Development of the 2026 Pacific halibut (*Hippoglossus stenolepis*) stock assessment

PREPARED BY: IPHC SECRETARIAT (I. STEWART & A. HICKS; 15 APRIL 2026)

PURPOSE

To provide the IPHC's Scientific Review Board (SRB) with a response to recommendations and requests from SRB026 (IPHC-2025-SRB026-R) and SRB027 (IPHC-2025-SRB027-R) and to provide the Commission with an update on progress toward the 2026 stock assessment update.

INTRODUCTION

The International Pacific Halibut Commission (IPHC) conducts an annual coastwide stock assessment of Pacific halibut (*Hippoglossus stenolepis*). The most recent full assessment was completed in 2025 (IPHC-2026-SA01, IPHC-2026-SA02, and detailed model evaluation presented during SRB026: IPHC-2025-SRB026-07). That analysis followed updates in 2023 (IPHC-2024-SA01) and 2024 (IPHC-2025-SA01). The 2026 stock assessment is planned as an update analysis, updating data sources but not revisiting all data processing and model development.

Major topics explored during the 2025 full assessment included:

1) updating the time-series information for the Pacific Decadal Oscillation, used as a covariate to the stock-recruitment relationship,
2) improving the bootstrapping approach to pre-model calculation of maximum effective sample sizes to include ageing imprecision (Hulson and Williams 2024),
3) re-tuning the process and observation error components of assessment models to achieve internal consistency,
4) and updating the maturity ogive to reflect the recent histology-based estimates produced by the IPHC's Biological and Ecosystem Sciences Branch.

The results of the final 2025 assessment (IPHC-2026-AM102-10), including stock projections and the harvest decision table (IPHC-2026-AM102-12), were generally consistent with those from previous stock assessments.

Starting from the final 2025 stock assessment, this document addresses requests and recommendations made during SRB026 and SRB027 and prepares for the 2026 updated analysis.

TIME-SERIES AND SOFTWARE UPDATES

In order to provide comparability between the final 2025 results and all subsequent steps working toward the final 2026 stock assessment, this evaluation began with a bridging analysis. First, each of the four assessment models was extended by one year, including projected 2026 mortality from all sources based on the mortality limits set during AM102 (IPHC-2026-AM102-R). Extending the time-series without adding any new data does not affect the historical time-series' estimates but allows for a simple stepwise evaluation of the effects of adding data (including updating from the projected to actual fishery harvest) and any other changes to the models prior to the final version used for management.

Next, the Stock Synthesis (SS) software was updated from the version used for the 2025 stock assessment (3.30.23.1; Methot Jr 2024) to the newest release available (3.30.24.2; Methot Jr et al. 2026). The changes to the software between these two versions had no effect on the Pacific halibut stock assessment (the results were identical to the final 2025 assessment). There was a small improvement in run time for each model (in both safe and optimized modes) but no changes were noted in convergence performance, or other technical aspects of the software update relevant to this assessment.

SRB REQUESTS AND RECOMMENDATIONS

There are two pending SRB recommendations specific to the stock assessment from SRB026 and SRB027:

1) SRB026 (para. 26):

   *"The SRB **RECOMMENDED** that a candidate state space assessment model (e.g. WHAM) be developed for Pacific halibut and presented by SRB032, tentatively scheduled for June 2028. Progress toward this modelling framework may also be presented at interim SRB meetings."*

2) SRB027 (para. 16):

   *"The SRB **RECOMMENDED** that the analysis of projection performance be expanded to include plotting receiver operating characteristic (ROC) curves and evaluating the area under the curve (AUC) to understand the predictive performance of probabilistic advice from the stock assessment projections. This approach is commonly used as a threshold-independent metric of performance in applications such as species distribution modelling."*

Request 1 – State space model development

As discussed in previous documents, the IPHC continues to rely on SS for its annual tactical stock assessment modelling; however we recognize that future development of this software will be limited. During 2024, Secretariat staff explored the capabilites of R-Template Model Builder (RTMB; Kristensen et al. 2016), via a training course hosted by Fisheries and Oceans Canada. TMB forms the basis of most state-space models currently used for stock assessment (e.g., SAM, WHAM; Nielsen and Berg 2014; Nielsen et al. 2021; Stock and Miller 2021), provides a more efficient Auto-Differentiation (AD) algorithm than Automatic Differentiation Model Builder (ADMB; Fournier et al. 2012) and extremely efficient capabilities for modelling random effects and sparse matrices. As the Pacific halibut stock assessment models include time-varying processes (i.e. recruitment, selectivity, and catchability) it would be ideal to treat them as random effects, rather than using the penalized likelihood curently employed.

The IPHC development and review process benefits from the use of a generalized software platform that is widely used and tested, rather than a customized platform that may have more features but may be much more difficult to maintain and review effectively. Secretariat staff followed the December 2025 Center for the Advancement of Population Assessment Modelling (CAPAM) workshop on developing the next generation of tuna models. This workshop comprised an extensive discussion of needed model features for those species, but also included extensive consideration of exisiting stock assessment software used around the world. The requirements for a Pacific halibut modelling platform are reletively complex compared to many stock assessments and not all features are supported by all exisiting state-space platforms (Table 1). Assessment development in prepartion for the next full stock assessment (2028) will include monitoring of development and testing of each of the platforms currently in mangement use.

Table 1. Generalized state-space stock assessment platforms and existing capabilities for features needed in the Pacific halibut assessment.

Model	Reference	Sex-specific dynamics	Multiple aging error matrices	Environmental covariates to S-R function	Time varying selectivity	Prior distribution for M	Used for assessments informing management
SAM	Nielsen and Berg 2014	No	No	Yes	Yes	Yes	Yes
WHAM	Stock and Miller 2021	No	No	Yes	Yes	No	Yes
FIMS[1]	None	No	No	No	No	Yes	No
SPoRC	Cheng et al. 2026	Yes	No	No	Yes	Yes	No
CEATTLE	Adams et al. 2022	Yes	Yes	Yes	Yes	Yes	Yes

[1]FIMS is the generalized NOAA Fisheries stock assessment platform currently under development and intended to replace stock synthesis at some point in the future (https://noaa-fims.github.io/about/).

Development of the IPHC's stock assesment is highly dependent on the type of management procedure in use by the Commission. Currently, the IPHC Harvest Strategy Policy requires an full stock assessment every three years with an update stock assessment occurring in intervening years. The stock assesment analysis conducted each fall in order to provide annual management information is based on the current year's data and must be stable and simple enough to be completed in less than two weeks. If a management procedure based on modelled survey trends, or a multi-year procedure was to be adopted in the future, it may be unneccesary to conduct annual stock assessments. That type of procedure and timeline could allow for the development of more complex stock assessment ensembles/models (including fully Bayesian analyses), given extended development time between assessments. Therefore, any updates to the current IPHC management procedure, developments in the MSE process, and strategic planning for the stock assessment modelling platform should be considered together: the long-term focus should be on selecting the most efficient tools to meet management needs as they continue to evolve and ensuring that the IPHC relies on appropriate models that can be efficiently applied and reviewed and would remain stable for any future transitions in secretariat staff.

Recommendation 2 – ROC curves

IPHC decision making since 2013 has been informed by a 'Harvest decision table' illustrating the risk-benefit trade-off between fishery yield and a suite of stock and fishery metrics (Table 2). An emerging question after 14 years of this process is "How have the estimated risk probabilities in the table performed?". The only metric that has been maintained through every year of the decision table and the one most frequently evaluated as part of the decision-making process is the probability of stock decline in the next year (row a). Therefore, we focused the receiver-operating characteristic (ROC) performance analysis on that metric.

Table 2. Harvest decision table provided for AM102 to inform mortality limits for 2026-2028. Columns correspond to yield alternatives and rows to risk metrics. Values in the table represent the probability, in "times out of 100" (or percent chance) of a particular risk.

		2026 Alternative			*Status quo -10%*	*Status quo -5%*	*Status quo*	*Status quo +5%*	*Status quo +10%*	*$F_{46\%}$*	*3-Year Surplus / $F_{43\%}$*	*MEY proxy*	*Overfishing limit*	
		Total mortality (M lb)	0.0	21.9	28.6	30.1	31.6	33.1	34.6	37.0	40.8	45.1	53.7	
		TCEY (M lb)	0.0	20.0	26.8	28.2	29.7	31.2	32.7	35.1	39.0	43.3	51.9	
		2026 fishing intensity	$F_{100\%}$	$F_{62\%}$	$F_{54\%}$	$F_{52\%}$	$F_{51\%}$	$F_{49\%}$	$F_{48\%}$	$F_{46\%}$	$F_{43\%}$	$F_{40\%}$	$F_{35\%}$	
		Fishing intensity interval	--	47-77%	39-71%	37-70%	36-69%	34-68%	33-67%	31-65%	28-62%	26-59%	22-54%	
Stock Trend (spawning biomass)	in 2027	is less than 2026	<1	3	10	12	15	18	22	28	40	54	80	a
		is 5% less than 2026	<1	<1	1	1	2	2	3	4	8	14	32	b
	in 2028	is less than 2026	<1	2	8	10	13	16	19	26	38	54	82	c
		is 5% less than 2026	<1	<1	2	3	4	5	7	10	17	28	55	d
	in 2029	is less than 2026	<1	3	11	14	18	22	27	35	50	68	91	e
		is 5% less than 2026	<1	1	5	6	8	11	13	19	30	46	77	f
Stock Status (Spawning biomass)	in 2027	is less than 30%	24	25	26	26	26	26	26	26	26	26	27	g
		is less than 20%	<1	<1	<1	1	1	1	1	1	1	1	2	h
	in 2028	is less than 30%	14	22	23	24	24	24	24	25	25	26	27	i
		is less than 20%	<1	<1	<1	<1	<1	1	1	1	1	2	3	j
	in 2029	is less than 30%	5	17	20	21	22	22	23	23	24	25	27	k
		is less than 20%	<1	<1	<1	<1	1	1	1	1	2	3	6	l
Fishery Trend (TCEY)	in 2027	is less than 2026	0	<1	11	16	20	25	30	37	49	60	75	m
		is 10% less than 2026	0	<1	4	9	10	14	18	25	35	47	65	n
	in 2028	is less than 2026	0	<1	11	15	20	24	29	37	50	61	78	o
		is 10% less than 2026	0	<1	4	10	10	14	18	25	36	49	68	p
	in 2029	is less than 2026	0	1	11	15	10	25	30	39	53	65	82	q
		is 10% less than 2026	0	<1	5	10	11	15	19	26	39	53	73	r
Fishery Status (Fishing intensity)	in 2026	is above $F_{43\%}$	0	<1	13	18	23	27	32	39	50	60	73	s

ROC curves are used in a wide range of fields to determine the performance of testing or detection methods (e.g. Zweig and Campbell 1993). In the context of spawning biomass trends, we are interested in the frequency of accurately projecting a decline in the next year's spawning biomass (true positive; P_t) vs the frequency of projecting a decline when it does not occur (false positive; P_f). Possible outcomes are illustrated in Table 3.

Table 3. Frequency table of projected and actual estimated trends used to calculate the receiver-operating characteristic curves.

		Spawning biomass projection	
		Declining	Not declining
Actual SB trend	Declining	P_t	N_f
	Not Declining	P_f	N_t

Following the standard approach for ROC calculations, two quantities are needed for comparison: 1) the proportion of years with true positive projections of declining spawning biomass are calculated as:

$$\boldsymbol{Proportion\ true\ postive} = \frac{\boldsymbol{P_t}}{(\boldsymbol{P_t} + \boldsymbol{N_f})}$$

and 2) the proportion of years with false positive projections of declining spawning biomass are calculated as:

$$\boldsymbol{Proportion\ false\ postive} = \frac{\boldsymbol{P_f}}{(\boldsymbol{P_f} + \boldsymbol{N_t})}$$

Since the harvest decision table does not represent a strict "test" but rather reports the estimated probability of spawning biomass decline, we compare incremental values (from 5% to 95%) for the estimated probability of decline as different testing/detection methods. Further, we differentiate among actual spawning biomass declines (any decline, >1% decline, >2% decline, >3% decline). Interpretation of ROC curves is based on the area-under-the-curve (AUC), integrating over the proportion of true positive outcomes on the y-axis and the proportion of false positive outcomes on the x-axis. The relative importance of detection of true positive outcome and false negative outcomes is a value judgement that depends on the application. However, a value of 0.5 represents a random outcome and a value of 1.0 a perfect test.

We use the estimated time-series of spawning biomass from the 2025 stock assessment as the 'true' trend and compare that to the projected probability of decline reported in the decision tables created in each year from 2013 through 2025 (both the decision making and the spawning biomass estimate occur at the beginning of the year; Table 4).

Table 4. Current estimated spawning biomass (SB) time-series compared to probability of decline estimated at the time for each year from 2013-2026.

Year	Current estimated SB (Mlb)	Percent change to next year	Estimated probability of decline at the time
2013	216.5	1.8%	84%
2014	220.4	3.0%	67%
2015	227.1	1.9%	30%
2016	231.4	-1.3%	29%
2017	228.3	-6.0%	71%
2018	214.7	-5.8%	93%
2019	202.2	-9.6%	84%
2020	182.9	-7.4%	95%
2021	169.4	-6.0%	65%
2022	159.2	-2.9%	59%
2023	154.7	-1.0%	38%
2024	153.2	1.3%	40%
2025	155.1	7.0%	25%
2026	166.0	NA	14%

The Pacific halibut spawning biomass is estimated to have increased over 2013-2016, declined over 2017-2024 and then increased to 2026. The annual decision table generally favored a higher probability for stock decline when that was the subsequent outcome (2017-2022) but sometimes estimated a relatively high probability when the stock did not decline (2013-2014). All annual spawning biomass changes have been less than 10%, with about half (6/13) greater than 5%.

Based on estimated ROC curves, the harvest decision table is generally a reliable tool to project spawning biomass trends with skill well above random for all 4 levels of decline evaluated (Figure 1). At a 50% estimated probability of any decline, the true positive frequency is 0.75 and the false positive frequency is 0.4 (Figure 1; red point). For a decline greater than 1%, the true positive frequency goes up to 0.86 and the false positive frequency goes down to 0.33 (Figure 1; green point). For a decline greater than 2%, the true positive frequency goes up again to 1.00 (all declines of this magnitude were projected with at least 50% probability) and the false positive frequency goes down to 0.29 (Figure 1; blue point). For a decline greater than 3%, the true positive frequency remains at 1.00 but the false positive frequency goes up slightly to 0.38 (Figure 1; purple point). If an estimated probability threshold of at least 60% is used, the true positive frequency remains at 1.00 and the false positive frequency goes down to 0.25 (Figure 1; open purple point); this represents the best performance of any threshold evaluated. In simple terms, if at least a 3% decline in spawning biomass is considered important, a high probability of stock decline (> 50/100) estimated in the harvest decision table has correctly identified this in all years it has occurred.

While it would be helpful to have a larger sample size for this type of analysis, the stock assessment models themselves are changing over time such that even comparisons made across the existing time series may not accurately reflect the performance of future stock assessments and harvest decision tables.

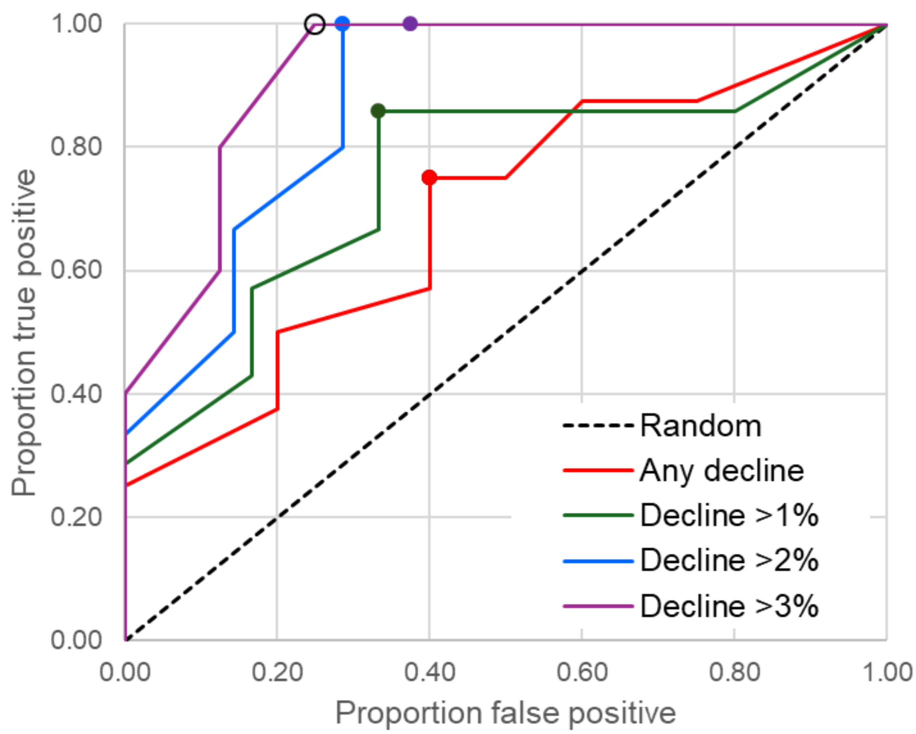

Figure 1. ROC curves for different levels of spawning biomass decline (colors) and different estimated probability of decline cutoffs (points defining each line). Solid circles represent an estimated probability of decline for each curve of 50%. Open circle indicates the best performing cutoff for declines of >3% decline (highest true positive rate and lowest false positive rate; probability of decline of 60%).

RECOMMENDATION/S

That the SRB:

a) **NOTE** paper IPHC-2026-SRB028-07 which provides a response to requests from SRB026 and SRB027, and an update on model development for 2026.

b) **REQUEST** any analyses to support the final 2026 stock assessment.

c) **REQUEST** any analyses to be provided at future SRB meetings as part of the longer-term stock assessment development.

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.

Cheng, M.L.H., Goethel, D.R., Cunningham, C.J., Hulson, P.J.F., Ianelli, J.N., and Omori, K.L. 2026. The SPoRC Stock Assessment Package: A Generalized Next-Generation Platform to Assess Spatial, Age and Sex-Structured Populations. Fish and Fisheries. doi:10.1111/faf.70082.

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.

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.

IPHC. 2025a. Report of the 26th session of the IPHC Scientific Review Board (SRB). Meeting held in Seattle, WA, USA, 10-12 June 2025. IPHC-2025-SRB026-R. 17 p.

IPHC. 2025b. Report of the 27th session of the IPHC Scientific Review Board (SRB027). IPHC-2025-SRB027-R. 18 p.

IPHC. 2026. Report of the 102nd session of the IPHC annual meeting (AM102). Bellevue, WA, USA 19-22 January 2026. IPHC-2026-AM102-R. 54 p.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software **70**(5). doi:10.18637/jss.v070.i05.

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., and Doering, K. 2026. Stock synthesis user manual version 3.30.24.2. NOAA Fisheries. Seattle, WA. 264 p.

Nielsen, A., and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research **158**: 96–101. doi:10.1016/j.fishres.2014.01.014.

Nielsen, A., Hintzen, N.T., Mosegaard, H., Trijoulet, V., Berg, C.W., and Subbey, S. 2021. Multi-fleet state-space assessment model strengthens confidence in single-fleet SAM and provides fleet-specific forecast options. ICES Journal of Marine Science **78**(6): 2043–2052. doi:10.1093/icesjms/fsab078.

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. 2026. Overview of data sources for the Pacific halibut stock assessment, harvest policy, and related analyses. IPHC-2026-SA-02. 59 p.

Stewart, I., and Hicks, A. 2026a. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2025. IPHC-2026-SA-01. 38 p.

Stewart, I., and Hicks, A. 2026b. Stock projections and the harvest decision table for 2026-2028. IPHC-2026-AM102-12. 6 p.

Stewart, I., Hicks, A., Webster, R., and Wilson, D. 2026. Data overview and stock assessment for Pacific halibut (*Hippoglossus stenolepis*) at the end of 2025. IPHC-2026-AM102-10. 17 p.

Stock, B.C., and Miller, T.J. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research **240**. doi:10.1016/j.fishres.2021.105967.

Zweig, M.H., and Campbell, G. 1993. Receiver-operating characteristic (ROC) plots: a fundamental tool in clinical medicine. Clinical Chemistry **39**(4): 561–577.