-49). Among the fishery fleets, the Region 4 data were most heavily down-weighted from the bootstrapped input sample sizes (Table 18).

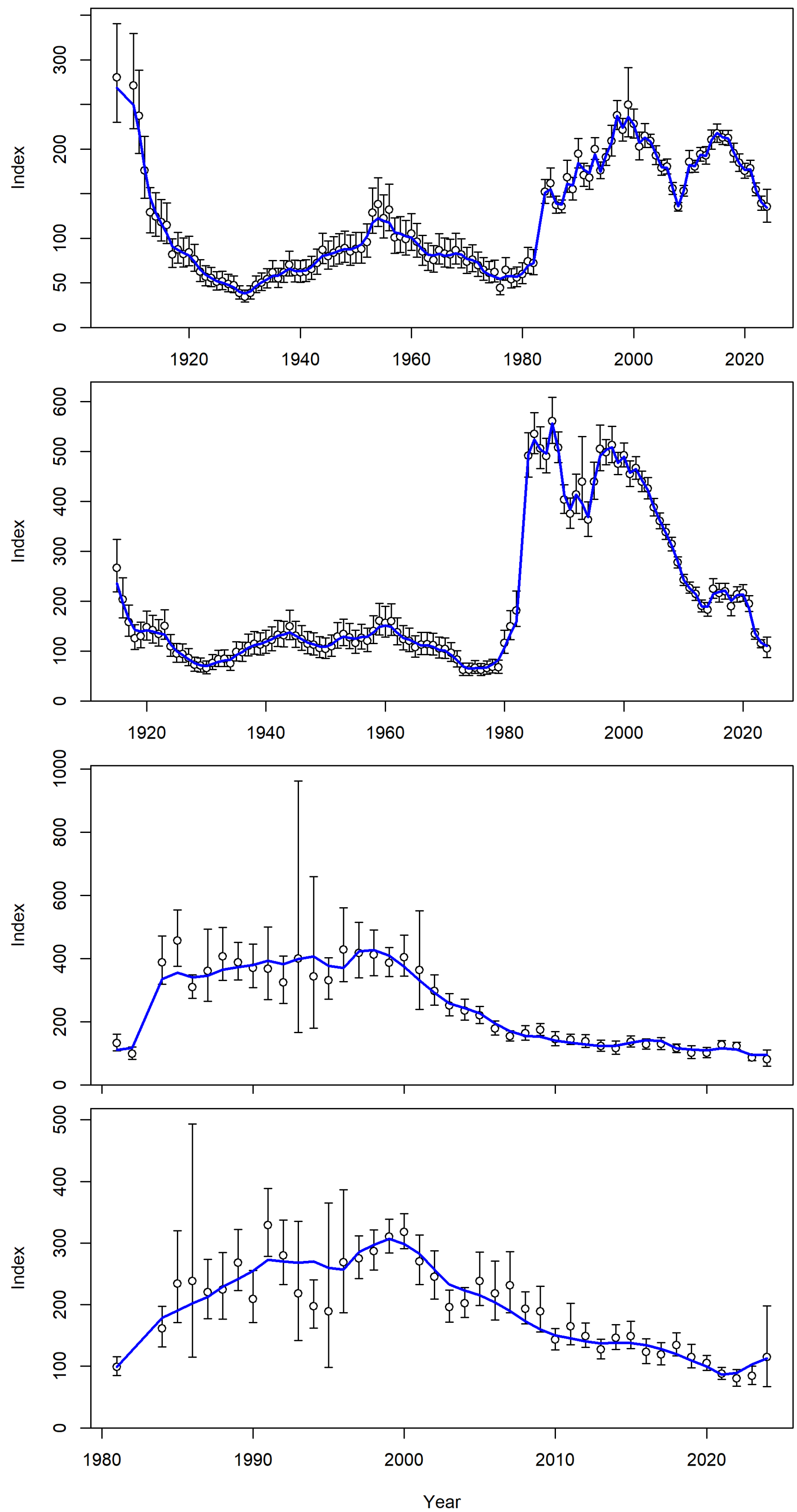


Figure 43. Fit to fishery trends in Biological Regions 2, 3, 4, and 4B (top to bottom) in the AAF long model.

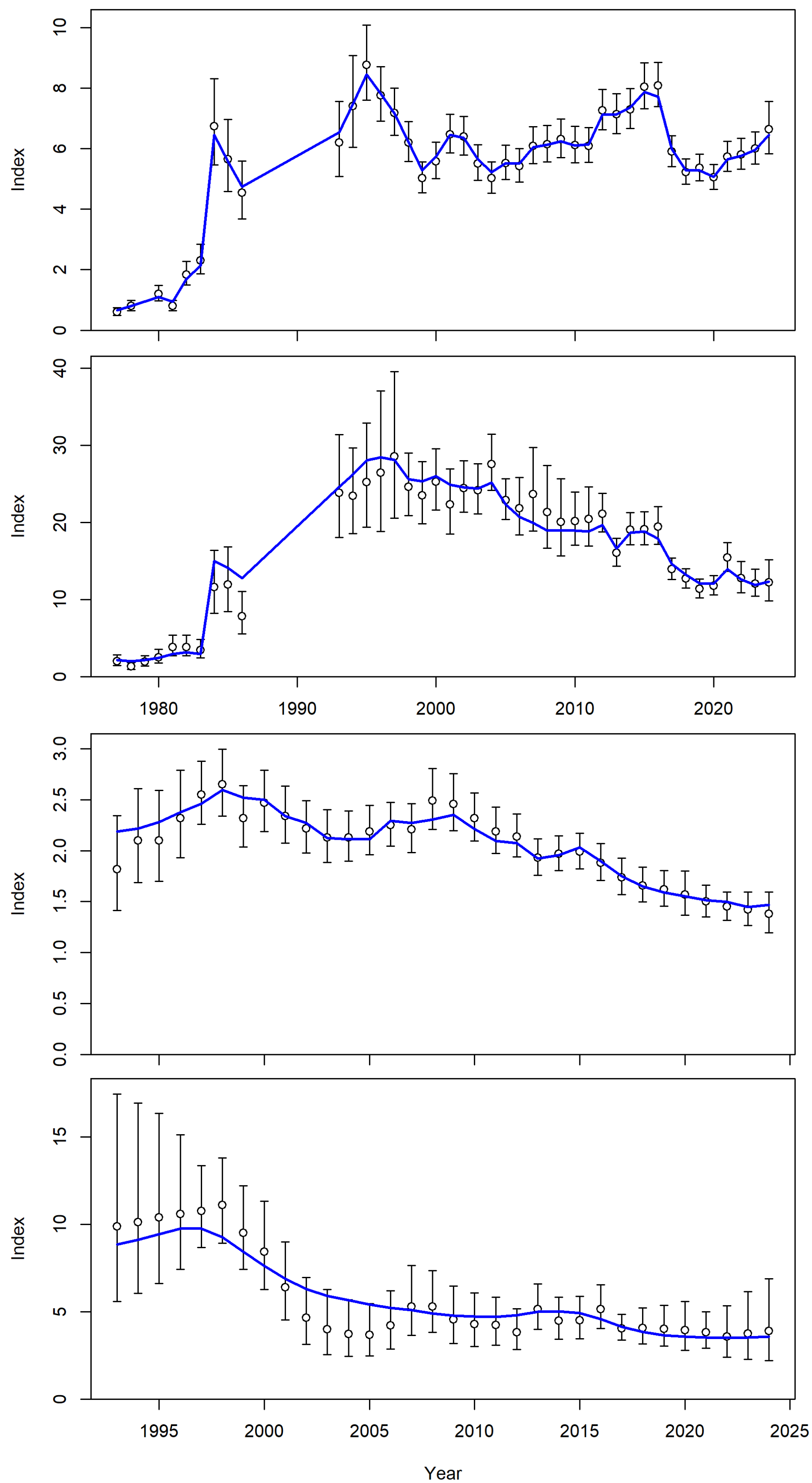


Figure 44. Fit to FISS trends in Biological Regions 2, 3, 4, and 4B (top to bottom) in the AAF long model.

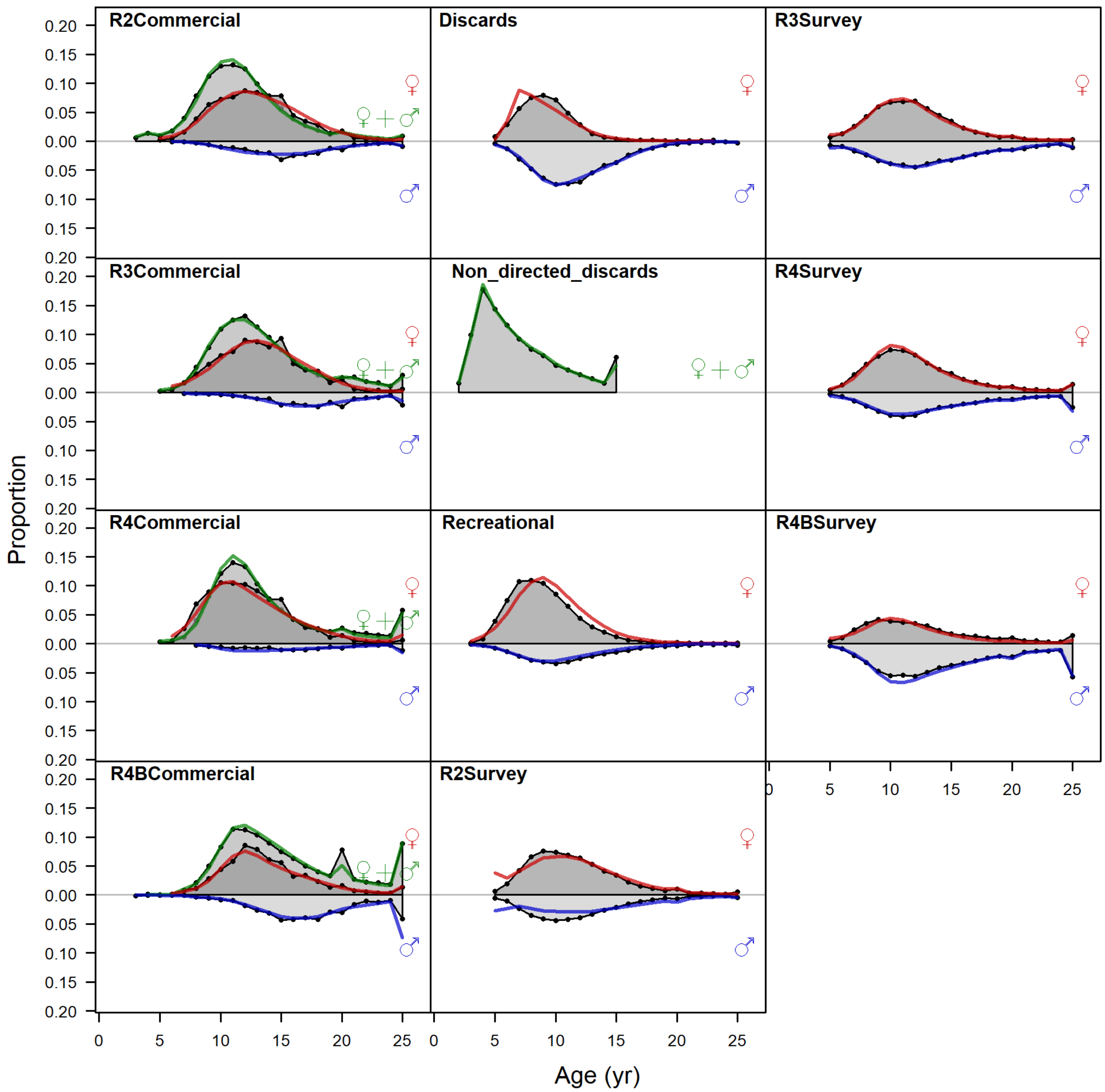


Figure 45. Aggregate fit to age data for each model fleet in the AAF long model.

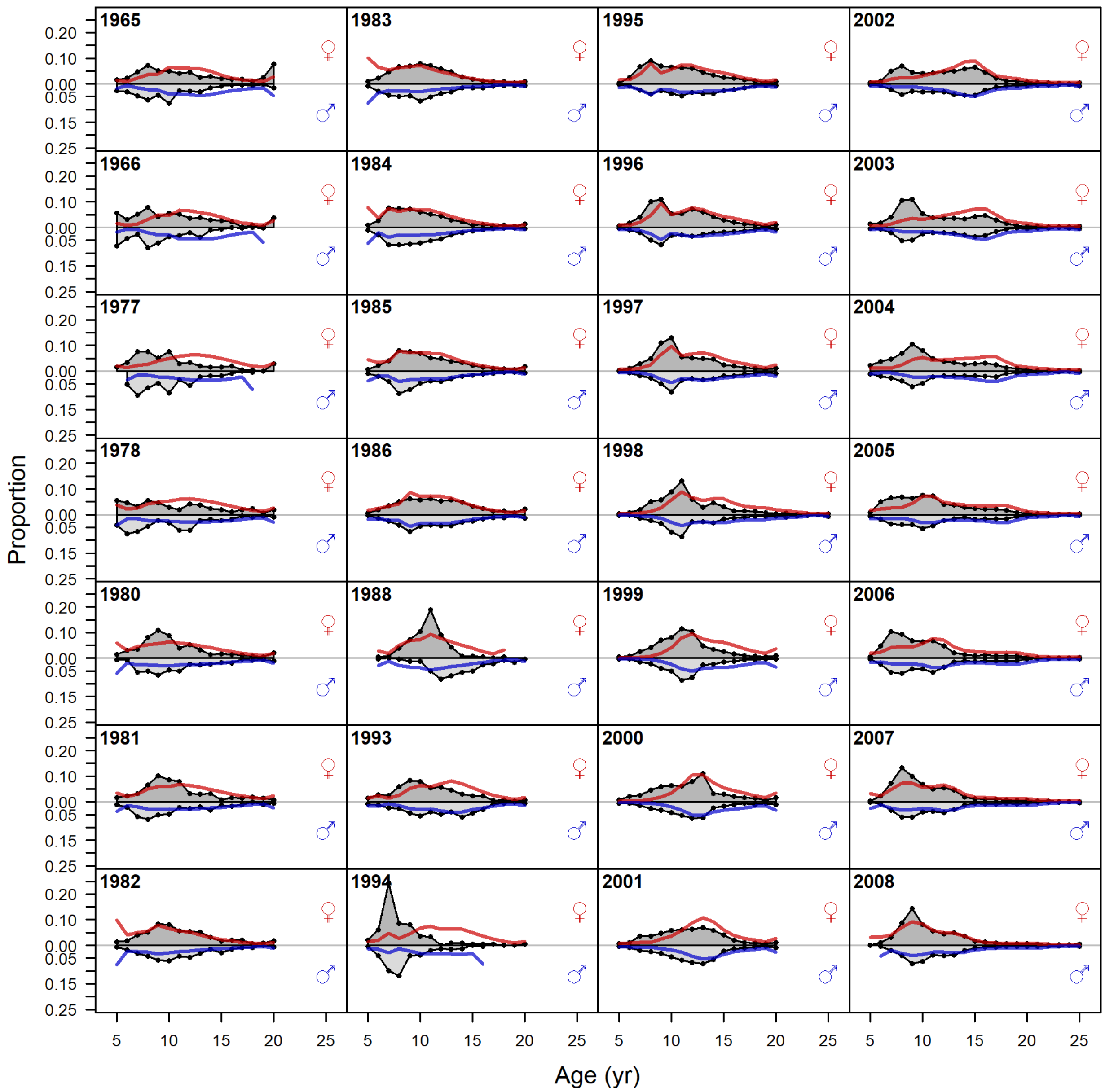


Figure 46. Fit to 1965-2008 age data from the Region 2 FISS in the AAF long model.

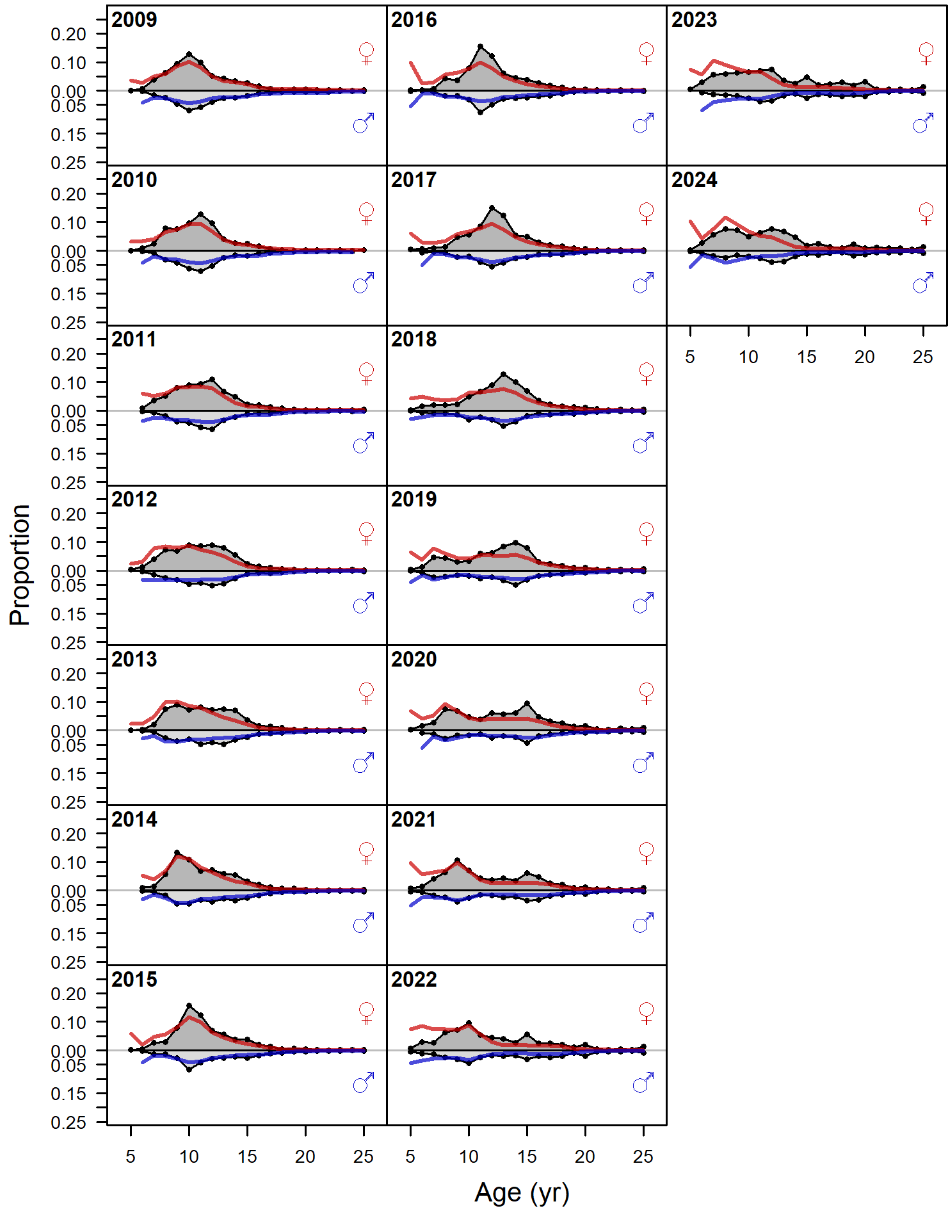


Figure 47. Fit to recent (2009+) age data from the Region 2 FISS in the AAF long model.

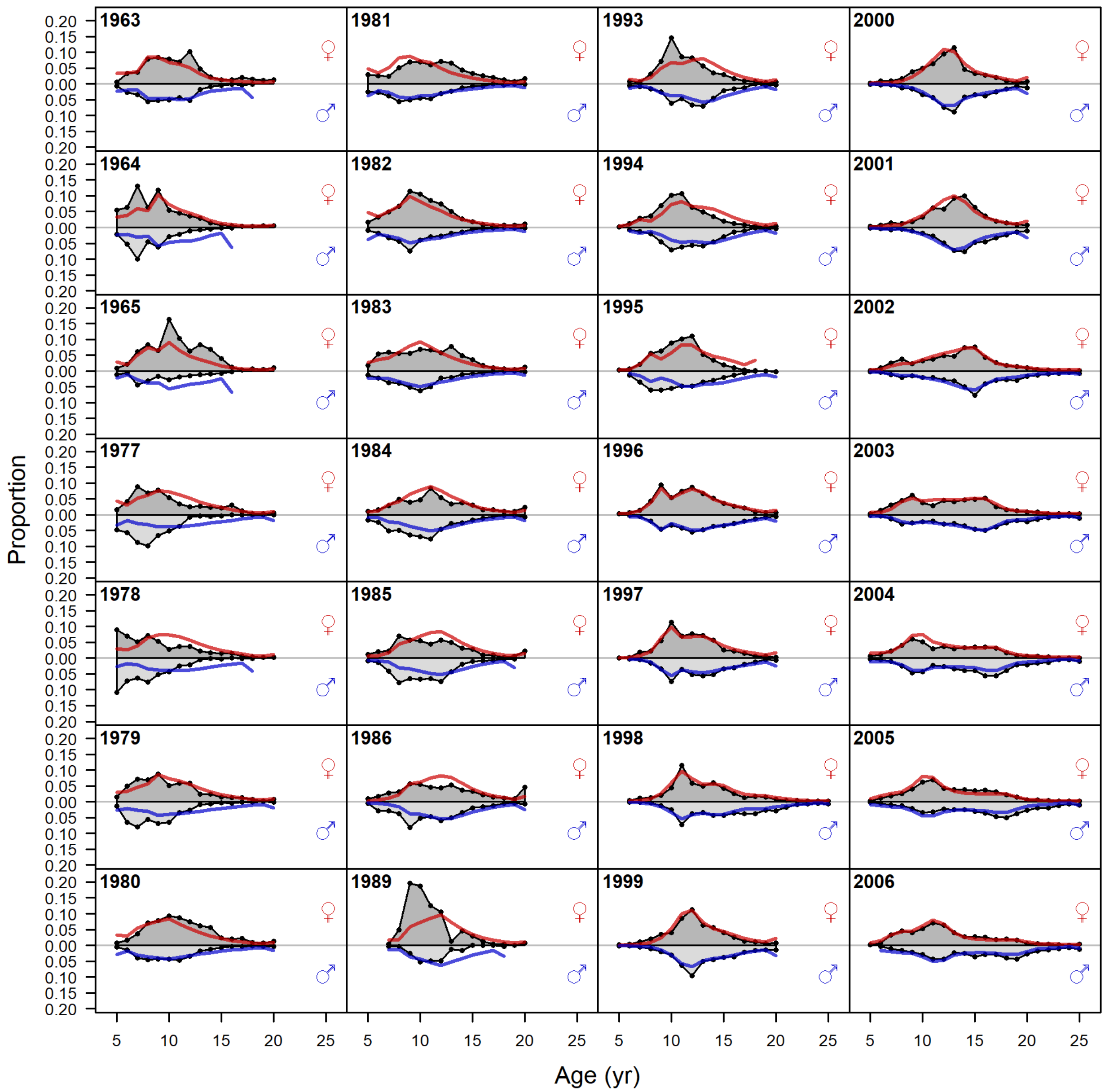


Figure 48. Fit to early age data from the Biological Region 3 FISS in the AAF long model.

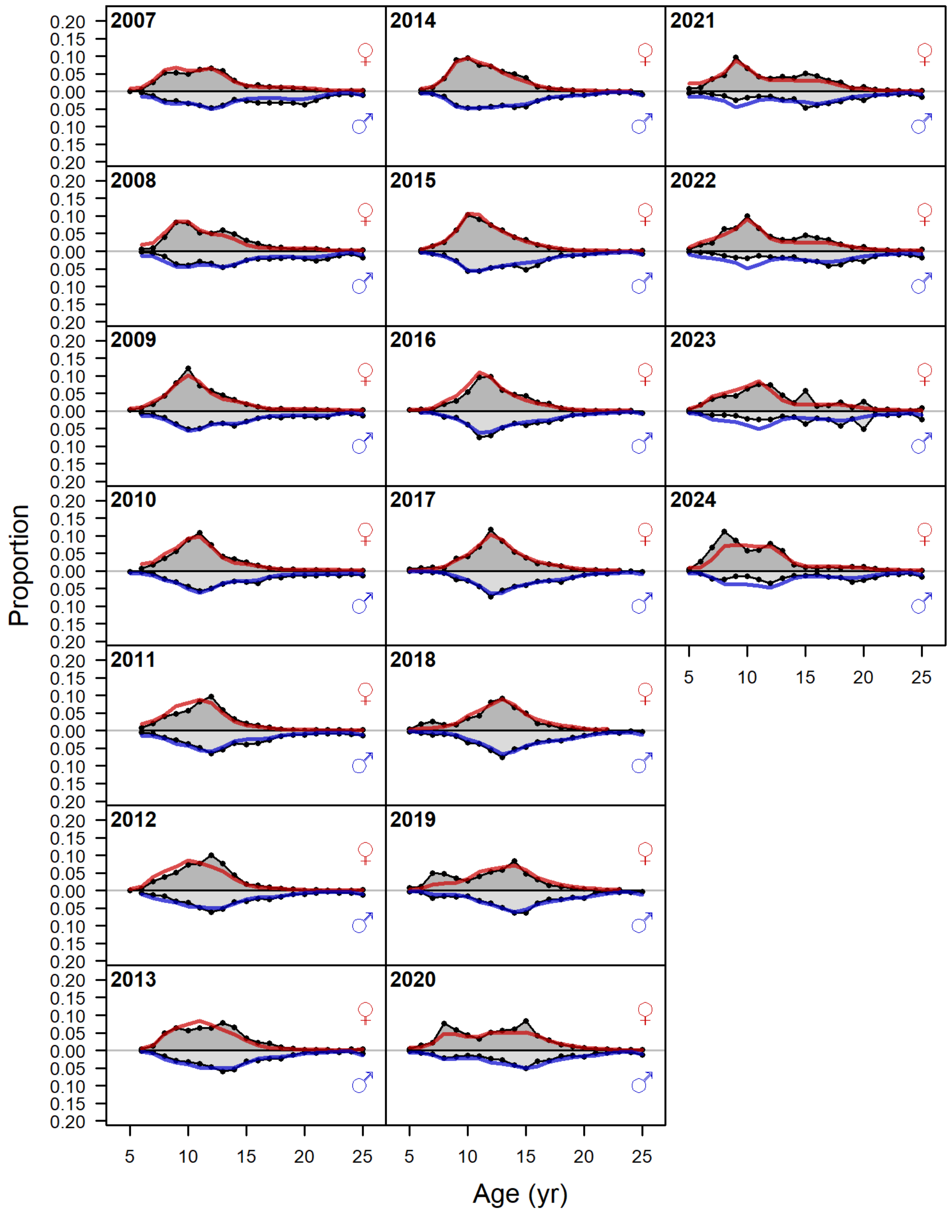


Figure 49. Fit to later age data from the Biological Region 3 FISS in the AAF long model.

Similar to the AAF short model, FISS selectivity was estimated to be asymptotic for Biological Regions 4 and 4B. Peak male selectivity in the commercial fishery landings was also estimated to be asymptotic. All fleets with data extending past the transition from J to circle hooks (1984) showed a strong offset in the unconstrained deviation in catchability for that year (Table 19). Discard and recreational selectivity estimates were similar in the AAF long model to those estimated in the coastwide long model. Non-directed discard selectivity was estimated to be domed, again illustrating the trade-off between domed fleets in the AAF models and asymptotic selectivity over the entire time-series in the coastwide models. This likely interacts with the estimation of natural mortality, producing slightly lower values in the AAF long model (0.186 for females, and 0.163 for males) than in the coastwide long model (Table 19). The environmental link coefficient was estimated to be slightly weaker (0.430) than in the coastwide long model, although the 95% interval still did not approach zero indicating a highly significant relationship (Table 19). The AAF long model produced intermediate estimates of recent recruitment and female spawning biomass compared to the other three model (Table 19). This result is consistent with the intermediate estimates of male and female *M* from this model.

Summary of strengths and weaknesses for the AAF long model:

Strengths:

- Includes uncertainty in female and male *M*
- Includes extensive historical data
- Uses environmental information to inform recruitment
- Fits the fishery and survey indices well
- Fits both the Regions 2, 3 and 4B fishery age data well
- Fits Region 4 and 4B FISS age data well
- Internally consistent data weighting

Weaknesses:

- Highest technical overhead (complexity) of the four models
- Most challenging model to check and ensure reliable convergence
- Relies heavily on only fishery trends over the historical period
- Implicitly assumes stationarity in some processes (e.g., the stock-recruitment function, *M*) over the long historical period
- Fit Biological Regions 2 and 3 survey age data poorly

Sources of uncertainty

The four models evaluated here represent, within the set itself, significant sources of uncertainty in how to treat the data (partitioning by fleets or aggregating to a single series), as well as how to treat the time-series (emphasizing the recent dynamics or including more historical information). Further, the differing assumptions of fixed vs. estimated female natural mortality rate and the treatment of environmental covariates to the stock-recruitment relationship are also embedded in the differences observed among the four model results. These factors lead to important differences in both scale and trend. In aggregate, the four models together reflected much more uncertainty than would any single model. However, it is notable that the data remain generally informative of a similar population scale and recent trend for both spawning biomass and recruitment.

Sensitivity analyses

Many alternative model configurations were evaluated during model development, but only a subset of these is reported here. Several of the bridging steps from the 2024 models to the 2025 preliminary models described here also represent sensitivity analyses. The focus of the analyses described below was on model behavior and understanding; sensitivity analyses specifically intended to highlight the importance of ongoing research (e.g., whale depredation, maturity ogives, etc.) are produced each year as part of the final stock assessment (Stewart and Hicks 2025).

The large differences in the scale of the spawning biomass in the historical period between the two long time series models represent importantly differing assumptions about the connectivity of the stock via spatial availability (Figure 50). Specifically, domed selectivity for Biological Regions 2 and 3 in the long AAF model implicitly assumes that older fish (located in northern and western areas) were historically less available and therefore not mobile enough to be readily available to those fisheries. Conversely, in the coastwide long model the assumption of asymptotic selectivity implies a high degree of availability and therefore connectivity between all geographic components in the population. Sensitivity analyses in the 2015 assessment indicted that these two models could be made much more similar by adjusting the degree of domed selectivity (Stewart and Martell 2016). The use of both models encompasses the range of uncertainty that exists over this aspect of the historical population dynamics, thus the primary sensitivity in the stock assessment is included in the ensemble results.

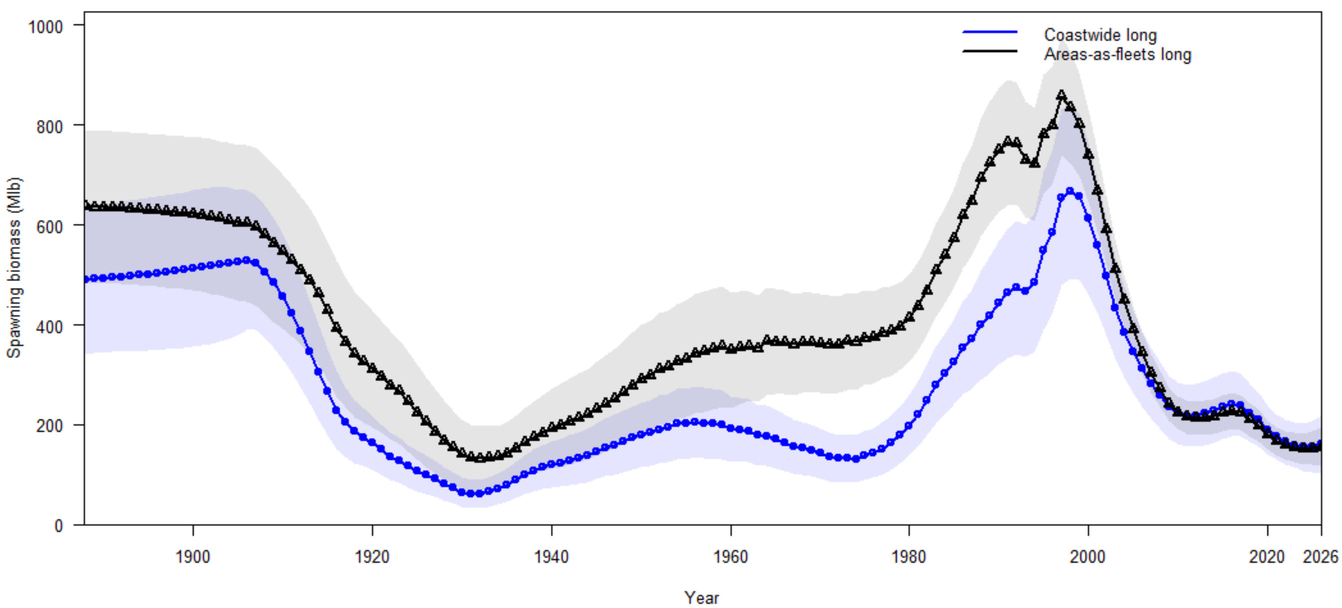


Figure 50. Comparison of the spawning biomass for the long coastwide and AAF models.

The treatment of the PDO in the two long time-series models was explored extensively as part of the 2022 full stock assessment. As described above, the current approach classifies the PDO into a series of binary ‘regimes’, and then estimates a coefficient describing the effect of these regimes on the equilibrium recruitment used in the stock-recruitment relationship. There is still considerable variability remaining in the annual recruitment deviations and the overall effect on the estimates of recruitment from the use of this covariate primarily occur at the end of the time-series when there is little other information to inform recruitment estimates. Due to increasing evidence that the environmental and oceanographic conditions associated with the PDO may be changing (Litzow et al. 2020), it is possible that at some point the use of the PDO as a covariate will no longer provide an improvement to the Pacific halibut models. To explore how this might affect model results, a sensitivity analysis of the two long time-series models was conducted by removing the PDO entirely. Results indicate an increased biomass in the very early part of the modelled period, but either little change (AAF long) or a slight increase in spawning biomass (CW long) in the most recent years (Figure 51).

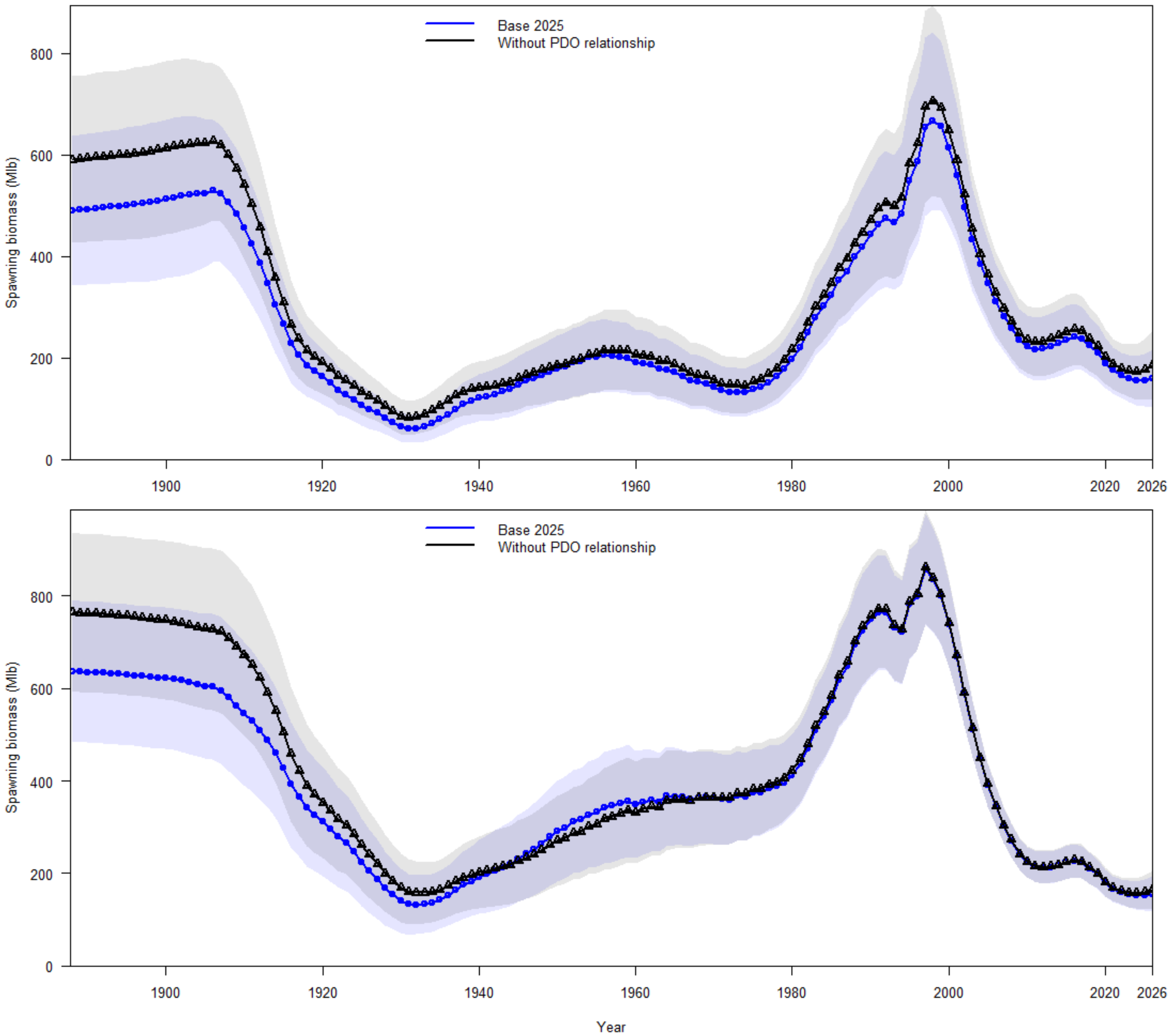


Figure 51. Comparison of the spawning biomass for the long coastwide (top panel) and long AAF (bottom panel) models with and without the PDO relationship included.

To further explore the sensitivity of the stock-recruitment relationship in the Pacific halibut models a sensitivity analysis to the value for steepness (*h*) was also performed. Each of the four models uses a fixed value of 0.75 as the base case. Previous assessments and other supporting analyses have found that this choice provides for modest feedback between spawning biomass and subsequent recruitment but does not have a strong effect on the modelled dynamics. In contrast, a fixed steepness is known to have a very important effect on reference points that rely on the stock-recruitment relationship (e.g. MSY; Mangel et al. 2013). For this reason, the MSE operating models used by the IPHC include additional variability in steepness beyond what is used in the stock assessment. A sensitivity to higher and lower values of steepness showed that the coastwide short model (Figure 52) and the AAF short model (Figure 53) spawning biomass estimates were largely unchanged for alternative values of steepness. Recruitments tended to be estimated slightly higher at the end of the time-series for lower values of steepness, indicating there may be a small effect on forecasts. The coastwide long model showed the greatest sensitivity to steepness with a slightly larger spawning biomass at lower steepness, albeit with a very similar trend over time (Figure 54). The spawning biomass from the AAF long model only differed at the beginning of the modelled period and was almost identical in the most recent years (Figure 55). As currently configured, this sensitivity analysis indicated that the assumed value for steepness was not critically important to the stock assessment results over the range of values considered plausible for flatfish (Myers et al. 1999).

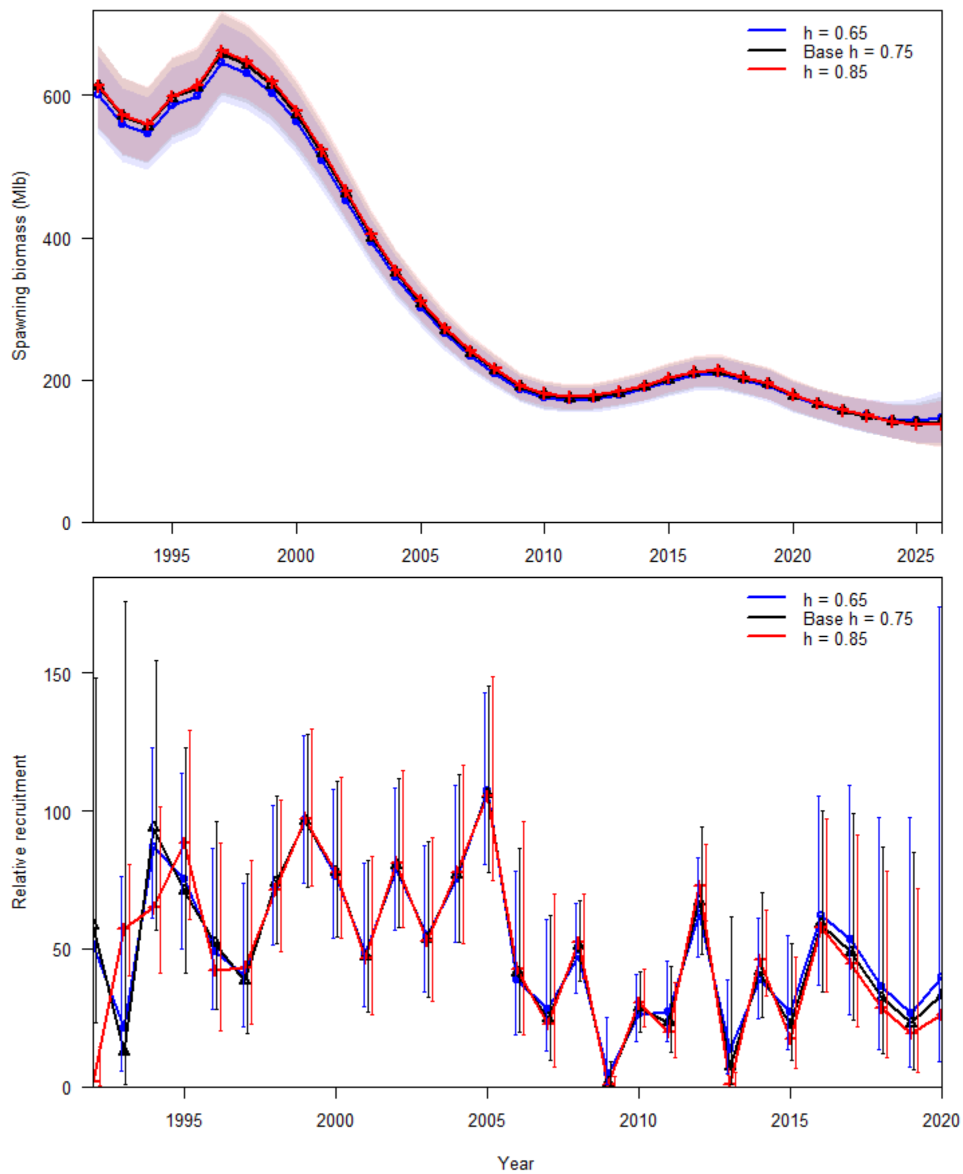


Figure 52. Sensitivity of the CW short model to alternative values of steepness.

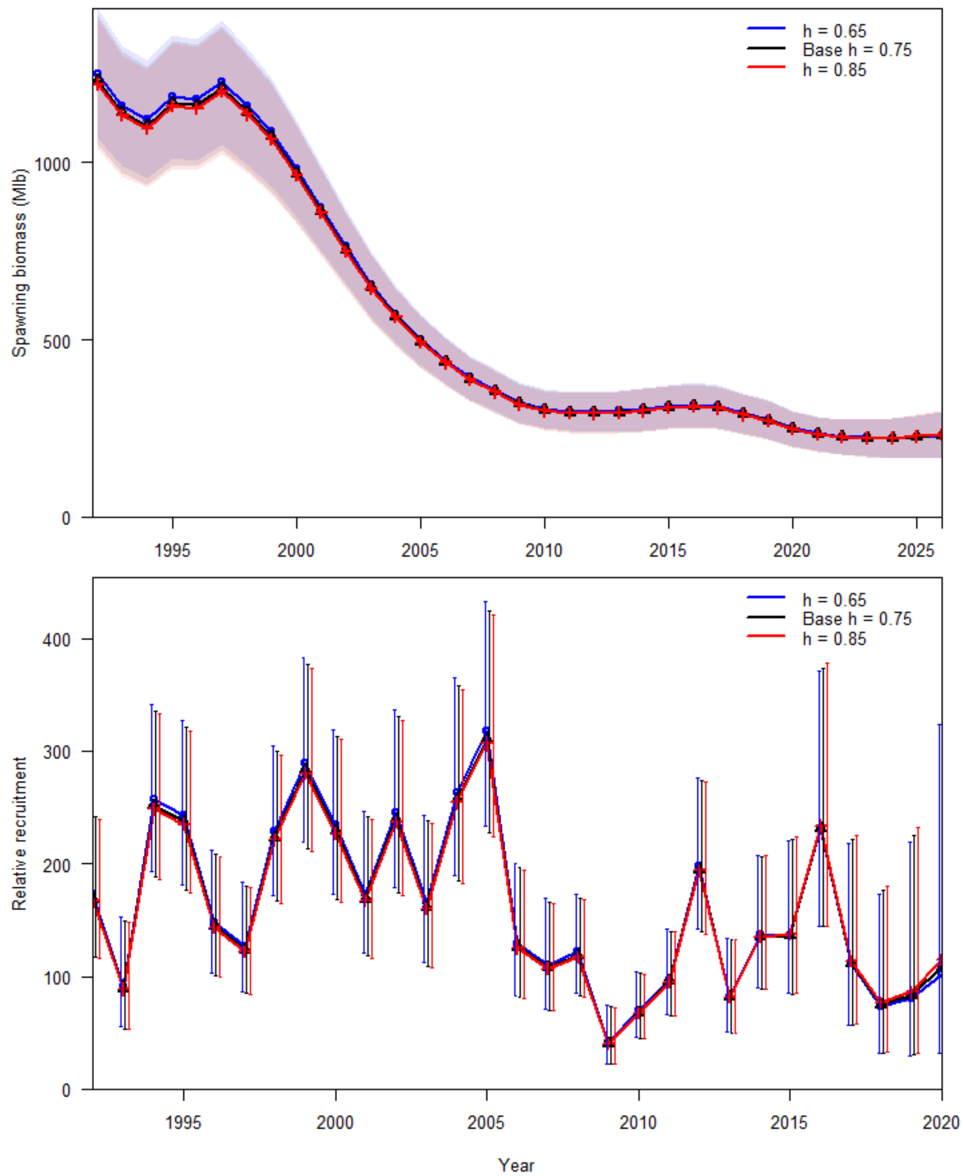


Figure 53. Sensitivity of the AAF short model to alternative values of steepness.

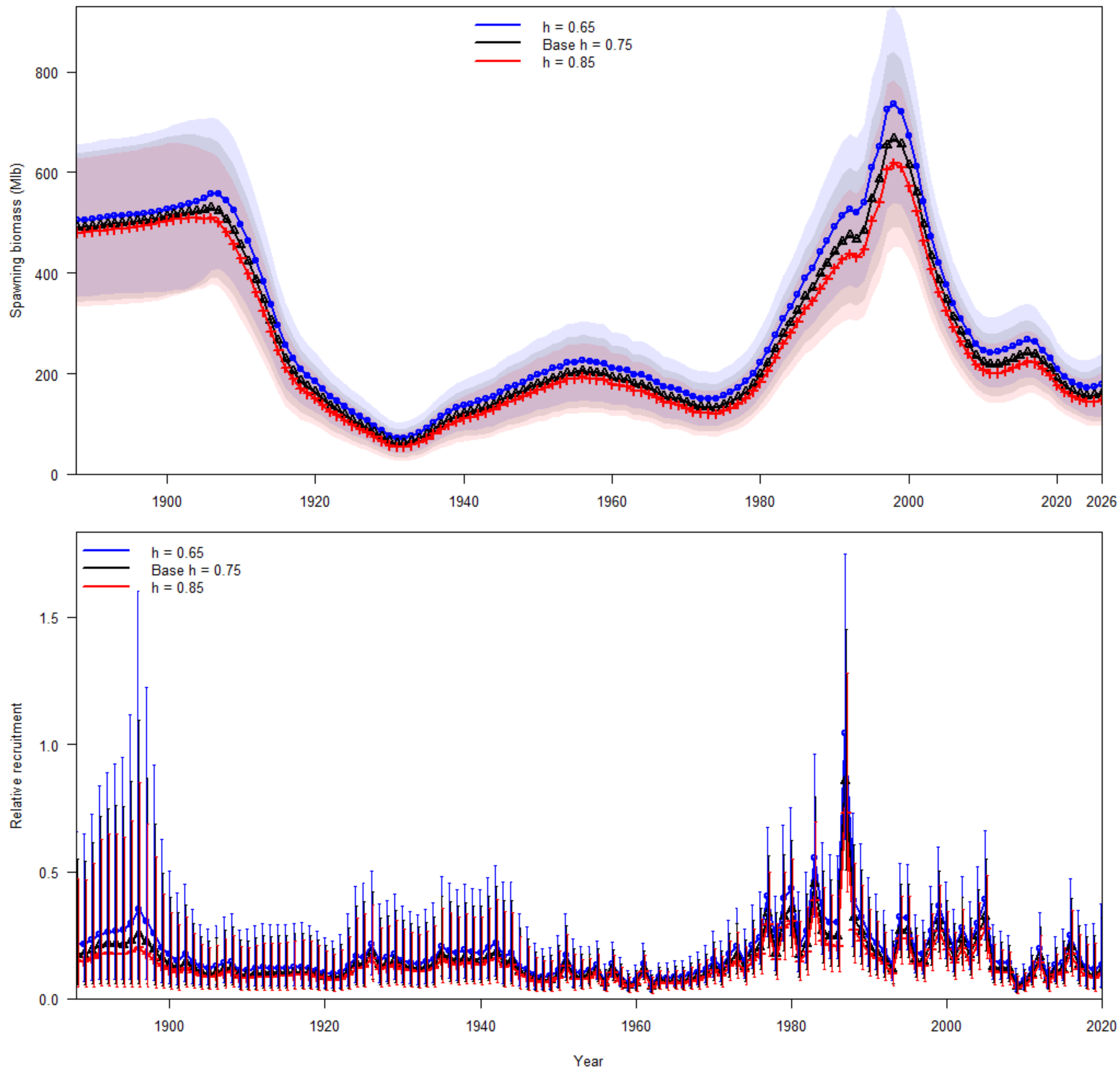


Figure 54. Sensitivity of the CW long model to alternative values of steepness.

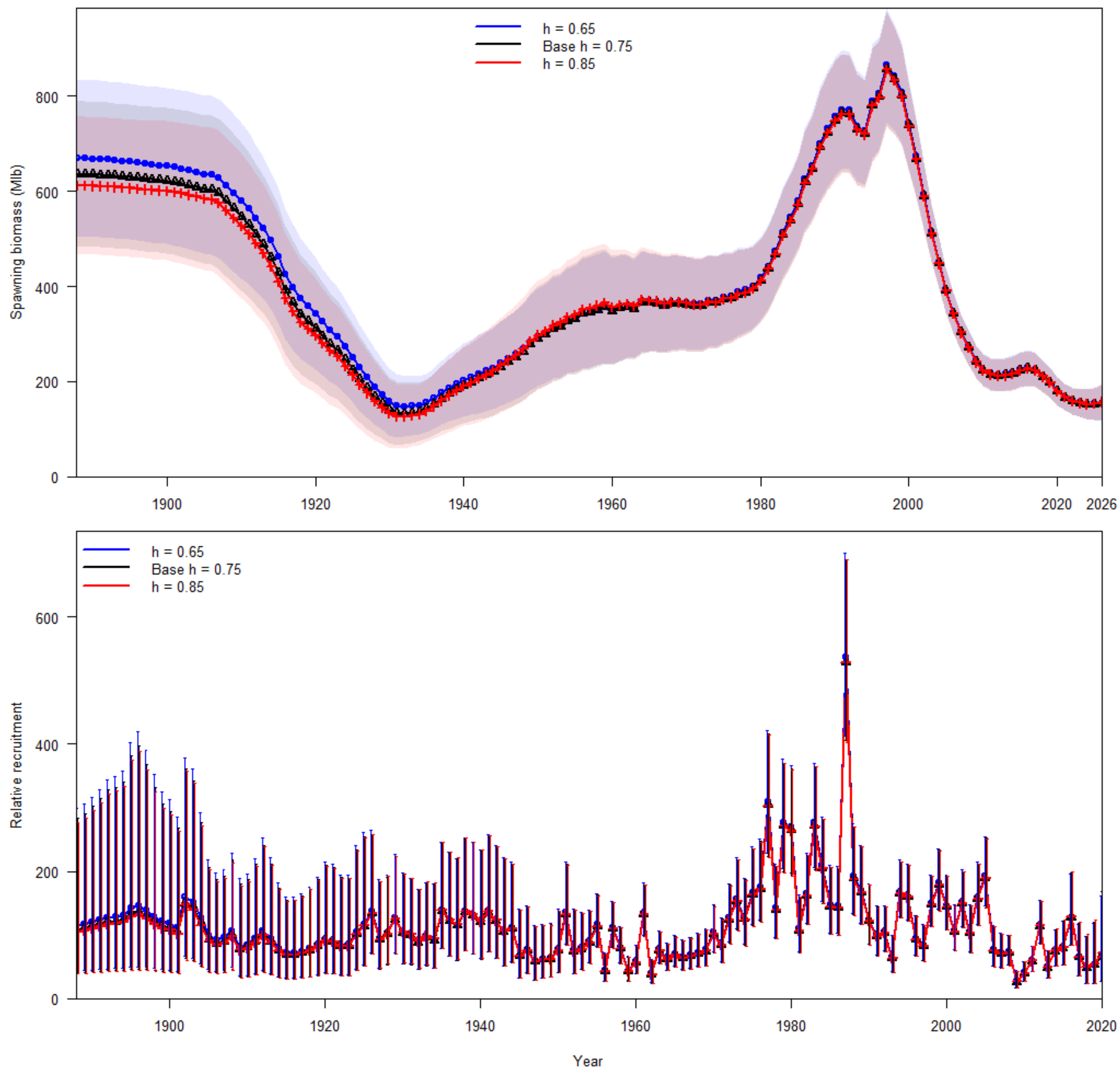


Figure 55. Sensitivity of the AAF long model to alternative values of steepness.

The next sensitivity analysis focused on the fixed value of natural mortality used in the coastwide short model. Previous stock assessments have shown that the scale of the estimated spawning biomass and recruitment is very sensitive to natural mortality, and the coastwide short model is the only model where this value is not estimated, a topic of substantial exploration in the 2022 assessment (Stewart and Hicks 2022). Models were fit assuming a fixed value higher (0.18) and lower than the base case (0.15) natural mortality for female Pacific halibut (the value for males is estimated). Results were consistent with previous assessments showing larger biomass estimates for higher values of female natural mortality, but little difference in estimated spawning biomass trends or relative recruitment strengths (Figure 56). Extensive exploration of this model and the potential for estimating this parameter did not indicate a model configuration that produced a reliable value - all tended to favor much higher estimates at whatever upper bound was specified. The estimation of relative male:female selectivity parameters reduction of time-varying processes, as well as different (non-iterated) values for the standard deviation of recruitment variability all produced similar behaviour. At this time it was concluded that natural mortality was not able to be reliably estimated in the coastwide short model. As discussed in the 2022 stock assessment, it would be possible to include uncertainty in the fixed value of natural mortality used in this model via inclusion and appropriate weighting of alternative values in the stock assessment ensemble.

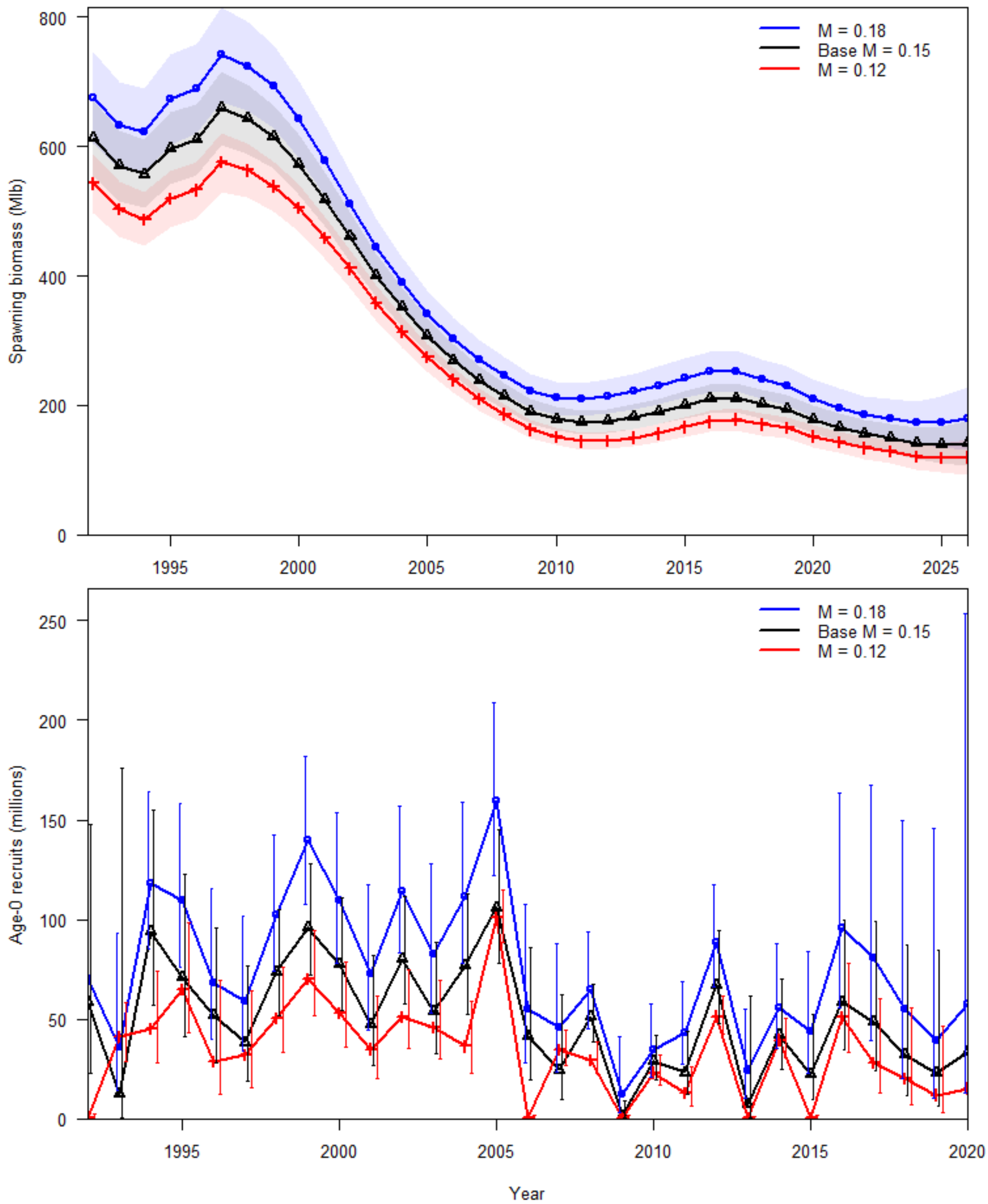


Figure 56. Sensitivity of the coastwide short model to alternative fixed values of female *M*.

Sensitivity to the revised maturity ogive used as the base case for this preliminary stock assessment was included as a step in the bridging analysis described above. The modelled ogive was truncated below the youngest age for which a female Pacific halibut has been observed (age 7). This choice was made to avoid assuming even a small fraction of the much more numerous younger ages was mature without clear evidence suggesting this might be the case. To explore how sensitive the models might be to this choice, alternative models were run without truncating the ogive and allowing a small fraction of the age-5 and age-6 females to be mature (Figure 8). All four models estimated a larger spawning biomass with the updated maturity ogive compared to the historical curve shifted toward older fish (Figure 57-60). The non-truncated ogive had little effect on model results, with only a very slight increase in the estimated spawning biomass.

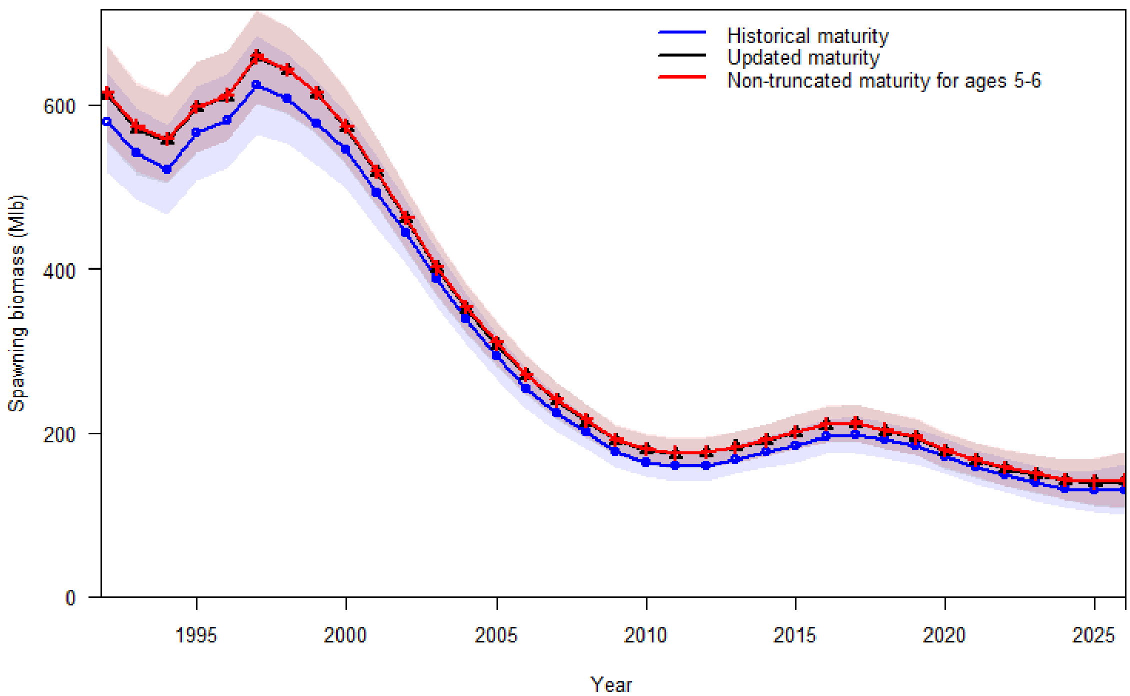


Figure 57. Sensitivity of the coastwide short model to the historical, updated, and non-truncated updated maturity ogives.

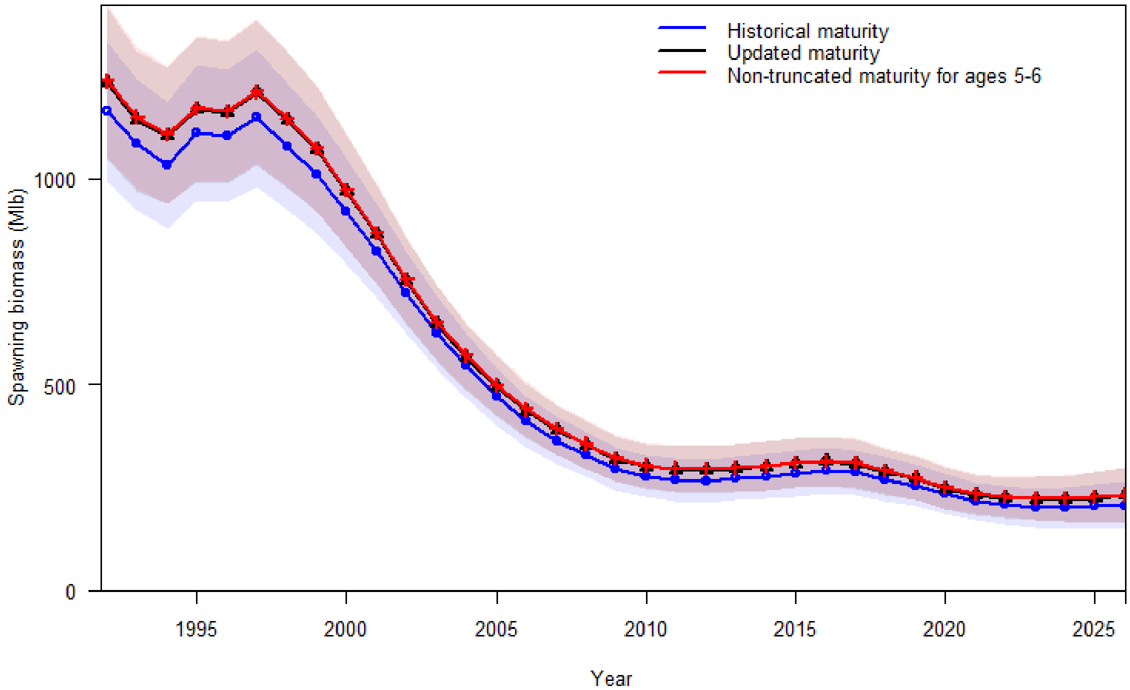


Figure 58. Sensitivity of the AAF short model to the historical, updated, and non-truncated updated maturity ogives.

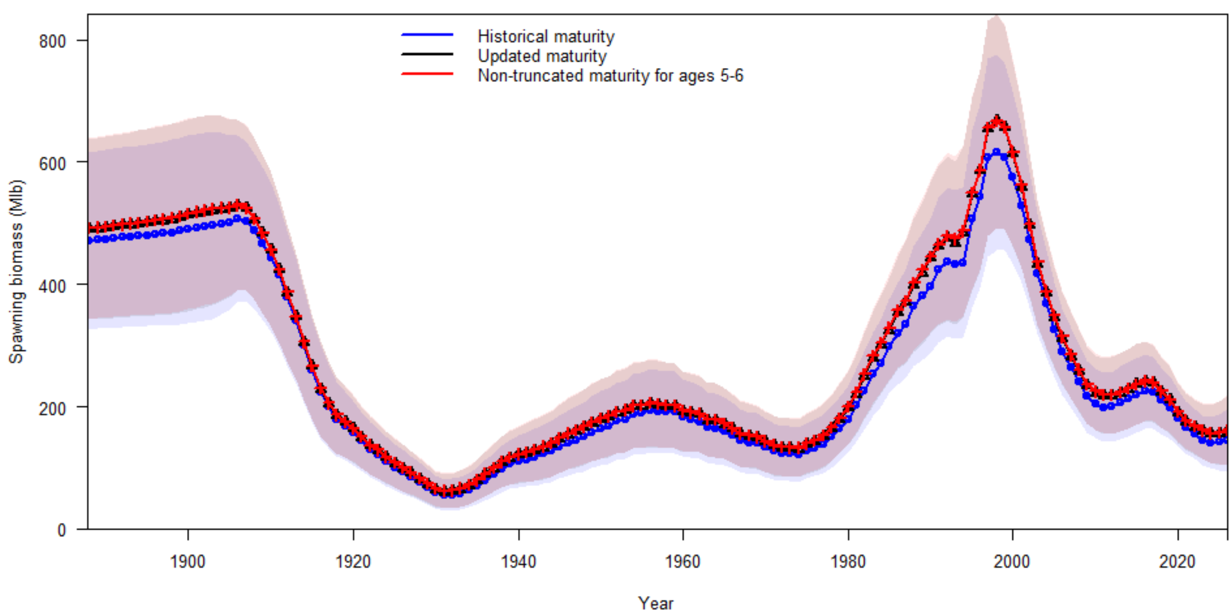


Figure 59. Sensitivity of the coastwide long model to the historical, updated, and non-truncated updated maturity ogives.

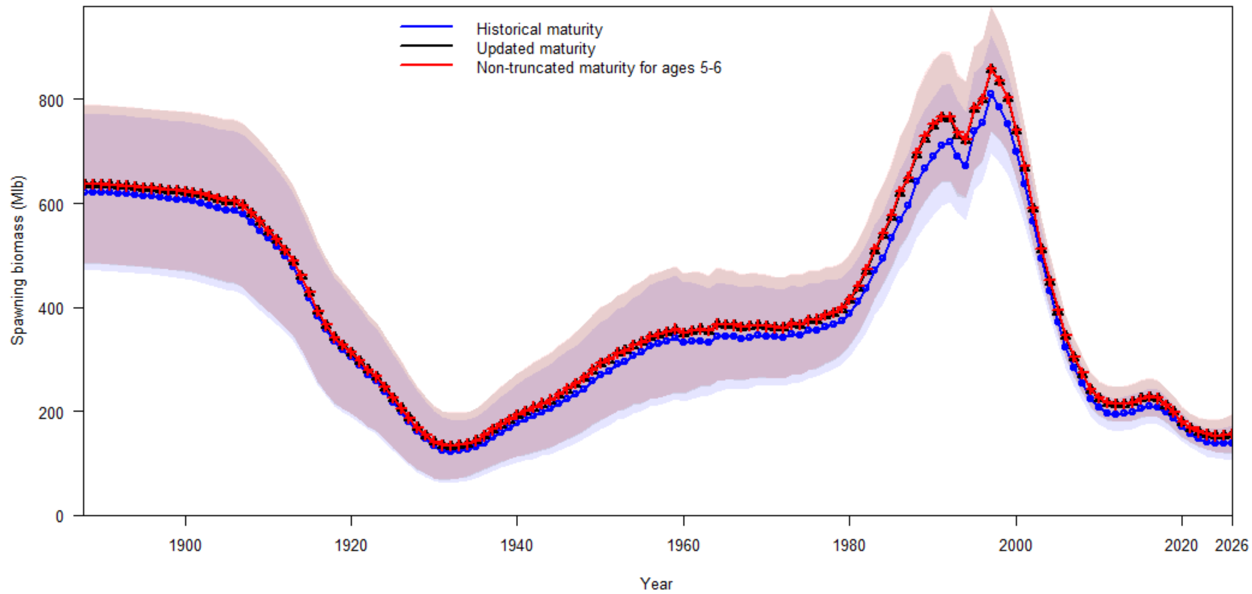


Figure 60. Sensitivity of the AAF long model to the historical, updated, and non-truncated updated maturity ogives.

Additional sensitivity analyses were explored but did not or were not intended to produce reliable models for consideration. The first of these represented an effort to address the lack of fit to Region 2 FISS age composition data. Lack of fit in the coastwide short model showed a clear pattern of large positive residuals at younger ages in the early time-series (until the mid-2000s) and large positive residuals for the older ages in the latter part of the time series (Figure 40). Models were fit to each part (early vs late) of the FISS age composition data separately (but only one of the two periods at the same time) and achieved much improved fit with differing selectivity. However, a similar fit was not produced even when selectivity was allowed to change greatly over this same period. This indicated that there was a catchability component: it appears that spatial availability, particularly for Region 2, may have shifted over time to a degree that cannot be fully captured with changes in selectivity alone. Future modelling could consider allowing time varying catchability (but this would greatly reduce any information in the survey index) or further explore explicitly spatial models.

An additional sensitivity explored a question often raised during public interactions: Could the recent low recruitment since 2006 be explained by increased whale depredation? One hypothesis is that the estimated reduced recruitment may be a function of increased whale depredation on these year classes as they are entering the Pacific halibut fisheries. To explore this hypothesis, the commercial landings and discards were inflated by 50% in each of the four models beginning in 2010 (around the time the 2006 cohort were first entering the catch and being discarded). Models were run with only this change, retaining all other data, but re-estimating all model parameters. The results showed a slightly larger estimated spawning biomass and virtually no change in the time-series of relative recruitment (Figure 61). This result is not unexpected, as the relative recruitment strengths are largely dictated by the compositional data which were unchanged in this sensitivity analysis. A similar analysis, with a three-fold increase in commercial catch produced a much larger spawning biomass but little change in relative recruitment. This sensitivity provides a response to stakeholder concerns and supports the conclusion that recent low productivity is not a direct result of unobserved mortality due to whale depredation on the directed commercial halibut fishery.

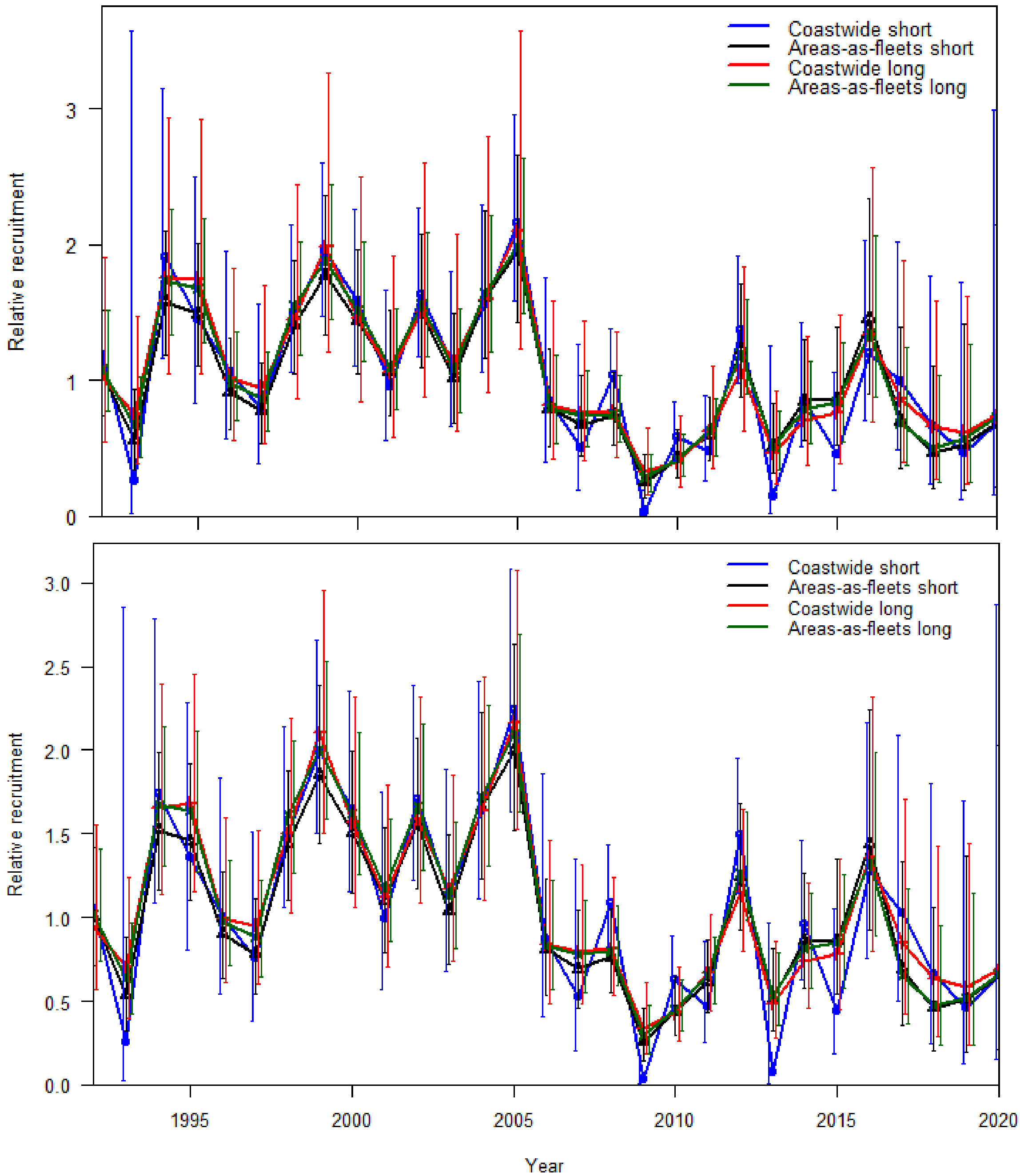


Figure 61. Relative recruitment estimates (divided by the mean of each model) for the preliminary stock assessment (upper panel) and an alternative model assuming whale depredation on the commercial catch (landings + discards) of 50% (lower panel).

Retrospective analyses

The halibut model used from 2006 until 2011 was plagued by a very strong retrospective pattern, both in the scale of the most recent stock size estimates as well as the trend in those estimates (Stewart and Martell 2014; Stewart et al. 2013a). The solution to this problem was additional flexibility for process error (temporal variability) in the selectivity curves for both the fishery and survey representing not just gear (or 'contact') selectivity but also spatial availability.

Retrospective analyses were conducted for these preliminary 2025 models by sequentially removing the terminal eight years of data from the model (a seven-year retrospective, since the terminal year currently contains no information other than mortality projections). Limiting this approach to the most recent eight years of data allows the models to be informed by at least one year of commercial fishery sex-ratio data (2017).

The coastwide short model showed very little retrospective change as the terminal years of data were removed (Figure 62). The AAF short model showed a trend toward higher biomass estimates with a similar trend as data were sequentially removed (Figure 63). This indicates an updating of information informing scale in this model with the most recent observations. Somewhat differently, the coastwide long model showed some increase for some of the most recent years but did not show a strong increase across the entire time-series (Figure 64). Finally, the AAF long model showed a positive retrospective pattern that had changes in both the scale and recent trend (Figure 65). These patterns were more pronounced than those observed in the 2022 stock assessment, but much less pronounced than those found in the 2019 assessment (Stewart and Hicks 2019b). To explore whether the changes in scale could be related to estimates of natural mortality, each retrospective estimate of this parameter from the three models estimating it were compared; however, there were no clear trends (Figures 66).

A further retrospective analysis is based on comparing the spawning biomass estimates among the actual stock assessments conducted since 2012. This type of 'across assessment' retrospective looks at the performance of the stock assessment ensemble as new data and model changes have evolved over time and best reflects the changes actually incorporated into management supporting information. The terminal spawning biomass estimated from most of these assessments are nearly identical to the time series from the preliminary 2025 analysis (Figures 67). However, as has been the focus of much discussion in the 2023 and 2024 stock assessments, the terminal estimates from those analyses both showed a downward revision from the previous year (Stewart and Hicks 2024, 2025). Supplementary and bridging model runs in both of those assessments indicated that the commercial fishery data were providing most of the downward trend; when those data were removed the FISS and other information was very consistent with the previous year's results. This could be due to changes in the fishery, loss of information and/or bias in the FISS, or other unmodelled processes.

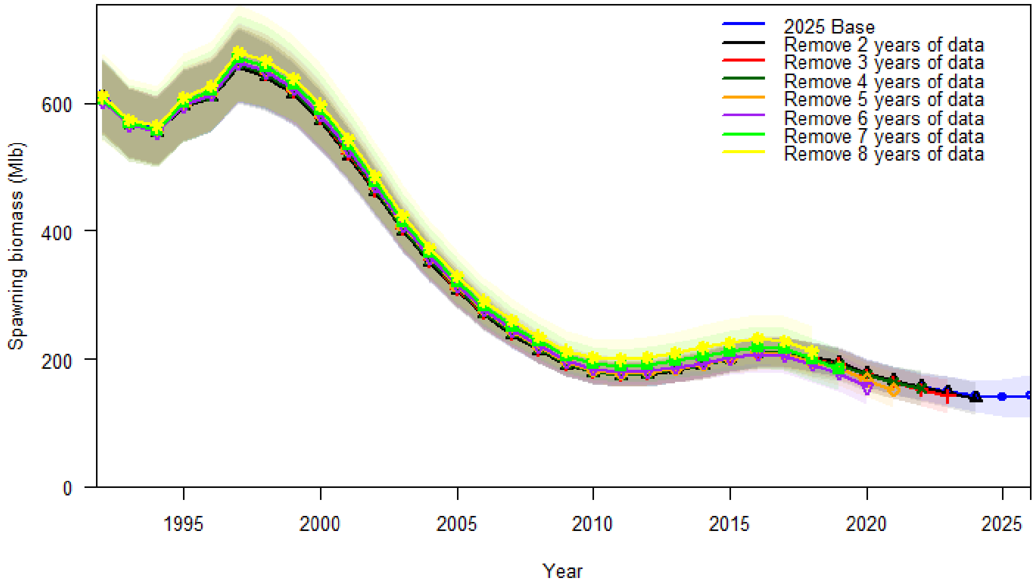


Figure 62. Seven-year retrospective analysis of spawning biomass (there are no data available for 2025 at this time so two years are removed for the first comparison) based on the coastwide short model.

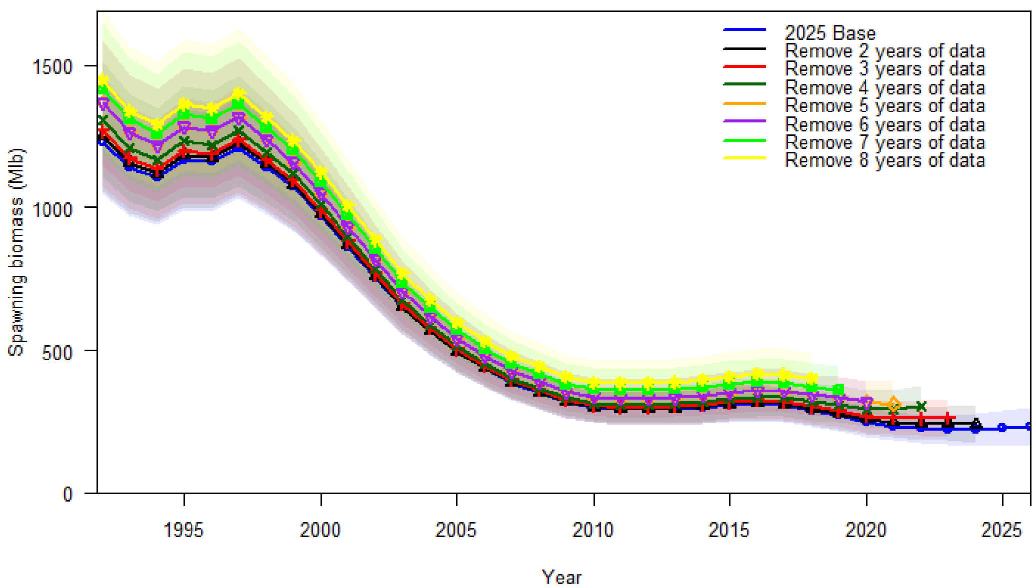


Figure 63. Seven-year retrospective analysis of spawning biomass (there are no data available for 2025 at this time so two years are removed for the first comparison) based on the AAF short model.

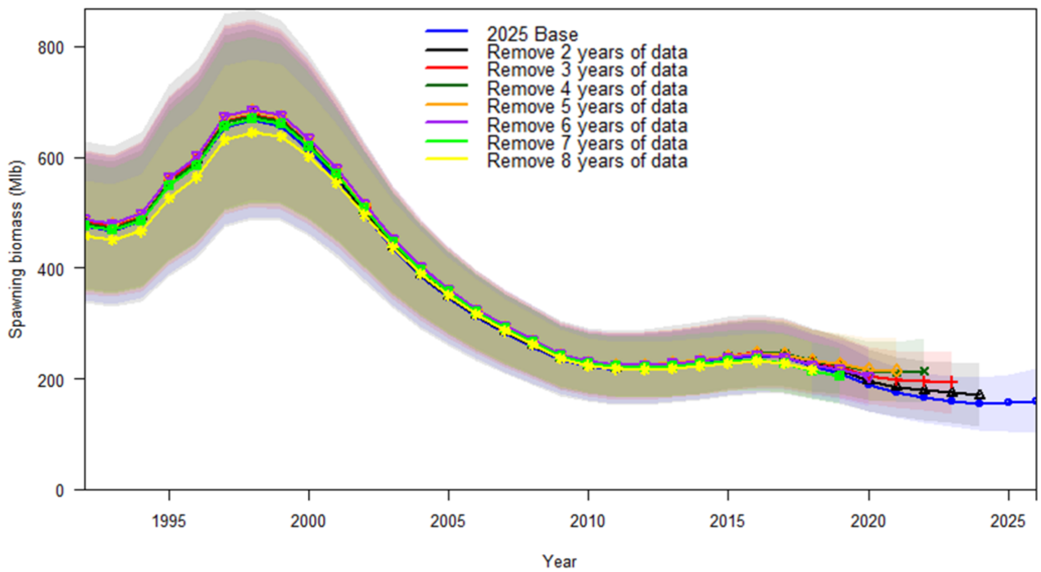


Figure 64. Seven-year retrospective analysis of spawning biomass (there are no data available for 2025 at this time so two years are removed for the first comparison) based on the coastwide long model. Time-series is truncated in 1992 so that differences in the terminal years are more visible.

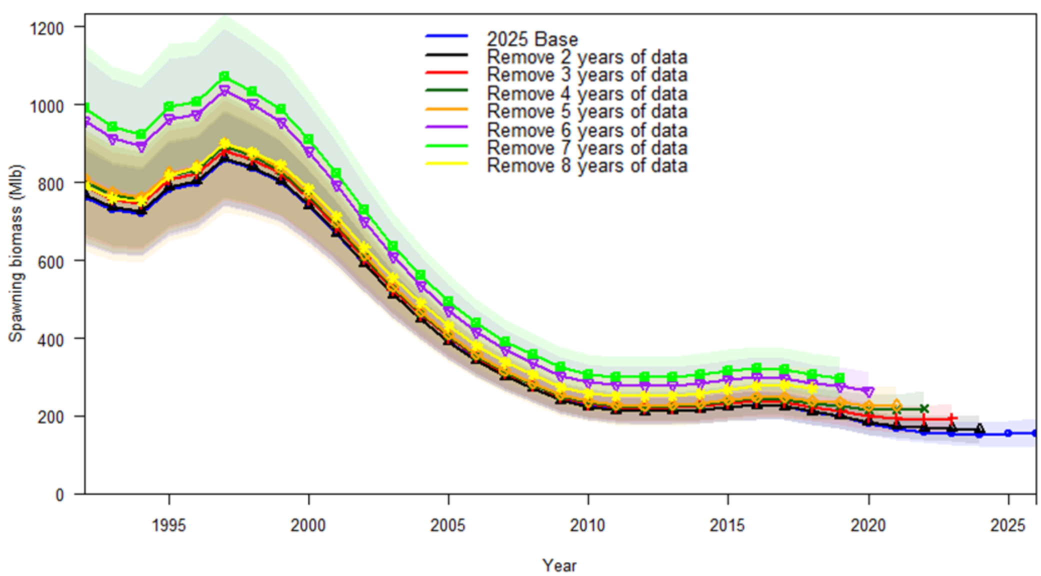
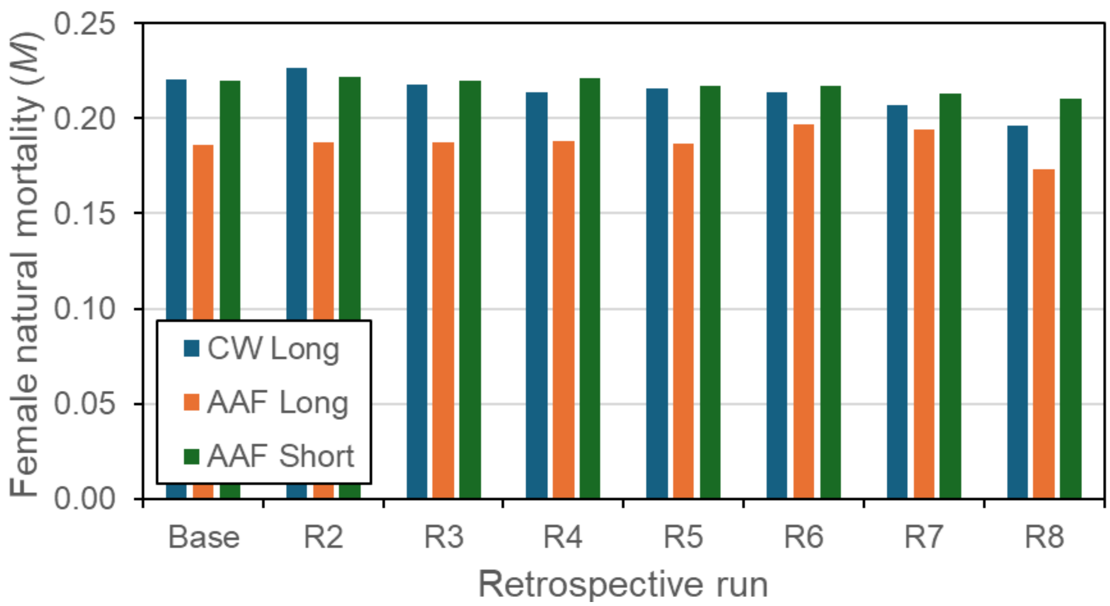
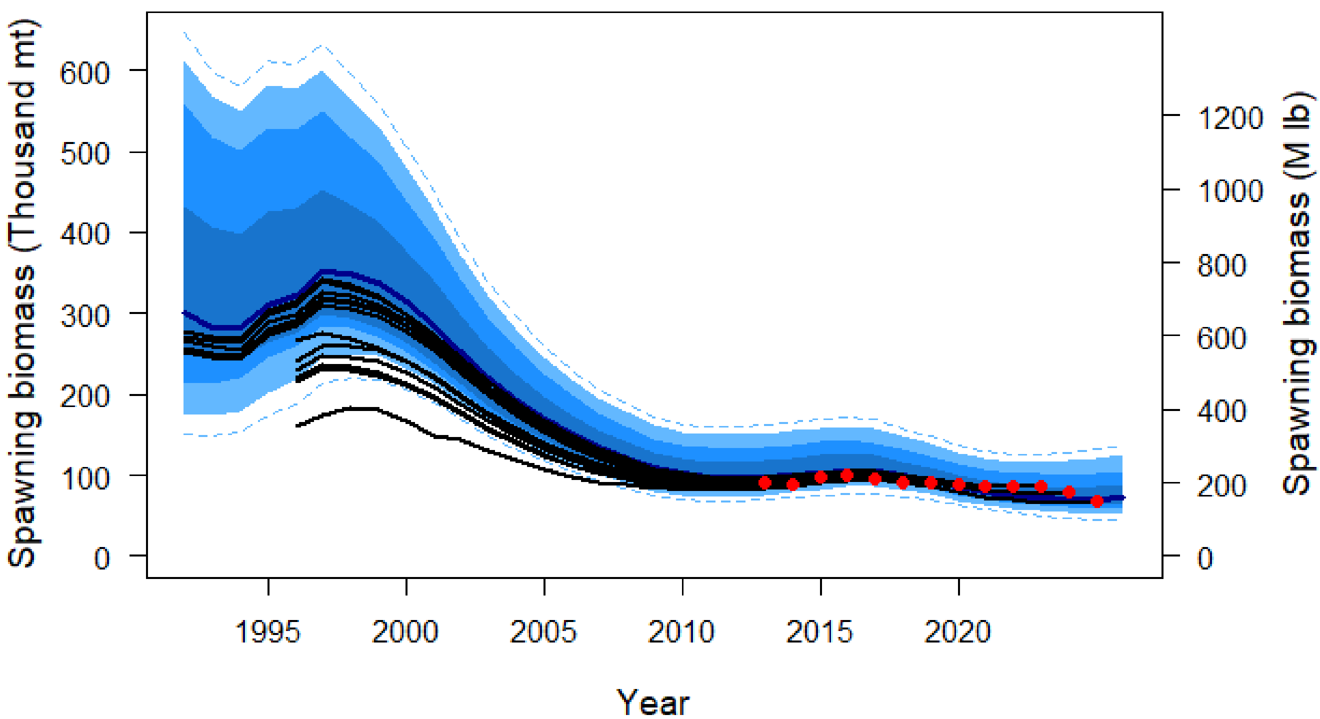


Figure 65. Seven-year retrospective analysis of spawning biomass (there are no data available for 2025 at this time so two years are removed for the first comparison) based on the AAF long model. Time-series is truncated in 1992 so that differences in the terminal years are more visible.



Figures 66. Estimates of female natural mortality (M) over the 7-year retrospective analyses for the three models where this parameter is estimated. The base model includes all data through 2024 (there is no data from 2025 at this preliminary stage); each of the other estimates represent models with two (R2) to seven (R8) years of data removed.



Figures 67. Retrospective analysis of spawning biomass across stock assessments conducted from 2012 to 2024. Red points indicate the terminal estimate from each stock assessment; shaded region indicates the uncertainty around the median ensemble estimate (solid blue line) from the preliminary 2025 stock assessment.

Bayesian analysis

The 2019 stock assessment included a substantial evaluation of Bayesian integration for the short coastwide model (Stewart and Hicks 2019b). This effort did not produce substantially different results from the maximum likelihood and asymptotical variance methods (Fournier et al. 2012) routinely employed. However, there are a number of potential benefits to using an explicitly Bayesian approach, including better characterization of uncertainty (Magnusson et al. 2012) and a more directly interpretable characterization of the probability distributions. There is also the potential for differences in the results of Bayesian analyses due to the right-skewed nature of some distributions for key parameter and management-related quantities in complex fisheries models (Stewart et al. 2013b).

In aggregate, the 2019 results suggested that the asymptotic distributions were a reasonable approximation for the full posterior distributions in these models, and also that the process of regularizing the selectivity parameters and removing some deviations to improve integration did not having an appreciable effect on the solution. This is generally consistent with studies of process error where overparameterizing (adding the capability for variation when it wasn't present) was generally found to be unbiased, and therefore preferable to underparameterizing when temporal variability was present (e.g., Martell and Stewart 2014; Stewart and Monnahan 2017).

Additional Bayesian analysis was not included in this (or the 2022) assessments. However, if a multi-year assessment approach was to become part of a future management procedure for the IPHC more time could be devoted to exploring Bayesian models.

Other uncertainty considerations

There are many important sources of uncertainty not captured in the four models included in this ensemble. These include myriad alternative structural assumptions such as spatially-explicit population dynamics, connection with Russian waters, alternative stock-recruitment functions, time-varying mortality, different data weighting approaches, and many others. There are also several tractable sources of projection uncertainty that are not in the current approach, including uncertainty in projected weight-at-age (although the sensitivity of this was investigated at SRB request in 2016 and found to be low) and uncertainty in the realized mortality associated with limits set by the Commission.

Within the modelled time-series there are also data-related uncertainties that could be addressed via a range of alternative approaches. Uncertainty in the time series of mortality for these models is not currently captured, as they are treated as inputs and assumed to be known without error. In previous assessments, sensitivity analyses have been conducted to the degree of discard mortality in the commercial fishery, potential effects of unobserved whale depredation, as well as to the magnitude of total bycatch mortality. In concept, these types of uncertainties could be explicitly included in the models; however, full estimation of catch in statistical catch-at-age models generally requires other stabilizing assumptions, so direct integration of this uncertainty may still prove challenging. Additional sources of uncertainty and avenues for development are identified in the Research Priorities section below.

The ensemble

Model-integrated quantities are used as the primary stock assessment output for management use, as well as the basis for decision table probabilities (Stewart and Hicks 2025). All quantities of management interest are integrated for the recent time period (1992+), for which all four sets of model results are available. These quantities include: spawning biomass, relative spawning biomass, and the Spawning Potential Ratio (SPR; summarized as fishing intensity, $F_{XX\%}$, where the *XX%* represents *SPR*). Decision table quantities are divided into four categories: stock trend (which is the only set of metrics that are independent of any harvest strategy related assumptions), stock status, fishery trend, and fishery status. Integration is performed for all these quantities using the basic approach outlined below.

Methods

The basic approach to model integration remains unchanged from the 2015 and subsequent analyses. A sample of random draws is created from the output from each of the models included in the ensemble. For the spawning biomass time-series, the estimates and associated standard deviations for female spawning biomass from each of the four models were extracted from the report file. A vector of length n is created for each model (m), where the relative weight (w_m) is simply the relative fraction of the total draws across all models comprised by n_m:

$$w_m = \frac{n_m}{\sum_m n_m}$$

This approach allows for easily adjusted weighting of models. Routine reporting of results uses $\sum_m n_m$ for all models equal to twenty million; this has been found to produce negligible Monte-Carlo error even in the tails of extremely skewed distributions, creating robust and stable reporting of all quantities of interest with a smooth distribution. Although this choice could potentially be optimized for each statistic of interest, current integration code (in *R*) does not represent a constraining step in the analysis.

The harvest strategy employs a control rule that reduces the coastwide SPR target linearly from the interim 'reference level' at $SB_{30\%}$ to zero at $SB_{20\%}$. Since the 2019 assessment this calculation uses a dynamic estimate of 'unfished' biomass calculated for each year of the time-series. This calculation replays the entire time-series, without the fishing mortality, assuming the same parameter values (including recruitment deviations) but accounting for the different level of spawning biomass projected for each year and its effect on subsequent expected (pre-deviation) recruitment in each year. Since 2020 the dynamic unfished biomass calculation has been included simultaneously with variance calculations of all model parameters and outputs and (importantly) includes the covariance in the estimated and unfished dynamic spawning biomass in the variance of the IPHC's reference points and other outputs.

Evaluation of weighting based on predictive skill

All Pacific halibut assessments since 2014 have relied on equal weighting of all four models. However, weighting based on several potential approaches has been considered since the 2015 stock assessment (Stewart and Martell 2016). Briefly, these have included:

AIC – but this is known to be highly dependent on data weighting, and can only be applied in cases where the same data sets are being fit by all models under consideration

Strength of retrospective patterns – perhaps relative to a 'null' distribution for a statistic like Mohn's rho (Mohn 1999) based on simulation (Hurtado-Ferro et al. 2015); while helpful to diagnose model performance, it does not necessarily indicate a 'good' model, as evidenced by the fact that a static prediction will have no retrospective pattern at all.

Fit to the FISS index – without an AIC-type correction, there is no penalty for overparameterized models

Expert opinion – this is subjective, and the tendency has been to revert to equal weighting in the absence of strong evidence to the contrary.

Mean Absolute Standardized Error (MASE; Hyndman and Koehler 2006) – a measure of predictive skill.

Most recently in the 2022 full stock assessment, the MASE statistic was extensively evaluated (Stewart and Hicks 2022). After considerable exploration and review the SRB recommended against moving forward with MASE-based model weighing (IPHC 2022). Model weighting has not yet been revised or explored further for 2025.

Preliminary results for 2025

Comparison of the spawning biomass estimates from the four stock assessment models comprising the ensemble shows that the 95% intervals from any single model are substantially narrower than the aggregate (Figure 68). All four models indicate a similar overall trajectory, including the small increase in biomass over 2011-2016 and subsequent decrease as the effects of reduced recruitment subsequent to 2006 (Figure 61; upper panel) graduate through to the spawning biomass. The AAF long model provides the largest estimate of the beginning of year 2026 spawning biomass; however, that distribution still contains the estimates from the other three models (Figure 69).The differing estimates of natural mortality in each of the four models result in recruitment (at age-0) of differing scales; however when divided by the average for each model trends in recruitment are very similar across all models (Figure 70).

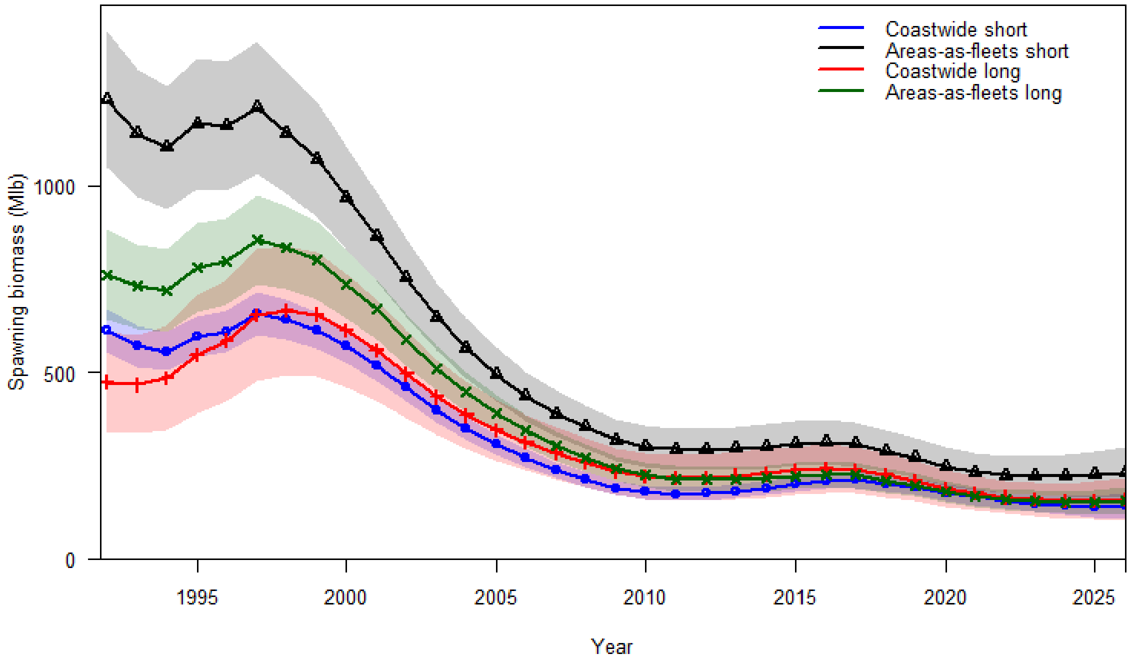


Figure 68. Comparison of spawning biomass time series (shaded regions indicate asymptotic approximations to the 95% confidence interval) from each of the preliminary models contributing to the 2025 preliminary ensemble.

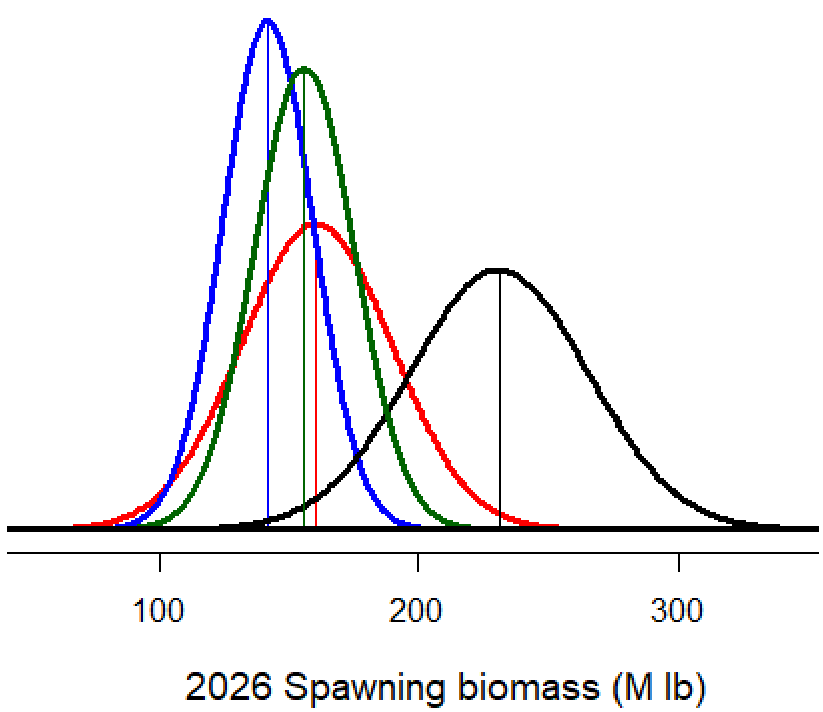


Figure 69. Comparison of terminal (2026) spawning biomass estimates (pdfs) from each of the preliminary models contributing to the 2025 preliminary ensemble. Vertical lines represent the median value from each model.

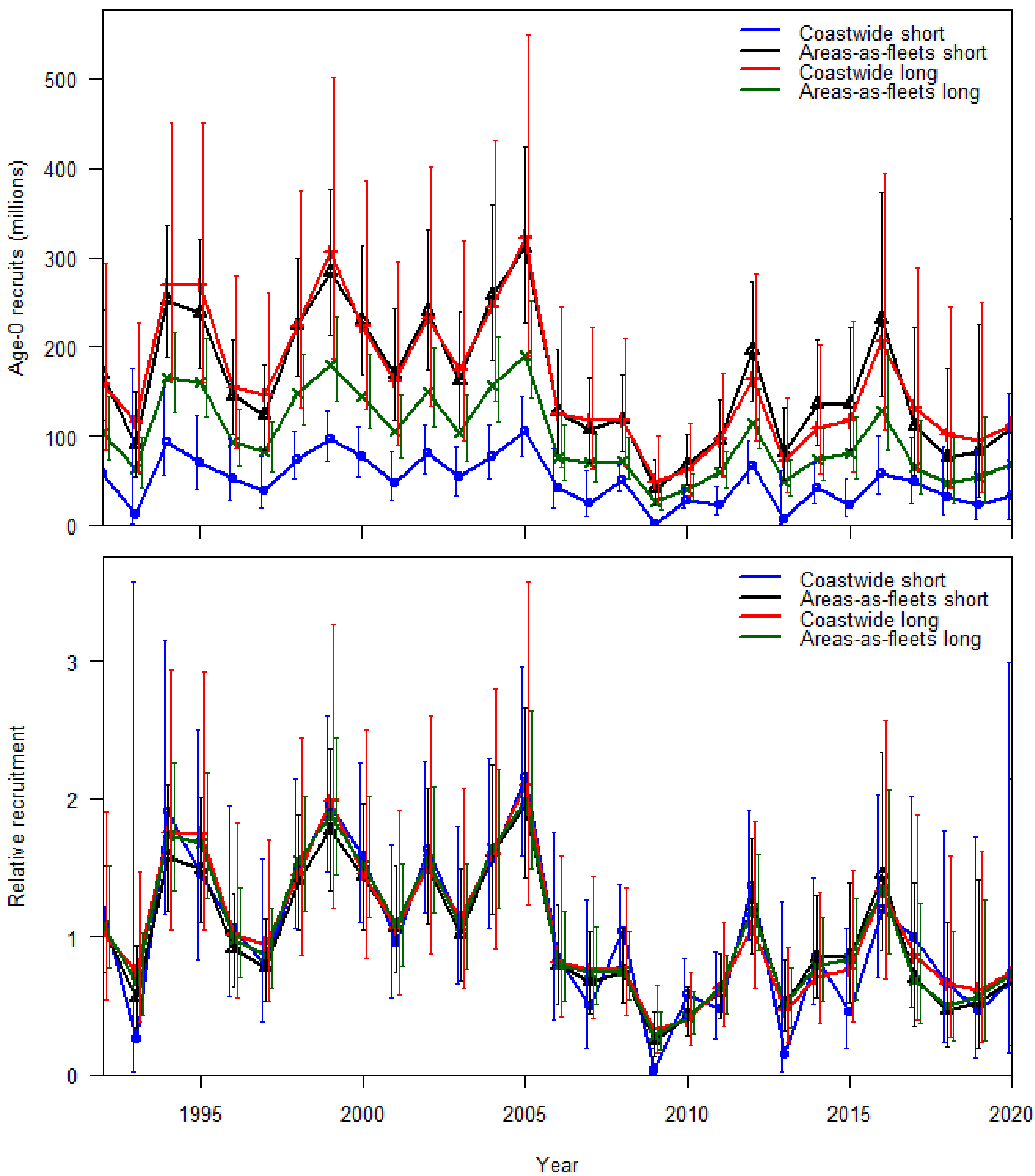


Figure 70. Comparison of recruitment time series (upper panel; vertical lines indicate asymptotic approximations to the 95% confidence interval) and relative recruitment series (each standardized to its mean; lower panel) from each of the preliminary models contributing to the 2025 ensemble.

Future development

Several extensions to this preliminary assessment will be possible for the final 2025 analysis. These include:

- Responses to suggestions and comments generated from SRB026 and SRB027.
- Addition of all 2025 data, extending existing time series (mortality, indices, ages, etc.).
- The sex-ratio of the 2024 commercial fisheries landings based on the IPHC's genetic assay will be available by late summer.

In addition to the list of research priorities (longer list below), there are several potential avenues for development within and among the four models included in the ensemble.

The updated bootstrapping performed for this assessment provides a strong basis for objective interannual and among fleet weighting of age composition data. Both alternative likelihoods, including those already evaluated to some degree for this assessment over the last several years (e.g., the Dirichlet-multinomial, logistic normal) and alternative calculations of composition residuals (e.g., One-Step-Ahead (OSA) residuals; Thygesen et al. 2017; Trijoulet et al. 2023) are strong candidates for further investigation. A considerable effort exploring the properties of OSA residuals was made as part of this 2025 stock assessment, and a draft manuscript has been produced. Incorporation of that approach may be possible for the next Pacific halibut assessment.

Other avenues for development include changes to the ensemble approach itself. The 2019 assessment explored expanding the number of models included in the ensemble to better capture the uncertainty in M that was missed through using a fixed value in the two (at that time) short time-series models. By estimating M for the short AAF model in the 2022 stock assessment, the integration of uncertainty was improved. Upcoming assessments may need to explore whether the fixed value of 0.15 in the coastwide short model is still appropriate given the increasing weight of evidence that M for Pacific halibut is higher.

As ensemble changes are evaluated, both weighting and technical efficiency should be considered. Technical costs of adding additional models to the ensemble include additional time spent running these additional models rather than exploring other sensitivities and identifying clear effects of newly available data during the very short assessment analysis period each fall. Pragmatically, there may be relatively little to be gained from increasing the ensemble in this manner beyond slightly smoother integrated distributions. As the IPHC's management procedure evolves, to potentially include multi-year assessments, there may be additional latitude for increased model and ensemble complexity.

The current ensemble is based on maximum likelihood estimates and asymptotic approximations to the posterior distributions for model parameters and derived quantities. Bayesian posteriors represent a conceptually more appealing basis for probability distributions, and could better capture the full range and potential asymmetries in the distributions for model quantities (Magnusson et al. 2012; Stewart et al. 2013b). Bayesian integration may also allow for statistically correct treatment of variance parameters (such as the sigmas governing recruitment variability and selectivity or catchability process error), as would use of true random effects methods. Although it would be technically preferable to regularize and run all four assessment models as Bayesian analyses, at present this is technically infeasible given the tight time-line between data availability and the deadline for the annual stock assessment. The analysis time difference between minimization and full posterior integration, even using the most efficient methods available for the coastwide short model (see section above), is still too large. However, if the IPHC were to move to a more formal management procedure and/or to a multi-year mortality limit-setting process, the stock assessment could be conducted at a pace that would allow much greater reliance on Bayesian models.

Research priorities

The development of the IPHC's research priorities has been closely tied to the needs of the stock assessment and harvest strategy policy analyses, such that the IPHC's research projects will provide data, and hopefully knowledge, about key biological and ecosystem processes that can then be incorporated directly into analyses supporting the management of Pacific halibut. Research priorities for the Pacific halibut stock assessment are delineated into three broad categories: improvements in basic biological understanding (including fishery dynamics), investigation of existing data series and collection of new information, and technical development of models and modelling approaches. The highest priority items in each of these categories are highlighted in the 5-year research plan and are expected to be the primary focus of ongoing efforts. However, it is helpful to maintain a longer list of items to inform future prioritization, to create a record of data and research needs, and to foster opportunistic and/or collaborative work on these topics when possible.

Biological understanding and fishery yield

Key areas for improvement in biological understanding include:

- *Highest priority*: Updating the fecundity-weight relationship and the presence and/or rate of skip spawning.
- *Highest priority*: The relative role of potential factors underlying changes in size-at-age is not currently understood. Delineating between competition, density dependence, environmental effects, size-selective fishing and other factors could allow improved prediction of size-at-age under future conditions.
- Movement rates among Biological Regions at the adult, juvenile and larval stages remain uncertain and likely variable over time. Long-term research to inform these rates could lead to a spatially explicit stock assessment model for future inclusion into the ensemble.
- Improved understanding of recruitment processes and larval dynamics could lead to covariates explaining more or the residual variability about the stock-recruit relationship than is currently accounted for via the binary indicator used for the Pacific Decadal Oscillation.

Data related research

This section represents a list of potential projects relating specifically to existing, and new data sources that could benefit the Pacific halibut stock assessment:

- *Highest priority*: Continued collection of sex-ratio from the commercial landings will provide valuable information for determining relative selectivity of males and females, and therefore the scale of the estimated spawning biomass, and the level of fishing intensity as measured by SPR.
- *Highest priority*: Evaluation of the magnitude of marine mammal depredation and tools to reduce it.
- A space-time model could be used to calculate weighted FISS and/or commercial fishery age-composition data. This might alleviate some of the lack of fit to existing data sets that is occurring not because of model misspecification but because of incomplete spatial coverage in the annual FISS sampling which is accounted for in the generation of the index, but not in the standardization of the composition information.
- The work of Monnahan and Stewart (2015) modelling commercial fishery catch rates could be used to provide a standardized fishery index for the recent time-series that would be analogous to the space-time model used for the FISS.
- There is a vast quantity of archived historical data that is currently inaccessible until organized, electronically entered, and formatted into the IPHC's database with appropriate meta-data. Information on historical fishery landings, effort, and age samples would provide a much clearer (and more reproducible) perception of the historical period.
- Additional efforts could be made to reconstruct estimates of subsistence harvest prior to 1991.
- Discard mortality estimates for the IPHC Regulatory Area 2B recreational fishery are currently unavailable, but there is an estimation system in place. Further work to develop these estimates would be preferable to the use of proxy rates from IPHC Regulatory Area 2C.
- NMFS observer data from the directed Pacific halibut fleet in Alaska could be evaluated for use in updating discard mortality rates and the age-distributions for discard mortality. This may be more feasible if observer coverage is increased and if smaller vessels (< 40 feet LOA, 12.2 m) are observed in the future. Post-stratification and investigation of observed vs. unobserved fishing behavior may be required.
- Historical bycatch length frequencies and mortality estimates should be reanalyzed accounting for sampling rates in target fisheries and evaluating data quality over the historical period.

- There are currently no comprehensive variance estimates for the sources of mortality used in the assessment models. In some cases, variance due to sampling and perhaps even non-sampling sources could be quantified and used as inputs to the models via scaling parameters or even alternative models in the ensemble.

Technical development

There are a variety of technical explorations and improvements that could benefit the stock assessment models and ensemble framework. Larger changes (such as entirely new data sets) naturally fit into full assessment analyses; however, incremental changes may be possible during updated assessments when and if new information or methods become available. Specifically, development is intended to occur in time for initial SRB review (generally in June), with primarily only refinements made for final review (October), such that untested approaches are not being implemented during the annual stock assessment itself. Technical research priorities include:

- *Highest priority*: Maintaining consistency and coordination between MSE, and stock assessment data, modelling and methodology.
- *Highest priority*: Exploration of state-space models for Pacific halibut allowing for direct estimation of the variance in time-varying processes.
- *Highest priority*: Continued exploration into the estimation of M in the short coastwide model.
- Continued refinement of the ensemble of models used in the stock assessment. This may include investigation of alternative approaches to modelling selectivity that would reduce relative down-weighting of certain data sources (see section above), evaluation of additional axis of uncertainty (e.g., steepness, as explored above), or others.
- Exploration of methods for better including uncertainty in directed and non-directed discard mortalities in the assessment (now evaluated only via alternative mortality projection tables or model sensitivity tests) in order to better include these sources uncertainty in the decision table. These could include explicit discard/retention relationships, including uncertainty in discard mortality rates, and allow for some uncertainty directly in the magnitude of mortality for these sources.
- Bayesian methods for fully integrating parameter uncertainty may provide improved uncertainty estimates within the models contributing to the assessment, and a more natural approach for combining the individual models in the ensemble (see section above).
- Alternative model structures, including a growth-explicit statistical catch-at-age approach and a spatially explicit approach may provide avenues for future exploration. Efforts to develop these approaches thus far have been challenging due to the technical complexity and data requirements of both. Previous reviews have indicated that such

efforts may be more tractable in the context of operating models for the MSE, where conditioning to historical data may be much more easily achieved than fully fitting an assessment model to all data sources for use in tactical management decision making.

Acknowledgements

IPHC datasets comprise a wide array of sources based on extensive sampling and reporting efforts by state and national agencies in the U.S. and Canada. The IPHC's annual stock assessment benefits from the hard work of all of its current and former employees providing high-quality data sets as comprehensive as any used for fisheries analysis. The Scientific Review Board and national science advisors have provided extensive guidance and constructive criticism of the treatment of data sources, the individual models and the stock assessment ensemble. Ray Webster leads, or contributes to, many of the supporting data analyses on which the assessment is based.

References

Adams, G.D., Holsman, K.K., Barbeaux, S.J., Dorn, M.W., Ianelli, J.N., Spies, I., Stewart, I.J., and Punt, A.E. 2022. An ensemble approach to understand predation mortality for groundfish in the Gulf of Alaska. Fisheries Research **251**. doi:10.1016/j.fishres.2022.106303.

Bosley, K.M., Schueller, A.M., Goethel, D.R., Hanselman, D.H., Fenske, K.H., Berger, A.M., Deroba, J.J., and Langseth, B.J. 2021. Finding the perfect mismatch: Evaluating misspecification of population structure within spatially explicit integrated population models. Fish and Fisheries. doi:10.1111/faf.12616.

Clark, W.G. 2003. A model for the world: 80 years of model development and application at the international Pacific halibut commission. Natural Resource Modeling **16**(4): 491-503.

Clark, W.G. 2004a. Statistical distribution of IPHC age readings. IPHC Report of Assessment and Research Activities 2003. p. 99-110.

Clark, W.G. 2004b. Nonparametric estimates of age misclassification from paired readings. Canadian Journal of Fisheries and Aquatic Sciences **61**: 1881-1889.

Clark, W.G., and Hare, S.R. 2002. Effects of climate and stock size on recruitment and growth of Pacific halibut. North American Journal of Fisheries Management **22**: 852-862.

Clark, W.G., and Hare, S.R. 2006. Assessment and management of Pacific halibut: data, methods, and policy. International Pacific Halibut Commission Scientific Report No. 83, Seattle, Washington. 104 p.

Clark, W.G., Hare, S.R., Parma, A.M., Sullivan, P.J., and Trumble, R.J. 1999. Decadal changes in growth and recruitment of Pacific halibut (*Hippoglossus stenolepis*). Canadian Journal of Fisheries and Aquatic Sciences **56**: 242-252.

Cox, S.P., Ianelli, J., and Mangel, M. 2014. IPHC Scientific Review Board Meeting. June 2014., 3 p.

Cox, S.P., Ianelli, J., and Mangel, M. 2017. 4.7 Reports of the IPHC Scientific Review Board (SRB). IPHC Report of Assessment and Research Activities 2016. p. 443-452.

Du, J. 2014. Uncertainty and ensemble forecast. National Weather service. 42 p.

Forsberg, J.E., and Stewart, I.J. 2015. Re-ageing of archived otoliths from the 1920s to the 1990s. IPHC Report of Assessment and Research Activities 2014. p. 405-428.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software **27**(2): 233-249.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences **68**: 1124-1138.

Francis, R.I.C.C. 2016. Revisiting data weighting in fisheries stock assessment models. Fish. Res. doi:10.1016/j.fishres.2016.06.006.

Goethel, D.R., and Berger, A.M. 2017. Accounting for spatial complexities in the calculation of biological reference points: effects of misdiagnosing population structure for stock status indicators. Canadian Journal of Fisheries and Aquatic Sciences **74**(11): 1878-1894. doi:10.1139/cjfas-2016-0290.

Hamel, O.S. 2014. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science **72**(1): 62-69. doi:10.1093/icesjms/fsu131.

Hamel, O.S., and Cope, J.M. 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. Fisheries Research **256**. doi:10.1016/j.fishres.2022.106477.

Hamel, O.S., Ianelli, J.N., Maunder, M.N., and Punt, A.E. 2023. Natural mortality: Theory, estimation and application in fishery stock assessment models. Fisheries Research **261**. doi:10.1016/j.fishres.2023.106638.

Hamill, T.M., Brennan, M.J., Brown, B., DeMaria, M., Rappaport, E.N., and Toth, Z. 2012. NOAA's Future Ensemble-Based Hurricane Forecast Products. Bulletin of the American Meteorological Society **93**: 209-220.

Hicks, A., and Stewart, I. 2025. Updates to the IPHC MSE and a review of coastwide management procedures. IPHC-2025-MSAB021-07. 18 p.

Hulson, P.-J.F., and Williams, B.C. 2024. Inclusion of ageing error and growth variability using a bootstrap estimation of age composition and conditional age-at-length input sample size for fisheries stock assessment models. Fisheries Research **270**. doi:10.1016/j.fishres.2023.106894.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., and Punt, A.E. 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science **72**(1): 99-110. doi:10.1093/icesjms/fsu198.

Hutniczak, B., Tran, H., and Kong, T. 2025. Fisheries data overview. IPHC-2025-AM101-08 Rev_2. 15 p.

Hyndman, R.J., and Koehler, A.B. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting **22**(4): 679-688. doi:10.1016/j.ijforecast.2006.03.001.

Ianelli, J., Holsman, K.K., Punt, A.E., and Aydin, K. 2016. Multi-model inference for incorporating trophic and climate uncertainty into stock assessments. Deep Sea Research Part II: Topical Studies in Oceanography **134**: 379-389. doi:10.1016/j.dsr2.2015.04.002.

Ianelli, J., Fissel, B., Stienessen, S., Honkalehto, T., Siddon, E., and Allen Akselrud, C. 2021. Assessment of the walleye pollock stock in the Eastern Bering Sea. Chapter 1. NPFMC Bering Sea and Aleutian Islands SAFE. 171 p.

IPHC. 2022. Report of the 21st session of the IPHC Scientific review board (SRB021). IPHC-2022-SRB021-R.

IPHC. 2025. Report of the 101st session of the IPHC Annual Meeting (AM101). Vancouver, BC, Canada, 27-31 January 2025. IPHC-2025-AM101-R. 52 p.

Karp, M.A., Blackhart, K., Lynch, P.D., Deroba, J., Hanselman, D., Gertseva, V., Teo, S., Townsend, H., Williams, E., and Yau, A. 2018. Proceedings of the 13th National Stock Assessment Workshop: Model Stability, Model Complexity, & Ensemble Modelling. U.S. Dept. of Commerce, NOAA. NOAA Tech. Memo. NMFS-F/SPO-189. 49 p.

Kong, T., Tran, H., and Prem, C. 2022. Fisheries data overview (2021). IPHC-2022-AM098-06 Rev_1. 15 p.

Li, Y., Bence, J.R., and Brenden, T.O. 2014. An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. ICES Journal of Marine Science **72**(1): 70-81. doi:10.1093/icesjms/fsu057.

Litzow, M.A., Hunsicker, M.E., Bond, N.A., Burke, B.J., Cunningham, C.J., Gosselin, J.L., Norton, E.L., Ward, E.J., and Zador, S.G. 2020. The changing physical and ecological meanings of North Pacific Ocean climate indices. Proceedings of the National Academy of Sciences of the United States of America **117**(14): 7665-7671. doi:10.1073/pnas.1921266117.

Magnusson, A., Punt, A.E., and Hilborn, R. 2012. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. Fish and Fisheries. doi:10.1111/j.1467-2979.2012.00473.x.

Mangel, M., MacCall, A.D., Brodziak, J., Dick, E.J., Forrest, R., Pourzand, R., and Ralston, S. 2013. A perspective on steepness, reference points, and stock assessment. Canadian Journal of Fisheries and Aquatic Sciences **70**(6): 930-940.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.R., and Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society **78**(6): 1069-1079.

Martell, S., and Stewart, I. 2014. Towards defining good practices for modeling time-varying selectivity. Fisheries Research **158**: 84-95. doi:10.1016/j.fishres.2013.11.001.

Maunder, M.N., Crone, P.R., Valero, J.L., and Semmens, B.X. 2015. Growth: theory, estimation, and application in fishery stock assessment models. CAPAM Workshop Series Report 2. 57 p.

McGilliard, C.R., and Palsson, W. 2021. Assessment of the rex sole stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE. Chapter 6. 83 p.

McGilliard, C.R., Punt, A.E., Methot, R.D., Hilborn, R., and Jacobson, L. 2014. Accounting for marine reserves using spatial stock assessments. Canadian Journal of Fisheries and Aquatic Sciences: 1-19. doi:10.1139/cjfas-2013-0364.

Methot Jr, R.D. 2024. Stock Synthesis User Manual Version 3.30.23.1. NOAA Fisheries. Seattle, WA. December 5, 2024. 272 p.

Methot Jr, R.D., Wetzel, C.R., Taylor, I.G., Doering, K.L., and Johnson, K.F. 2021a. Stock synthesis user manual version 3.30.18. NOAA Fisheries. Seattle, WA. 234 p.

Methot Jr, R.D., Wetzel, C.R., Taylor, I.G., Doering, K.L., and Johnson, K.F. 2021b. Stock synthesis user manual version 3.30.17. NOAA Fisheries, Seattle, WA. 238 p.

Methot, R.D., and Taylor, I.G. 2011. Adjusting for bias due to variability in estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences **68**: 1744-1760.

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

Methot, R.D., and Wetzel, C.R. 2013b. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Appendix A: Technical description of the Stock Synthesis assessment program. Fisheries Research **142**: 26 p.

Methot, R.D., Wetzel, C.R., and Taylor, I.G. 2019. Stock Synthesis User Manual Version 3.30.13. NOAA Fisheries. Seattle, WA. 213 p.

Methot, R.D., Wetzel, C.R., Taylor, I.G., Doering, K.L., Perl, E.F., and Johnson, K.F. 2024. Stock Synthesis user manual version 3.30.22.1. NOAA Fisheries, Seattle, Washington. January 31, 2024. 256 p.

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

Myers, R.A., Bowen, K.G., and Barrowman, N.J. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences **56**: 2404-2419.

Piner, K.R., and Wischnioski, S.G. 2004. Pacific halibut chronology of bomb radiocarbon in otoliths from 1944 to 1981 and a validation of ageing methods. Journal of Fish Biology **64**: 1060-1071.

Punt, A.E., Haddon, M., and Tuck, G.N. 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia. Fisheries Research **168**: 85-99. doi:10.1016/j.fishres.2015.04.002.

Stewart, I., and Hicks, A. 2019a. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2018. IPHC-2019-AM095-09. 26 p.

Stewart, I., and Hicks, A. 2019b. 2019 Pacific halibut (*Hippoglossus stenolepis*) stock assessment: development. IPHC-2019-SRB014-07. 100 p.

Stewart, I., and Hicks, A. 2020. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2019. IPHC-2020-SA-01. 32 p.

Stewart, I., and Webster, R. 2022. Overview of data sources for the Pacific halibut stock assessment, harvest policy, and related analyses. IPHC-2022-SA-02. 56 p.

Stewart, I., and Hicks, A. 2022. Development of the 2022 Pacific halibut (*Hippoglossus stenolepis*) stock assessment. IPHC-2022-SRB020-07. 128 p.

Stewart, I., and Hicks, A. 2024. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2023. IPHC-2024-SA-01. 37 p.

Stewart, I., and Hicks, A. 2025. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2024. IPHC-2025-SA-01. 40 p.

Stewart, I., and Webster, R. 2025. Overview of data sources for the Pacific halibut stock assessment, harvest policy, and related analyses. IPHC-2025-SA-02. 57 p.

Stewart, I., Hicks, A., and Hutniczak, B. 2021. Evaluation of directed commercial fishery size limits in 2020. IPHC-2021-AM097-09. 28 p.

Stewart, I.J. 2014. Overview of data sources for the Pacific halibut stock assessment and related analyses. IPHC Report of Assessment and Research Activities 2013. p. 95-168.

Stewart, I.J., and Hamel, O.S. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences **71**(4): 581-588. doi:10.1139/cjfas-2013-0289.

Stewart, I.J., and Martell, S.J.D. 2014. A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. Fisheries Research **158**: 40-49. doi:10.1016/j.fishres.2013.09.012.

Stewart, I.J., and Martell, S.J.D. 2015. Reconciling stock assessment paradigms to better inform fisheries management. ICES Journal of Marine Science **72**(8): 2187-2196. doi:10.1093/icesjms/fsv061.

Stewart, I.J., and Martell, S.J.D. 2016. Appendix: Development of the 2015 stock assessment. IPHC Report of Assessment and Research Activities 2015. p. A1-A146.

Stewart, I.J., and Monnahan, C.C. 2017. Implications of process error in selectivity for approaches to weighting compositional data in fisheries stock assessments. Fisheries Research **192**: 126-134. doi:10.1016/j.fishres.2016.06.018.

Stewart, I.J., and Hicks, A.C. 2018. Interannual stability from ensemble modelling. Canadian Journal of Fisheries and Aquatic Sciences **75**: 2109-2113. doi:10.1139/cjfas-2018-0238.

Stewart, I.J., Wallace, J.R., and McGilliard, C.R. 2009. Status of the U.S. yelloweye rockfish resource in 2009. Pacific Fishery Management Council, Portland, OR. 236 p.

Stewart, I.J., Leaman, B.M., Martell, S., and Webster, R.A. 2013a. Assessment of the Pacific halibut stock at the end of 2012. IPHC Report of Assessment and Research Activities 2012. p. 93-186.

Stewart, I.J., Hicks, A.C., Taylor, I.G., Thorson, J.T., Wetzel, C., and Kupschus, S. 2013b. A comparison of stock assessment uncertainty estimates using maximum likelihood and Bayesian methods implemented with the same model framework. Fisheries Research **142**(0): 37-46. doi:http://dx.doi.org/10.1016/j.fishres.2012.07.003.

Stokes, K. 2019. Independent peer review for the 2019 IPHC stock assessment. August 2019. 31 p. https://www.iphc.int/uploads/pdf/sa/2019/stokes_2019-independent_peer_review_for_the_2019_iphc_stock_assessment.pdf.

Taylor, I.G., Doering, K.L., Johnson, K.F., Wetzel, C.R., and Stewart, I.J. 2021. Beyond visualizing catch-at-age models: Lessons learned from the r4ss package about software to support stock assessments. Fisheries Research **239**. doi:10.1016/j.fishres.2021.105924.

Thompson, G.G., and Lauth, R.R. 2012. Chapter 2: Assessment of the Pacific cod stock in the Eastern Bering Sea and Aleutian Islands area. *In* NPFMC Bering Sea and Aleutian Islands SAFE. p. 245-544.

Thorson, J.T. 2019. Perspective: Let's simplify stock assessment by replacing tuning algorithms with statistics. Fisheries Research ***217***: 133-139.

Thorson, J.T., Johnson, K.F., Methot, R.D., and Taylor, I.G. 2016. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fish. Res. doi:10.1016/j.fishres.2016.06.005.

Thygesen, U.H., Albertsen, C.M., Berg, C.W., Kristensen, K., and Nielsen, A. 2017. Validation of ecological state space models using the Laplace approximation. Environmental and Ecological Statistics **24**(2): 317-339. doi:10.1007/s10651-017-0372-4.

Trijoulet, V., Albertsen, C.M., Kristensen, K., Legault, C.M., Miller, T.J., and Nielsen, A. 2023. Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research **257**. doi:10.1016/j.fishres.2022.106487.

Ualesi, K., Jack, T., Rillera, R., and Coll, K. 2025. IPHC Fishery-Independent Setline Survey (FISS) design and implementation in 2024 IPHC-2025-AM101-09. 9 p.

Waterhouse, L., Sampson, D.B., Maunder, M., and Semmens, B.X. 2014. Using areas-as-fleets selectivity to model spatial fishing: Asymptotic curves are unlikely under equilibrium conditions. Fisheries Research **158**: 15-25. doi:10.1016/j.fishres.2014.01.009.

Webster, R. 2018. Space-time modelling of fishery independent setline survey data. IPHC-2018-AM094-07. 27 p.

Webster, R., and Stewart, I. 2022. Revision of the IPHC length-weight relationship. IPHC-2022-AM098-INF07. 29 p.

Webster, R.A. 2025. Space-time modelling of survey data. IPHC-2025-AM101-10. 9 p.

Webster, R.A., Clark, W.G., Leaman, B.M., and Forsberg, J.E. 2013. Pacific halibut on the move: a renewed understanding of adult migration from a coastwide tagging study. Canadian Journal of Fisheries and Aquatic Sciences **70**(4): 642-653. doi:10.1139/cjfas-2012-0371.

Webster, R.A., Soderlund, E., Dykstra, C.L., and Stewart, I.J. 2020. Monitoring change in a dynamic environment: spatio-temporal modelling of calibrated data from different types of fisheries surveys of Pacific halibut. Canadian Journal of Fisheries and Aquatic Sciences **77**: 1421-1432.

Appendices

Appendix A: Supplementary material

In addition to this document, supplementary material is available electronically, including:

1) Stock synthesis input files for each of the assessment models included in the proposed ensemble: data file, weight-at-age file, control file with model configuration, starter and forecast files with additional settings. Each of these files has been extensively annotated to aid in locating the various sections, as well as identifying which options and features were implemented or are irrelevant for the configuration.

2) Output from each of the stock assessment models: a sub-directory of all plotting and diagnostic output from each model created by the r4ss package (the entire set can be loaded at once via opening the "_SS_output.html" file), and the raw report (text) file from each model. The report file has not been annotated and contains some information not relevant to the Pacific halibut model configurations; content and formats can be determined from the stock synthesis user manual (Methot Jr et al. 2021a) and technical documentation (Methot and Wetzel 2013a).

3) Copies of the primary software documentation including the general modelling approach implemented in stock synthesis (Methot and Wetzel 2013b), the technical documentation (Methot and Wetzel 2013a) and the current user manual (Methot et al. 2024). From these documents, detailed model equations, data configurations, and control settings can be evaluated for the specific features implemented in the models for Pacific halibut.

4) The overview of data sources (Stewart and Webster 2025) and the stock assessment results (Stewart and Hicks 2025) from the 2024 stock assessment.

5) The documentation from the development of the most recent (2022) full stock assessment (Stewart and Hicks 2022).

6) Recent background papers describing the bootstrapping method employed for fishery and FISS age compositions (Hulson and Williams 2024; Stewart and Hamel 2014), the history of the halibut stock assessment (Stewart and Martell 2014), an evaluation of data weighting and process-error considerations (Stewart and Monnahan 2017), the general rationale for the ensemble approach (Stewart and Martell 2015), and the stability properties of ensemble assessments (Stewart and Hicks 2018).

7) A full record of the historical stock assessment documentation from 1978 to the present can be found on the IPHC's web site (https://www.iphc.int/management/science-and-research/stock-assessment). Individual Scientific Review Board reports and presentations (2013-2024) are available through the IPHC's meetings webpage (https://www.iphc.int/iphc-meetings).

INTERNATIONAL PACIFIC HALIBUT COMMISSION

IPHC Secretariat MSE Program of Work (2025) and an update on development of a Harvest Strategy Policy

PREPARED BY: IPHC SECRETARIAT (A. HICKS & I. STEWART; 8 MAY 2025)

PURPOSE

To provide the SRB with an update of the IPHC Management Strategy Evaluation (MSE) and the Harvest Strategy Policy (HSP).

1 INTRODUCTION

Rapid investigation of different questions is possible with the fully developed MSE framework. The operating models (OMs) in this framework were conditioned using the 2022 stock assessment and will be reconditioned after the 2025 full stock assessment to reflect new understanding of the Pacific halibut population and fishery dynamics. Given that new OMs will be available in 2026, major investigations of Management Procedures (MPs) will be done after then. Checking for exceptional circumstances and investigations the effect of recruitment and weight-at-age on outcomes are presented. Additionally, a brief update on the development of a Harvest Strategy Policy (HSP) is provided.

2 EXCEPTIONAL CIRCUMSTANCES

Two exceptional circumstances are considered for inclusion in the HSP.

- The coastwide all-sizes FISS WPUE or NPUE from the space-time model is above the 97.5^{th} percentile or below the 2.5th percentile of the simulated FISS index for two or more consecutive years.
- The realised coastwide mortality is above the 97.5^{th} percentile or below the 2.5^{th} percentile of the simulated realised coastwide mortality for two or more consecutive years.

Exceptional circumstances would be reviewed by the SRB to determine if one should be declared. In the event that an exceptional circumstance is declared, the following actions are to be completed. These actions have been recently updated to include how the Commission interacts with the process.

- Review the MSE simulations to determine if the OM can be improved and MPs should be re-evaluated.
- Consult with the SRB and MSAB to identify why the exceptional circumstance occurred, what can be done to resolve it, and determine a set of MPs to evaluate with an updated OM.
- Present these recommendations to the Commission for a Commission recommendation whether to update the OM and re-evaluate the reference MP and alternative MPs.

- Further consult with the SRB and MSAB after simulations are complete to recommend a new MP to the Commission.
- Present these results to the Commission to identify whether a new MP is appropriate and the HSP should be updated.

Three quantities were examined for the years 2023 and 2024 since the OM was tuned to data through 2022. The observations from these years were compared to the simulations using an SPR=43% and a 30:20 control rule with an annual stock assessment. The observed coastwide NPUE from the space-time model in 2023 and 2024 was within the 95% prediction interval from the simulated coastwide FISS NPUE (Figure 1). The coastwide realised Total Mortality in 2023 and 2024 was within the 95% prediction interval from the simulated coastwide TM (Figure 2). Therefore, the exceptional circumstances are not triggered. Furthermore, the 95% prediction interval of spawning biomass from the ensemble stock assessment is within the projected 95% interval from the MSE simulations (Figure 3). Given these results, the simulations remain relevant, but there are a number of improvements to the understanding of Pacific halibut that make it useful to recondition the OM after the upcoming full stock assessment.

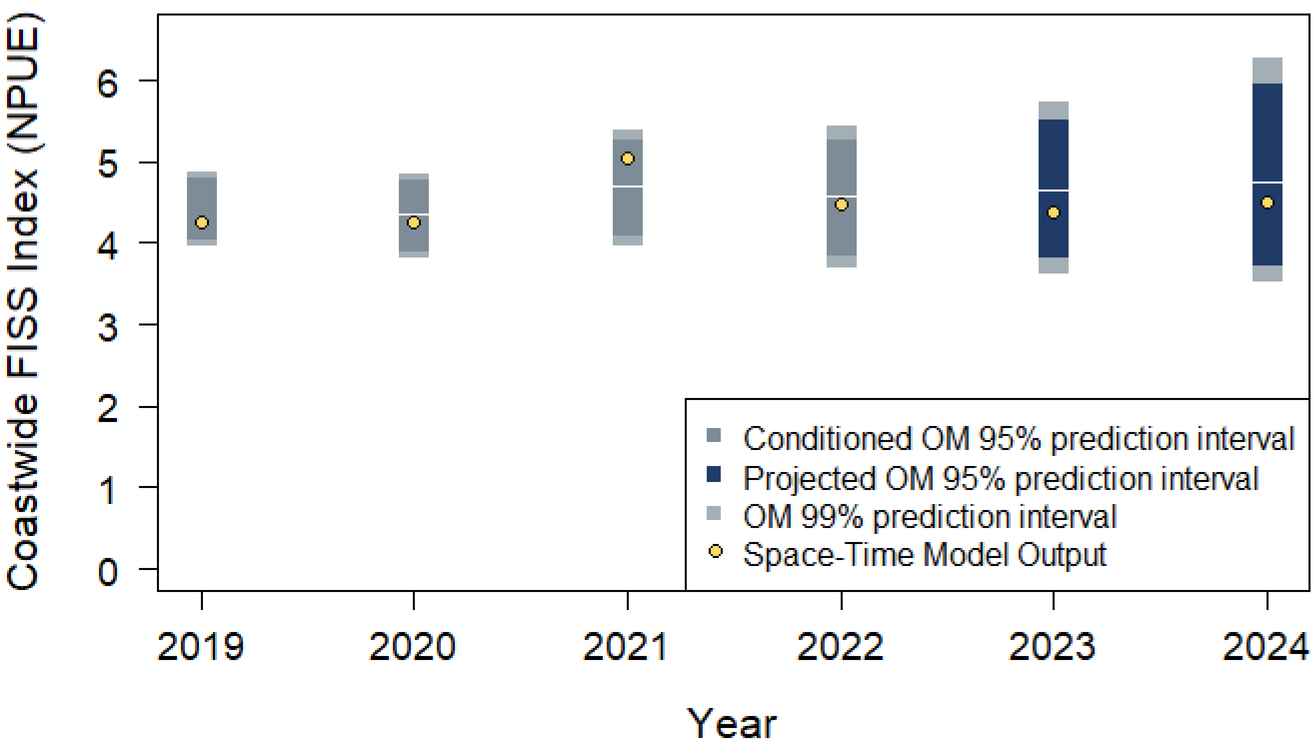


Figure 1. Simulated coastwide FISS NPUE (blue) from MSE simulations with an annual assessment, SPR=43%, and a 30:20 control rule. Space-time model output of FISS NPUE is shown in yellow.

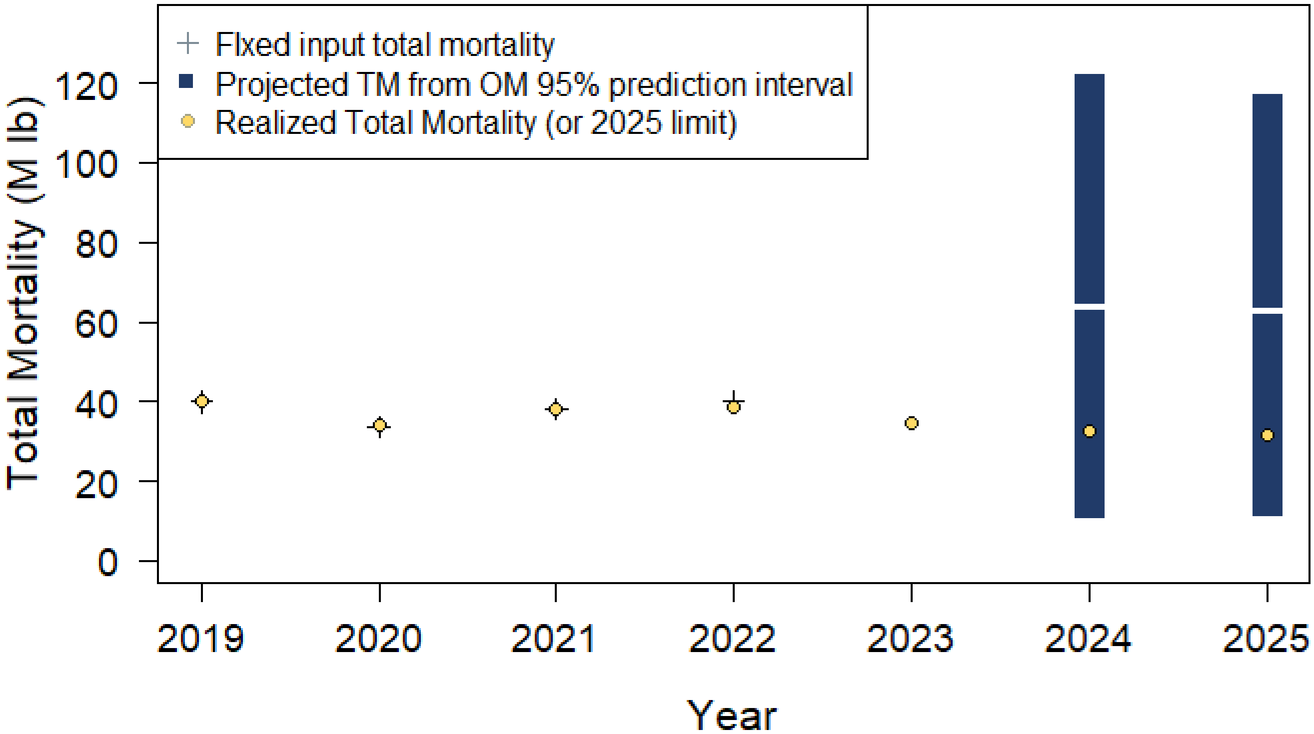


Figure 2. Simulated coastwide Total Mortality (TM, blue) from MSE simulations with an annual assessment, SPR=43%, and a 30:20 control rule. Realized TM is shown in yellow.

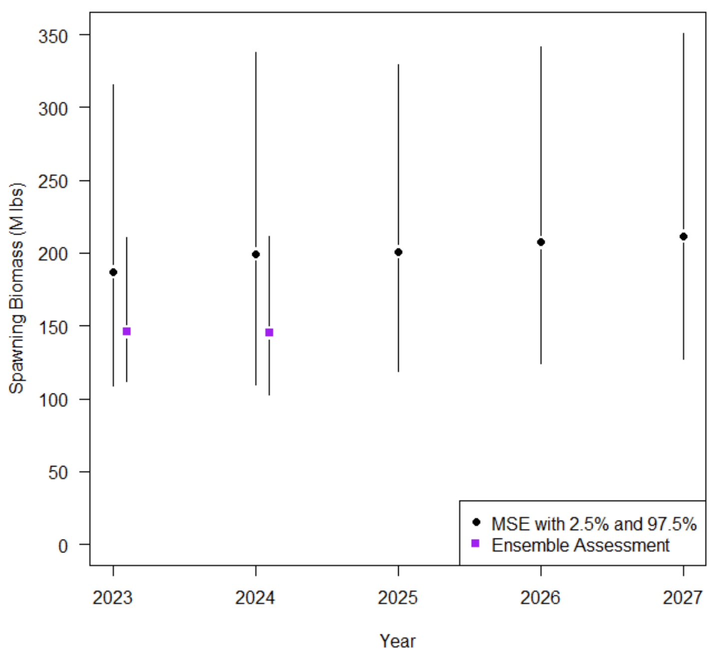


Figure 3. Simulated coastwide spawning biomass (blue) from MSE simulations with an annual assessment, SPR=43%, and a 30:20 control rule. Estimated spawning biomass with uncertainty from the ensemble assessment is shown in magenta.

3 EFFECTS OF WEIGHT-AT-AGE AND RECRUITMENT REGIMES

Pacific halibut exhibit high variability in weight-at-age and recruitment. Over the past 100 years, the average weight of an age 12 Pacific halibut has ranged from below 20 pounds in recent years to near 40 pounds in the mid-1970's (Figure 4). In the last ten years, the weight of the oldest fish has been declining or stable, but the weight of younger fish has been increasing. Recruitment is variable as well, and 1987 was one of the largest recruitments on record, as estimated in both 'long' assessment models (Figure 5). The two "long time-series" models in the IPHC stock assessment (IPHC-2025-SA-01) estimated a link between the Pacific Decadal Oscillation (PDO, Mantua et al. (1997)) and average unfished equilibrium recruitment (R_0), with an estimated average recruitment more than 50% greater during a positive PDO . Previous analyses (Clark and Hare 2002; Stewart and Martell 2016) have also shown that a positive PDO phase is correlated with enhanced productivity, while productivity decreases in negative PDO phases. Although the PDO is strongly correlated with historical recruitments, it is unclear whether the effects of climate change and other recent anomalous conditions in both the Bering Sea and Gulf of Alaska are comparable to those observed in previous decades (Litzow et al. 2020).

To investigate the effects of these low and high weight-at-age and recruitment regimes, different scenarios were defined from past observations and the population was projected 70 years with an SPR of 43%, assuming constant weight-at-age and average recruitment defined by the scenario. Three levels were developed for weight-at-age: low weight-at-age was defined from a five-year period in the 2010s, high weight-at-age was defined from a five-year period in the 1970s, and current weight-at-age was defined as the most recent five-years (Figure 4). These three weight-at-age levels show different patterns and although the low weight-at-age and current weight-at-age scenarios were both low in general, they differed between the weight of young fish and older fish. The current weight-at-age scenario had larger young fish but smaller older fish. High and low recruitment regimes were defined based on the stock assessment estimates of average recruitment in positive and negative PDO regimes. The PDO also affects movement and distribution of newly recruited (age-0) Pacific halibut. Overall, there were six scenarios crossing current, low, and high weight-at-age with low and high PDO.

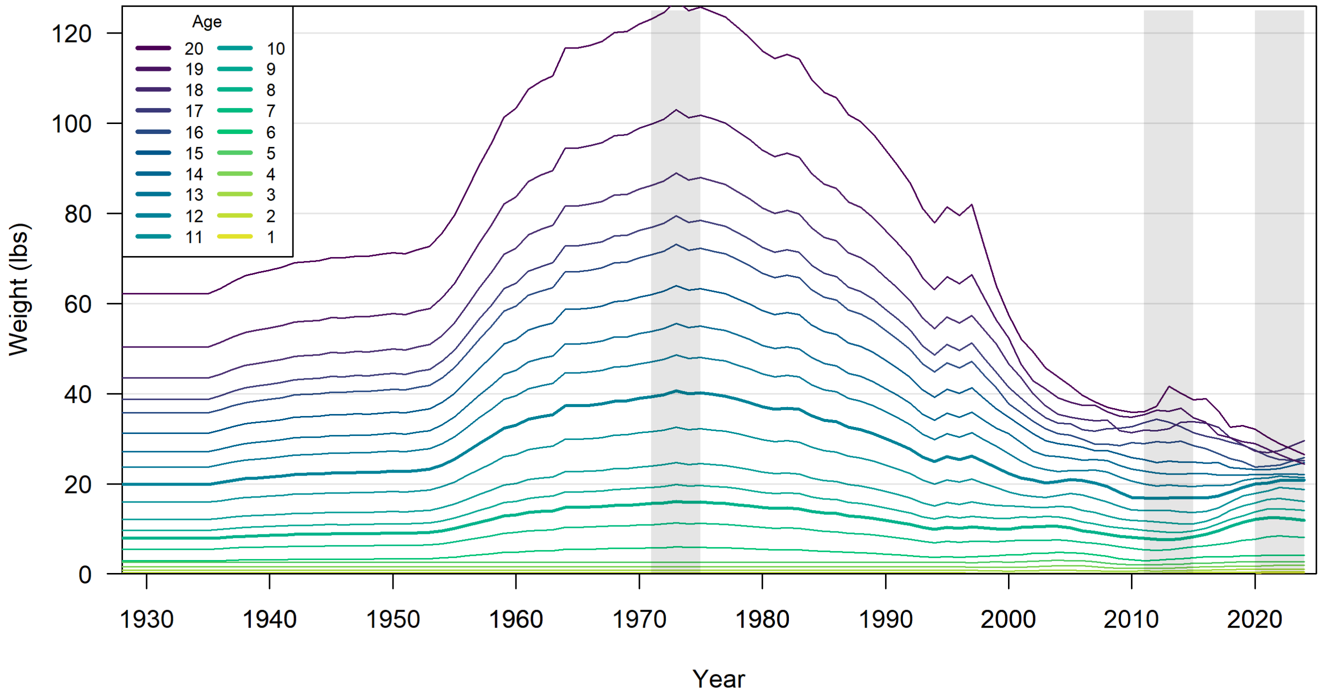


Figure 4. Average historical weight of Pacific halibut for ages one to twenty. Gray bands show three blocks of five years classified as high (1970s), low (2010s) and current (recent).

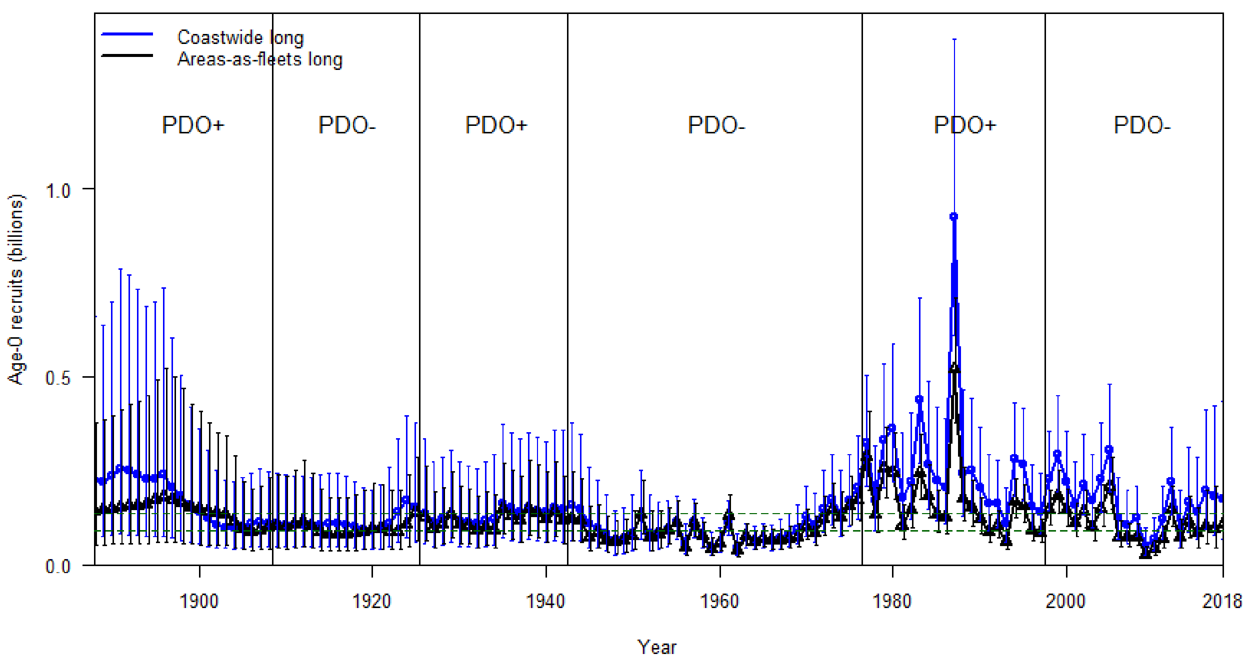


Figure 5. Trend in historical recruitment strengths (by birth year) estimated by the two long time-series stock assessment models, including the effects of the Pacific Decadal Oscillation (PDO) regimes. Figure reproduced from IPHC-2025-SA-01.

The spawning biomass differed substantially across different scenarios, but the high weight-at-age scenarios showed a considerable higher spawning biomass than the others (Figure 6). The sudden increase in the spawning biomass when the projections began indicates that weight-at-age is an important driver to the spawning biomass in the current year and future years. Average recruitment had a significant effect as well, but affected the spawning biomass in the longer term since the fish must age into the spawning biomass and was more prevalent with higher weight-at-age. For a given recruitment regime, the current weight-at-age scenario resulted in a smaller spawning biomass than the low weight-at-age scenario. This indicates the importance of the older fish in the spawning biomass.

Simulated TCEYs showed the same pattern for high weight-at-age, but different patterns for low and current weight-at-age scenarios. Weight-at-age and recruitment both had a profound effect on the TCEY with the high weight-at-age and high recruitment scenario supporting TCEYs near 120 Mlb and the high weight-at-age and low recruitment scenario supporting TCEYs near 75 Mlb. The low and current weight-at-age scenarios resulted in TCEYs in the range of 30 to 60 Mlb, on average. The TCEY showed a different pattern in the low and current weight-at-age scenarios when compared to the spawning biomass. The TCEY was higher for the current weight-at-age scenario while the spawning biomass was higher for the low weight-at-age scenario. Young Pacific halibut are more influential to the TCEY than to the spawning biomass because some are selected by the fishery before they become mature.

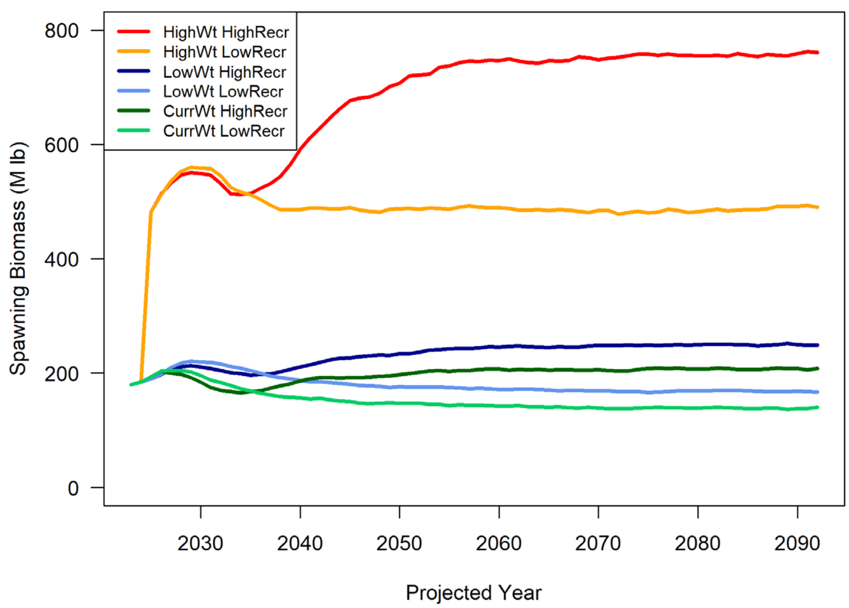


Figure 6. Simulated projections of spawning biomass assuming six different regimes for combinations of weight-at-age and recruitment and an SPR of 43%. Each projection held the weight-at-age and average recruitment at the defined level for all projected years.

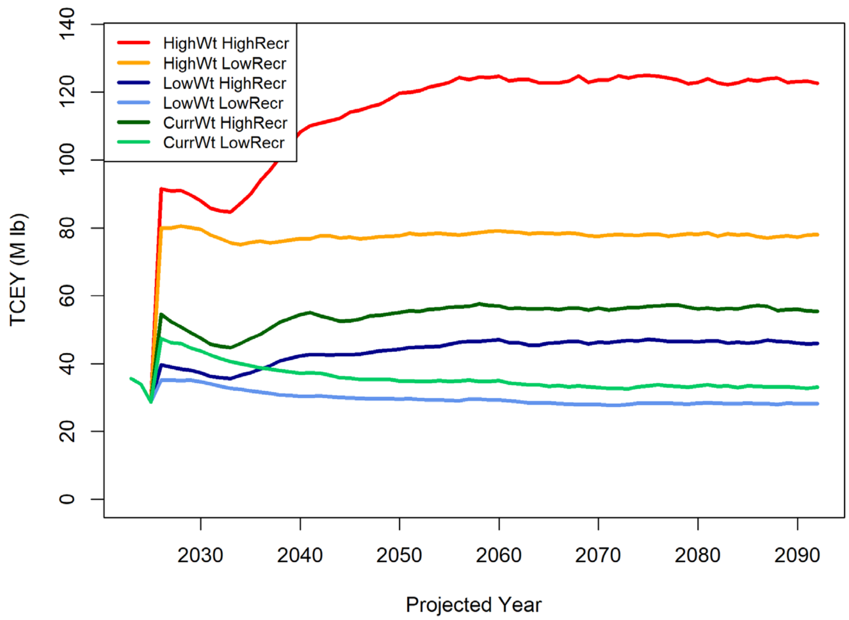


Figure 7. Simulated projections of the TCEY assuming six different regimes for combinations of weight-at-age and recruitment and an SPR of 43%. Each projection held the weight-at-age and average recruitment at the defined level for all projected years.

4 HARVEST STRATEGY POLICY (HSP)

A workshop with Commissioners occurred in April 2025 to discuss potential changes to the draft HSP and how to move it forward for adoption. Many edits were suggested and the next steps are for the Commission to review the edits, possibly hold another work session, consider a new draft at the Work Meeting, and move the HSP forward for adoption at the next Interim Meeting or Annual Meeting.

The hierarchical nature of the objectives was discussed at the workshop. In particular, the concepts of and trade-offs between the sustainability of the stock, maximising yield, and minimising yield were considered. Figure 8 shows the hierarchical nature of the objectives and new wording to identify how the trade-offs are considered.

A second important discussion was the inclusion of a timeframe for specific events such as stock assessments and re-evaluation of MPs. The IPHC currently operates off a schedule of three-years for full stock assessments, with update stock assessments in the intervening two years, and the MSE OM is updated following each full stock assessment to maintain consistent approaches and paradigms. Therefore, MPs are re-evaluated at a minimum of three years after implementation, and shall not exceed two cycles (six years as shown in Table 1). An exceptional circumstance may trigger a re-evaluation of the MP.

LONG-TERM OVERARCHING OBJECTIVES DEFINING ACCEPTABLE MPs

1. SUSTAINABILITY

GENERAL OBJECTIVE	MEASURABLE OBJECTIVE
KEEP FEMALE SPAWNING BIOMASS ABOVE A LIMIT TO AVOID CRITICAL STOCK SIZES	Maintain the long-term coastwide female relative spawning biomass above a biomass limit reference point ($RSB_{20\%}$) at least 95% of the time

2. OPTIMISE FISHING ACTIVITIES AND OPPORTUNITIES

GENERAL OBJECTIVE	MEASURABLE OBJECTIVE
MAINTAIN SPAWNING BIOMASS AT OR ABOVE A LEVEL THAT SUPPORTS OPTIMAL FISHING ACTIVITIES AND OPPORTUNITIES	Maintain the long-term coastwide female relative spawning biomass at or above a biomass threshold reference point ($RSB_{36\%}$) 50% or more of the time.

SHORT-TERM MANAGEMENT OBJECTIVES INFORMING A REFERENCE MP

3. OPTIMISE YIELD

GENERAL OBJECTIVE	MEASURABLE OBJECTIVE
PROVIDE DIRECTED FISHING YIELD WHILE LIMITING VARIABILITY IN MORTALITY LIMITS	Balance the trade-off between maximising the sustainable average coastwide mortality limit and minimising annual changes in the coastwide mortality limit.

Figure 8. Priority objectives for the long-term sustainable management of Pacific halibut that support optimal yield and fisheries opportunities. The hierarchy of the objectives is shown by the arrows. The green colour indicates a conservation goal while the blue colours indicate fishery goals.

Table 1. Stock assessment, MSE, exceptional circumstances check, review, and decision processes on an annual basis. Year 1 could correspond to 2025, 2028, 2031, and so on. Upper case 'Y' indicates that the task is done, a lower case 'x' indicates that the task may be done. 'EC' refers to Exceptional Circumstance and 'FISS' to Fishery-Independent Setline Survey.

Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Example Year	2025	2026	2027	2028	2029	2030	2031	2032
FISS coastwide index	Y	Y	Y	Y	Y	Y	Y	Y
Full stock assessment	Y			Y			Y	
Update stock assessment		Y	Y		Y	Y		Y
Commission TCEY decision	Y	Y	Y	Y	Y	Y	Y	Y
MSE OM updated		Y			x			Y
MP re-evaluated		Y			x			Y
Exceptional circumstances checked	Y		Y	Y	x[1]	Y	Y	
- Consult with SRB and MSAB			x	x	x	x	x	
- Present to Commission			x	x	x	x	x	
- Re-evaluate MP due to EC			*	*	Y[2]	x*	x*	
Update HSP			x			x		

[1] The exceptional circumstance would be checked only if a new MSE OM was not updated.

[2] The MP would be re-evaluated as part of the normal three-year cycle due to an exceptional circumstance occurring in two sequential years.

* An exceptional circumstance can be declared after two sequential instances, thus re-evaluation of an MP would have a delay, unless recommended by the Commission outside of the normal process.

Recommendation/s

That the SRB:

1) **NOTE** paper IPHC-2025-SRB026-08 which details testing for exceptional circumstances, recent work done using the management strategy evaluation framework, and progress on the Harvest Strategy Policy.
2) **REQUEST** any topics to add to the 2025-2026 MSE Program of Work.

References

Clark, W.G., and Hare, S.R. 2002. Effects of climate and stock size on recruitment and growth of Pacific halibut. North American Journal of Fisheries Management **22**: 852-862.

Litzow, M.A., Malick, M.J., Bond, N.A., Cunningham, C.J., Gosselin, J.L., and Ward, E.J. 2020. Quantifying a Novel Climate Through Changes in PDO - Climate and PDO - Salmon Relationships. Geophysical Research Letters **47**(16). doi:10.1029/2020gl087972.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society **78**(6): 1069-1079.

Stewart, I.J., and Martell, S.J.D. 2016. Appendix: Development of the 2015 stock assessment. International Pacific Halibut Commission.

2026-28 FISS design evaluation

PREPARED BY: IPHC SECRETARIAT (R. WEBSTER, I. STEWART, K. UALESI, T. JACK & D. WILSON; 09 MAY 2025)

PURPOSE

To present the SRB with potential FISS designs for 2026-28, including a preliminary cost evaluation of the 2026 designs.

BACKGROUND

The IPHC's Fishery-Independent Setline Survey (FISS) provides data used to compute indices of Pacific halibut density for use in monitoring stock trends, estimating stock distribution, and as an important input in the stock assessment. Stock distribution estimates are based on the annual mean weight per unit effort (WPUE) for each IPHC Regulatory Area, computed as the average of WPUE of all Pacific halibut and for O32 (greater than or equal to 32" or 81.3cm in length) Pacific halibut estimated at each station in an area. Mean numbers per unit effort (NPUE) is used to index the trend in Pacific halibut density for use in the stock assessment models. Annual FISS designs are developed by selecting a subset of stations for sampling from the full 1890-station FISS footprint (Figure 1).

In recent years, financial constraints due to reduced catch rates, lower sales prices and higher costs have led to the implementation of FISS designs with reduced spatial footprints. Effort has been concentrated in IPHC Regulatory Areas 2B, 2C, 3A and 3B, with limited sampling in other areas (Figures 2 and 3). The Base Block Design (described below) was presented to the Commission at the September 2024 Work Meeting and the 14th Special Session of the IPHC (SS014, IPHC-2024-SS014-03) as a more efficient approach to annual sampling in the core of the stock compared to recent designs based on random selection of FISS stations. For 2025, high projected financial costs for this design meant that it was not viable to undertake without substantial supplementary funding. Therefore, IPHC Secretariat staff developed a "fiscally viable" design for 2025 that would have reduced spatial coverage for the third year in a row but at a projected loss that could be covered by revenue, supplementary funding and IPHC reserve funds. Following SS014, the final 2025 FISS design was approved via inter-sessional agreement (IPHC-2024-CR-030, IPHC-2024-CR-031; Figure 3). This design included sampling of FISS charter regions in IPHC Regulatory Areas 3A, 3B, 4A and 4B that were unsampled in either 2023, 2024 or both.

FISS history 1993-2019

The IPHC has undertaken FISS activity since the 1960s. However, methods were not standardized to a degree (e.g., the bait and gear used) that allows for simple combined analyses until 1993. From 1993 to 1997, the annual design was a modification of a design developed and implemented in the 1960s, and involved fishing triangular clusters of stations, with clusters located on a grid (IPHC 2012). Coverage was limited in most years and was generally restricted to IPHC Regulatory Areas 2B through 3B. The modern FISS design, based on a grid with 10 nmi (18.5 km) spacing, was introduced in 1998, and over the subsequent two years was expanded to include annual coverage in parts of all IPHC Regulatory Areas within the depth ranges of 20-

275 fathoms (37-503 m) in the Gulf of Alaska and Aleutian Islands, and 75-275 fathoms (137-503 m) in the Bering Sea (IPHC 2012). Annually-fished stations were added around islands in the Bering Sea in 2006, and in the same year, a less dense grid of paired stations was fished in shallower waters of the southeastern Bering Sea, providing data for a calibration with data from the annual National Marine Fishery Service (NMFS) bottom trawl survey (Webster et al. 2020).

Through examination of commercial logbook data and information from other sources, it became clear by 2010 that the historical FISS design had gaps in coverage of Pacific halibut habitat that had the potential to lead to bias in estimates derived from its data. These gaps included deep and shallow waters outside the FISS depth range (0-20 fathoms and 275-400 fathoms), and unsurveyed stations on the 10 nmi grid within the 20-275 fathom depth range within each IPHC Regulatory Area. This led the IPHC Secretariat to propose expanding the FISS to provide coverage of the unsurveyed habitat with United States and Canadian waters. In 2011 a pilot expansion was undertaken in IPHC Regulatory Area 2A, with stations on the 10 nmi grid added to deep (275-400 fathoms) and shallow (10-20 fathoms) waters, the Salish Sea, and other, smaller gaps in coverage. The 10-fathom limit in shallow waters was due to logistical difficulties in standardized fishing of longline gear in shallower waters. A second expansion in IPHC Regulatory Area 2A was completed in 2013, with a pilot survey in California waters between the latitudes of 40 and 42°N.

The full expansion program began in 2014 and continued through 2019, resulting in the sampling of the entire FISS design of 1890 stations in the shortest time logistically possible. The FISS expansion program allowed us to build a consistent and complete picture of Pacific halibut density throughout its range in Convention waters. Sampling the full FISS design has reduced bias as noted above, and, in conjunction with space-time modelling of survey data (see below), has improved precision and fully quantified the uncertainty associated with estimates based on partial annual sampling of the species range. It has also provided us with a complete set of observations over the full FISS design (Figure 1) from which an optimal subset of stations can be selected when devising annual FISS designs. This station selection process began in 2019 for the 2020 FISS and continues with the current review of design proposals for 2024-26. Note that in the Bering Sea, the full FISS design does not provide complete spatial coverage, and FISS data are augmented with calibrated data from National Marine Fisheries Service (NMFS) and Alaska Department of Fish and Game (ADFG) trawl surveys (stations can vary by year – 2019 designs are shown in Figure 1). Both supplementary surveys have been conducted approximately annually in recent years.

Rationalized FISS, 2020-25

Following the 2011-2019 program of FISS expansions, rationalized FISS designs were approved for 2020 based on random selection of over 50% of stations in the core of the stock (IPHC Regulatory Areas 2B, 2C, 3A and 3B) and sampling of all stations in selected subareas of the remaining IPHC Regulatory Areas. For the latter areas, sampling priorities were determined based on maintaining precise estimates of area-specific indices of density and ensuring low bias in index estimators. That year, the COVID19 pandemic led to a reduced FISS with sampling only in the core areas. The 2021-22 FISS sampling proceeded largely as designed, although with planned stations in western IPHC Regulatory 4B in 2022 unsampled due to a lack of viable charter bids. In some charter regions in the core areas, 100% of stations were sampled in order to achieve revenue goals (see below). The 2023 FISS design had more limited spatial coverage, with almost no FISS sampling outside of the core areas due to large projected revenue losses from designs that included extensive sampling in IPHC Regulatory Areas 2A, 4A, 4B and 4CDE.

Limited sampling was carried out in northern IPHC Regulatory 2A, while planned stations around the IPHC Regulatory Area 4A/4B boundary were not sampled due to a lack of charter bids. The adopted 2024 FISS design (IPHC-2024-AM100-R) included high sampling rates in IPHC Regulatory Areas 2B and 2C, a small number of charter regions in IPHC Regulatory Areas 3A and 3B, and sampling of the southern shelf edge and Bering Sea islands in IPHC Regulatory Area 4CDE (Figure 2). The 2025 design includes stations in IPHC Regulatory Areas 3A and 3B that complement coverage in recent years (Figure 3), along with stations in IPHC Regulatory Areas 2A, 4A and 4B that have not been sampled for three or more years and is therefore expected to reduce the potential for bias in most IPHC Regulatory Areas relative to recent years.

Space-time modelling

In 2016, a space-time modelling approach was introduced to estimate time series of weight and numbers-per-unit-effort (WPUE and NPUE), and to estimate the stock distribution of Pacific halibut among IPHC Regulatory Areas. This represented an improvement over the largely empirical approach used previously, as it made use of additional information within the survey data regarding the degree of spatial and temporal correlation in Pacific halibut density, along with information from covariates such as depth (see Webster 2016, 2017). It also allowed a more complete accounting of uncertainty; for example, prior to the use of space-time modelling, uncertainty due to unsurveyed regions in each year was ignored in the estimation. Prior to the application of the space-time modelling, these unsampled regions were either filled in using independently estimated scalar calibrations (if fished at least once), or catch-rates at unsampled stations were assumed to be equal to the mean for the entire Regulatory Area. The IPHC's Scientific Review Board (SRB) has provided supportive reviews of the space-time modelling approach (e.g., IPHC-2018-SRB013-R), and the methods have been published in a peer-review journal (Webster et al. 2020). Similar geostatistical models are now routinely used to standardize fishery-independent trawl surveys for groundfish on the West Coast of the U.S. and in Alaskan waters (e.g., Thorson et al. 2015 and Thorson 2019). The IPHC space-time models are fitted through the R-INLA package in the R software (R Core Team, 2024).

FISS DESIGN OBJECTIVES (Table 1)

Note that the secondary objective was revised at AM101 (IPHC-2025-AM101-R, para. 61).

Primary objective: *To sample Pacific halibut for stock assessment and stock distribution estimation.*

The primary purpose of the annual FISS is to sample Pacific halibut to provide data for the stock assessment (abundance indices, biological data) and estimates of stock distribution for use in the IPHC's management procedure. The priority of the current rationalized FISS is therefore to maintain or enhance data quality (precision and bias) by establishing baseline sampling requirements in terms of station count, station distribution and skates per station.

Secondary objective: *Cost effectiveness.*

The FISS is intended to be cost-effective without compromising the scientific integrity of the design. Any implemented design must consider logistics and cost together with scientific integrity.

Tertiary objective: *Minimize removals and assist others where feasible on a cost-recovery basis.*

Consideration is also given to the total expected FISS removals (impact on the stock), data collection assistance for other agencies, and IPHC policies.

Table 1 Prioritization of FISS objectives and corresponding design layers.

Priority	**Objective**	**Design Layer**
Primary	Sample Pacific halibut for stock assessment and stock distribution estimation	Minimum sampling requirements in terms of: • Station distribution • Station count • Skates per station
Secondary	Cost effectiveness without compromising the scientific integrity of the FISS design.	Logistics, cost, scientific integrity: operational feasibility and cost/revenue, and scientific needs. With an aspirational target reserve of US$2,000,000
Tertiary	Minimize removals and assist others where feasible on a cost-recovery basis.	Removals: minimize impact on the stock while meeting primary priority Assist: assist others to collect data on a cost-recovery basis IPHC policies: ad-hoc decisions of the Commission regarding the FISS design

Annual design review, endorsement, and finalisation process

Since completion of the FISS expansions in 2019, a review process has been developed for annual FISS designs created according to the above objectives:

- Step 1: The Secretariat presents preliminary design options based on the primary objective (Table 1) to the SRB for three subsequent years at the June meeting based on analysis of prior years' data. Commencing in 2024, this has included preliminary cost projections based on prior year fiscal details (revenue) and current year vessel contract cost updates;
- Step 2: Updated design options for the following year that account for both primary and secondary objectives (Table 1) are reviewed by Commissioners at the September work meeting, recognising that revenue and cost data from the current year's FISS are still preliminary at this time;
- Step 3: At their September meeting, the SRB reviews design options accounting for both primary and secondary objectives (Table 1) for comment and advice to the Commission (recommendation);
- Step 4: Designs are further modified to account for updates based on secondary and tertiary objectives before being finalized during the Interim and Annual meetings and the period prior to implementation:
    - Presentation of FISS designs for 'endorsement' by the Commission occurs at the annual November/December Interim Meeting;
    - Ad-hoc modifications to the design for the current year (due to unforeseen issues arising) are possible at the Annual Meeting of the Commission;

- The endorsed design for current year is then modified (if necessary) to account for any additional tertiary objectives or revision to inputs into the evaluation of secondary objectives prior (i.e. updated cost estimates) and logistical considerations raised by the operators of contracted vessels prior to summer implementation (February-April).

Consultation with industry and stakeholders occurs throughout the FISS planning process, at the Research Advisory Board meeting (late November) and particularly in finalizing design details as part of the FISS charter bid process, when stations can be added and other adjustments made to provide for improved logistical efficiency. We also note the opportunities for direct stakeholder input during public meetings (Interim and Annual Meetings).

Note that while the review process examines designs for the next three years, revisions to designs for the second and third years are expected during subsequent review periods as additional data are collected. Having design proposals available for three years instead of the next year only assists the Secretariat with medium-term planning of the FISS, and allows reviewers (SRB, Commissioners) and stakeholders to see more clearly the planning process for sampling the entire FISS footprint over multiple years.

POTENTIAL DESIGNS FOR 2026-28

BASE BLOCK DESIGN

At AM101, Secretariat staff presented the Base Block design for 2025 and subsequent years based a rotational block design (IPHC-2025-AM101-14). Instead of the random selection of FISS stations used in past designs for IPHC Regulatory Areas 2B, 2C, 3A and 3B, the design features the sampling of complete FISS charter regions in each area. Sampled regions are rotated over two or three years depending on area. This type of design was first proposed in 2019 (IPHC-2019-IM095-07 Rev_1, Figure 4) to complement the similar subarea design proposed and adopted for areas at the ends of the stock (2A, 4A and 4B). Block designs are potentially more efficient from an operational perspective than a randomized design, as they involve less running time between stations, possibly leading to cost reductions on a per station basis.

The Base Block designs shown in Figures 4 to 6 for 2026-28 were revised from the designs presented to Commissioners at AM101 to account for the Commission-approved 2025 design; in particular, charter regions not selected in IPHC Regulatory Areas 3A and 3B in 2025 were prioritized for sampling in 2026. The Base Block design ensures that all charter regions in the core areas are sampled over a three-year period, while prioritizing coverage in other areas based on minimizing the potential for bias and maintaining CVs below 25% for each IPHC Regulatory Area. The Base Block designs also include some sampling in all IPHC Biological Regions in each year, ensuring that trend and biological data from across the spatial range of Pacific halibut are available to the stock assessment and for stock distribution estimation.

Using samples generated from the fitted 2024 space-time models as simulated data for 2025-28, we projected the coefficient of variation (CV, a relative measure of precision) for mean O32 WPUE for each year of the design by area. As CVs are generally greater in the terminal year of the time series and that year is usually the most relevant for informing management, the CV values in Table 2 are for the final year of the modelled time series. For example, the values for 2027 were found by fitting the model to the data for 1993-2027, with simulated data used for 2025-27.

Table 2. Projected coefficients of variation (CVs, %) of mean O32 WPUE for the Base Block design by terminal year of time series and IPHC Regulatory Area and Biological Region.

Regulatory Area	2026	**Year** 2027	2028
2A	21	22	14
2B	11	7	10
2C	6	6	6
3A	8	7	8
3B	11	15	11
4A	18	22	13
4B	15	16	17
4CDE	9	9	8
Biological Region			
Region 2	6	5	5
Region 3	7	7	7
Region 4	9	10	7
Region 4B	15	16	17
Coastwide	4	4	4

Projected terminal year CVs for the Base Block design are 25% or less for all IPHC Regulatory Areas. In the core areas (2B, 2C, 3A and 3B), CVs are projected to be 15% or less (Table 2). All Biological Region CVs, except that of Region 4B, are at most 10%, while the coastwide CV is projected to be 4% in all years. The Base Block design is therefore expected to maintain precise estimates of indices of Pacific halibut density and abundance across the range of the stock. At the same time, the rotating nature of the sampled blocks means that almost all FISS stations are sampled within a 5-year period (2-3 years within the core areas) resulting in low risk of missing important stock changes and therefore a low risk of large bias in estimates of trend and stock distribution.

The 'global average' research survey CVs has been estimated to be approximately ~20%; however, this value includes estimated observation and process error (based on lack of fit in the stock assessments), and so is larger than the survey-only observation CVs projected in this report (Francis et al. 2003). In NOAA Fisheries trawl survey results in the Bering Sea (roughly analogous to one Biological Region for Pacific halibut), commercially important species showed a range of average annual model-based CVs, including: Pacific cod (5%), Walleye pollock (7%), Northern rock sole (6%), and yellowfin sole (5%) over 1982-2019 (DeFilippo et al. 2023). These values are comparable to the projected 5-9% CVs for IPHC Biological Regions that would be expected from the base block design (with the exception of Biological Region 4B), but lower than corresponding values for the Core Block and Reduced Core designs.

REDUCED LOSS DESIGN

The Base Block design is projected to result in a substantial operating loss (Table 3) and would require supplementary funding to be viable. As an alternative, the Secretariat staff has developed a preliminary design that would result in a net operating loss of approximately $500,000 (Figure 7). This Reduced Loss design includes revenue positive charter regions in

IPHC Regulatory Areas 2B and 2C and maintains a subsample of 30 stations in each of three other revenue-negative charter regions from the Base Block design in IPHC Regulatory Areas 2B and 3A. The three regions with partial sampling were prioritized as they are among the regions not sampled in the last two to three years.

Table 3 gives preliminary net revenue projections for Base Block and Reduced Loss designs. Projections include the following assumptions:

1. Designs are optimized for numbers of skates, with 4, 6 or 8 skate-sets used, depending on projected catch rates and bait costs.
2. 2026 Pacific halibut price and landings do not change from values realized in 2024.

Costs do not include the costs associated with IPHC's Seacat water column profilers. As in 2025, at this stage we anticipate such costs to be covered by the IPHC's General Fund.

Cost estimates are largely based on information from the 2024 FISS and outcomes of the 2025 charter bidding process, and it is important to note there is high uncertainty in the any catch and cost projections for 2026 this far in advance. Final cost and accounting information will be available at the end of the 2025 fiscal year and will be used to refine these preliminary projections at that time.

Table 3. Comparison of preliminary projected net revenue for the 2026 Base Block and Reduced Loss designs.

Design	Projected net revenue
Base Block	−$1,818,000
Reduced Loss	−$536,000

INTERMEDIATE DESIGNS

Here we present several intermediate designs that could be considered if supplementary funding became available or if greater losses might be considered acceptable to the Commission (Table 4). As before, revenue estimates are very preliminary and subject to change as inputs are revised following the 2025 FISS season.

Option 1 in Table 4 is the Reduced Loss design (Figure 5), and Options 2 through 6 successively add stations or charter regions based on scientific priorities. Option 2 (Figure 8) samples the same charter regions as Option 1, but the partial regions are now fully sampled, reducing the risk of bias within those regions. IPHC Regulatory Area 4B is added in Option 3 (Figure 9), which is therefore the least expensive of the options in Table 4 that includes sampling of some kind in all Biological Regions (assuming the NOAA trawl survey provides coverage in Region 4). Option 4 (Figure 10) improves spatial coverage in Biological Region 3 by adding a charter region in IPHC Regulatory Area 3B, while Option 5 (Figure 11) adds FISS sampling to Region 4 with a charter region in IPHC Biological Region 4A. Option 6 (Figure 12) includes all charter regions from the Base Block design (Option 7, Figure 4), together with three revenue-positive regions in Biological Region 2 that are not part of the Base Block Design.

Table 4. Comparison of 2026 preliminary revenue projections for the Reduced Loss design, the Base Block design and design options providing intermediate coverage. For each design, the final column shows the difference in projected revenue from the design in the previous row.

Design	Sampled IPHC Regulatory Areas (with number of FISS charter regions)	Projected net revenue ($US)	Difference ($US)
Option 1: Reduced Loss	2B(2 full, 2 partial), 2C(3), 3A(1 partial)	−536,000	
Option 2	2B(4), 2C(3), 3A(1)	−556,000	−20,000
Option 3	2B(4), 2C(3), 3A(1), 4B(1)	−860,000	−304,000
Option 4	2B(4), 2C(3), 3A(1), 3B(1), 4B(1)	−1,012,000	−152,000
Option 5	2B(4), 2C(3), 3A(1), 3B(1), 4A(1), 4B(1)	−1,240,000	−228,000
Option 6	2B(4), 2C(3), 3A(4), 3B(2), 4A(1), 4B(1)	−1,740,000	−500,000
Option 7: Base Block	2B(2), 2C(2), 3A(4), 3B(2), 4A(1), 4B(1)	−1,818,000	−78,000

DISCUSSION

The **Base Block** design has a projected net loss of around $1,818,000 and therefore will rely on supplementary funding for implementation. Unlike the Base Block design, the preliminary **Reduced Loss** design does not have extensive spatial coverage, with sampling concentrated in regions of greatest Pacific halibut density in IPHC Biological Region 2, only 30 FISS stations in Biological Region 3, and no FISS sampling in Biological Regions 4 and 4B. Such a design comes with a greater risk of bias relative to the Base Block design due to the increased chance of stock changes being unobserved. Despite the uncertainty being properly propagated, of increasing concern is the potential for the space-time model expectations to move toward the long term mean in the absence of new data. This increased uncertainty in the index of abundance is likely to cause the assessment model to rely more heavily on the commercial fishery catch-per-unit-effort index, as was the case in 2024. Given current spatial variability and uncertainty in the magnitude of younger year classes (2016 and younger), the limited biological information from the core of the stock distribution (Biological Region 3) makes it unclear whether the stock assessment will detect a major change in year class abundance, either up or down. Although the stock assessment methods can remain unchanged, a greater portion of the actual uncertainty in stock trend and demographics will not be able to be quantified due to missing FISS data from a large fraction of the Pacific halibut stock's geographic range.

The implications for the assessment would be of increasing concern if designs like the Reduced Loss design were implemented beyond 2026 due to increasing uncertainty and risk of bias in stock trend estimates and the unrepresentativeness of the biological samples. Further, as was evident at AM100 and AM101, reduced FISS designs that do not fully inform stock distribution with annual sampling in all IPHC Regulatory areas lead to reduced stakeholder confidence in the FISS results and in the aggregate scientific information from the stock assessment. As it did with the relatively conservative mortality limits set for 2025, this may have a strong effect on the perception of risk and on decision making by the Commission if reduced survey designs continue to be consecutively implemented.

RECOMMENDATION

That the Scientific Review Board **NOTE** paper IPHC-2025-SRB026-09, which presents an evaluation of design options for 2026-28, including a preliminary option accounting for the secondary FISS objective of cost effectiveness.

References

DeFilippo, L., Kotwicki, S., Barnett, L., Richar, J., Litzow, M.A., Stockhausen, W.T., and Palof, K. 2023. Evaluating the impacts of reduced sampling density in a systematic fisheries-independent survey design. Frontiers in Marine Science **10**. doi:10.3389/fmars.2023.1219283.

Francis, R.I.C.C., Hurst, R.J., and Renwick, J.A. 2003. Quantifying annual variation in catchability for commercial and research fishing. Fishery Bulletin **101**: 293-304.

IPHC 2012. IPHC setline charters 1963 through 2003 IPHC-2012-TR058. 264p.

IPHC 2018. Report of the 13th Session of the IPHC Scientific Review Board (SRB) IPHC-2018-SRB013-R. 17 p.

IPHC 2024. Report of the 100th Session of the IPHC Annual Meeting (AM100) IPHC-2024-AM100-R. 55 p.

IPHC 2024. IPHC Circular 2024-030, Subject: For decision – FISS 2025 design. 4 p.

IPHC 2024. IPHC Circular 2024-031, Subject: For information – Intersessional decisions 2024-ID009 - ID010 FISS 2025 design. 1 p.

IPHC 2025. Report of the 101st Session of the IPHC Annual Meeting (AM101) IPHC-2025-AM101-R. 52 p.

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science 72(5): 1297-1310. doi:10.1093/icesjms/fsu243.

Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210: 143-161. doi:10.1016/j.fishres.2018.10.013.

Webster, R. A. 2016. Space-time modelling of setline survey data using INLA. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2015: 552-568.

Webster, R. A. 2017. Results of space-time modelling of survey WPUE and NPUE data. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016: 241-257.

Webster, R. 2019. Space-time modelling of IPHC Fishery-Independent Setline Survey (FISS) data. IPHC-2019-IM095-07 Rev_1. 19 p.

Webster, R. A., Soderlund, E, Dykstra, C. L., and Stewart, I. J. (2020). Monitoring change in a dynamic environment: spatio-temporal modelling of calibrated data from different types of fisheries surveys of Pacific halibut. Can. J. Fish. Aquat. Sci. 77(8): 1421-1432.

Webster, R., Stewart, I., Ualesi, K., Jack, T. and Wilson, D. (2024). 2025 and 2026-29 FISS Designs. IPHC-2024-SS014-03. 21 p.

Webster, R., Stewart, I., Ualesi, K., Jack, T. and Wilson, D. (2025). 2025 and 2026-29 FISS Designs. IPHC-2025-AM101-14. 15 p.

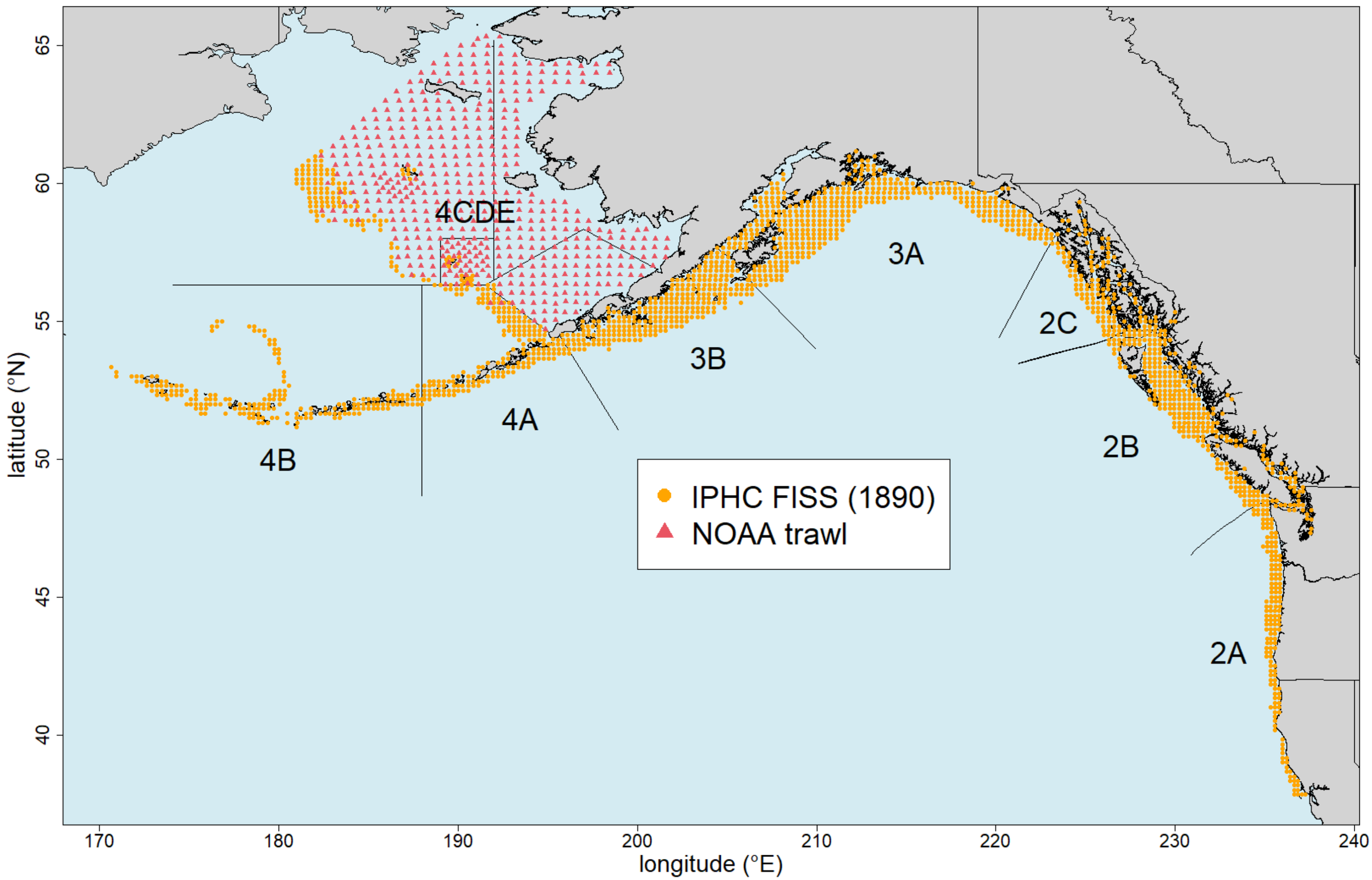


Figure 1. Map of the full 1890 station FISS design, with orange circles representing stations available for inclusion in annual sampling designs. Red triangles represent the locations NOAA trawl stations used to provide complementary data for Bering Sea modelling.

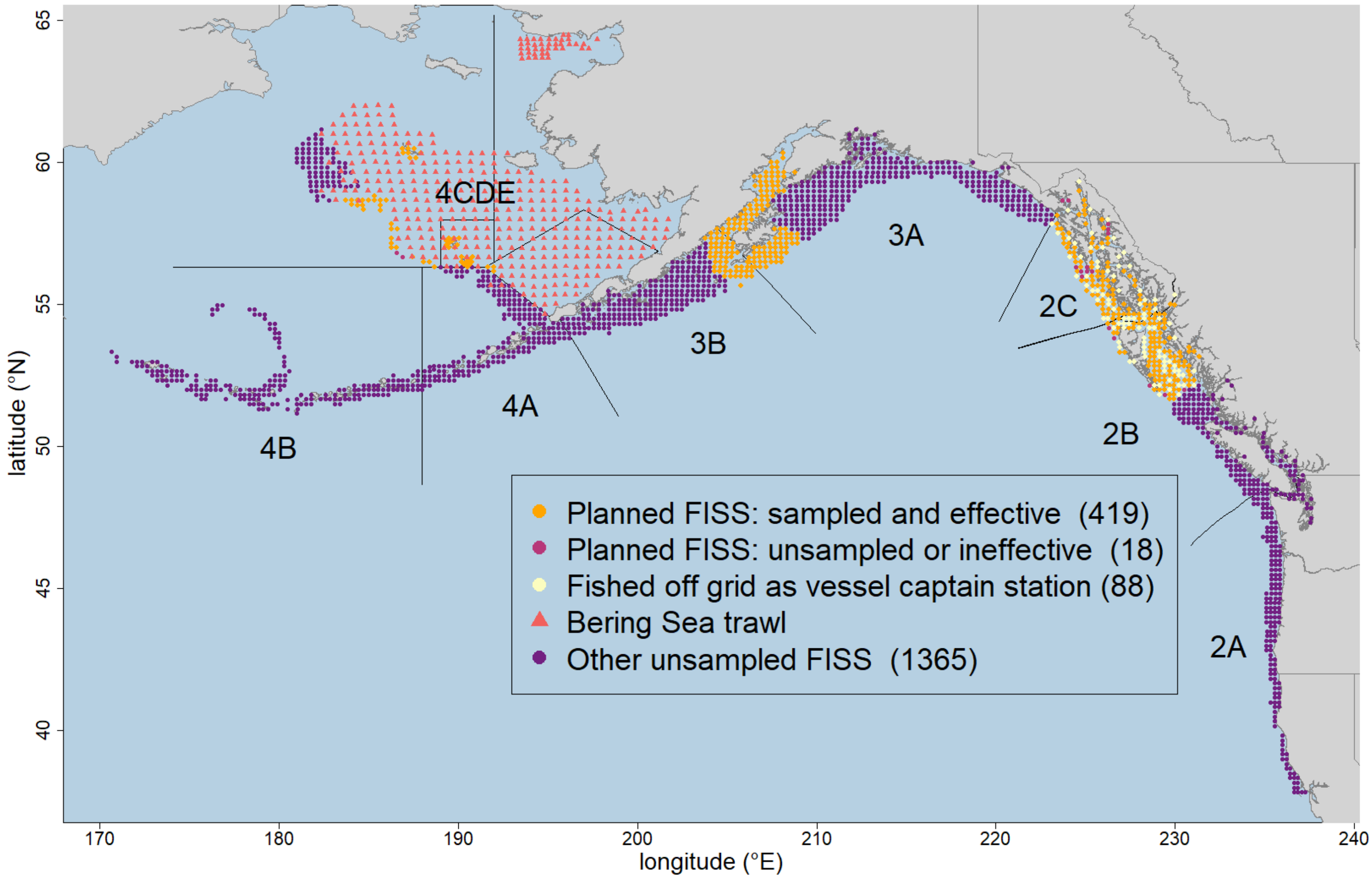


Figure 2. Map of implemented 2024 sampled FISS design showing stations with data used in modelling (orange circles for FISS, red triangles for trawl), along with planned but ineffective FISS stations, FISS grid stations fished off grid as vessel captain stations and other unsampled FISS stations.

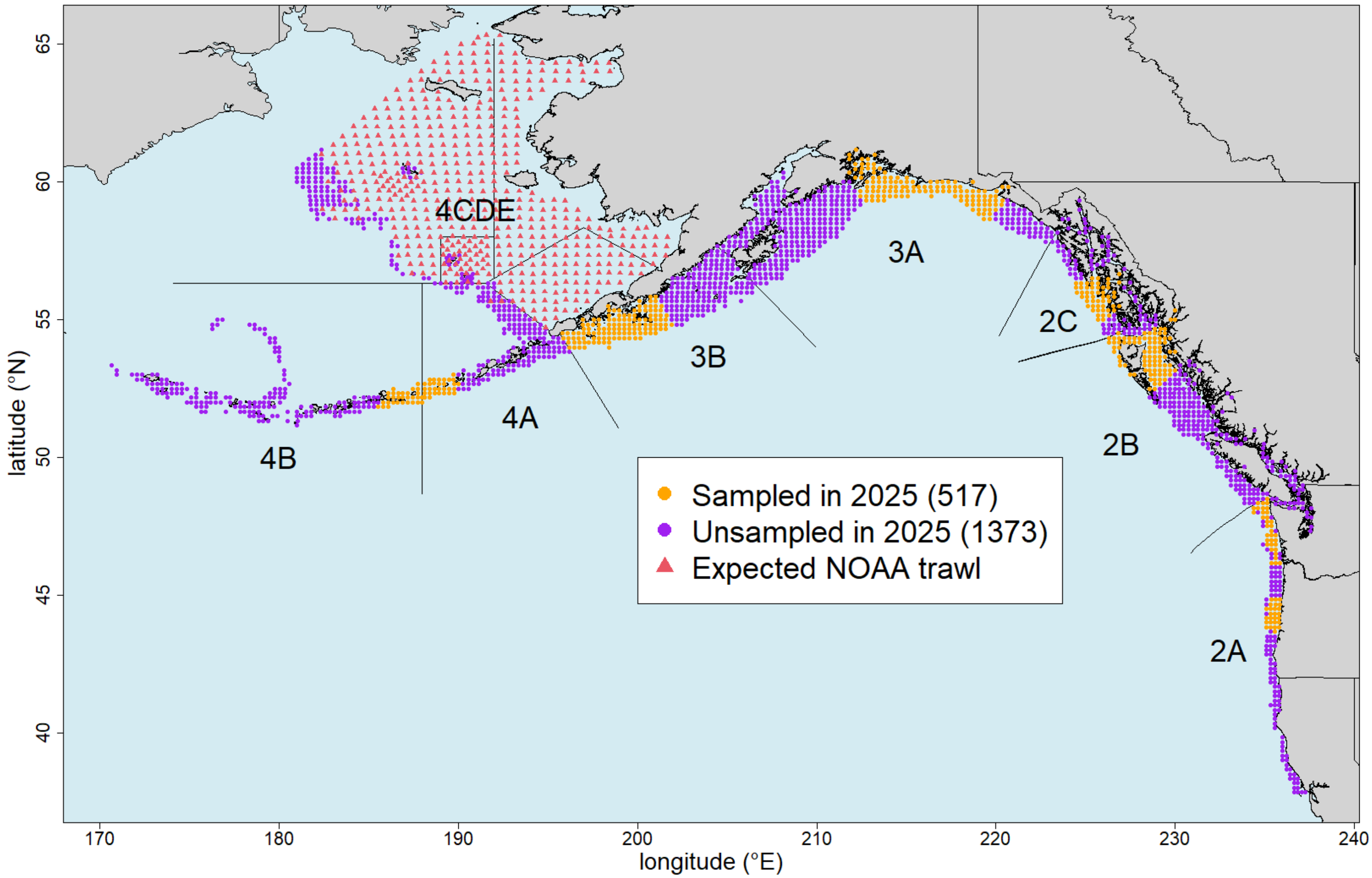


Figure 3. Adopted 2025 FISS design, with planned FISS stations shown as orange circles.

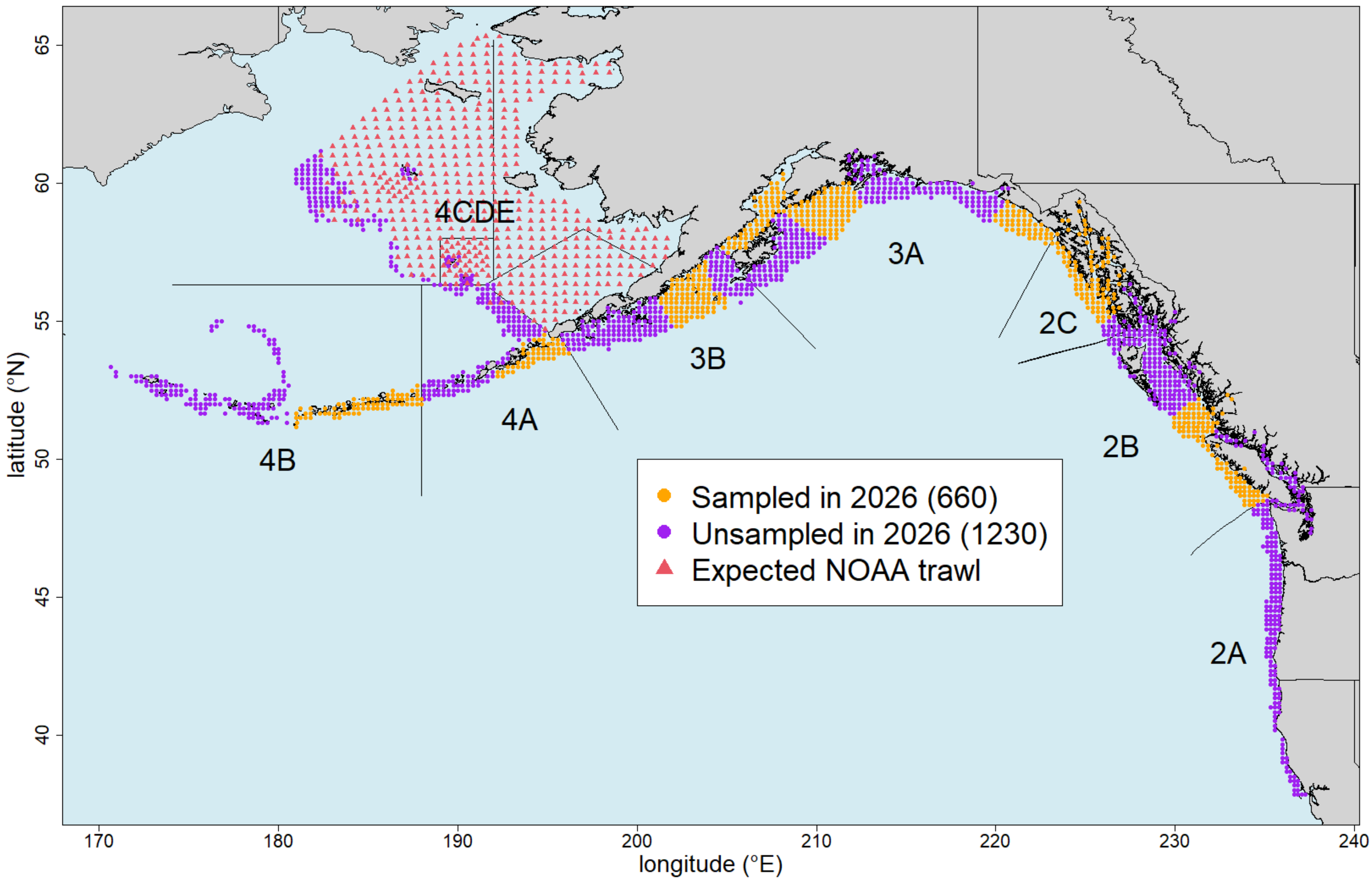


Figure 4. Base Block design for 2026 (orange circles). Design is based on fishing 2-4 complete blocks of stations (charter regions) in the core areas (2B, 2C, 3A and 3B) and previously implemented subareas elsewhere.

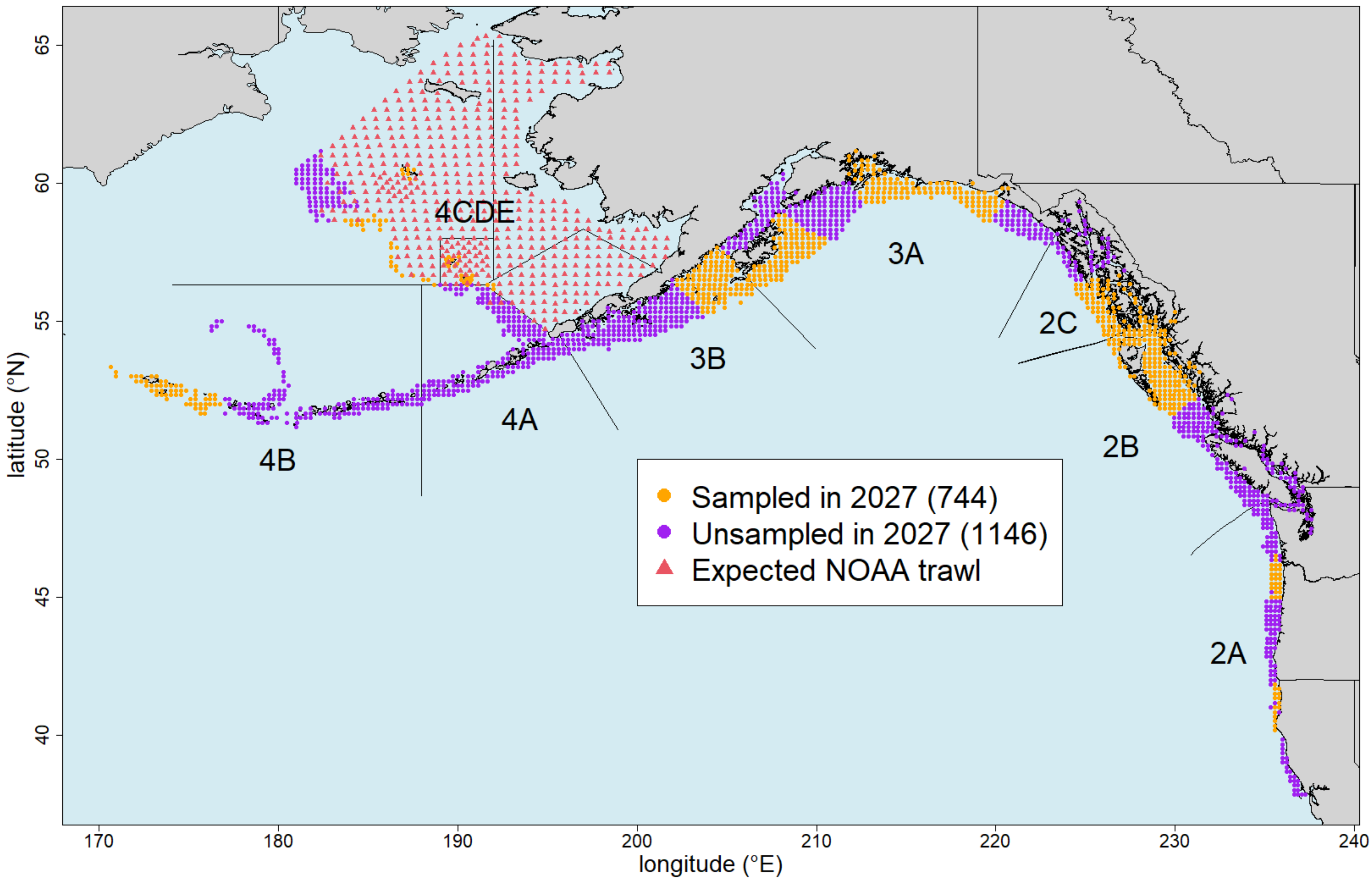


Figure 5. Base Block design for 2027 (orange circles). Design is based on fishing 2-4 complete blocks of stations (charter regions) in the core areas (2B, 2C, 3A and 3B) and previously implemented subareas elsewhere.

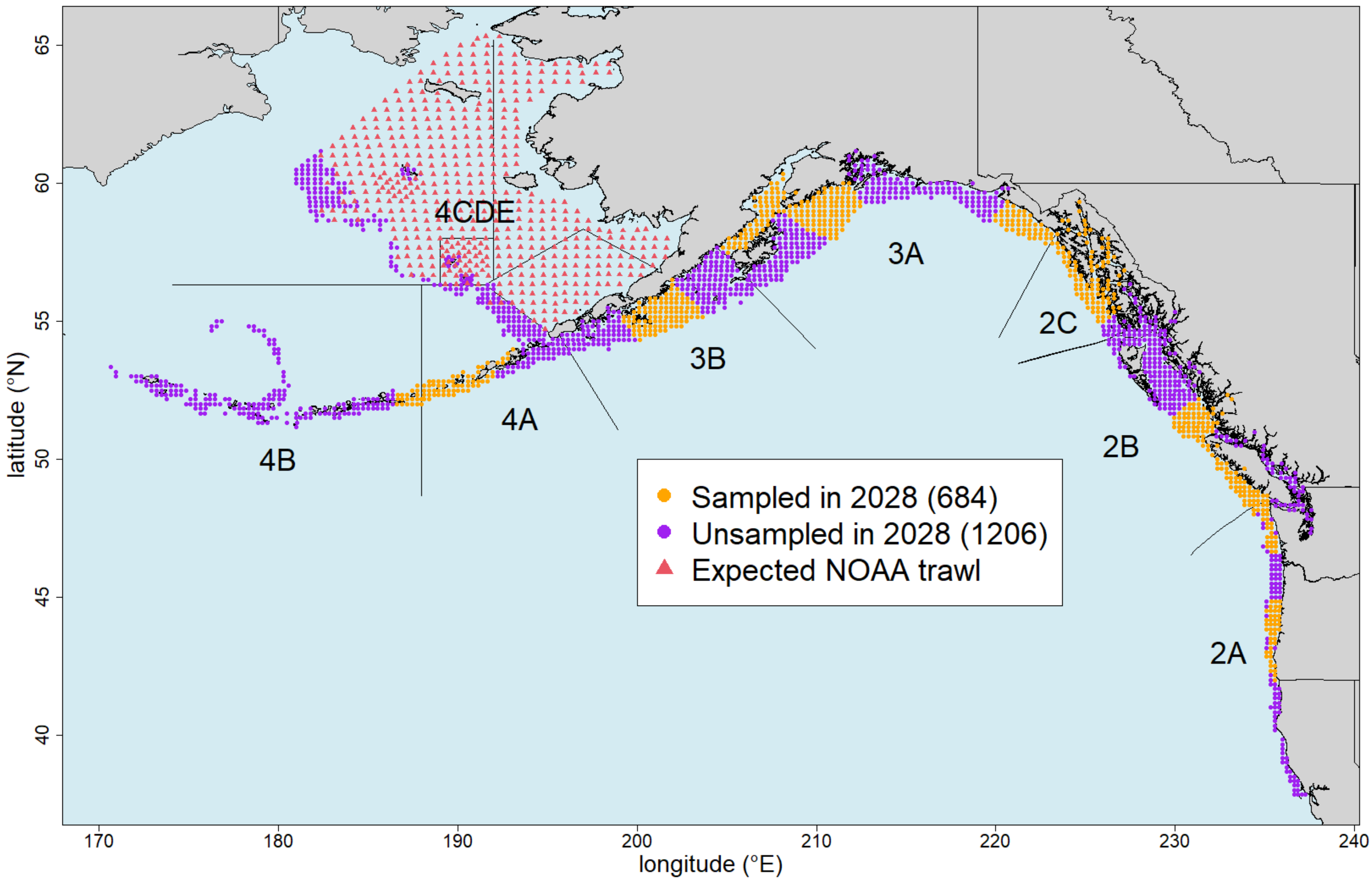


Figure 6. Base Block design for 2028 (orange circles). Design is based on fishing 2-4 complete blocks of stations (charter regions) in the core areas (2B, 2C, 3A and 3B) and previously implemented subareas elsewhere.

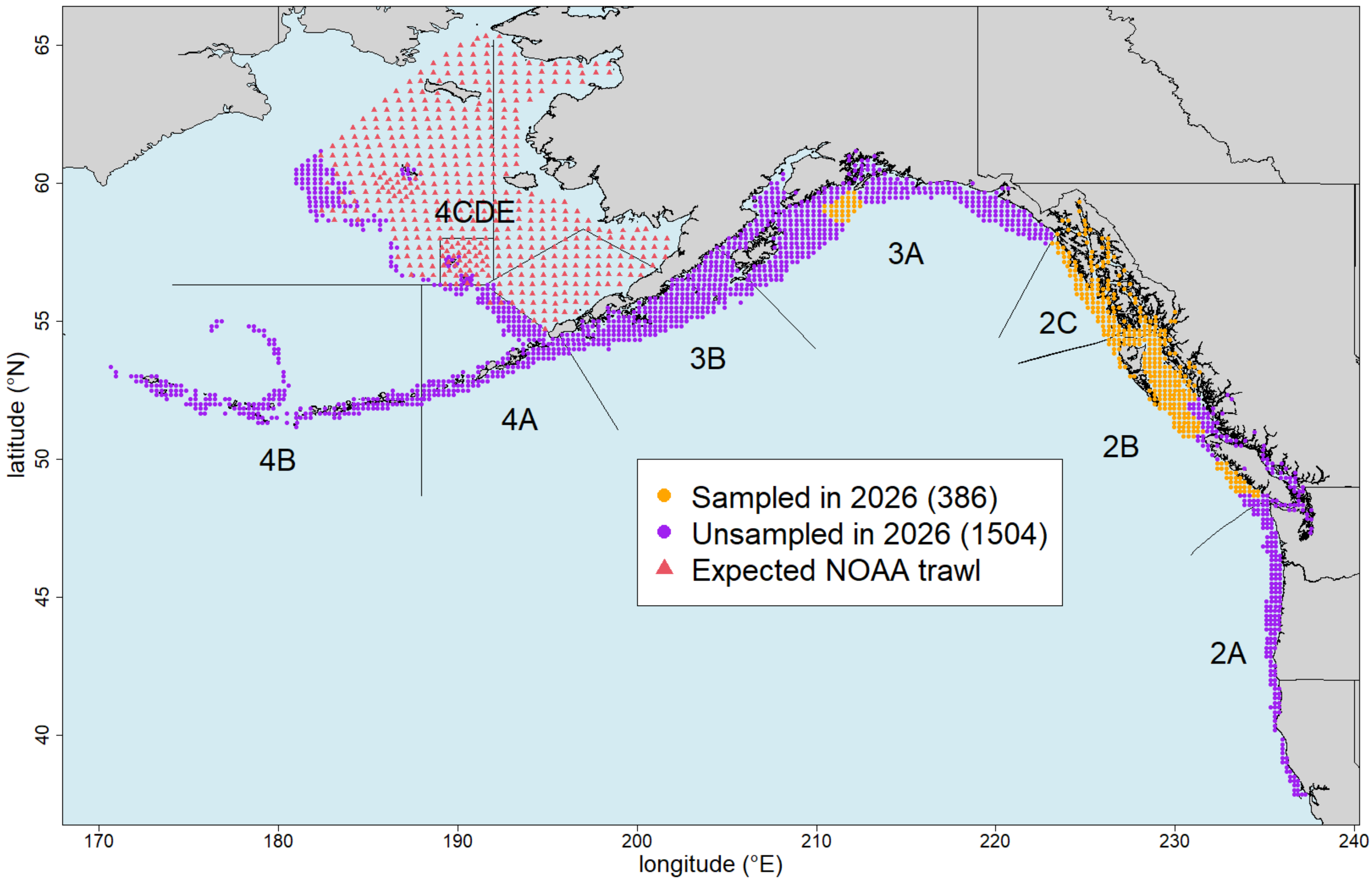


Figure 7. Preliminary Reduced Loss design for 2026 (orange circles). Note that stations in partially-sampled charter regions (2B and 3A) are only for the purpose of illustrating the spatial extent of the design. Actual stations to be fished within partially-sampled charter regions will be selected at a later date based on the priorities in Table 1.

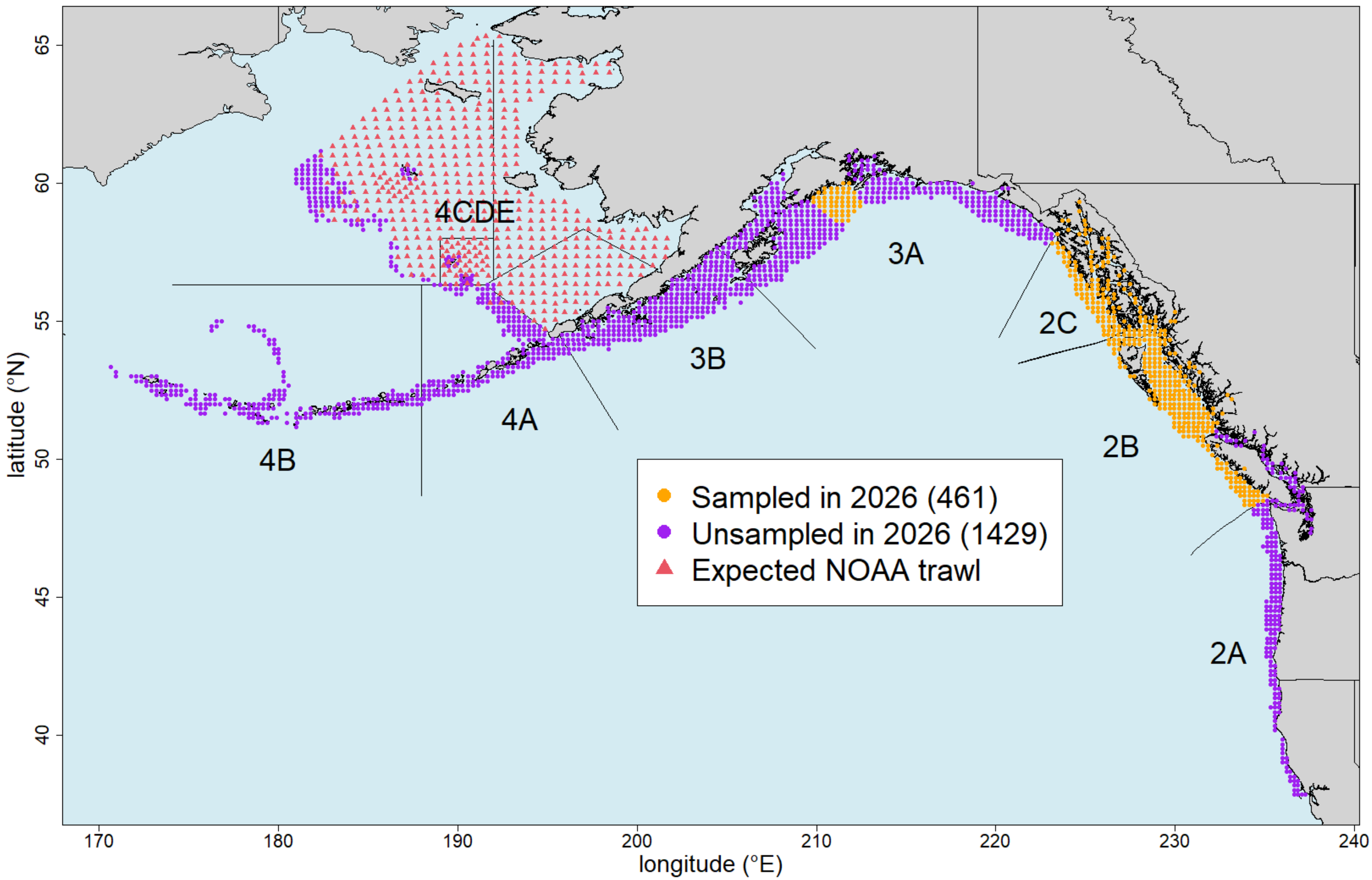


Figure 8. Preliminary Option 2 design for 2026 (orange circles).

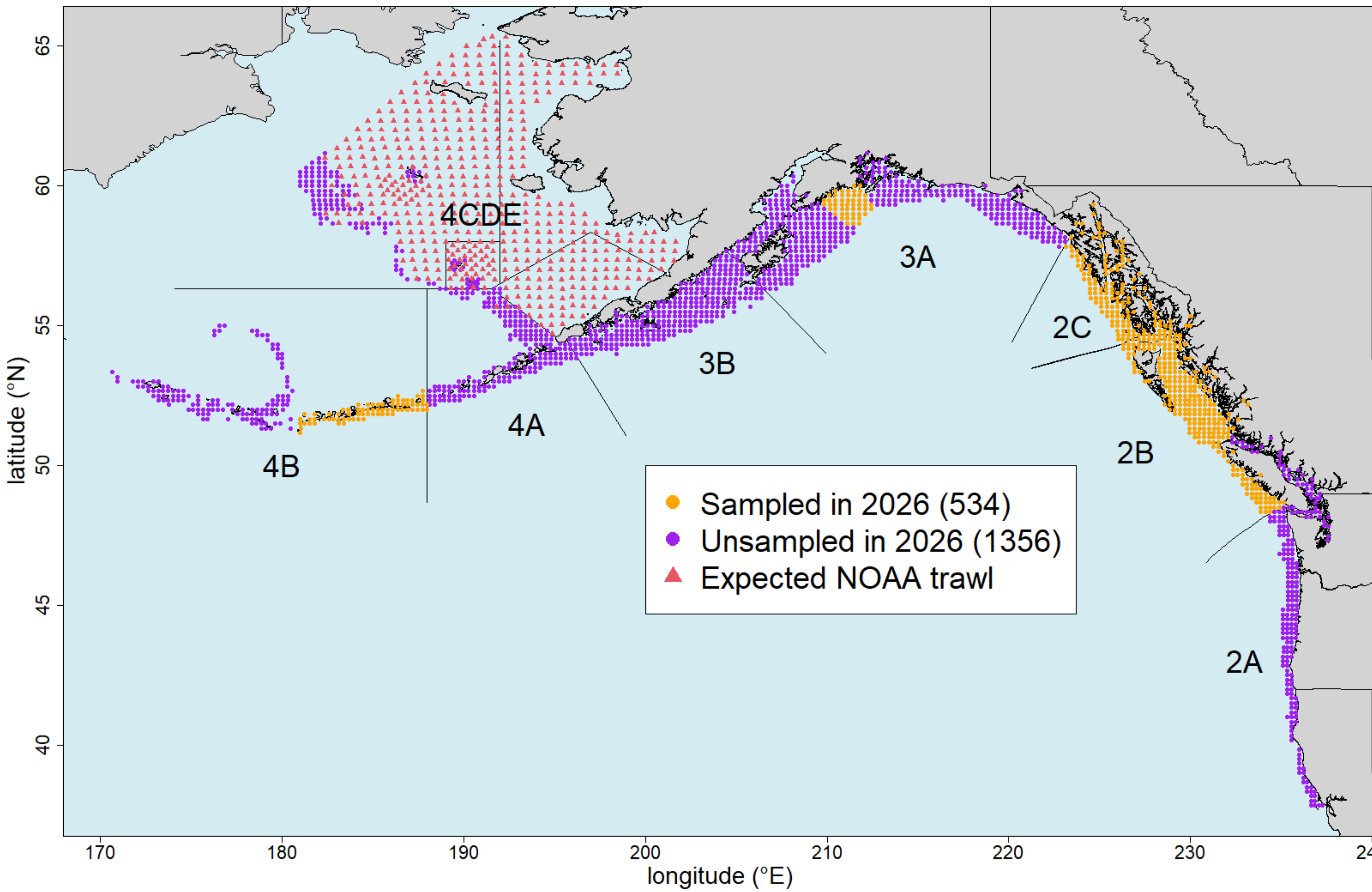


Figure 9. Preliminary Option 3 design for 2026 (orange circles).

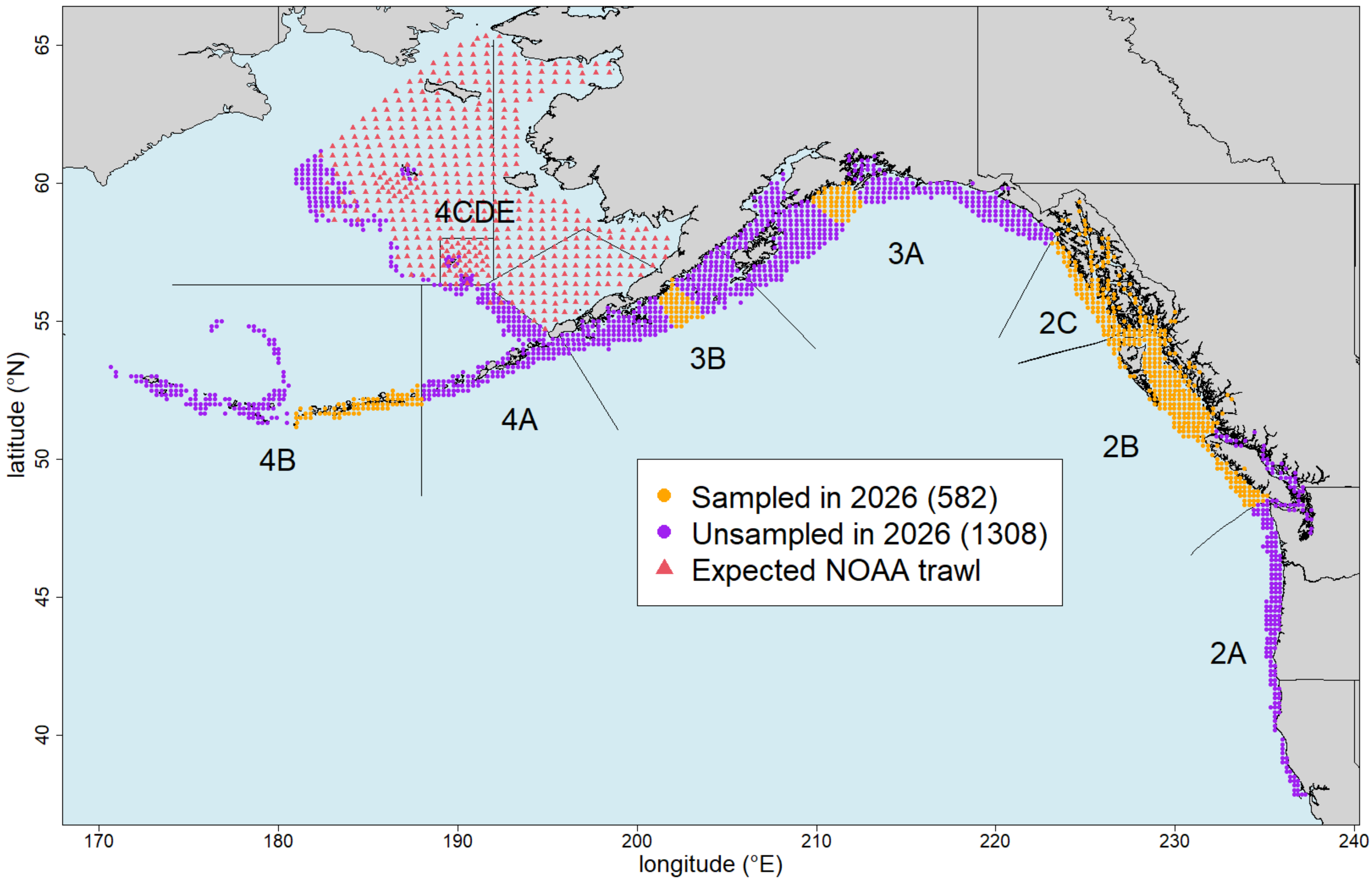


Figure 10. Preliminary Option 4 design for 2026 (orange circles).

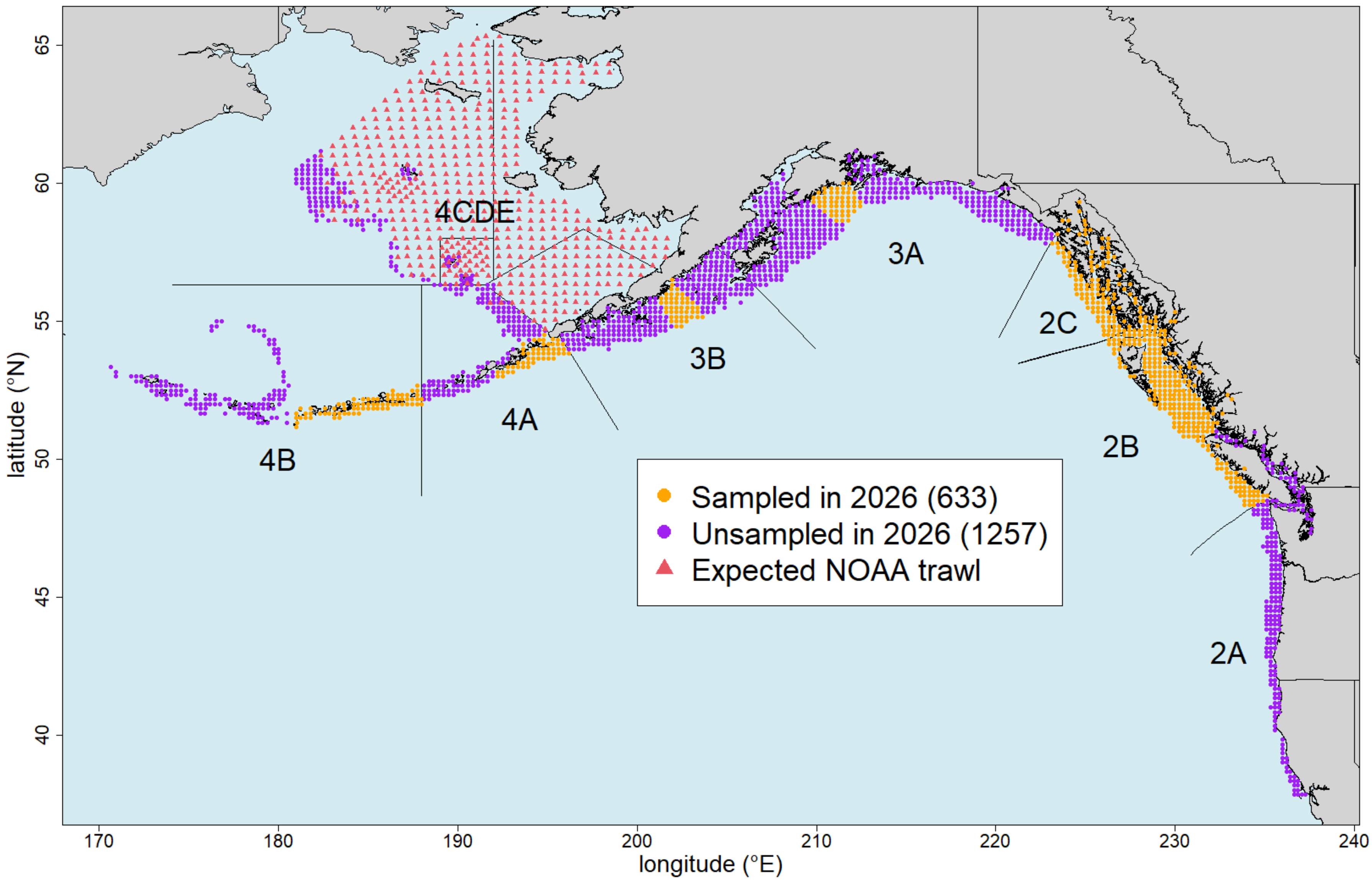


Figure 11. Preliminary Option 5 design for 2026 (orange circles).

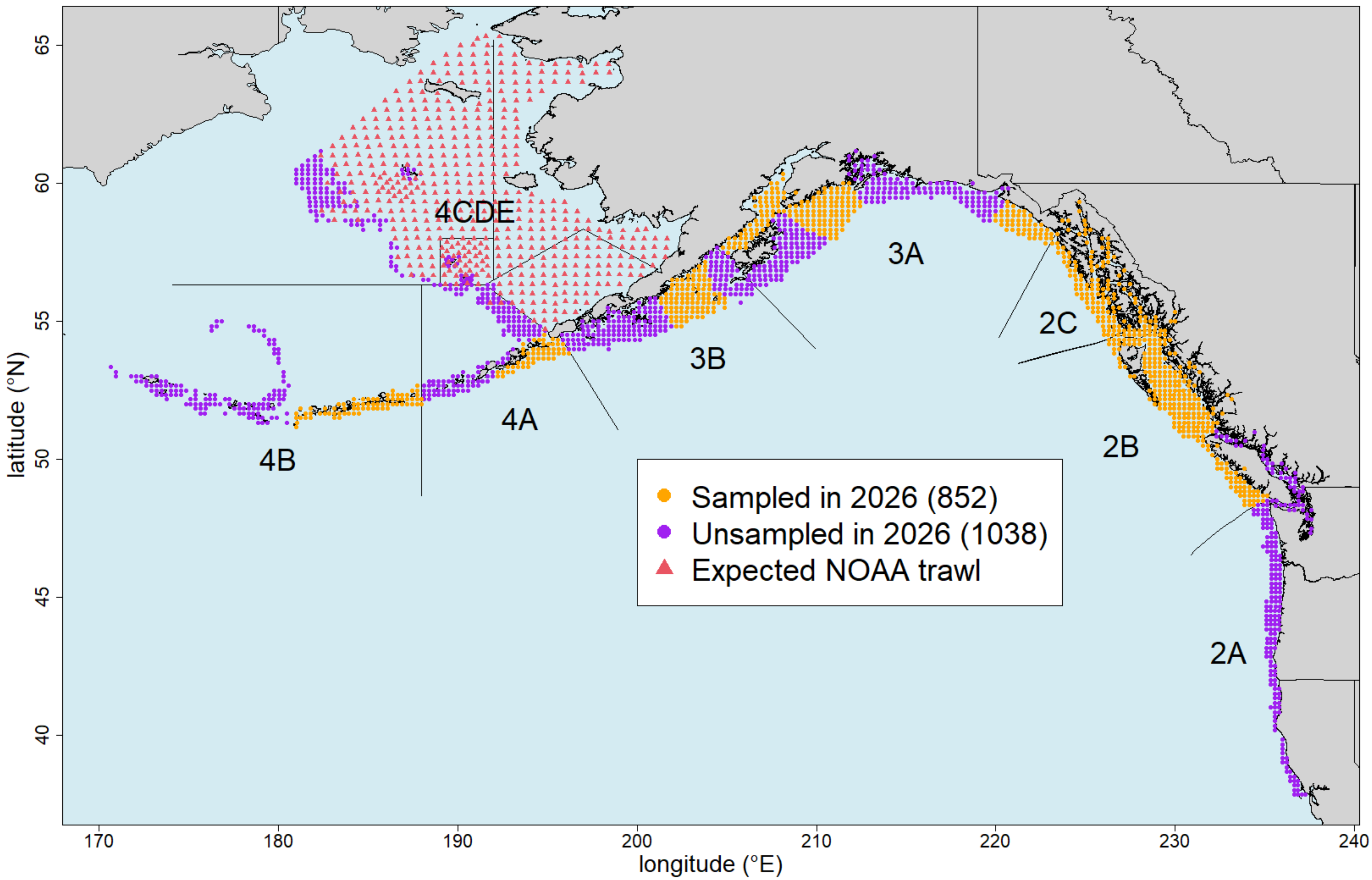


Figure 12. Preliminary Option 6 design for 2026 (orange circles).

Using artificial intelligence (AI) for supplementing Pacific halibut age determination from collected otoliths

PREPARED BY: IPHC SECRETARIAT (B. HUTNICZAK, J. FORSBERG, K. SAWYER VAN VLECK, & K. MAGRANE; 5 MAY 2025)

PURPOSE

This document summarizes the information available on the use of artificial intelligence (AI) for determining the age of fish from images of collected otoliths and provides an update on the exploratory work of implementing an AI-based age determination model for Pacific halibut.

The progress summarized in this document includes:

- Testing various deep learning architectures to identify the optimal approach given the available otolith images.
- Evaluating model generalization by comparing age predictions from a model trained on images from one year to those from a different year.
- Assessing differences in model performance between images of processed (sectioned and baked) and unprocessed (surface) otoliths.
- Utilizing confidence intervals derived from deep ensemble techniques to assess the model's capability in identifying ambiguous or noisy samples.
- Evaluating the model's performance in predicting the geographic region of sample collection.

The purpose of this document is twofold. First, it provides essential background information to support ongoing efforts in establishing a comprehensive database of otolith images with expert-provided labels for future ageing use. Second, it provides an update on the viability of an AI-based modeling approach for supplementing current Pacific halibut ageing protocol, while also outlining the remaining steps and requirements necessary for operational implementation.

BACKGROUND

Otoliths are crystalline calcium carbonate structures, mostly in the form of aragonite, found in the inner ear of fish. They contain growth rings, that are often compared to tree growth rings. By analyzing the growth patterns in otoliths, scientists estimate the age of fish (Campana, 1999; Campana & Neilson, 1985), supporting the estimation of fish population demographics and population dynamics (Campana & Thorrold, 2001). In turn, fish age is a key input to stock assessment models that inform management decisions related to fish exploitation (Methot & Wetzel, 2013). It is estimated that the number of otoliths from captured fish that are read annually worldwide is on the order of one million (Campana & Thorrold, 2001).

The current method for determining ages of most fish species relies on manually extracting, preparing (embedding, sectioning), and reading otoliths. The simplest approach to reading the otolith is to immerse it in a clear liquid, such as water or alcohol solution, illuminate it from above, and view it against a dark background, using a stereo microscope. This method is suitable only for otoliths that are relatively thin with all annual bands visible from the surface. For species such as Pacific halibut, as the growth rate of the fish slows down, the outer growth bands become increasingly compressed and difficult to read from the surface of the whole otolith. To correctly determine the number of annual bands in such cases, otoliths are typically viewed in cross section which allows viewing the bands that are not visible from the surface view. In addition, the contrast between the growth rings can be enhanced through the baking process. Pacific halibut otoliths are aged using the 'break and bake' technique.

This manual ageing process is expensive, time-consuming,[1] and can be subject to bias[2] as well as imprecision due to variations in age estimations between readers and within readers over time. Recent advances in imaging technologies and machine learning suggest that AI can assist in this process by automating the analysis of otolith images[3] and identifying and measuring the growth rings to determine age. AI algorithms can be trained on a large dataset of otolith images with known ages to learn the patterns and variations in growth rings. Once trained, the AI model can analyze new otolith images and predict the age of the fish based on the identified patterns in the image.

Using AI for age determination of Pacific halibut could improve consistency and replicability of age estimates, as well as provide time and cost savings to the organization, providing age data for reliable management advice. However, it's important to note that the AI model's accuracy depends on the quality and diversity of the training data, as well as the expertise of the scientists involved in training and validating the model. Regular validation and calibration with manual age determinations may be necessary to ensure the accuracy and reliability of the AI predictions. Thus, the proposed approach explores integrating AI-based age determination and traditional ageing methods for maximum accuracy of the estimates.

MODEL

Model framework

The proposed model framework (Figure 1) includes a continuous process of training the model using available labelled data (aged otoliths), querying the model to select the next sample, labeling or relabeling the selected sample, and enriching the model with newly labelled samples.

This model relies on automatized ageing that is supplementing the expert-derived age estimates continuously improving the model in the *Label* phase and the *Enrich* phase.

[1] While the actual reading may account only for a fraction of the total cost and time required to process the otolith from collection to age determination, skilled readers require years of training, which should be considered when conducting a cost-benefit analysis.

[2] While the count of annual rings on Pacific halibut otoliths was found to provide unbiased age estimate using validation against bomb radiocarbon isotopes (Piner & Wischniowski, 2004), an earlier oxytetracycline (OTC) mark-recapture study indicated biases among age readers (Blood, 2003). In the 1980s, the IPHC applied injections with the antibiotic oxytetracycline (OTC) during routine tagging operations to evaluate validity of ageing method (IPHC, 1985). Upon injection, the OTC is absorbed by the fish's bony structure, including the otoliths, and leaves a mark that is easily seen when viewed under an ultraviolet light. When an OTC-injected tagged fish is recovered, the otoliths are removed and examined under the ultraviolet light. By comparing the number of annuli laid since the OTC mark to the fish recovery, the accuracy of the age readings can be determined.

[3] Although the idea of taking pictures of Pacific halibut otoliths is not new. See 1960 report by G. Morris Southward, *Photographing Halibut Otoliths for Measuring Growth Zones* (Southward, 1962).

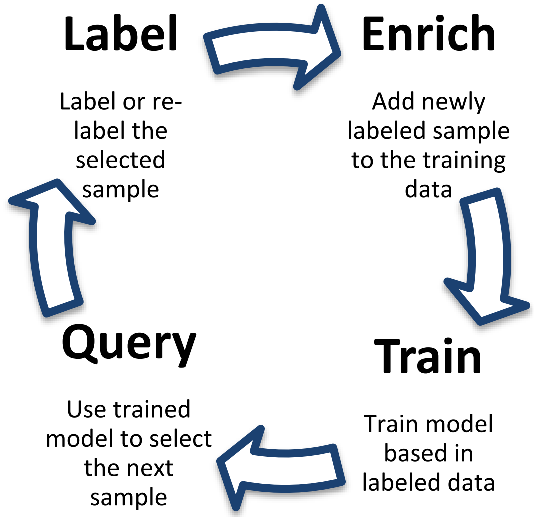


Figure 1. Model framework.

Modeling approach

Previous literature (see perspective piece by Malde et al., 2020) suggests adapting a pre-trained convolutional neural network (CNN) designed for image classification to estimate age using otolith images obtained via microscope camera. This type of model is trained on a large collection of images of otoliths previously aged by human readers. Moen et al. (2018) presents the first case of the use of deep learning and CNN to estimate age from images of whole otoliths of Greenland halibut (*Reinhardtius hippoglossoides*).[4]

Artificial neural networks (ANNs) are computational structures inspired by biological neural networks. They consist of simple computational units referred to as neurons, organized in layers. The neuron parameters (or weights) are estimated by training the model using supervised learning. This process consists of two steps: forward propagation, where the network makes a prediction based on the input; and back propagation, where the network learns from its mistake by calculating the gradient of a loss function, and then uses the gradient to update the neuron weights. The ANNs approach has been used for fish ageing by Robertson & Morison (1999) and Fablet & Le Josse (2005) with a limited success.

The neural networks approach significantly improved in recent years with the increase in the number of layers, applying an approach often referred to as deep learning. Deep learning neural networks are known for their generality. With sufficient training data, they can be used to classify raw data (e.g., an array of pixels) directly, without explicit design of low-level features. The deep learning algorithm lower layers learn to distinguish between primitive features automatically, typically identifying sharp edges or color transitions. Subsequent layers then learn to recognize more abstract features as combinations of lower layer features, and finally merge this information to provide a high-level classification.

In CNNs (LeCun et al., 1998; Simonyan & Zisserman, 2015), the layers are structured as stacks of filters, each recognizing increasingly abstract features in the data. Convolutional layers may be understood as an efficient way to transform an input image into another image, highlighting meaningful patterns, learned from data during training. The training is sequential, meaning the output of each layer is the input of the next layer, and the useful features are learned in the

[4] CNN was also applied for other tasks related to fisheries management, e.g. fish species identification (Allken et al., 2019).

various layers during training. This approach is very effective for many image analysis problems, where objects are often recognized independent of their location. During network training, the performance is monitored over sequential epochs. Epochs represent the number of times that the training dataset is passed forward and backward through the network to refine model weights. Whenever the validation loss decreases, the trained model is saved, ending up with the network that corresponds to the minimum loss and highest accuracy on the validation set. The trained network is then evaluated on the testing set.

In the CNN model, age prediction from otolith images can be formulated either as a classification task - where age is treated as a categorical variable - or as an image regression task, which involves predicting a continuous numerical value. Although treating fish age as a discrete parameter is a common method for identifying individual year classes, i.e., grouping fish by spawning year (Moen et al., 2018), this approach has proven less effective for Pacific halibut. As a long-lived species with a wide distribution of age classes, Pacific halibut pose a challenge for classification-based methods. The oldest Pacific halibut on record have been aged at 55 years (Keith et al., 2014).

Software and architectural options

The proposed approach builds on prior work by Moen et al., (2018) and Moore et al., (2019), who implemented CNNs for otolith-based fish age estimation using the TensorFlow and Keras libraries. TensorFlow remains one of the most widely used and well-supported frameworks for deep learning, and Keras provides a high-level API that simplifies TensorFlow model development.

The approach utilizes a transfer-learning technique to develop a CNN for otolith age estimation. Transfer learning is the process of repurposing a machine learning model that has been pre-trained for another, related, task. Specifically, it starts with the InceptionV3 model from Google, pre-trained on the ImageNet database. ImageNet database contains over 14 million annotated images classified into 1,000 categories. By loading CNN layers with publicly available pre-trained weights rather than random initialization, transfer learning significantly enhances model performance.

To adapt this model specifically for Pacific halibut ageing, modifications included scaling the input layer to match otolith images' resolution[5] and changing the output from multi-dimensional class probabilities to a single numeric output for regression.[6] Thus, the architecture employed follows the pattern: Input → InceptionV3 (feature extractor) → Regressor → Output, optimized

[5] Resolution is the total number of pixels along an image's width and height, expressed as pixels per inch (PPI). The Inception v3 model processes images that are 299 x 299 pixels in size. The original images (2548 × 2548 pixels) were first resized to 400 × 400 pixels prior to input into the model. This intermediate resizing step preserves more visual detail than a direct downscaling to 299 × 299 and allows for subsequent data augmentation operations (such as cropping, flipping, or rotation) to be applied more effectively before the final resize to the model's required input size.

[6] Alternatively, Politikos et al. (2021) replaced the last layer with a feed-forward network with two hidden layers replacing the default 1000-categories output layer with a fully-connected layer with six hidden nodes, corresponding to a limited number of age categories [Age-0 – Age-5+], with the last one representing fish of age 5 and older, In this case, the network outputs probabilities using the softmax function, a function that performs multi-class classification and transforms the outputs to represent the probability distributions over a list of potential outcomes. The IPHC uses in its stock assessment bins Age-2 – Age 25+ for the current age data and Age-2 - Age-20+ for the historical surface read ages. The adoption of a larger number of age categories prompted the decision to incorporate a regression layer in place of class probabilities.

using stochastic gradient descent (SGD) to minimize mean squared error (MSE) between model predictions and expert annotations.[7]

A similar approach, although adopting classification approach, was applied for ageing Greek Red Mullet (*Mullus barbatus*) (Politikos et al., 2022) and the associated code is available on GitHub (github.com/dimpolitik/DeepOtolith). The available open-source code was adapted to test the approach for Pacific halibut.

In addition to the InceptionV3 architecture, alternative architectures were explored to identify potentially superior performance or efficiency advantages. These included EfficientNet variants (EfficientNetB4, EfficientNetB5, EfficientNetV2 S/M/L) and ConvNeXt. EfficientNet architectures are known for their balanced approach to scaling depth, width, and resolution, optimizing computational efficiency and accuracy. EfficientNetV2 further refines this by introducing progressive training and improved scaling techniques. ConvNeXt architectures, inspired by transformer models, incorporate modifications to convolutional structures, achieving competitive accuracy with a simplified design and potentially improved model interpretability.

While TensorFlow/Keras has been the primary framework used in the current implementation, future work may explore alternative frameworks such as PyTorch (originally developed by Meta), which offers flexible dynamic computation graphs and growing adoption in the deep learning research community.

Performance metrics and achieved accuracy

Performance of the CNN to correctly assign ages (rounded output of the regression layer) to otolith images in the test set is assessed via the root mean squared error (RMSE) and the percentage of correctly predicted ages, as well as predictions within ±1 year tolerance. Moen et al., (2018) also suggest calculating coefficient of variation (CV).[8]

Moen et al., (2018), for Greenland halibut, achieved MSE for the left and right otoliths and pair of 3.27, 2.71 and 2.99, respectively. Age was correctly estimated for 48 out of the 164 tested otolith-pairs (29%). In addition, 63 cases (38%) were estimated to be one year off the read age. There was also a clear tendency for the system to predict a lower age for older individuals, when compared to human readers. The variance of the predictions also increased with the age of the otolith.

The model developed by Moore et al. (2019), for prediction of age of snapper using CT scans,[9] gave the same age as the human reader for 47% of otoliths in a test dataset, with a further 35% of ages estimated within 1 year of the human reader estimate of age (n=687). For hoki, the model gave the same age as the human reader for 41% of individuals (n=882).

The age model for Greenland halibut by Politikos et al., (2022) gave RMSE of 1.69 years between age prediction and age reading by experts (n=8,218, 26 age categories). For Greek

[7] In practice, the neural network minimizes the MSE of normalized age values, i.e., age values divided by the maximum age provided as input.

[8] The CV of the predicted age at true age is the primary input to the IPHC stock assessment. It is generally modelled as a parametric function of age accounting for the complex joint probability that both estimates can be incorrect (Punt et al., 2008).

[9] CT scanning uses X-ray technology to produce image slices through objects, which can be reconstructed into virtual, three-dimensional (3D) images that can be rotated and viewed in any orientation (Moore et al., 2019). Such images may provide more accurate estimates, but the cost of this approach is prohibitive at (based on trial conducted in New Zealand) $1,500 per day, with scan timed for an individual otolith between 40 min to one hour. However, as the technology progresses, this approach may provide an option for fully automating the entire ageing process by scanning a whole fish (e.g., along a conveyor belt). Deep learning methods (i.e., CNN) developed for age determination from surface images could serve as a base for age determination from CT scans.

red mullet, correct age was predicted for 69.2% individuals, with an additional 28.2% being within 1 year of error (n=5,027).

Benson et al., (2023), using near-infrared spectroscopy of otoliths, supplemented by geospatial and biological data routinely collected on the survey, estimated age of walleye pollock. For the optimal multimodal CNN model, an RMSE of 0.83 for the training set and an RMSE of 0.91 for the test set indicated that at least 67% of estimated ages were predicted within ±1 year of age compared to traditional microscope-based ages.

However, it should be noted that neither the traditional ageing methods for Pacific halibut are perfectly accurate. Within- and between-reader agreement in age assignment is generally 60%-70% complete agreement, 80% to 90% within one year, and 100% within 3 years. The IPHC Secretariat's publications report on % agreement (see *Technical Report No. 46* and *No. 47*).

Use of auxiliary data

The accuracy and precision of age predictions from otolith images using neural networks could potentially be enhanced by incorporating auxiliary data into the modeling process (Moen et al., 2018). For example, the geographic location where fish are captured could offer valuable supplementary information to the model. Past IPHC work suggests a good deal of spatial variation in Pacific halibut growth ring patterns. This points to the importance of good spatial coverage in the training sample.

The project plans to explore the integration of spatial covariates, such as latitude, longitude, or defined regulatory areas, to refine age predictions. Inclusion of these spatial factors could help the neural networks better interpret and account for region-specific growth patterns that influence otolith formation. Other available auxiliary data include collection year, which could be applied to account for variation between cohorts and prevalent environmental conditions throughout the aged fish life histories, and the collection dates, which provide insights into seasonal variation to the interpretation of the otolith edge.

Database

The IPHC annually ages a considerable number of otoliths (see *Appendix A* for details). Since 1925, over 1.5 million otoliths have been aged and stored for potential future use. Otoliths collected by the IPHC for ageing purposes undergo additional processing. Otoliths are sectioned (broken in half) and baked to enhance the contrast between the growth rings. These stored and previously aged otoliths serve as a valuable resource for creating a database of images for training purposes. To optimize model training, the selection of otoliths included in the model covers a broad spectrum of fish sizes, ages, sexes, and collection locations.

Before photographing, processed otoliths were placed in a monochrome tray featuring an elongated groove designed to keep the otolith upright and immersed in water. The pictures were taken with AmScope 8.5MP eyepiece cameras,[10] under consistent lighting conditions and magnification. The input database includes images of standardized size, 2,548 by 2,548 pixels, which are later resized to the desired resolution based on the model's specification.[11]

[10] The camera fits in one of the microscope eyepieces, eliminating the need to purchase a separate camera mount for the microscope.

[11] Moen et al. (2018) used images 400 by 400 pixels, which required the input layer to be scaled to match the Inception V3 requirements (299 by 299 pixels). Ordoñez et al. (2020), using the same set of images, built a CNN with images resized to 224 by 224 pixels, the default input of the VGG-19 model. Higher resolution images offer the flexibility to adapt the model in the future to more detailed and complex image analysis tasks, potentially improving the accuracy and effectiveness of image recognition capabilities.

It is important to note that it may not be necessary to image the otoliths at resolutions sufficient for human viewers to resolve, because the CNN may be able to arrive at an age estimate without directly counting bands (Moore et al., 2019).

Figure 2 shows an example of a range of images used in the CNN training dataset.

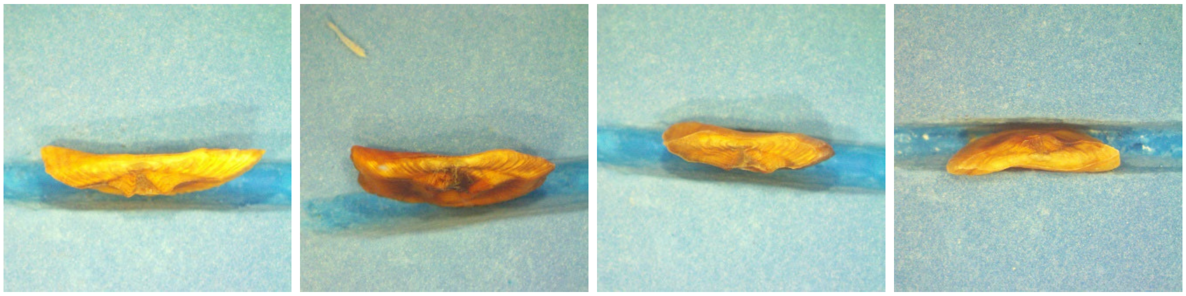


Figure 2. Examples of Pacific halibut otolith images taken for inclusion in the training set.

In addition, the IPHC is in the process of creating complimentary database comprising labelled images of otoliths captured prior to processing to conduct a cost-benefit analysis of using processed versus unprocessed otoliths for AI-based age determination. Example images are provided in Figure *3*. In their research, Politikos et al. (2022) utilized digital images of otoliths that were not subject to any additional processing in the laboratory, immersed in water and placed under a stereomicroscope on a white background with transmitted light. However, it is important to note that even if results indicate that breaking and baking is not necessary for age determination using AI, a subsample chosen for the Label and Enrich phases would have to be fully processed for age determination with traditional methods by an expert reader.

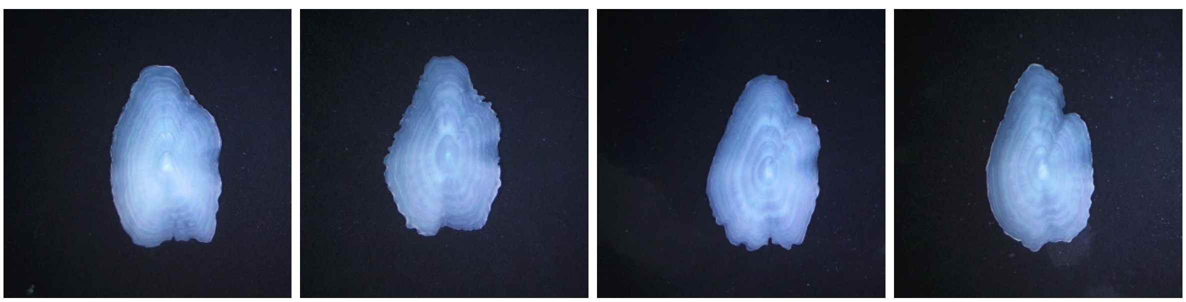


Figure 3. Examples of Pacific halibut otolith images taken for inclusion in the training set.

Presorting otoliths

The adopted procedure excludes broken otoliths, applying manual presorting at the image-taking stage. Presorting has also occurred at the collection stage when crystalized otoliths[12] are omitted when collecting samples.

Ongoing research [Dimitris Politikos, personal communication] is investigating the initial stage of the aging process, specifically assessing whether an otolith is of sufficient quality for age determination. This research is relevant for cases involving crystallized or broken otoliths and aims to potentially eliminate the need for subjective decisions by samplers regarding the usability of otoliths for age determination. This approach implements a two-stage classification system. In the first stage, the model assesses the otolith's suitability for ageing; in the second, it

[12] Crystalized otoliths have an altered composition – specifically, where the aragonite in the otolith is partially or mostly replaced by vaterite, a phenomenon known as otolith crystallization. Crystallized otoliths are not suitable for ageing.

determines the age. Th algorithm-driven presorting could also incorporate expert knowledge for handling problematic otoliths.

In developing the model, the training dataset can be strategically supplemented with images of samples that represent a group of otoliths with which the original model struggles the most (Query phase).[13]

Image collection

The image collection is associated with labels storing:

1. Otolith reference number – using referencing system already in place;
2. Image name and location – exact path for image access;
3. Resolved age – human reader derived age (**rsvage**);
4. Year collected – to account for variation between cohorts and prevalent environmental conditions;
5. Date collected – to account for the ‘edge effect’ reflecting seasonal changes;
6. Geospatial characteristics of the collection site (latitude, longitude and IPHC Regulatory Area) – to capture regional variation;
7. Resolved sex – to determine whether otolith characteristics (possibly not directly visible to human eye) could be used for sex determination.[14]

Uncertainty estimates

To further refine accuracy in a production setting, a mixed-method approach can be applied. This approach involves selecting a subset of otolith images - e.g., 10% or 20 % - for re-examination by human experts, focusing specifically on cases where the AI model expresses low confidence in its predictions. These selections would be guided by model-derived uncertainty estimates. The newly relabeled samples can then be incorporated into the training set for annual fine-tuning, contributing to ongoing model improvement in a resource-efficient and targeted manner.

In practice, this strategy would allow human experts to focus on “difficult” otoliths—those with high uncertainty—while automating the processing of “easy” ones with high model confidence. This hybrid workflow enhances throughput without compromising the accuracy and consistency necessary for applications such as stock assessment, where minimizing systematic bias is critical.[15]

Two approaches were considered for quantifying model uncertainty:

- **Monte Carlo dropout** (Gal & Ghahramani, 2016): This technique involves performing multiple forward passes through the model with dropout layers activated during inference. The resulting variability in predictions across passes is used to estimate confidence intervals. Monte Carlo Dropout is computationally efficient and easy to implement, and it provides a useful proxy for identifying ambiguous or noisy samples. This form of per-sample uncertainty is also referred to as training dynamics or soft loss tracing.

[13] About 1% of otoliths are partly crystallized and are assigned ages. The same is true for broken otoliths that are aged (1%)

[14] IPHC is currently using genotyping for Pacific halibut sex determination.

[15] If there is a strong junction in the relative precision between old and younger fish due to the change in methods this may require a nonparametric approach to ageing imprecision. If an AI method is biased as a function of age (standard for surface reading methods) and the break and bake method is unbiased, integrating the methods may prove challenging.

- **Deep ensembles** (Lakshminarayanan et al., 2017): This approach involves training multiple independently initialized models and aggregating their predictions to form a consensus output. The variance across ensemble members serves as an estimate of prediction uncertainty. Deep ensembles are generally more robust than Monte Carlo Dropout, especially in identifying out-of-distribution samples and capturing both model and data uncertainty. Their main advantage lies in their improved predictive performance and better-calibrated confidence intervals, though at the cost of increased computational resources.

Together, these tools support the design of a semi-automated, quality-controlled ageing protocol that leverages the strengths of both AI and human expertise.

PRELIMINARY RESULTS

Comparison of model architectures

Several modern CNN architectures were systematically evaluated to determine the most suitable approach for ageing Pacific halibut using otolith images. The architectures tested included:

- **InceptionV3**: A widely used CNN known for its balanced computational efficiency and accuracy.
- **EfficientNet (B4, B5, V2 S/M/L)**: Architectures optimized for scaling model depth, width, and resolution uniformly, enhancing computational efficiency and predictive accuracy.
- **ConvNeXt**: Inspired by transformer-based models, ConvNeXt utilizes modified convolutional operations aiming to simplify model complexity while maintaining competitive performance.

Each architecture was adapted via transfer learning, leveraging publicly available pre-trained weights from the ImageNet database, and subsequently fine-tuned specifically for the task of Pacific halibut age prediction. Adaptations involved resizing input images to match each architecture's requirements and adjusting the output layer to perform regression predicting age as a continuous numeric value.

The models were evaluated using standardized procedures to ensure valid and robust comparisons. The main evaluation criteria included:

- RMSE, percentage of exact age matches, and percentage within ±1 year tolerance between predicted ages and expert-provided ages for a test set of images collected within the same year as those used for training (without image overlap).
- RMSE, percentage of exact age matches, and percentage within ±1 year tolerance for a second test set comprising images collected five years after the training images, providing an assessment of temporal generalization.

The evaluation involved multiple experimental runs to ensure robustness. Selection of model run configurations and evaluation results are provided in *Appendix 2*.

The comparative evaluation revealed significant performance differences among tested CNN architectures. Despite their advanced theoretical advantages - such as better scalability, computational efficiency, and deeper learning capabilities - EfficientNet and ConvNeXt models underperformed relative to the simpler InceptionV3 architecture. Several configurations of EfficientNet and ConvNeXt exhibited limited learning, with predictions regressing toward the mean age of the test dataset. This outcome suggests that these more complex models struggled to extract meaningful age-related features from the otolith images, likely due to a combination of insufficient training data and overfitting driven by model complexity.

In contrast, the InceptionV3 architecture consistently derived more accurate and reliable predictions, suggesting that its simpler structure is more suitable given the current limitations in dataset size and variability. However, the selected final InceptionV3 configuration presented in this update demonstrates substantial improvements compared to previously evaluated models (IPHC-2024-SRB025-10). Driven by the goal of improved temporal generalization, the new model applies more aggressive image augmentation strategies,[16] an adaptive learning rate and better tuned training parameters. These methodological enhancements contribute to improved model performance and predictive reliability.

Selected model evaluation

The selected model configuration utilized 2,799 images of otoliths collected during the 2019 IPHC fishery-independent setline survey (FISS). The 2019 FISS represents a comprehensive sampling effort expected to reflect regional variability in Pacific halibut otolith characteristics. As such, it provides a robust foundation for initial model development and evaluation.

The images were divided into training, validation, and test datasets. The training set (1,665) was used for training purposes. The validation set (294) was used to evaluate the model during the training process, allowing for adjustments without using the test set, which was reserved for the final evaluation. The test dataset (30%, 840) was used to assess the performance of the model after training, providing an unbiased evaluation of its generalization capability to new, unseen data. Additionally, a separate set of 2,704 images of otoliths collected during the 2024 FISS was used to verify model performance on additional unseen data, testing the temporal generalization of the model configurations. All images were resized to 400x400 pixels. Images of broken otoliths were excluded.

The selected model employed a maximum of 600 training epochs, with early stopping patience set to 80 epochs. A learning rate reduction was triggered if validation loss plateaued for 40 epochs, reducing the rate by a factor of 0.6. The initial learning rate was set at 0.0002, and training was performed using a batch size of 16. A comprehensive suite of image augmentation techniques (e.g., rotation, zoom, flipping, brightness variation) was applied to improve generalization and robustness.

To enhance model reliability and quantify uncertainty, a deep ensemble approach was adopted. The model was trained 15 times, each with a different random seed. Ensemble outputs were averaged to produce final predictions and calculate prediction uncertainty. Detailed results for individual ensemble members are provided in *Appendix C*.

Across ensemble runs, the model trained for an average of 288 epochs (208 effective epochs with early stopping set at 80). It achieved a normalized MSE of 0.00016 on the validation set and 0.00188 on the test set. When results were rounded to the nearest integer age, the average RMSE for the test set was 1.80. On average, the ensemble predicted the exact age correctly for 30.3% of test images, and an additional 41.7% were within ±1 year of the manually assigned age, resulting in a total agreement within 1 year for over 70% of cases.

Figure *4* illustrates the evolution of model accuracy over training epochs for one representative run. Figure *5* shows a comparison between manually derived ages and AI-predicted ages across the ensemble. Figure *6* compares the age composition estimated manually with that derived from the ensemble model predictions.

[16] Rotation range=360, width shift range=0.1, height shift range=0.1, brightness range=[0.95, 1.05], and zoom range=[0.98, 1.02].

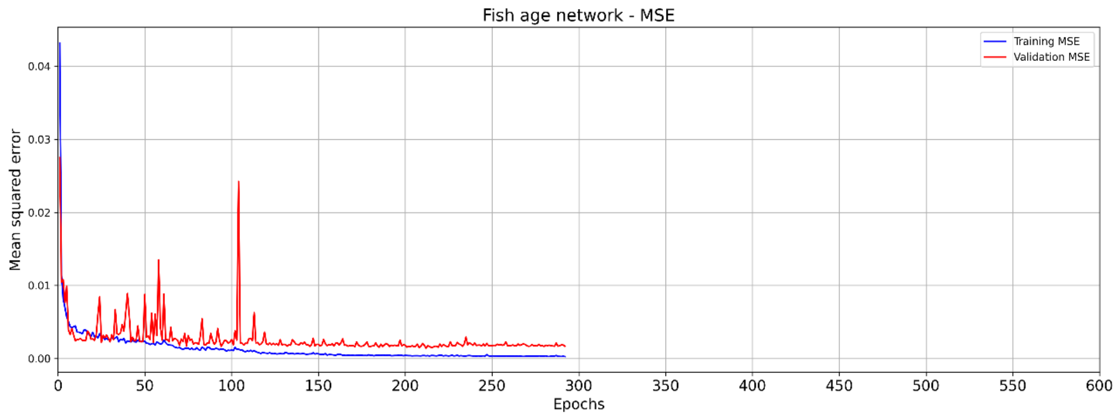


Figure 4. Age accuracy (measured as normalized age MSE) throughout the training process (example for seed 19).

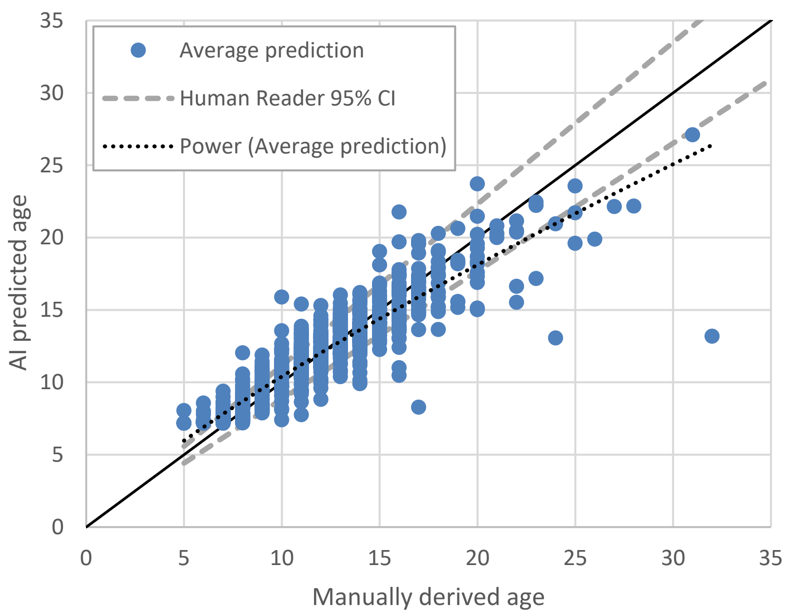


Figure 5. Comparison between manually derived age with AI predicted age.

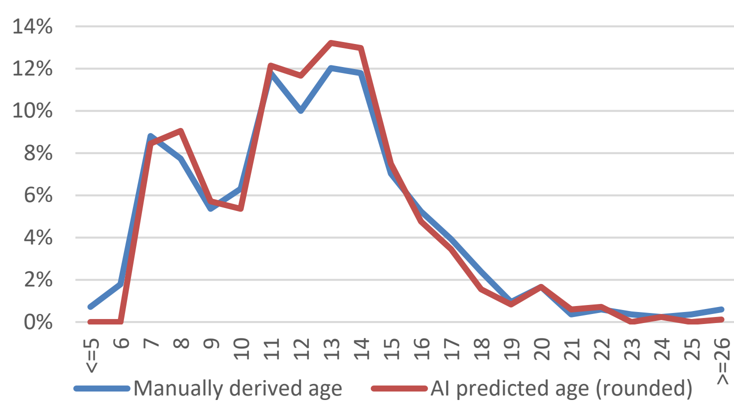


Figure 6. Comparison between manually derived age with AI predicted age – age composition.

It is important to note that statistically significant bias was observed mainly in age categories 21+ (increase from 16+ reported in IPHC-2024-SRB025-10). The number of observations for older age categories remains low despite an overall increase in sample size (Figure 7). This suggests that the saturation point for achieving optimal accuracy in older age categories may not yet have been reached, and the model could benefit from further improvement by adding more images representing older age categories to the training set. Currently, only 2.6% of the otoliths (74 samples) used in the model were from fish aged 21 or older.

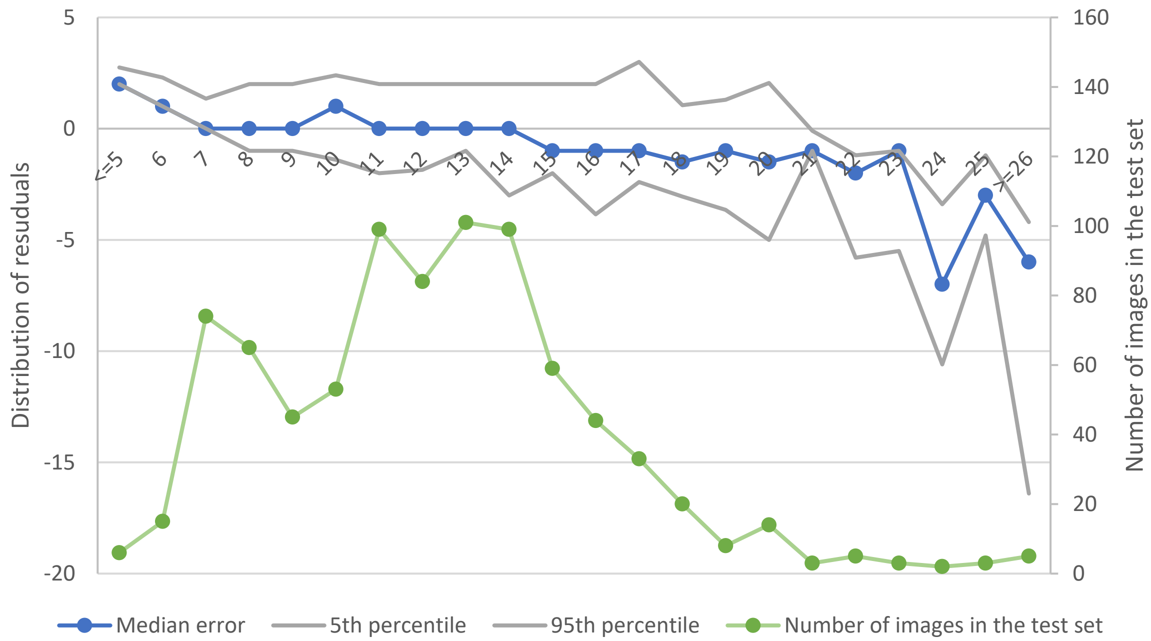


Figure 7. Distribution on residuals and number of images by age in the test set.

Testing temporal generalization

The performance of the model trained on the 2019 FISS sample declined when applied to otolith images collected during the 2024 survey, reflecting the challenges of temporal generalization. On average, the root mean squared error (RMSE) increased to 2.562, representing an approximate 42% increase compared to the 2019 test set. Furthermore, the proportion of predictions within ±1 year of the manually assigned age dropped by 16.7 percentage points, indicating a decline in predictive accuracy.

However, the use of a deep ensemble approach enabled a more nuanced evaluation of model reliability. Specifically, the ensemble framework provided per-sample uncertainty estimates (measured as the standard deviation across model predictions), which helped distinguish between confidently and less confidently predicted samples. This enabled stratification of predictions by uncertainty level.

Figure 8 shows the cumulative proportion of 2024 test samples for which the ensemble prediction falls within ±1 year of the manually assigned age, as a function of increasing prediction uncertainty (measured by the standard deviation across the ensemble). The curve confirms that predictions with lower uncertainty levels tend to be more accurate. For the least uncertain subset of the test data (e.g., the first ~20%), accuracy within ±1 year exceeds 80%, while this metric gradually declines as predictions with higher uncertainty are included. By the time the entire sample is considered, accuracy drops to approximately 59%.

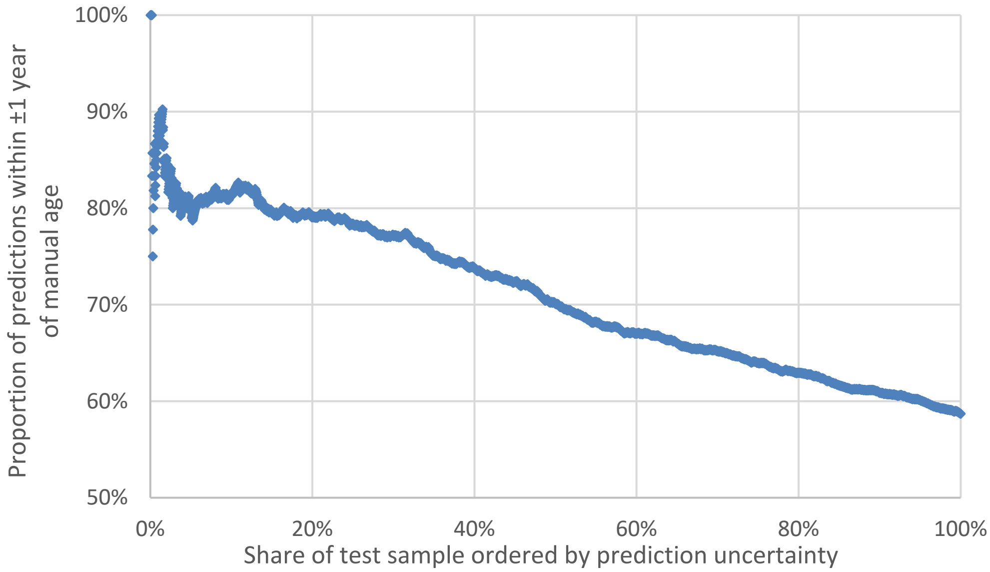


Figure 8: Proportion of ensemble predictions within ±1 year of manual age as a function of cumulative share of the test sample, ordered by prediction uncertainty (standard deviation).

Fine-tuning the model

To assess the impact of fine-tuning on model generalization across years, the ensemble originally trained on 2019 FISS images was fine-tuned using a randomly selected 20% subset of otoliths collected in 2024. The model was then evaluated on the remaining unseen 80% of 2024 images. Fine-tuning yielded measurable improvements: the average RMSE across ensemble runs decreased from 2.562 to 2.396, and the proportion of predictions within ±1 year of the manually assigned age increased from 55.4% to 57.6%.

In a separate analysis, the fine-tuning subset was selected based on uncertainty rather than random sampling. Specifically, 20% of 2024 images with the highest standard deviation across ensemble predictions - interpreted as the most ambiguous or noisy samples - were used for fine-tuning. This targeted approach led to further gains in predictive accuracy. When evaluated on the remaining 80%, the model achieved an RMSE of 2.150.

Predicting region of collection

In September 2024, the SRB made the following recommendation:

> The SRB R**ECOMMENDED** that the Secretariat investigate using the AI to identify region of collection. Otolith shape is sometimes used as a tool for understanding mixing and stock structure and the AI may have skill in identifying region of origin (and thus mixing and migration rates) from otolith images. (IPHC–2024–SRB025–R, par. 47)

In response, the InceptionV3 architecture model was rewritten to perform classification task, predicting IPHC Regulatory Areas (categorical label) from otolith images. The model was trained on the 2019 FISS dataset, and performance was evaluated using three test scenarios:[17]

- Test set from 2019 (same year as training data):

[17] Each model was run three times to account for variability due to random initialization.

The model achieved strong performance, with overall accuracy between 90% and 95%. Misclassifications were minimal and typically involved geographically adjacent areas.

(See Figure 8a: Confusion matrix – 2019 test set)

- Test set from 2024 (no fine-tuning):

When applied directly to otoliths collected in 2024, the model's predictive accuracy dropped sharply. Most images from multiple regulatory areas were misclassified as belonging to IPHC Regulatory Area 2C, suggesting a model bias toward centrally-located region.

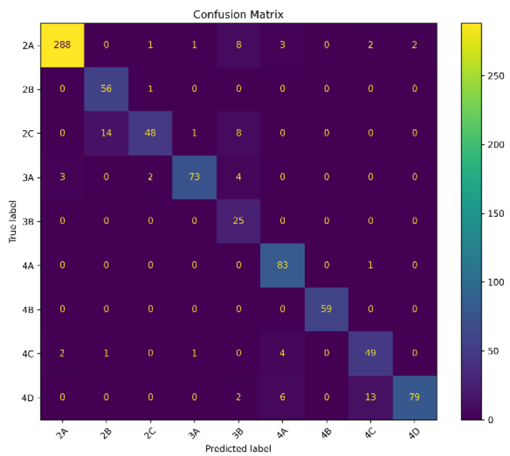
(See Figure 8b: Confusion matrix – 2024 test set without fine-tuning)

- 2024 test set with 20% samples used for fine-tuning:

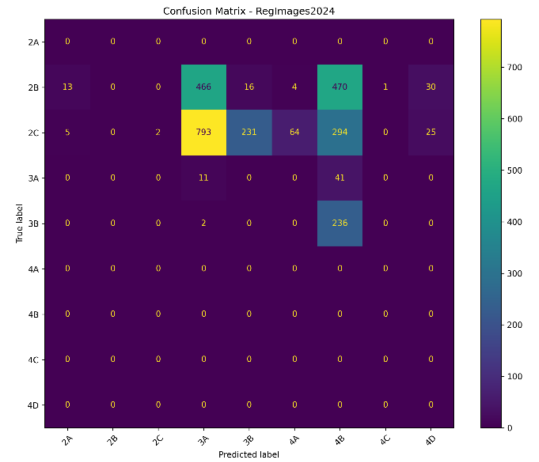
To improve temporal generalization, the model was fine-tuned using a 20% subset of the 2024 dataset, then evaluated on the remaining 80%. This approach substantially improved classification accuracy, yielding correct results for 88.4% samples. Predictions for Regulatory Areas 2B and 2C were particularly improved, with confusion concentrated around adjacent boundaries.

(See Figure 8c: Confusion matrix – 2024 test set with fine-tuning on 20% samples)

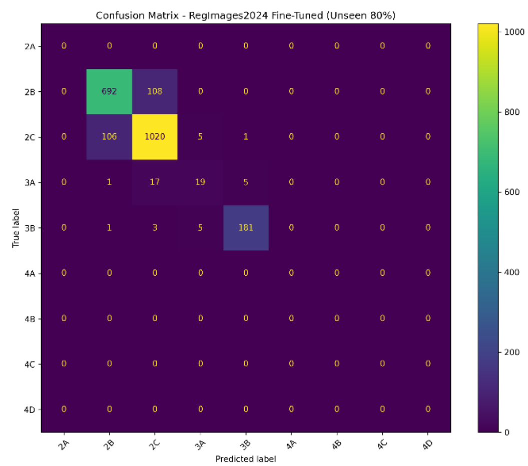
In addition, regional prediction was also evaluated using surface images (i.e., unprocessed otoliths). These models achieved promising results, with overall accuracy ranging between 87% and 91%, when trained on full sample of surface images (5,557 images). However, this evaluation was limited to data from a single year. As no multi-year dataset of surface images was available, it was not possible to assess the model's robustness or generalization across time for surface-based classification.



Panel a: 2019 test set



Panel b: 2024 test set without fine-tuning



Panel c: 2024 test set with fine-tuning on 20% samples

Figure 9: Confusion matrices representing results from predicting IPHC Regulatory Areas (categorical label) from otolith images.

Surface images

This analysis examined whether otolith images captured prior to processing (surface images) can be used to reliably predict fish age using AI models, and how their performance compares to the use of images of processed otoliths. The goal was to evaluate both the viability and potential accuracy of surface images as a practical alternative.

Three configurations were tested:

1. **BB match**: The model was trained using 2,696 sectioned and baked otolith images collected during the 2024 FISS, for which matching surface images were also available (5 runs).
2. **Surface match**: The model was trained on the same selection of 2,696 surface images (5 runs) to allow a direct comparison under identical input conditions (sample size and age distribution).
3. **Surface ALL**: A model was trained using the full set of 5,557 available surface images, maximizing data size (3 runs).

The comparative analysis of otolith surface images and images of processed otoliths (see Table 1) demonstrated that surface images are a viable alternative for AI-based age prediction.

When models were trained on matched datasets, predictive performance using surface images was comparable to that of processed otoliths images, with similar test set MSE and R^2 values. Furthermore, the model trained on the full set of 5,557 available surface images achieved strong results, with an average test MSE of 0.00298. These findings suggest that surface images, when available in sufficient quantity, can potentially match models based on processed otoliths. This highlights the potential to streamline future otolith ageing workflows by relying on unprocessed images without compromising predictive accuracy. However, it is important to note that this evaluation was limited to data from a single year. In the absence of a multi-year surface image dataset, it was not possible to assess the temporal robustness or generalization capability of the surface-image-based models.

Table 1: Average results of model configurations used to assess viability of surface images for AI-based ageing.

	BB match	**Surface match**	**Surface ALL**
Epochs trained	231	223	229
Validation MSE	0.00273	0.00298	0.00284
Test MSE	0.00315	0.00297	0.00298
R^2	0.79	0.80	0.79
Run time (VM)	159	164	345

CONCLUSIONS

The ongoing advancement of AI technologies in the field of marine science offers considerable potential to enhance the efficiency of age determination of Pacific halibut using otolith images. Preliminary results presented here suggest that convolutional neural networks (CNNs), particularly when implemented using a deep ensemble approach, could provide predictive accuracy that supports their use as a supplement- or in some cases, a potential alternative - to the current manual ageing protocol.

Among the models tested, the InceptionV3 architecture outperformed newer and more complex architectures such as EfficientNet and ConvNeXt. This outcome likely reflects the relatively limited size and variability of the training dataset, which favors architectures with fewer parameters and less sensitivity to overfitting. While deeper models may eventually outperform simpler ones with more data and advanced tuning, InceptionV3 currently offers the most robust and consistent performance for this application.

These results also highlight the practical value of the deep ensemble framework. In addition to improving predictive performance, ensemble-based models provide per-sample uncertainty estimates that can be used to identify potentially unreliable predictions. This enables a mixed-method protocol in which low-confidence predictions (e.g., those with high standard deviation across ensemble members) can be flagged for expert review, while high-confidence outputs may be accepted directly - streamlining the ageing workflow while maintaining accuracy.

Results also showed that model performance deteriorates when predictions are made on data collected in years different from the training sample (i.e., temporal generalization is limited). However, modest fine-tuning with current-year data improved predictive performance, reducing RMSE of predictions and increasing accuracy within ±1 year of expert labels. When fine-tuning was focused specifically on uncertain samples - those with the highest variance across ensemble predictions - performance gains were even better. These findings confirm that targeted fine-tuning, guided by uncertainty, is an effective strategy for adapting models to new data while minimizing manual ageing need.

Surface images also showed promise as a practical input for ageing models. When trained on matched datasets, models using unprocessed surface images performed comparably to those using sectioned and baked otoliths. These findings point to the possibility of eliminating otolith processing steps for AI-based ageing in the future, though further multi-year evaluation is needed to confirm long-term robustness.

Despite promising progress, important limitations remain. Statistically significant bias was observed in predictions for the oldest age categories (21+), which remain underrepresented in the training dataset. Only 2.6% of otoliths used in the main model were from fish aged 21 or older, suggesting that improved model accuracy for older fish will require supplementing database in a targeted manner with images from older fish. Expanding the dataset to improve representation across all age classes especially older individuals will be essential to reduce residual bias and ensure model reliability across the full biological age range.

Finally, it is crucial to emphasize that AI-based ageing models must continue to rely on human experts, both for validation and for providing high-quality training data that reflect temporal, spatial, and environmental variability. As environmental conditions and stock structure continue to change, integrating expert oversight and continual model updating will remain a critical part of accurate AI implementation for ageing process.

LITERATURE

Allken, V., Handegard, N. O., Rosen, S., Schreyeck, T., Mahiout, T., & Malde, K. (2019). Fish species identification using a convolutional neural network trained on synthetic data. *ICES Journal of Marine Science*, *76*(1), 342–349. https://doi.org/10.1093/icesjms/fsy147

Benson, I. M., Helser, T. E., Marchetti, G., & Barnett, B. K. (2023). The future of fish age estimation : deep machine learning coupled with Fourier transform near-infrared spectroscopy of otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, *00*, 1–13. https://doi.org/dx.doi.org/10.1139/cjfas-2023-0045

Blood, C. L. (2003). I . Age validation of Pacific halibut II . Comparison of surface and break-and-burn otolith methods of ageing Pacific halibut. *IPHC Technical Report*, *47*.

Campana, S. E. (1999). Chemistry and composition of fish otoliths: Pathways, mechanisms and applications. *Marine Ecology Progress Series*, *188*, 263–297. https://doi.org/10.3354/meps188263

Campana, S. E., & Neilson, J. D. (1985). Microstructure of Fish Otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(5), 1014–1032. https://doi.org/10.1139/f85-127

Campana, S. E., & Thorrold, S. R. (2001). Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences*, *58*(1), 30–38. https://doi.org/10.1139/f00-177

Fablet, R., & Le Josse, N. (2005). Automated fish age estimation from otolith images using statistical learning. *Fisheries Research*, *72*(2–3), 279–290. https://doi.org/10.1016/j.fishres.2004.10.008

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *Proceedings of the 33rd International Conference on Machine Learning*.

IPHC. (1985). Annual Report 1984. In *IPHC Annual Report*.

Keith, S., Kong, T., Sadorus, L. L., Stewart, I. J., & Williams, G. (2014). The Pacific halibut: biology, fishery, and management. *IPHC Technical Report*, *59*. https://doi.org/10.1042/bj0490062

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient Based Learning Applied to Document Recognition. *Proc. of the IEEE*.

Malde, K., Handegard, N. O., Eikvil, L., & Salberg, A. B. (2020). Machine intelligence and the data-driven future of marine science. *ICES Journal of Marine Science*, *77*(4), 1274–1285. https://doi.org/10.1093/icesjms/fsz057

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

Moen, E., Handegard, N. O., Allken, V., Albert, O. T., Harbitz, A., & Malde, K. (2018). Automatic interpretation of otoliths using deep learning. *PLoS ONE*, *13*(12), e0204713.

Moore, B. R., Maclaren, J., Peat, C., Anjomrouz, M., Horn, P. L., & Hoyle, S. (2019). Feasibility of automating otolith ageing using CT scanning and machine learning. *New Zealand Fisheries Assessment Report*, *58*.

Ordoñez, A., Eikvil, L., Salberg, A. B., Harbitz, A., Murray, S. M., & Kampffmeyer, M. C. (2020). Explaining decisions of deep neural networks used for fish age prediction. *PLoS ONE*, *15*(6), 1–19. https://doi.org/10.1371/journal.pone.0235013

Piner, K. R., & Wischniowski, S. G. (2004). Pacific halibut chronology of bomb radiocarbon in otoliths from 1944 to 1981 and a validation of ageing methods. *Journal of Fish Biology*, *64*(4), 1060–1071. https://doi.org/10.1111/j.1095-8649.2004.0371.x

Politikos, D. V, Petasis, G., Chatzispyrou, A., Mytilineou, C., & Anastasopoulou, A. (2021). Automating fish age estimation combining otolith images and deep learning: The role of multitask learning. *Fisheries Research*, *242*, 106033. https://doi.org/https://doi.org/10.1016/j.fishres.2021.106033

Politikos, D. V, Sykiniotis, N., Petasis, G., Dedousis, P., Ordoñez, A., Vabø, R., Anastasopoulou, A., Moen, E., Mytilineou, C., Salberg, A. B., Chatzispyrou, A., & Malde, K. (2022). DeepOtolith v1.0: An Open-Source AI Platform for Automating Fish Age Reading from Otolith or Scale Images. *Fishes*, *7*(3), 1–11. https://doi.org/10.3390/fishes7030121

Punt, A. E., Smith, D. C., KrusicGolub, K., & Robertson, S. (2008). Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(9), 1991–2005. https://doi.org/10.1139/F08-111

Robertson, S. G., & Morison, A. K. (1999). A trial of artificial neural networks for automatically estimating the age of fish. *Marine and Freshwater Research*, *50*(1), 73–82. https://doi.org/10.1071/MF98039

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR 2015 - Conference Track Proceedings*.

Southward, G. M. (1962). Photographing Halibut Otoliths for Measuring Growth Zones. *Journal of the Fisheries Research Board of Canada*, *19*(2), 335–338. https://doi.org/10.1139/f62-018

APPENDIX A: COUNTS OF OTOLITHS AGED BY THE IPHC

Collection year	Ageing method	IPHC FISS*	Commercial (Market Sample)*	NOAA Trawl survey*	Tag recovery*	ADF&G recreational*	Clean collection
pre-1960	surface	70,984			10,068		
1960	surface	6,606			681		
1961	surface	4,727		4,576	842		
1962	surface	2,605		1,692	594		
1963	surface	8,257		2,209	440		
1964	surface	10,295	27,828	1,001	353		
1965	surface	5,169	27,252	1,186	493		
1966	surface	3,750	24,638	1,777	796		
1967	surface	6,325	29,797	2,271	1,151		
1968	surface	2,314	29,772	1,887	1,813		
1969	surface	1,510	23,361	1,019	1,869		
1970	surface	1,138	24,686	1,184	867		
1971	surface	2,702	16,374	2,294	732		
1972	surface	2,597	23,381	1,180	490		
1973	surface	1,747	16,683	893	244		
1974	surface	1,021	11,569	1,189	128		
1975	surface	1,212	14,128	1,136	131		
1976	surface	1,843	14,103	969	72		
1977	surface	1,853	13,514	1,102	83		
1978	surface	1,933	11,434	1,309	61		
1979	surface	2,021	7,219	730	93		
1980	surface	5,022	10,317	717	168		
1981	surface	7,942	8,267	460	129		
1982	surface	5,720	9,644	443	208		
1983	surface	5,822	9,262	1,355	286		
1984	surface	6,508	10,233	1,089	455		
1985	surface	5,872	12,986	1,192	778		
1986	surface	5,139	12,426	1,120	1,020		
1987	surface	42	16,137		859		
1988	surface	1,179	17,154	98	761		
1989	surface	6,130	14,122		710		
1990	surface	2,201	14,800	4,802	397		
1991	surface	1,315	13,461	2,598	280		
1992	surface/BB	7,530	14,564	222	182		
1993	surface/BB	3,384	13,747		147		
1994	surface/BB	2,618	13,311		99		
1995	surface/BB	4,512	12,297	433			
1996	surface/BB	10,893	13,452	2,211			
1997	surface/BB	14,784	15,501	834	148		
1998	surface/BB	8,587	14,395	1,145	98		

1999	surface/BB	11,971	12,858	3,029	70	3,672	
2000	surface/BB	14,122	13,982	1,209	46	2,706	
2001	surface/BB	14,731	13,181	2,952	27	2,609	
2002	BB	13,635	17,932	761	24	2,349	
2003	BB	12,626	13,915	3,876	79	2,754	
2004	BB	14,474	11,798	897	450	3,288	
2005	BB	12,651	14,650	2,028	643	3,183	
2006	BB	14,976	13,399	2,621	679	3,179	
2007	BB	16,285	13,964	3,930	455	3,026	
2008	BB	15,545	13,460	1,527	304	1,500	
2009	BB	15,706	13,583	4,922	276	1,500	
2010	BB	14,080	16,106	1,915	21	1,500	625
2011	BB	14,451	11,391	4,592	26	1,500	676
2012	BB	17,896	12,902	1,639	9	1,500	1164
2013	BB	12,717	11,039	2,044	19	1,503	1020
2014	BB	16,194	12,606	1,476	22	1,500	1096
2015	BB	15,815	12,312	2,133	24	1,500	1072
2016	BB	15,113	11,618	742	21	1,502	902
2017	BB	12,565	10,821	1,384	15	1,500	756
2018	BB	12,935	11,013	576	39	1,499	798
2019	BB	17,716	10,711	1,640	34	1,497	925
2020	BB	10,323	10,568	-	34	1,413	577
2021	BB	12,253	11,051	1,444	38	1,500	547
2022	BB	9,702	10,942	1,902	39	2,334	519
2023	BB	8,506	10,932	(3,147)	(48)	(1,958)	462
2024	BB	5,770	10,474[1]	(1,058)	(61)	(1,542)	458

Notes:
- Star (*) indicates blind side otolith.
- BB stands for 'break and bake' approach.
- All otoliths reported in this table were aged with the exception of the clean collection.
- All aged otoliths are stored in glycerol/thymol solution.
- Some small fish from trawl survey collection are still aged by surface method; otoliths with surface age>4 are sectioned and baked.
- Sample data not entered prior to 1960 for FISS, 1964 for commercial, 1961 for NOAA trawl survey.
- Clean collection is not aged, stored dry, and include paired otoliths.
- Tribal otoliths are included in the Market Sample series.
- Additionally, there are 144 not aged 2A recreational otoliths, all from Hein Bank collected between 2004 and 2009.
- Sex information available since 2017 (typically ca. 1 year of lag).
- Trawl and recreational otoliths lag one year in ageing.
- In brackets, otoliths available for ageing but ageing not completed.

[1] Commercial otolith collection subsampled: 10,474 otoliths were collected, 7,057 were selected for ageing

APPENDIX B: SELECTION OF MODEL RUNS

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SETUP				**												**
Architecture	InceptionV3	InceptionV3	InceptionV3	InceptionV3	EfficientNetB4	EfficientNetB4	EfficientNetB4	EfficientNetB5	EfficientNetB5	EfficientNetB5	EfficientNetV2 S	EfficientNetV2 M	EfficientNetV2 L	ConvNeXt	ConvNeXt	InceptionV3
Max epochs	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
EarlyStopping patience	50	100	100	80	50	50	50	50	50	50	60	50	100	100	60	80
ReduceLROnPlateau	NA	NA	NA	40/r=0.6	NA	NA	NA	NA	NA	NA	30 /f=.8	30 /f=.8	50 / f=0.5	50 / f=0.9	30 /f=.8	40/r=0.6
Learning rate (initial)	0.0002	0.0004	0.0004	0.0002	0.0004	0.0002	0.0004	0.0004	0.0004	0.0004	0.0016	0.0004	0.0008	0.0016	0.0016	0.0002
Batch size	16	8	16	16	16	16	8	8	16	4	8	8	8	16	12	16
Image size	400	400	400	400	380	380	380	456	456	456	384	480	512	224	224	400
Dropout rate	0.2	0.2	0.2	0.2/0.25	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2/0.25
L2 parameter	0.025	0.025	0.025	.025	0.025	0.025	0.025	0.025	0.025	0.025	0.03	0.025	0.025	0.025	0.025	0.025
Augmentation[1]	NA	NA	NA	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full
RESULTS																
Validation MSE	0.00195	0.00167	0.00156	0.00170	0.00334	0.00372	0.00444	0.00414	0.00308	0.00375	0.00865	0.00223	0.00789	0.00856	0.00334	0.00163
Epochs trained	92	297	249	260	156	109	80	126	128	166	142	123	224	199	138	318
Test MSE	0.0023	0.0021	0.0020	0.0019	0.0032	0.0040	0.0044	0.0038	0.0030	0.0041	0.0087	0.0025	0.0087	0.0087	.0087	0.0019
R^2	*	*	*	.77	*	*	*	*	*	*	*	*	*	*	*	0.78
RMSE-unscaled	1.986	1.880	1.877	1.834	2.341	2.591	2.718	2.543	2.254	2.649	*	2.072	3.833	*	*	1.782
Correctly predicted	29.5%	33.6%	31.7%	31.7%	21.3%	15.6%	22.9%	31.1%	27.9%	26.9%	*	26.5%	19.3%	*	*	30.4%
Correctly predicted with ±1 year tolerance	75.6%	77.4%	78.8%	72.1%	55.4%	43.9%	63.9%	72.1%	75.3%	70.8%	*	75.6%	65.1%	*	*	74.4%
RUN parameters																
Machine[2]	DS	DS	DS	MM	QS	QS	QS	QS	QS	QS	QS	QS	QS	QS	QS	VM
Run time in hours	14.0	47.3	35.2	11	*	*	*	30.0	32.3	38.9	12.3	29.0	116.4	45.3	45	4
RESULTS for 2024																
RMSE-unscaled	2.852	2.864	2.970	2.779	3.057	3.274	*	*	*	*	*	2.801	*	*	*	2.696
Correctly predicted	18.0%	18.0%	19.3%	19.0%	17.7%	10.9%	*	*	*	*	*	15.7%	*	*	*	19.9%
Correctly predicted with ±1 year tolerance	52.5%	48.3%	50.4%	50.2%	46.4%	32.8%	*	*	*	*	*	48.9%	*	*	*	54.9%

Note: All models for randomly selected seed numbers – individual results would vary.

1: Full augmentation setup included rotation range=360, width shift range=0.1, height shift range=0.1, brightness range=[0.95, 1.05], and zoom range=[0.98, 1.02].

2: Machine setups were as follows:

- QS: 11th Gen Intel(R) Core(TM) i7-11700K @ 3.60GHz; 8 cores
- DS: 12th Gen Intel(R) Core(TM) i7-12700; 12 cores
- MM: AMD Ryzen 9 5900X; 12 cores
- VM: AMD EPYC 7V12 64-Core Processor with Nvidia Tesla T4 GPU

* Indicates values not recorded for the given run.

**Indicates models selected for further investigation.

APPENDIX C: DEEP ENSEMBLE INDIVIDUAL RESULTS

Model run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	**AVERAGE**
Epochs trained	194	557	172	159	318	235	263	338	204	380	192	483	292	174	364	**288**
Validation MSE	0.0017	0.0015	0.0017	0.0017	0.0016	0.0017	0.0015	0.0016	0.0018	0.0015	0.0017	0.0015	0.0014	0.0016	0.0016	**0.0016**
Test MSE	0.0020	0.0018	0.0021	0.0022	0.0019	0.0019	0.0019	0.0018	0.0021	0.0017	0.0020	0.0017	0.0018	0.0019	0.0018	**0.0019**
R^2	0.776	0.797	0.756	0.749	0.783	0.784	0.779	0.794	0.764	0.804	0.774	0.809	0.797	0.785	0.796	**0.783**
Rum time (VM, min)	148	418	133	123	240	179	203	256	156	286	148	369	223	134	276	**219**
RESULTS – TEST SET																
Test RMSE unscaled	1.819	1.742	1.908	1.960	1.782	1.786	1.817	1.757	1.876	1.719	1.856	1.693	1.741	1.814	1.745	**1.80**
Correctly predicted	30.0%	30.6%	28.9%	23.5%	30.4%	31.3%	32.0%	31.4%	28.7%	32.5%	30.6%	32.1%	33.6%	29.0%	30.4%	**30.3%**
Correctly predicted with ±1 year tolerance	72.0%	74.5%	69.8%	64.6%	74.3%	71.3%	73.3%	74.4%	69.5%	74.5%	69.2%	75.1%	72.6%	71.3%	74.2%	**72.0%**
RESULTS – 2024 IMAGES																
RMSE	2.509	2.472	2.598	2.844	2.514	2.539	2.631	2.498	2.613	2.477	2.660	2.548	2.481	2.519	2.518	**2.562**
Correctly predicted with ±1 year tolerance	56.8%	57.4%	55.4%	52.7%	55.9%	55.1%	55.2%	55.5%	54.0%	58.8%	52.1%	57.1%	56.3%	52.1%	56.0%	**55.4%**
RMSE – fine-tuned on 20% images	2.378	2.350	2.451	2.418	2.328	2.404	2.396	2.389	2.440	2.331	2.493	2.379	2.408	2.444	2.334	**2.396**
Correctly predicted with ±1 year tolerance– fine-tuned on 20% images	59.7%	58.0%	54.4%	56.2%	59.1%	56.5%	58.0%	57.5%	57.0%	59.7%	56.3%	58.8%	57.0%	57.1%	58.4%	**57.6%**
RMSE – fine-tuned on 20% images with highest standard deviation	2.151	2.105	2.142	2.211	2.069	2.133	2.159	2.108	2.270	2.073	2.280	2.084	2.116	2.260	2.089	**2.150**
Correctly predicted with ±1 year tolerance– fine-tuned on 20% images with highest standard deviation	56.3%	59.4%	58.7%	53.7%	60.9%	59.0%	57.6%	59.3%	52.1%	57.9%	51.6%	60.5%	59.1%	52.8%	60.2%	**57.3%**