Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2019

PREPARED BY: IPHC SECRETARIAT (I. STEWART & A. HICKS; 10 JANUARY 2020)

PURPOSE

To provide the Commission with a detailed report of the 2019 stock assessment analysis.

ABSTRACT

This stock assessment reports the status of the Pacific halibut (*Hippoglossus stenolepis*) resource in the International Pacific Halibut Commission (IPHC) Convention Area at the end of 2019. An overview of data sources is provided in a separate document on the IPHC's stock assessment webpage. A summary of both the data and assessment results, as well as management related information and Commission requests is provided both on the stock assessment webpage as well as in the meeting materials for the IPHC's 96th Annual Meeting (AM096; IPHC-2020-AM096-09).

Coastwide mortality (including all sizes of Pacific halibut) from all sources in 2019 was estimated to be 39.7 million pounds[1] (~18,000 t), up 3% from 2018. In addition to the estimated mortality, the assessment includes data from both fishery dependent and fishery independent sources, as well as auxiliary biological information. The 2019 modelled Fishery-Independent Setline Survey (FISS; see IPHC-2020-AM096-06 and IPHC-2020-AM096-07) detailed a coastwide aggregate Numbers-Per-Unit-Effort (NPUE) which showed a third consecutive year of decrease, down 4% from 2018, with individual Biological Regions ranging from a 5% increase (Region 2) to a 10% decrease (Region 3). The modelled survey Weight-Per-Unit-Effort (WPUE) of legal (O32) Pacific halibut, the most comparable metric to observed commercial fishery catch rates, was 5% lower than the 2018 estimate at the coastwide level, constituting the lowest value in the time series. Individual IPHC Regulatory Areas varied from a 26% increase (Regulatory Area 3B) to a 17% decrease (Regulatory Area 3A). The FISS sampling associated with the expansion in Region 3 (Regulatory Areas 3A, and 3B; the final year of the expansion project) resulted in lower estimated catch-rates in this Region compared to the rest of the coast, and reduced uncertainty in the index for both Region 3 and coastwide. Commercial fishery WPUE (based on extensive, but incomplete 2019 logbook records available for this assessment) increased 4% coastwide, with mixed performance across IPHC Regulatory Areas. A bias correction (to account for additional logbooks compiled after the fishing season) resulted in in an estimate of a 1% increase coastwide. Biological information (ages and lengths) from both the commercial fishery and FISS continue to show the 2005 year-class as the largest contributor (in number) to the fish encountered. In the FISS data, 2011 and 2012 cohorts (7 and 8 years old, following a series of weak cohorts from 2006-2010) represented the largest proportions in some IPHC Regulatory Areas for the total catch and the largest proportions coastwide for sublegal Pacific halibut. At the coastwide level, individual size-at-age continues to be very low relative to the rest of the time-series and there has been no clear trend across ages over the last several years. Sex-ratio data from the commercial fishery landings represented an important new source of information for the 2019 assessment. Data from sampled Pacific halibut in 2017 indicated a very high proportion female coastwide (82%), and a range from 65% in Biological Region 4B to 92% in Biological

[1] All weights in the document are 'net' weights; head-off and entrails removed approximately 75% of round weight. Mortality estimates reported in this document were current as of the end of October 2019, reflecting the information available for the stock assessment analysis.

Region 4. Data from 2018 reflected very similar patterns, with females comprising 80% of the coastwide commercial landings (by number). Updated trends indicate that population distribution (measured via the modelled FISS catch in weight of all Pacific halibut) has been decreasing in Biological Region 3 since 2004, and increasing in Biological Regions 2 and 4.

This stock assessment continues to be implemented using the generalized software stock synthesis (Methot and Wetzel 2013). The analysis consists of an ensemble of four equally weighted models: two long time-series models, reconstructing historical dynamics back to the beginning of the modern fishery, and two short time-series models incorporating data only from 1992 to the present, a time-period for which estimates of all sources of mortality and survey indices are available for all regions. For each time-series length, there are two models: one fitting to coastwide aggregate data, and one fitting to data disaggregated into the four geographic regions. This combination of models includes uncertainty in the form of alternative hypotheses about several important axes of uncertainty, including: natural mortality rates (estimated in the long time-series models, fixed in the short time-series models), environmental effects on recruitment (estimated in the long time-series models), and other model parameters. Results are based on the approximate probability distributions derived from the ensemble of models, thereby incorporating the uncertainty within each model as well as the uncertainty among models. The 2019 stock assessment was conducted in two phases: first, a preliminary assessment (Stewart and Hicks 2019a) underwent an external independent peer review, and a two-part review by the IPHC's Scientific Review Board (SRB; IPHC-2019-SRB014-R, IPHC-2019-SRB015-R), second, the preliminary assessment was updated to include all data through 2019. Overall, the inclusion of the 2017 sex-ratio data resulted in higher spawning biomass for all models, and the updated whale depredation data made little difference to the results. Extending the time-series back to 1992 in the two short models resulted in higher estimates of recruitment for 1994 and 1995. Regularizing and tuning the series had different effects on each model. The 2019 data revised the estimates of the 2012 year-class upward slightly, but had little effect on the overall time-series, and the 2018 sex-ratio data was very similar to the 2017 information included in the preliminary analysis and therefore produced little additional change.

The results of the 2019 stock assessment indicate that the Pacific halibut stock declined continuously from the late 1990s to around 2012. That trend is estimated to have been largely a result of decreasing size-at-age, as well as somewhat weaker recruitment strengths than those observed during the 1980s. The spawning biomass (SB) is estimated to have increased gradually to 2016, and then decreased to an estimated 194 million pounds (~87,850 t) at the beginning of 2020, with an approximate 95% confidence interval ranging from 133 to 248 million pounds (~60,500-112,500 t). The historical female spawning biomass estimated from the stock assessment ensemble was slightly larger than that estimated in previous assessments at the end of the time series, and considerably larger prior to the early 2000s, although the trend remains very similar in recent years using these updated data sources. Pacific halibut recruitment estimates show the largest recent cohorts in 1999 and 2005. Cohorts from 2006 through 2010 are estimated to be smaller than those from 1999-2005 which results in a high probability of decline in both the stock and fishery yield as these low recruitments become increasingly important to the age range over which much of the harvest and spawning takes place. Based on age data from the 2018 survey, this assessment estimated the 2011 and 2012 year-classes to be similar to those in 2000-04. This is consistent with the appearance of these cohorts in the 2018 assessment, although they remain below the level of the 1999 and 2005 year-classes even with second year of observation. The projected spawning biomass over the next 2-4 years includes the effects of these year classes maturing at ages 8-13.

The IPHC's interim management procedure uses a relative spawning biomass of 30% as a fishery trigger, to begin reducing the reference fishing intensity to the point when relative spawning biomass reaches a limit at 20%, where directed fishing is halted due to the critically low biomass condition. The 2019 assessment, after Scientific Review Board and external review, and following the developments in the IPHC's Management Strategy Evaluation (MSE) process, has updated this calculation to include recent biological conditions in a 'dynamic' calculation, measuring the effect of fishing on the spawning biomass. This avoids the potential situation where environmental and biological conditions could be conflated with fishing effects. The dynamic calculation is estimated to be 32% in 2020 (approximate credible interval: 22-46%) and relatively stable after dropping below 30% from 2009-2015. The probability that the stock is below the $SB_{30\%}$ level is estimated to be 46%, with less than a 1% chance that the stock is below $SB_{20\%}$. Based on the 2019 assessment, and including the higher proportion of females in the directed commercial landings than previously understood, the 2019 fishing intensity is estimated to correspond to an $F_{42\%}$ (credible interval: 29-57%). Stock projections were conducted using the integrated results from the stock assessment ensemble, details of IPHC Regulatory Area-specific catch sharing plans and estimates of mortality from the 2019 directed fisheries and other sources of mortality. The stock is projected to decrease with at least a 51% chance over the period from 2021-23 for all TCEYs greater than the "3-year surplus" of 18.4 million pounds (~8,350 t), corresponding to a projected SPR of 63% (credible interval 44-75%). At the reference level (a projected SPR of 46%) the probability of spawning biomass decline to 2021 is 89%, decreasing to 75% in three years, as the 2011 and 2012 cohorts mature. At the *status quo* TCEYs (38.61 million lb, (~17,500 t), the probability of spawning biomass declines is 97% and 87% for one and three years respectively. The one-year risk of the stock dropping below $SB_{30\%}$ ranges from 43% (at the 3-year surplus level) to 49% at the *status quo* TCEYs. Over three years these probabilities range from 37% to 50% depending on the level of mortality.

A stepwise bridging analysis from the 2018 to 2019 assessments, a summary of sensitivity and retrospective analyses conducted in the preliminary assessment, and a discussion of major sources of uncertainty are also included in this document.

INTRODUCTION

The stock assessment reports the status of the Pacific halibut (*Hippoglossus stenolepis*) resource in the IPHC Convention Area. As in recent stock assessments, the resource is modelled as a single stock extending from northern California to the Aleutian Islands and Bering Sea, including all inside waters of the Strait of Georgia and Puget Sound, but excludes known extremities in the western Bering Sea within the Russian Exclusive Economic Zone (Figure 1). The Pacific halibut fishery has been managed by the IPHC since 1923. Mortality limits for each of eight IPHC Regulatory Areas[2] are set each year by the Commission. The stock assessment provides a brief summary of recently collected data; a more detailed treatment of data sources included in the assessment and used for other analyses supporting harvest policy calculations is provided in a separate document on the IPHC's stock assessment webpage. Results include current model estimates of stock size and trend reflecting all available data. Specific management information is summarized via a decision table reporting the estimated risks associated with alternative management actions. Mortality tables projecting detailed summaries

[2] The IPHC recognizes sub-Areas 4C, 4D, 4E and the Closed Area for use in domestic catch agreements but manages the combined Area 4CDE.

for fisheries in each IPHC Regulatory Area (and reference levels indicated by the IPHC's interim management procedure) can be explored via the IPHC's mortality projection tool.

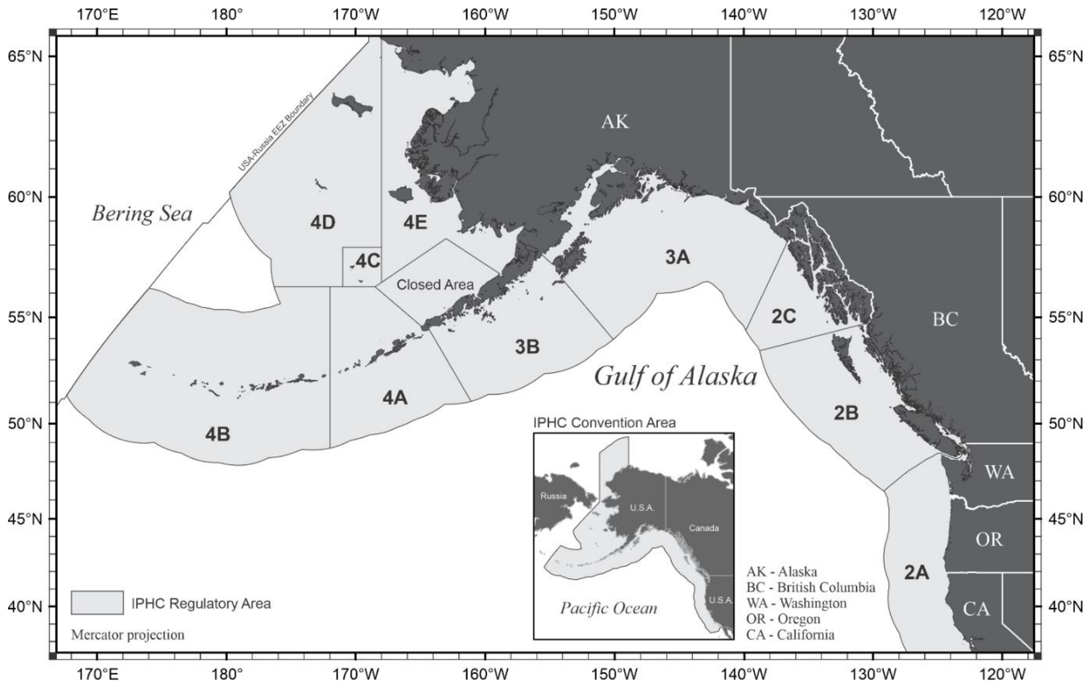

FIGURE 1. IPHC Convention Area (insert) and IPHC Regulatory Areas.

The IPHC's stock assessment and review process has developed from the first *ad hoc* meeting held in 2012 (Stewart et al. 2013) to a formal and documented process involving the SRB (https://www.iphc.int/the-commission/structure-of-the-commission) and periodic external independent peer review (https://www.iphc.int/management/science-and-research/stock-assessment). The IPHC's SRB meets two-three times per year, in June to review stock assessment development, in September to review progress in response to the June review and to finalize the model structure and methods to be used in conducting the year's stock assessment, and as needed in December to review any unexpected results and address any questions arising from the stock assessment. Within this annual review process two types of stock assessments are produced: 1) updated assessments where new data are added but the methods and model structures remain unchanged, and 2) full stock assessments occurring every three years in which model structure and methods are revised to reflect new data, approaches and comments from SRB and independent review. The 2015 stock assessment was a full analysis (Stewart and Martell 2016; Stewart et al. 2016), 2016-2018 were updated assessments (Stewart and Hicks 2018a, 2019b; Stewart and Hicks 2017), and the 2019 assessment is also a full analysis. New data sources including estimates of the sex-ratio of the directed commercial Pacific halibut landings for 2017, a revised modelled survey time series accounting for improved whale depredation criteria, and several improvements to the model structure and software were included in a preliminary assessment provided to the SRB and for the external review in June 2019 (IPHC-2019-SRB014-07). Requests from SRB014 were addressed at SRB015, which also included a discussion of the recommendations from the external independent peer review.

DATA SOURCES

Each year, the data sources used to support this assessment are updated to include newly available information, and refined to reflect the most current and accurate information available to the IPHC. Major reprocessing and development of supplementary data sources was conducted in 2013, 2015, and again in 2019 (Stewart and Hicks 2019a). All available information was finalized on 31 October 2019 in order to provide adequate time for analysis and modeling. As has been the case in all years, some data are incomplete, or include projections for the remainder of the year. These include commercial fishery WPUE, commercial fishery age composition data, and 2019 mortality estimates for all fisheries still operating after 31 October. All preliminary data series in this analysis will be fully updated as part of the 2020 stock assessment.

Data for stock assessment use are initially compiled by IPHC Regulatory Area, and then aggregated to four Biological Regions: Region 2 (Areas 2A, 2B, and 2C), Region 3 (Areas 3A, 3B), Region 4 (4A, 4CDE) and Region 4B and then coastwide. In addition to the aggregate mortality (including all sizes of Pacific halibut), the assessment includes data from both fishery dependent and fishery independent sources as well as auxiliary biological information, with the most spatially complete data available since the late-1990s. Primary sources of information for this assessment include modelled indices of abundance (IPHC-2020-AM096-07; based on the IPHC's annual fishery-independent setline survey (FISS; in numbers and weight) and other surveys), commercial Catch-Per-Unit-Effort (weight), and biological summaries from both sources (length-, weight-, and age-composition data). In aggregate, the historical time series of data available for this assessment represents a considerable resource for analysis. The range of relative data quality and geographical scope are also considerable, with the most complete information available only in recent decades (Figure 2). A detailed summary of input data used in this stock assessment can be found in a separate document on the IPHC's stock assessment webpage.

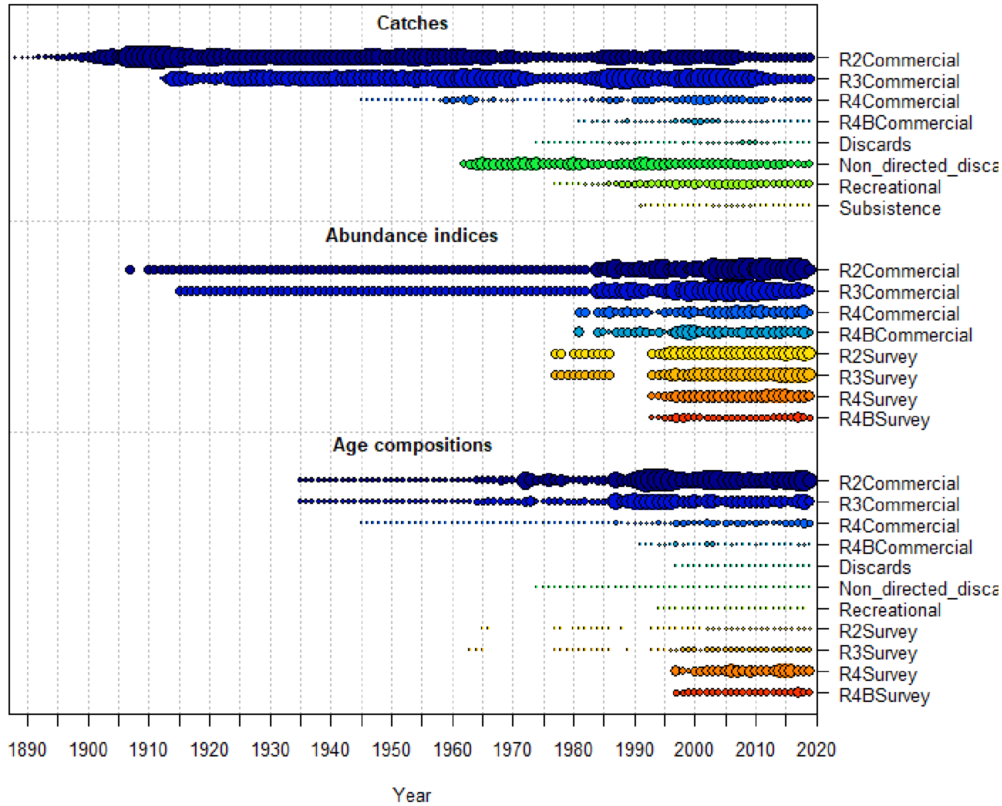

FIGURE 2. Overview of data sources. Circle areas are proportional to magnitude (catches) or the relative precision of the data (larger circles indicate greater precision for indices of abundance and age composition data).

Briefly, known Pacific halibut mortality consists of target commercial fishery landings and discard mortality (including research), recreational fisheries, subsistence, and bycatch mortality in fisheries targeting other species (where Pacific halibut retention is prohibited). Over the period 1920-2019 removals have totaled 7.2 billion pounds (~3.3 million metric tons, t), ranging annually from 34 to 100 million pounds (16,000-45,000 t) with an annual average of 63 million pounds (~29,000 t). Annual mortality was above this long-term average from 1985 through 2010, and has averaged 41 million pounds (~18,500 t) from 2016-19. Coastwide commercial Pacific halibut fishery landings (including research landings) in 2019 were approximately 24.3 million pounds (~11,000 t), up 3% from 2018. Discard mortality in non-directed fisheries was estimated to be 6.4 million pounds in 2019 (~2,900 t)[3], up 5% from 2018. The total recreational mortality (including estimates of discard mortality) was estimated to be 6.9 million pounds (~3,100 t), very close to the final estimate for 2018. Mortality from all sources increased by 3% to an estimated 39.7 million pounds (~18,000 t) in 2019.

The 2019 FISS detailed a coastwide aggregate NPUE (modelled via the space-time methodology) which showed a third consecutive year of decrease, down 4% from 2018 with 2017-19 each representing the lowest in the time-series. Biological Region 3 declined by 10% to the lowest estimate in the time-series while Biological Regions 2, 4, and 4B all increased slightly, but remain near historical lows. The 2019 modelled coastwide WPUE of legal (O32)

[3] The IPHC receives preliminary estimates of the current year's bycatch mortality in from the National Marine Fisheries Service Alaska Regional Office, Northwest Fisheries Science Center, and Fisheries and Oceans Canada in late October.

Pacific halibut, the most comparable metric to observed commercial fishery catch rates, was lower (5%) than 2018, down for the third consecutive year and at the lowest value in the time series. Individual IPHC Regulatory Areas varied from a 26% increase (Regulatory Area 3B) to a 17% decrease (Regulatory Area 3A). The FISS sampling associated with the expansion in Biological Region 3 resulted in lower estimated catch-rates in this Region compared to the rest of the coast, and reduced the uncertainty in the index both for Region 3 and coastwide. Commercial fishery WPUE (based on extensive, but incomplete logbook records available for this assessment) increased 4% coastwide, with mixed performance across IPHC Regulatory Areas. A bias correction (to account for additional logbooks compiled after the fishing season, standard practice in recent years) resulted in an estimate of a 1% increase coastwide. As in 2018, fisheries and gear types are reported separately to allow more detailed evaluation of fishery performance.

Biological information (ages and lengths) from the commercial fishery continue to show the 2005 year-class as the largest contributor (in number) to the fish encountered. In the FISS age-frequency data, 2011 and 2012 cohorts (7 and 8 years old, following a series of weak cohorts from 2006-10) represented the largest proportions in some IPHC Regulatory Areas for the total catch, and the largest proportions coastwide for sublegal female Pacific halibut. At the coastwide level, individual size-at-age continues to be very low relative to the rest of the time-series and there has been no clear trend across ages over the last several years. For the first time, direct estimates of the sex-ratio at age for the directed commercial fishery were available for the IPHC's stock assessment. Data from sampled Pacific halibut in 2017 indicated a very high proportion female coastwide (82%), and a range from 65% in Biological Region 4B to 92% in Biological Region 4. Data from 2018 reflected very similar patterns, with females comprising 80% of the coastwide commercial landings (by number).

Updated trends indicate that population distribution (measured via the modelled FISS catch in weight of all Pacific halibut) has been decreasing in Biological Region 3 since 2004, and increasing in Biological Regions 2 and 4. Survey data are insufficient to estimate stock distribution prior to 1993. It is therefore unknown how historical distributions, and the average distribution likely to occur in the absence of fishing mortality may compare with recent observations.

STOCK ASSESSMENT

Creating robust, stable, and well-performing stock assessment models for the Pacific halibut stock has historically proven to be problematic 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). Although recent modelling efforts have created some new alternatives, no single model satisfactorily approximates all aspects of the available data and scientific understanding. For 2019, an ensemble of four stock assessment models representing a two-way cross of short vs. long time series', and aggregated coastwide vs. Areas-As-Fleets (AAF) models was again used to explore the range of plausible current stock estimates. AAF models are commonly applied when biological differences among areas or sampling programs 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, the 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 coastwide level.

The ensemble approach recognizes that there is no “perfect” 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; Stewart and Hicks 2018b). 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, and provides a more realistic perception of uncertainty than any single model, and therefore a stronger basis for risk assessment.

This stock assessment continues to be implemented using the generalized software stock synthesis (Methot and Wetzel 2013). The analysis consists of an ensemble of four equally weighted models: two long time-series models, reconstructing historical dynamics back to the beginning of the modern fishery, and two short time-series models incorporating data only from 1992 to the present, a time-period for which estimates of all sources of mortality and survey indices are available for all regions. For each time-series length, there are two models: one fitting to coastwide aggregate data, and one fitting to data disaggregated into the four geographic regions. This combination of models includes uncertainty in the form of alternative hypotheses about several important axes of uncertainty, including: natural mortality rates (estimated in the long time-series models, fixed in the short time-series models), environmental effects on recruitment (estimated in the long time-series models), and other model parameters. This combination of models included a broad suite of structural and parameter uncertainty, including natural mortality rates (estimated in the long time-series models, fixed in the short time-series models), environmental effects on recruitment (estimated in the long time-series models), fishery and survey selectivity (by region in the AAF models) and other model parameters. These sources of uncertainty have historically been very important to the understanding of the stock, as well as the annual assessment results (Clark and Hare 2006; Clark et al. 1999; Stewart and Martell 2016). The benefits of the long time-series models include historical perspective on recent trends and biomass levels; however, these benefits come at a computational and complexity cost. The short time-series models make fewer assumptions about the properties of less comprehensive historical data, but they suffer from much less information in the short data series as well as little context for current dynamics.

As has been the case since 2012, the results of this stock assessment are based on the approximate probability distributions derived from the ensemble of models, thereby incorporating the uncertainty within each model (parameter or estimation uncertainty) as well as the uncertainty among models (structural uncertainty). This approach reduces the potential for abrupt changes in management quantities as improvements and additional data are added to individual models, and provides a more realistic perception of uncertainty than any single model, and therefore a stronger basis for risk assessment.

Each of the models in the ensemble was equally weighted, and differences in uncertainty within models propagated in the integration of results. In the future, it may be desirable to develop a method for weighting models based on the lack-of-fit to key data sources, retrospective patterns within models, as well as consistency of the results with biological understanding. Evaluation of alternative weighting approaches was presented to the IPHC Scientific Review Board (SRB) in

2015, 2016 and 2017 (Stewart 2017), but did not suggest a change to the equal weights that have been applied; therefore, that assumption has been retained. Additional models or variations of existing models for potential inclusion into the ensemble were evaluated in 2019, but none were determined to appreciably improve the characterization of overall uncertainty in stock trend, scale or fishing intensity. Future additions can be made as alternative approaches and refinements become available. Within-model uncertainty from each model was propagated through to the ensemble results via the maximum likelihood estimates and an asymptotic approximation to their variance. Point estimates in this stock assessment correspond to median values from the ensemble: with the simple probabilistic interpretation that there is an equal probability above or below the reported value.

The 2019 stock assessment was conducted in two phases: first, a preliminary assessment underwent an external independent peer review, and a two-part review by the IPHC's Scientific Review Board (SRB; IPHC-2019-SRB014-R, IPHC-2019-SRB015-R), second the preliminary assessment was updated to include all data through 2019. The 2019 stock assessment included a complete re-evaluation of all data sources and modelling choices. Although the basic ensemble approach and four structural models remain consistent with previous analyses, several key improvements were made including: extending the short time-series models back to 1992 to utilize the full modelled FISS index (beginning in 1993), additional flexibility in modelling fishery selectivity enabled by newly available sex-ratio at age data, and re-weighting the contributions of each type of data to the stock assessments based on the goodness of fit to index and age frequencies. The sex-ratio data were critically important to this assessment, as they allowed for direct estimation of parameters describing the scale of male selectivity in each of the individual models.

COMPARISON WITH PREVIOUS ASSESSMENTS

The transition from the 2018 stock assessment included five steps, beginning from the 2018 model using the newest version of the stock synthesis software, to update to the preliminary individual model results for 2019 (Stewart and Hicks 2019a) and then to the final estimates reported included in the 2019 ensemble:

1) Add the newly available sex-ratio data from the 2017 commercial fishery landings and estimate male selectivity scale parameters.
2) Extend the time series (for the two short models only) from 1996 to 1992 and add a stock-recruitment function to these models.
3) Replace the modelled FISS time-series with the series corrected for whale depredation.
4) Regularize and tune each model to be reliable and internally consistent given all the changes that had been made.
5) Add the 2018 sex-ratio data, estimates of 2019 mortality, extend all data sources through 2019, and adjust data-weighting to be consistent with the model fit as documented in Stewart and Hicks (2019) for the final assessment.

Each of these five steps is included in the bridging analysis provided in Figures 3-5. Overall, the inclusion of the 2017 sex-ratio data resulted in higher spawning biomass for all models, and the updated whale depredation data made little difference to the results. Extending the time-series back to 1992 in the two short models resulted in higher estimates of recruitment for 1994 and 1995. Regularizing and tuning the series had different effects on each model. The 2019 data revised the estimates of the 2012 year-class upward slightly, but had little effect on the overall time-series, and the 2018 sex-ratio data was very similar to the 2017 information included in the preliminary analysis and therefore produced little additional change. For the two coastwide models, the historical female spawning biomass estimated from the stock assessment ensemble was slightly larger than that estimated in previous assessments at the end of the time series, and for the Areas-As-Fleets models the biomass estimated at the end of the time-series was generally lower than the 2018 assessment at the end of the time-series, but higher historically.

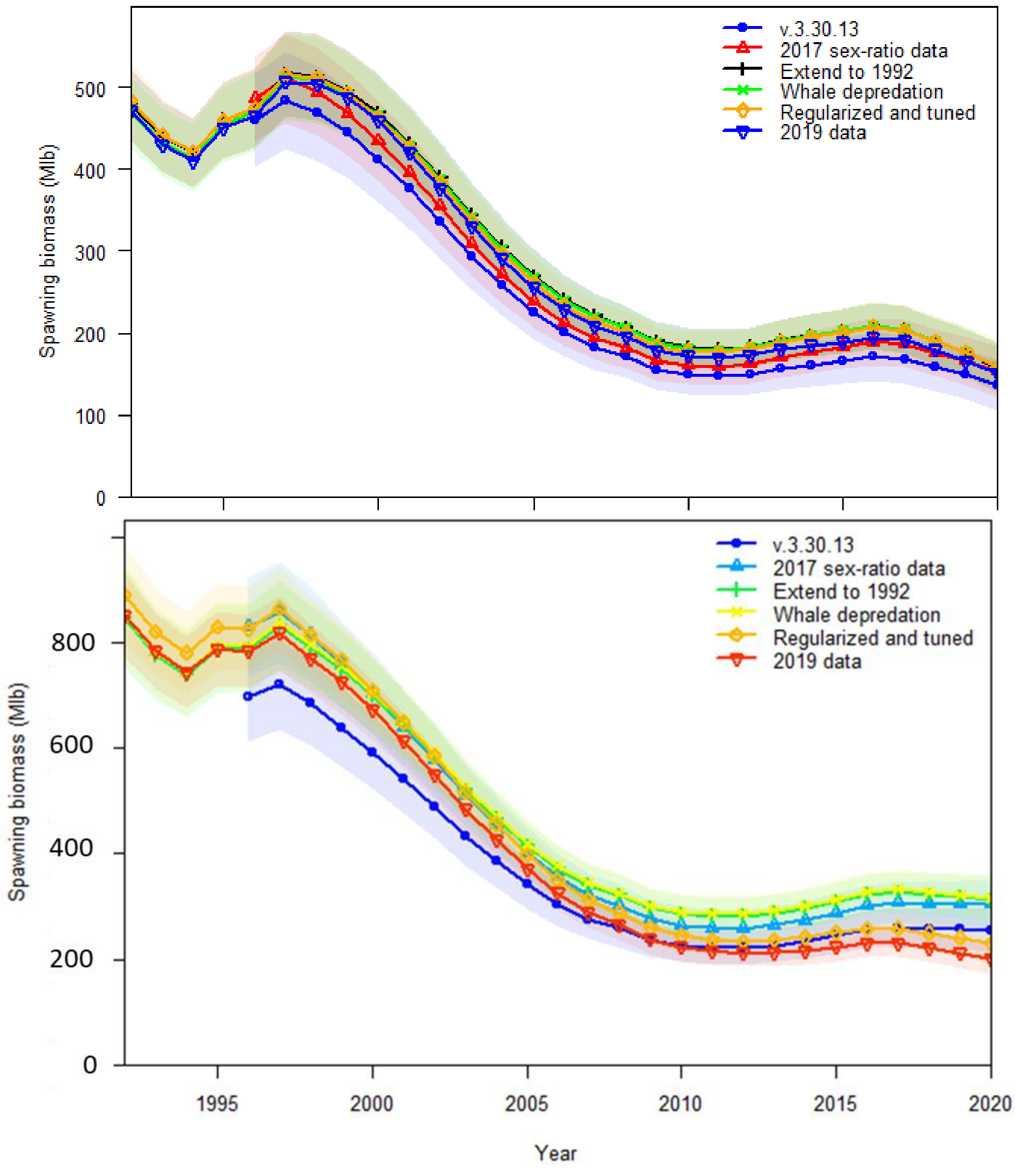

FIGURE 3. Bridging analysis showing the five steps between the 2018 and 2019 stock assessment model estimates of spawning biomass for the short coastwide model (top panel) and short Areas-As-Fleets model (bottom panel).

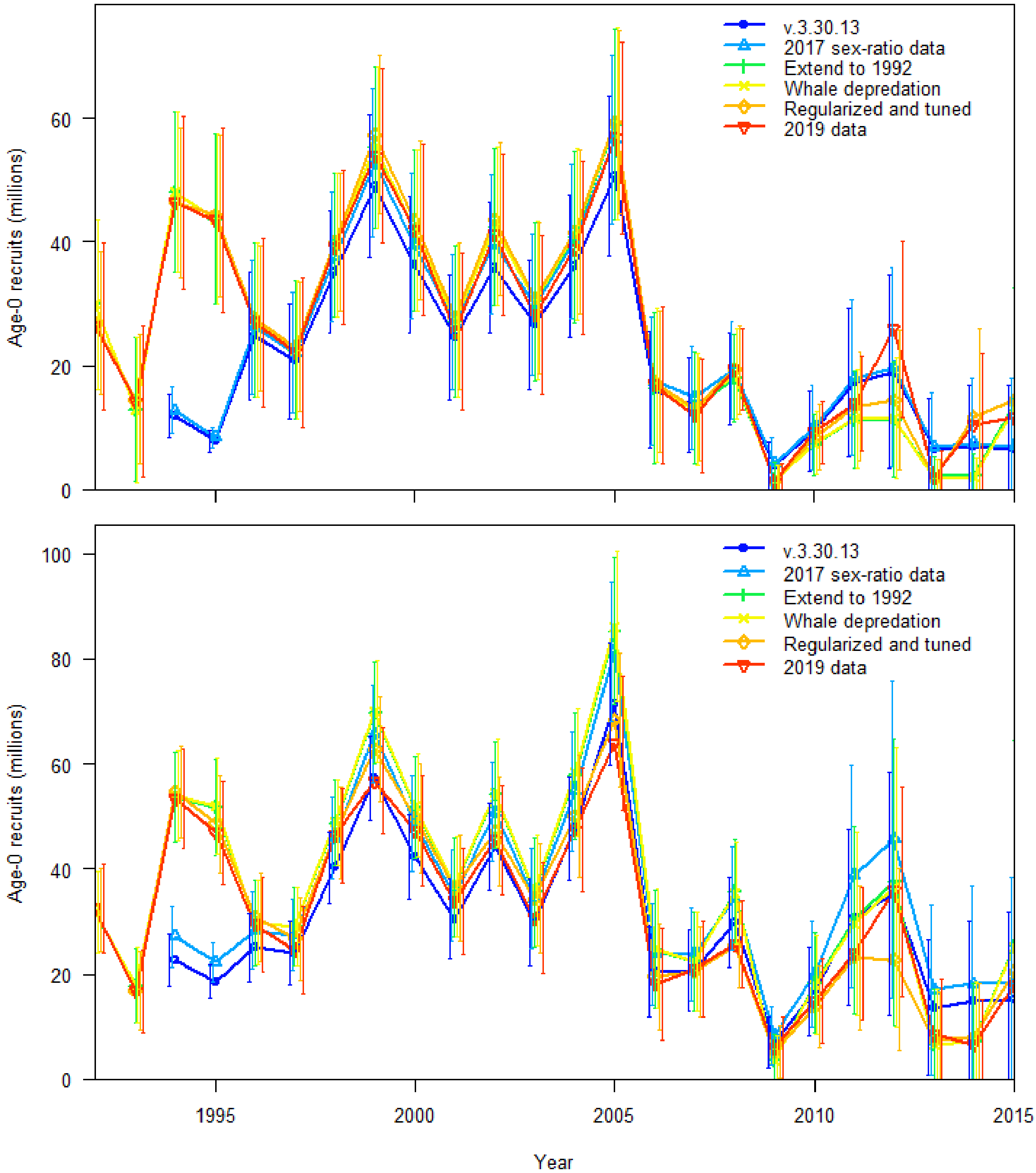

FIGURE 4. Bridging analysis showing the five steps between the 2018 and 2019 stock assessment model estimates of recruitment for the short coastwide model (top panel) and short Areas-As-Fleets model (bottom panel).

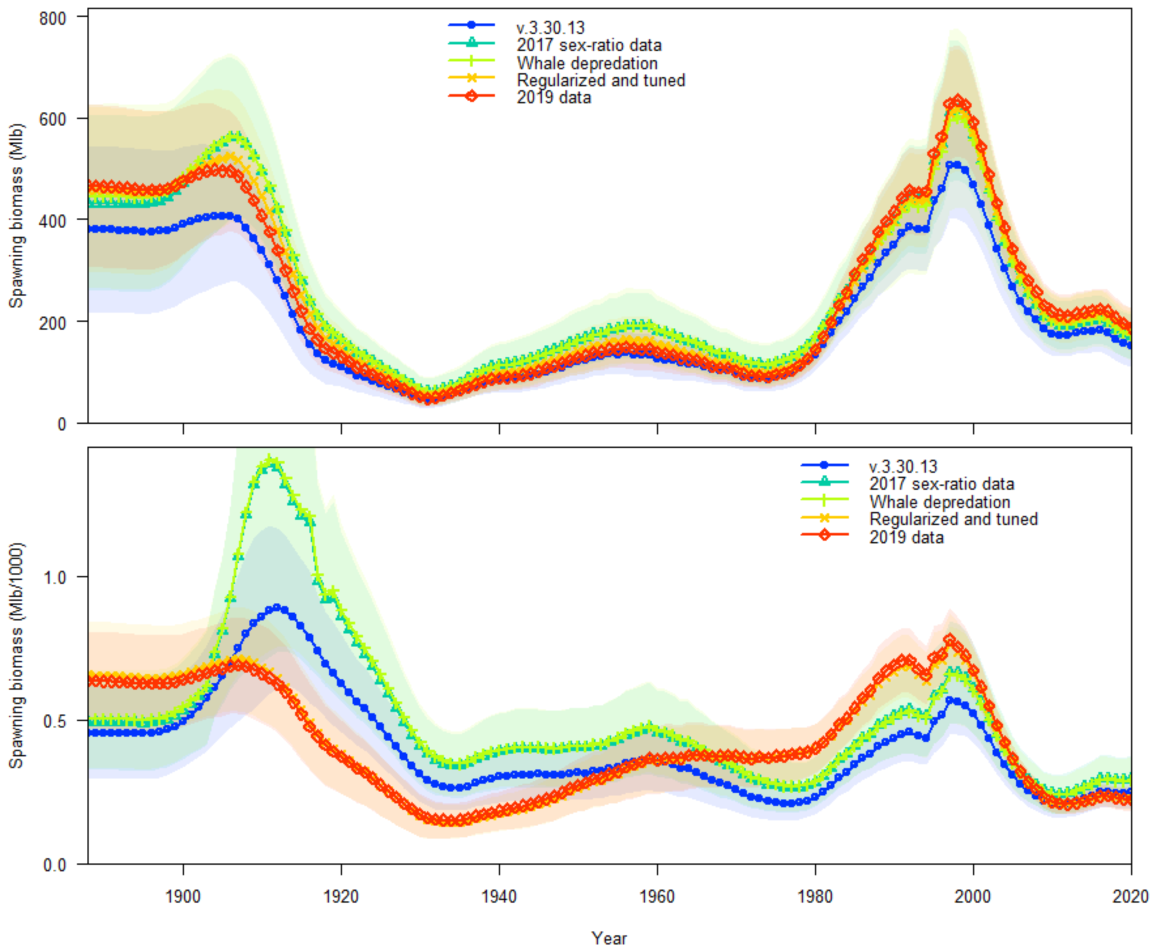

FIGURE 5. Bridging analysis showing the five steps between the 2018 and 2019 stock assessment model estimates of spawning biomass for the long coastwide model (top panel) and long Areas-As-Fleets model (bottom panel).

Comparison of this year's ensemble results with previous stock assessments indicates that the estimates of spawning biomass from the 2019 ensemble remain consistent with those from 2012-18. Each of the previous terminal assessment values lie inside the predicted 50% interval of the current ensemble (Figure 6). The 2019 assessment estimates a larger spawning biomass for the entire time-series, with the difference being more pronounced prior to 2010. The uncertainty is much greater prior to approximately 2005 reflecting the differences among the four individual models as well as the increased uncertainty in scale resulting from only two years of sex-ratio data to inform the models.

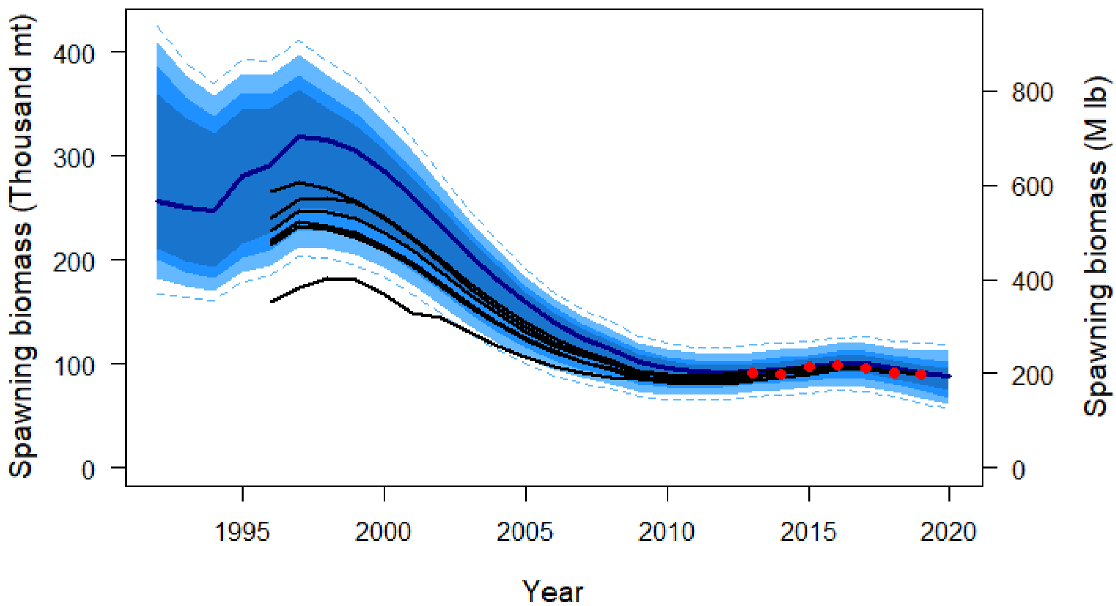

FIGURE 6. Retrospective comparison among recent IPHC stock assessments. Black lines indicate estimates of spawning biomass from assessments conducted from 2012-2018 with the terminal estimate shown as a point, the shaded distribution denotes the 2019 ensemble: the dark blue line indicates the median (or "50:50 line") with an equal probability of the estimate falling above or below that level; colored bands moving away from the median indicate the intervals containing 50/100, 75/100, and 95/100 estimates; dashed lines indicating the 99/100 interval.

BIOMASS, RECRUITMENT, AND REFERENCE POINT RESULTS

Ensemble

The results of the 2019 stock assessment indicate that the Pacific halibut stock declined continuously from the late 1990s to around 2011 (Figure 6, Table 1). That trend is estimated to have been largely a result of decreasing size-at-age, as well as somewhat weaker recruitment strengths than those observed during the 1980s. The spawning biomass (SB) is estimated to have increased gradually to 2016, and then decreased to an estimated 194 million pounds (~87,850 t) at the beginning of 2020, with an approximate 95% confidence interval ranging from 133 to 248 million pounds (~60,500-112,500 t; Figure 7). The differences among the individual models contributing to the ensemble are most pronounced prior to the early 2000s (Figure 8); however, current stock size estimates (at the beginning of 2020) also differ substantially among the four models (Figure 9). The differences in both scale and recent trend reflect the structural assumptions, e.g., higher natural mortality estimated in the long coastwide model and dome-shaped selectivity for Regions 2 and 3 in the AAF models.

Differences are also apparent in the recent recruitment estimates, which suggest larger recruitments in 1999 and 2005 than in other recent years (Figure 10, Table 1). All of these recent recruitments are much lower than the 1987 cohort, and in the two long time-series models they are at or below those in the late 1970s and early 1980s (Figure 11). Cohorts from 2006 through 2010 are estimated to be much smaller than those from 1999-2005 which results in a high probability of decline in both the stock and fishery yield as these low recruitments become increasingly important to the age range over which much of the harvest and spawning takes place. Based on age data from the 2019 survey, this assessment estimated the 2011 and 2012 year-classes to be similar to those in 2000-04. This is consistent with the appearance of these cohorts in the previous (2018) assessment, although they remain below the level of the 1999 and 2005 year-classes even with second year of observation. The projected spawning biomass over the next 2-4 years includes the effects of these year classes maturing at ages 8-13. Short-term trends in fishery yield are likely to decrease as Pacific halibut born in 2006 and later become increasingly important to the directed fisheries. The differing effects of these reduced recruitments on fishery yield (the effects are more delayed for spawning biomass, which largely comprises ages greater than 11) are illustrated in the estimated declines in age-8+ biomass, which start earlier and are more pronounced than those seen for spawning biomass (Table 1). Recruitment estimates after 2012 remain poorly informed by information from the fishery and survey data, and are therefore highly uncertain.

TABLE 1. Estimated recent median spawning biomass (SB; millions lbs) and fishing intensity (smaller values indicate higher fishing intensity) with approximate 95% credibility intervals, and age-0 recruitment (millions) and age-8+ biomass (millions lbs) from the individual models (CW=coastwide, AAF=Areas-As-Fleets) comprising the ensemble.

Year	SB	SB interval	Fishing intensity ($F_{XX\%}$)	Fishing intensity interval	Recruitment				Age-8+ biomass			
					CW Long	CW Short	AAF Long	AAF Short	CW Long	CW Short	AAF Long	AAF Short
1992	567	401-904	44%	30-56%	78.2	26.4	45.4	32.4	1,609	2,611	2,038	1,823
1993	553	387-832	44%	30-56%	53.3	14.3	28.6	16.7	1,510	697	1,883	1,674
1994	547	375-789	44%	31-57%	134.8	46.3	74.5	53.4	1,446	1,106	1,785	1,577
1995	619	415-837	53%	37-64%	125.9	43.5	68.7	46.9	2,013	1,049	2,223	1,905
1996	644	430-834	52%	37-64%	74.3	27.0	41.3	29.4	1,978	986	2,195	1,888
1997	703	469-875	46%	33-59%	67.2	22.2	36.3	24.5	2,042	1,304	2,264	1,953
1998	694	467-832	44%	32-58%	107.8	39.2	65.6	46.4	1,941	1,287	2,132	1,839
1999	673	451-792	41%	30-56%	141.0	53.9	80.7	56.7	1,786	1,344	1,959	1,693
2000	629	424-735	40%	29-57%	104.4	41.9	66.7	47.3	1,616	1,296	1,784	1,553
2001	576	387-671	38%	28-55%	75.6	25.5	50.0	33.8	1,423	1,205	1,586	1,382
2002	516	346-600	33%	26-51%	100.1	41.1	65.7	45.5	1,360	1,102	1,486	1,306
2003	453	303-527	30%	24-48%	77.9	28.2	47.6	30.7	1,296	974	1,386	1,219
2004	400	266-465	27%	23-45%	107.8	39.0	70.3	47.5	1,178	921	1,258	1,112
2005	350	233-406	25%	21-43%	137.9	56.9	90.6	63.8	1,057	866	1,126	997
2006	309	208-359	25%	21-43%	50.2	16.9	30.2	18.1	999	789	1,058	937
2007	277	190-322	25%	20-42%	45.4	12.0	33.7	21.0	995	703	1,031	910
2008	253	178-296	24%	19-41%	50.9	19.5	39.6	25.6	945	658	989	876
2009	226	162-267	25%	19-42%	23.2	1.4	16.2	5.9	853	651	907	803
2010	212	156-252	25%	19-42%	35.5	9.3	28.3	14.9	814	624	874	773
2011	205	154-244	28%	23-47%	72.5	13.9	54.3	24.0	766	560	827	729
2012	202	156-242	32%	28-51%	99.1	26.0	80.6	35.8	764	542	827	726
2013	206	162-247	34%	29-53%	44.9	2.1	35.6	8.4	807	513	871	762
2014	210	167-250	40%	32-56%	69.6	10.4	44.7	6.5	753	513	826	722
2015	216	170-256	41%	32-56%	77.8	11.6	47.6	17.9	704	552	794	694
2016	222	175-264	42%	31-56%	NA	NA	NA	NA	673	525	782	683
2017	221	172-264	41%	30-56%	NA	NA	NA	NA	611	493	735	635
2018	212	161-255	43%	31-58%	NA	NA	NA	NA	564	481	703	597
2019	203	148-250	42%	29-57%	NA	NA	NA	NA	579	432	730	583
2020	194	133-248	NA	NA	NA	NA	NA	NA	626	392	802	594

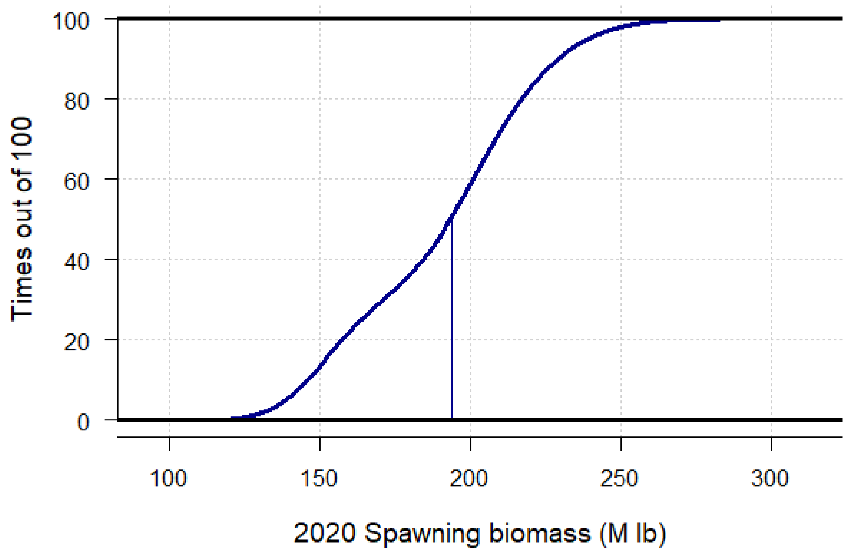

FIGURE 7. Cumulative distribution of the estimated spawning biomass at the beginning of 2020. Curve represents the estimated probability that the biomass is less than or equal to the value on the x-axis; vertical line represents the median (194 million pounds; ~87,850 t).

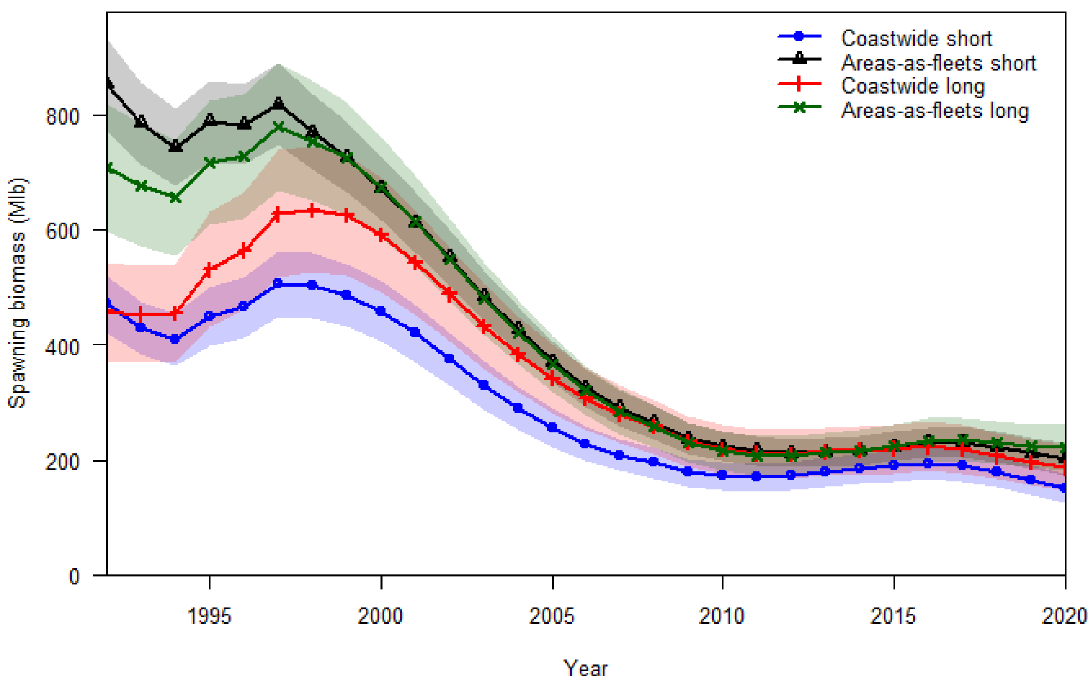

FIGURE 8. Estimated spawning biomass trends (1996-2020) based on the four individual models included in the 2018 stock assessment ensemble. Solid lines indicate the maximum likelihood estimates; shaded intervals indicate approximate 95% credible intervals.

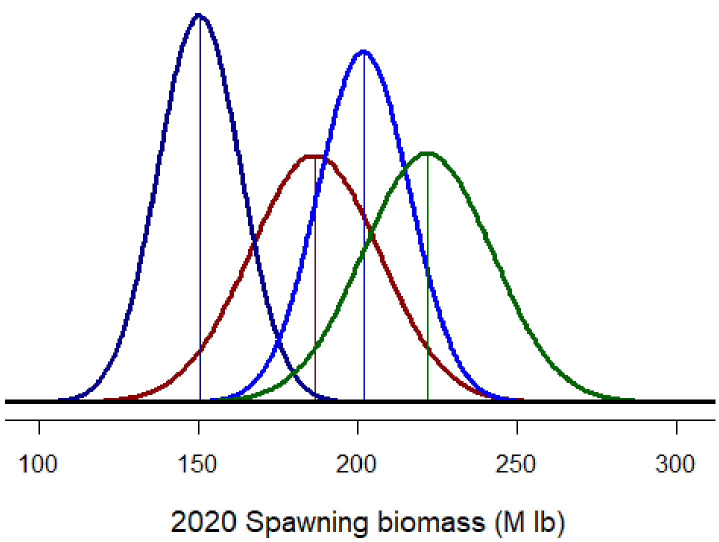

FIGURE 9. Distribution of individual model estimates for the 2020 spawning biomass. Vertical lines indicate the median values.

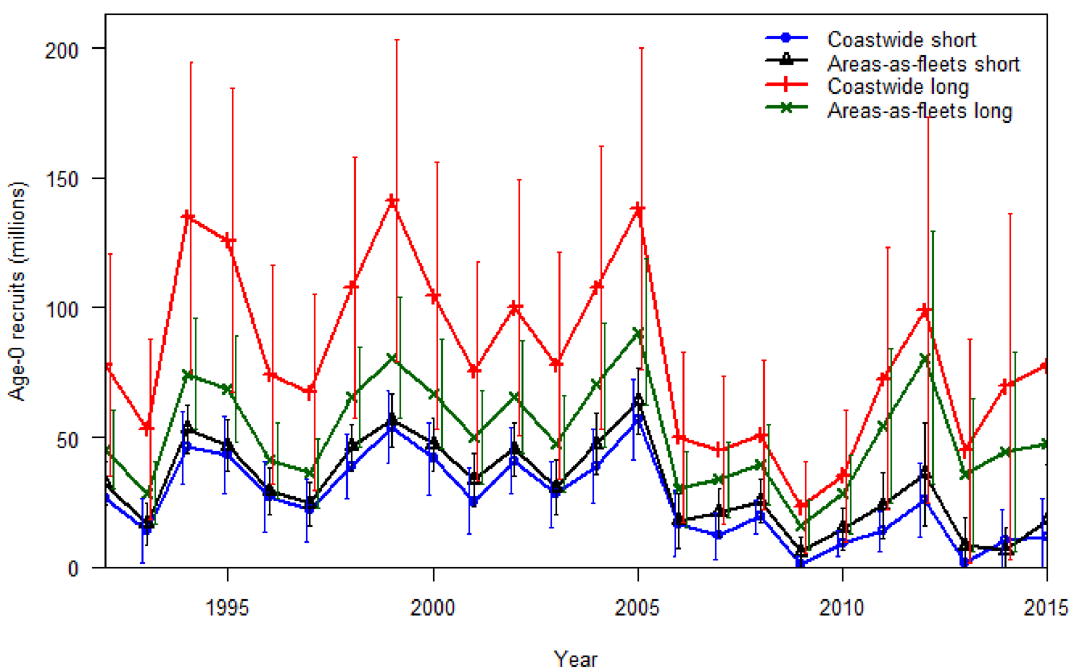

FIGURE 10. Estimated age-0 recruitment trends (1992-2015) based on the four individual models included in the 2019 stock assessment ensemble. Series indicate the maximum likelihood estimates; vertical lines indicate approximate 95% credible intervals.

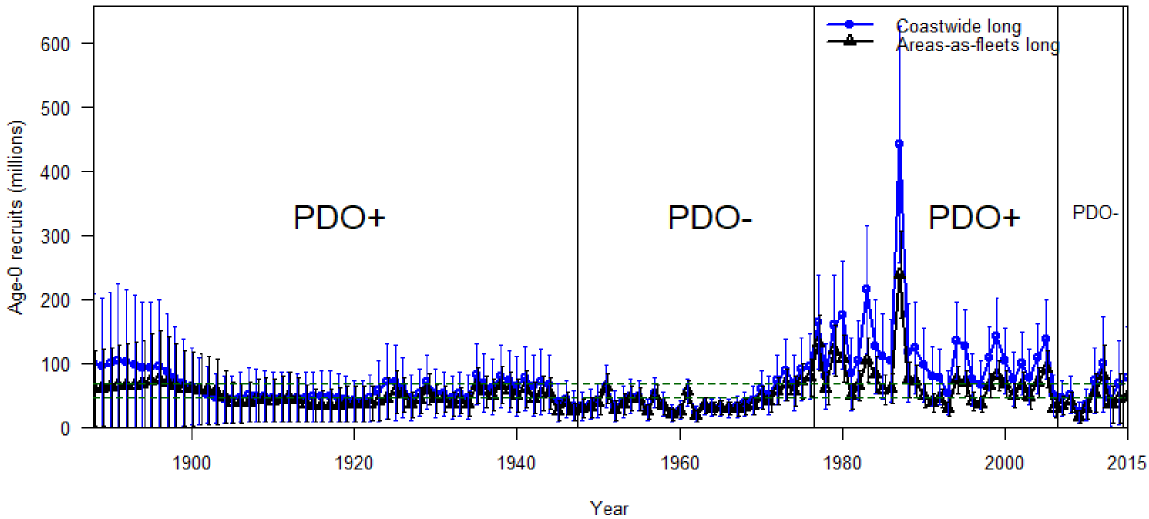

FIGURE 11. Trend in historical recruitment strengths (by birth year) estimated by the two long time-series models, including the effects of the Pacific Decadal Oscillation (PDO) regimes.

Ecosystem conditions

Average Pacific halibut recruitment is estimated to be higher (69 and 76% for the coastwide and AAF models respectively) during favorable Pacific Decadal Oscillation (PDO) regimes, a widely used indicator of productivity in the north Pacific. Historically, these regimes included positive conditions prior to 1947, poor conditions from 1947-77, positive conditions from 1978-2006, and poor conditions from 2007-13. Annual averages from 2014 through September 2019 have been positive; however, over this period many other environmental indicators, current and temperature patterns have been anomalous. Therefore, historical patterns of productivity related to the PDO may not be relevant to the most recent few years, and it will be years or decades before this can be verified via observed recruitment strengths.

Reference points

The IPHC's interim management procedure uses a relative spawning biomass of 30% as a trigger, to begin reducing the reference fishing intensity to a limit at 20%, where directed fishing is halted due to the critically low biomass condition. The relative spawning biomass has historically been calculated based on an arbitrary choice of 'good' weight-at-age and 'poor' recruitment levels estimated decades ago. The 2019 assessment, after Scientific Review Board (IPHC-2019-SRB014-07) and external review, and following the developments in the IPHC's Management Strategy Evaluation (MSE) process, has updated this calculation to include recent biological conditions. By using current weight-at-age and estimated recruitments influencing the current stock only, the 'dynamic' calculation measures the effect of fishing on the spawning biomass and provides a clear unfished reference point from which to calculate relative spawning biomass. This avoids the potential situation where environmental and biological conditions could be conflated with fishing effects.

Given the change in the calculation of these reference points from the fixed historical inputs to the dynamic calculation, a series of comparisons were made in order to clearly determine how much of the change in status from the 2018 assessment was due to the additional year of projection, the calculation methods, and the new data and updated models. The following reference points were constructed from the 2018 stock assessment and the final 2019 results:

- From the 2018 stock assessment: median relative biomass in 2019 (based on the previous reference points) was estimated to be 43% (95% interval from 27-63%), with a probability of being below $SB_{30\%}$ of 11%, and a probability of being below $SB_{20\%}$ of <1%.
- Extending the 2018 stock assessment time series, but not making any changes to the data or calculations: median relative biomass in 2020 (based on the previous reference points) was estimated to be 38% (95% interval from 22-51%), with a probability of being below $SB_{30\%}$ of 25%, and a probability of being below $SB_{20\%}$ of <1%.
- After updating the assessment to the preliminary 2019 configuration: median relative biomass in 2019 (based on the updated calculations) was estimated to be 32% (95% interval from 23-44%), with a probability of being below $SB_{30\%}$ of 38%, and a probability of being below $SB_{20\%}$ of <1%. The median relative spawning biomass at the beginning of 2020 was estimated to be 31% (95% interval from 20-44%), with a probability of being below $SB_{30\%}$ of 44%, and a probability of being below $SB_{20\%}$ of 2%.
- The median relative spawning biomass at the beginning of 2020 from the final 2019 stock assessment was estimated to be 32% (95% interval from 22-46%), with a probability of being below $SB_{30\%}$ of 46% (Figure 12), and a probability of being below $SB_{20\%}$ of <1%.

Thus, a portion of the change in status (from the beginning of 2019 based on the 2018 assessment to the beginning of 2020 based on the final 2019 assessment) is due to the change in reference points, and a portion of the change is due to the addition of new data and updating of the individual models comprising the ensemble (Table 2). The considerable uncertainty in these estimates leads to overlapping confidence intervals in all reference point comparisons.

TABLE 2. Comparison of 'historical' and 'dynamic' relative spawning biomass estimates from the 2018 and current 2019 stock assessments. Percentage indicates the relative spawning biomass estimated for that year with approximate 95% credible intervals in parentheses; $P(SB<SB_{XX\%})$ indicates the probability that the relative spawning biomass in that year is below the reference point (either 20 or 30%).

Year	**2018 Assessment** ('Historical' relative *SB*)	**2019 Assessment** ('Dynamic' relative *SB*)
2019	43% (27-63%) *P(SB<SB30%)* = 11% *P(SB<SB20%)* = <1%	32% (23-46%) *P(SB<SB30%)* = 44% *P(SB<SB20%)* = <1%
2020	38% (22-51%) *P(SB<SB30%)* = 25% *P(SB<SB20%)* = <1%	32% (22-46%) *P(SB<SB30%)* = 46% *P(SB<SB20%)* = <1%

The relative spawning biomass decreased continuously over the period 1992-2012, then increased gradually to just above the $SB_{30\%}$ fishery trigger after 2015 (Figure 13). This result reflects the greater effects of reduced recruitment, rather than fishing in the last few years. The probability that the stock is below the $SB_{30\%}$ level is estimated to be 46% at the beginning of 2020, with less than a 1% chance that the stock is below $SB_{20\%}$.

The IPHC's interim management procedure specifies a reference level of fishing intensity of a Spawning Potential Ratio (SPR) corresponding to an $F_{46\%}$; this equates to the level of fishing that would reduce the lifetime spawning output per recruit to 46% of the unfished level given current biology, fishery characteristics and demographics. In addition to the changes in relative spawning biomass from the 2018 to 2019 stock assessments, fishing intensity is estimated to be somewhat higher since 2003. Because the mortality inputs to the assessment models have not changed appreciably and the biomass is larger, this illustrates the effect of an increased fraction of females estimated to occur in the commercial landings, and therefore a greater effect of the lifetime spawning output of the stock. All sources of estimated mortality for 2018 correspond to a fishing intensity point estimate of $F_{42\%}$ (Table 1, Figure 14). Harvest levels of this magnitude are generally consistent with target rates for many similar stocks. The 95% interval of this distribution is considerable ($F_{57\%}$-$F_{29\%}$), reflecting the broad uncertainty in the estimated value. The recent time-series shows that the 2019 estimate corresponds to slightly higher fishing intensity than 2018, but similar to 2014-2017 (Figure 15). The updated estimates for 2014-2016, which was used as the reference level for the IPHC's interim management procedure reflect a value of $F_{42\%}$, a higher level of fishing intensity than estimated in 2017 ($F_{46\%}$). Comparing the relative spawning biomass and fishing intensity simultaneously over the recent historical period provides for an evaluation of trends conditioned on the currently defined reference points; this type of comparison is commonly called a 'phase' plot. The phase plot for Pacific halibut shows a counter-clockwise cycle, indicating that the relative spawning biomass increased in response to decreased fishing intensity from 2011-15 (Figure 16).

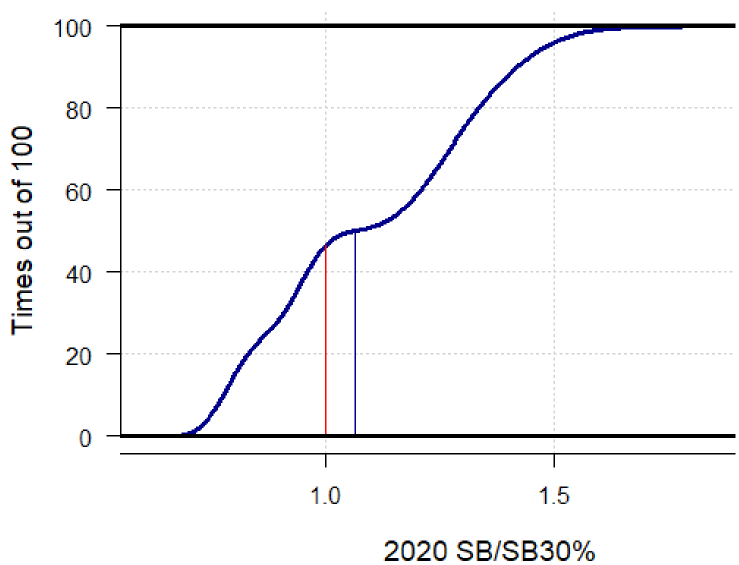

FIGURE 12. Cumulative distribution of 2020 ensemble spawning biomass estimates relative to the $SB_{30\%}$ reference point. Curve represents the estimated probability that the biomass is less than or equal to the value on the x-axis. Vertical lines denote the values corresponding to the fishery threshold in the IPHC's harvest policy (red; $SB_{30\%}$), and the median (blue; 32%).

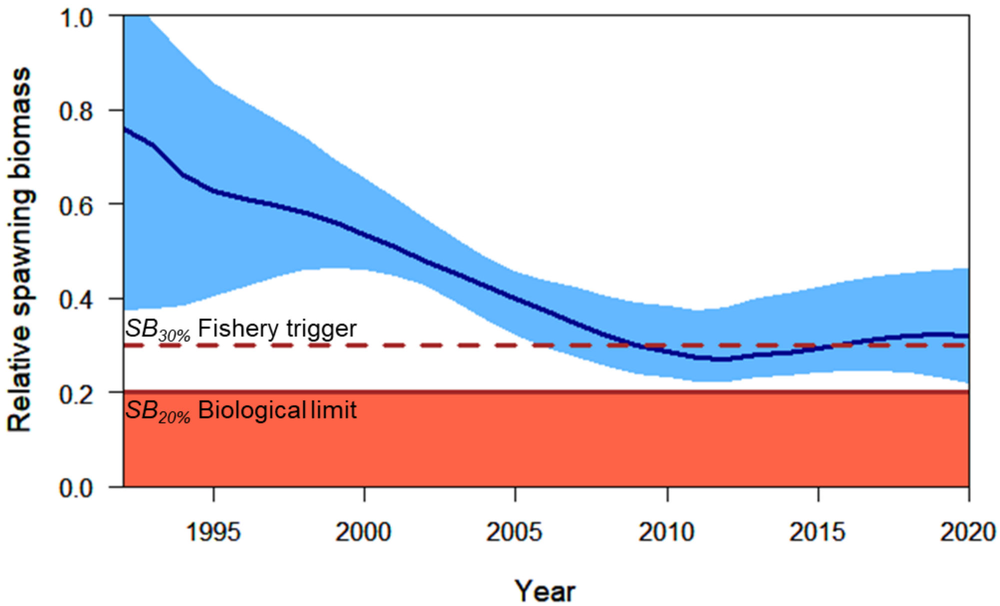

FIGURE 13. Estimated time-series of relative spawning biomass (compared to the unfished condition in each year) based on the median (dark blue line) and approximate 95% credibility interval (blue shaded area). IPHC management procedure reference points ($SB_{30\%}$ and $SB_{20\%}$) are shown as dashed and solid lines respectively, with the region of biological concern ($<SB_{20\%}$) shaded in red.

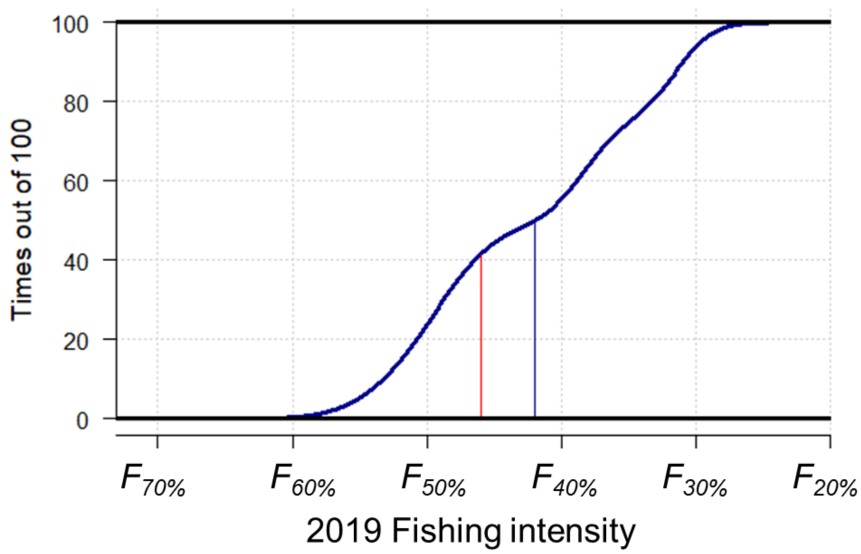

FIGURE 14. Cumulative distribution of the estimated fishing intensity (based on the Spawning Potential Ratio) estimated to have occurred in 2019. Curve represents the estimated probability that the fishing intensity is less than or equal to the value on the x-axis. Vertical lines indicates the reference ($F_{46\%}$; red) and the median value ($F_{42\%}$; blue).

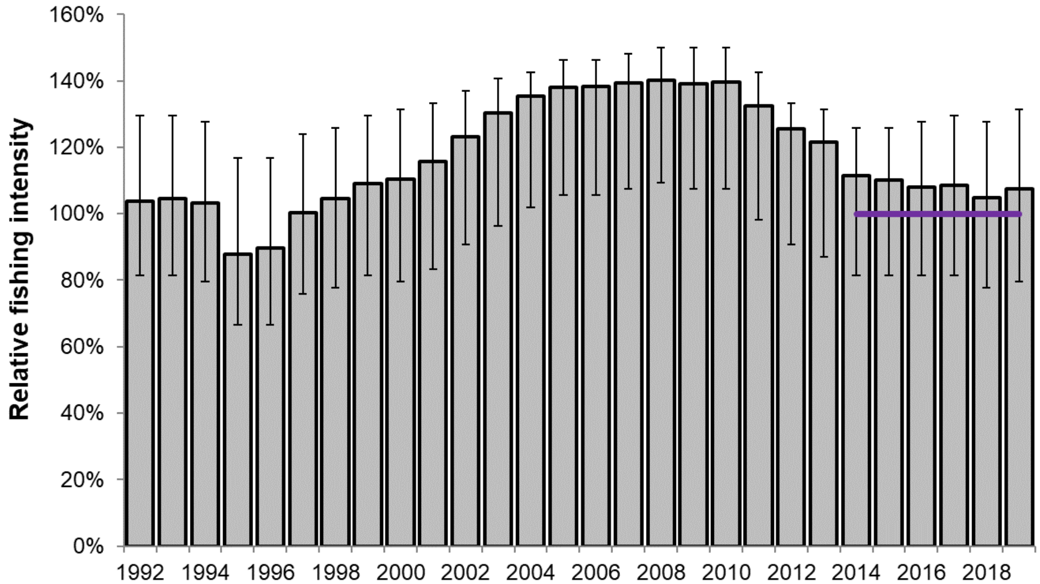

FIGURE 15. Recent estimated fishing intensity (1992-2019; based on the Spawning Potential Ratio) relative to the SPR=46% reference level (horizontal line). Vertical lines indicate approximate credible intervals from the stock assessment ensemble.

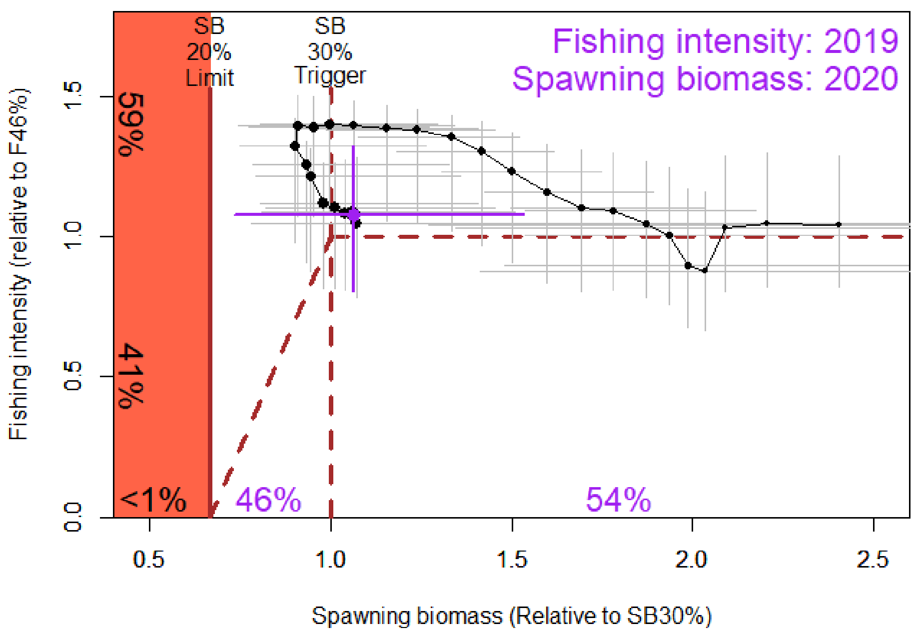

FIGURE 16. Phase plot showing the time-series (1992-2020) of estimated spawning biomass and fishing intensity relative to the reference points specified in the IPHC's interim management procedure. Dashed lines indicate the $F_{46\%}$ (horizontal) reference fishing intensity, with linear reduction below the $SB_{30\%}$ (vertical) trigger, the red area indicates relative spawning biomass levels below the $SB_{20\%}$ limit. Each year of the time series is denoted by a solid point (credible intervals by horizontal and vertical whiskers), with the relative fishing intensity in 2019 and spawning biomass at the beginning of 2020 shown as the largest point (purple). Percentages along the y-axis indicate the probability of being above and below $F_{46\%}$ in 2019; percentages on

the x-axis the probabilities of being below $SB_{20\%}$, between $SB_{20\%}$ and $SB_{30\%}$ and above $SB_{30\%}$ at the beginning of 2020.

Long time-series models

The two long time-series models provided different perceptions of current vs. historical stock sizes (Figure 17). The two long time-series models (coastwide and areas-as-fleets) show different results when comparing the current stock size to that estimated at the historical low in the 1970s. The AAF model estimates that recent stock sizes are below those levels, and the coastwide model above. Relatively large differences among models reflect both the uncertainty in historical dynamics as well as the importance of spatial patterns in the data and population processes, for which all of the models represent only simple approximations. Recent differences are likely attributable to the separation of signals from each region (particularly Region 2, with the longest time-series of data), and allowance for different properties in each region's fishery and survey. Historical differences appear to be due to the differing assumptions regarding connectivity between Regions 2-3 and Regions 4-4B during the early part of the 1900s when there are no data available from Regions 4-4B (Stewart and Martell 2016).

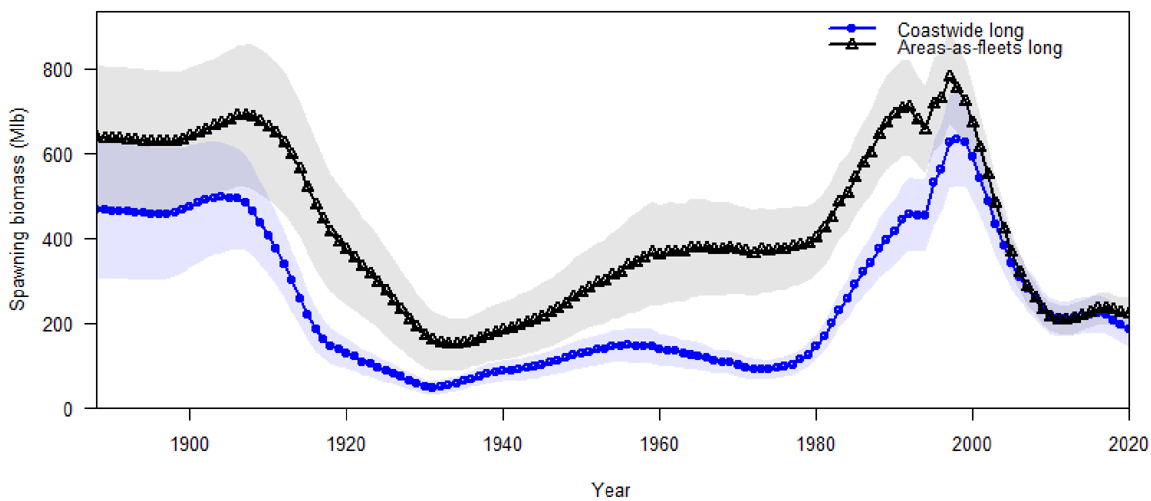

FIGURE 17. Spawning biomass estimates from the two long time-series models. Shaded region indicates the approximate 95% within-model credible interval. The red (black) series is the Areas-As-Fleets model and the blue (lower) series is the coastwide model.

MAJOR SOURCES OF UNCERTAINTY

This stock assessment includes uncertainty associated with estimation of model parameters, treatment of the data sources (e.g. short and long time-series), natural mortality (fixed vs. estimated), approach to spatial structure in the data, and other differences among the models included in the ensemble. Although this is an improvement over the use of a single assessment model, there are important sources of uncertainty that are not included.

The 2019 assessment utilizes two years (2017-18) of sex-ratio information from the directed commercial fishery landings. However, uncertainty in historical ratios, and the degree of variability likely present in those and future fisheries remains unknown. Additional years of data are likely to further inform selectivity parameters and cumulatively reduce uncertainty in stock size in the future. The treatment of spatial dynamics and movement rates among Biological Regions, which are represented via the coastwide and AAF approaches, has large implications for the current stock trend, as evidenced by the different results among the four models comprising the stock assessment ensemble. Further, movement rates for adult and younger Pacific halibut (roughly ages 2-6, which were not well-represented in the PIT-tagging study), particularly to and from Biological Region 4 (and especially to and from the Eastern Bering Sea), are important and uncertain components in understanding and delineating between the distribution of recruitment among biological Regions, and other factors influencing stock distribution and productivity. This assessment also does not include mortality, trends or explicit demographic linkages with Russian waters, although such linkages may be increasingly important as warming waters in the Bering Sea allow for potentially important exchange across the international border. Ongoing research to better understand the stock structure within the Convention Area as well as connectivity to Western North Pacific waters is ongoing. These investigations are particularly important for understanding the dynamics in IPHC Regulatory Area 4B, which is potentially the most demographically isolated of the eight Areas.

Additional important contributors to assessment uncertainty (and potential bias) include factors influencing recruitment, size-at-age, and some estimated components of the fishery removals. The link between Pacific halibut recruitment strengths and environmental conditions remains poorly understood, and although correlation with the Pacific Decadal Oscillation is currently useful, it may not remain so in the future. Therefore, recruitment variability remains a substantial source of uncertainty in current stock estimates due to the lack of mechanistic understanding and the lag between birth year and direct observation in the fishery and survey data (6-10 years). Reduced size-at-age relative to levels observed in the 1970s has been the most important driver of recent decade’s stock trends, but its cause also remains unknown. Like most stock assessments, mortality estimates are assumed to be accurate. Therefore, uncertainty due to discard mortality estimation (observer sampling and representativeness), discard mortality rates, and any other unreported sources of removals in either directed or non-directed fisheries (e.g., whale depredation) could create bias in this assessment.

Maturation schedules are currently under renewed investigation by the IPHC. Currently used historical values are based on visual field assessments, and the simple assumption that fecundity is proportional to spawning biomass and that Pacific halibut do not experience appreciable skip-spawning (physiologically mature fish which do not actually spawn due to environmental or other conditions). To the degree that maturity, fecundity or skip spawning may be temporally variable, the current approach could result in bias in the stock assessment trends and reference points. New information will be incorporated as it becomes available; however, it may take years to better understand these biological processes including the spatial and temporal variability inherent in them.

This stock assessment contains a broad representation of uncertainty in stock levels when compared to analyses for many other species. This is due to the inclusion of both within-model (parameter or estimation uncertainty) and among-model (structural) uncertainty. The distributions for spawning biomass and fishing intensity estimated at the beginning of 2020 and for 2019 reflects this, such that the small differences between the estimate from this and recent assessments are not statistically significant.

Since 2012, natural mortality has been an important source of uncertainty that is included in the stock assessment. In 2012, three fixed levels were used to bracket the plausible range of values. In 2013, the three models contributing to the ensemble included both fixed and estimated values of natural mortality. In the current ensemble, the models again span both fixed (0.15/year for female Pacific halibut) and estimated values. The female value estimated in the long AAF model (0.18) differs substantially from the value estimated in the coastwide model (0.22), and both estimates are highly correlated to the relative selectivity of males and females, which was updated based on the model structure improvements associated with fitting the newly available sex-ratio information from the directed commercial fishery (Stewart and Hicks 2019a). This discrepancy contributes to the difference in scale and productivity for the two models, but is not easily reconciled at present. Although this uncertainty is directly incorporated into the ensemble results, uncertainty in natural mortality in the two short models is not and remains an avenue for future investigation.

Due to the many remaining uncertainties in Pacific halibut biology and population dynamics, a high degree of uncertainty in both stock scale and trend will continue to be an integral part of an annual management process. Potential solutions include management procedures that utilize multi-year management approaches, which are being tested with the MSE framework.

SENSITIVITY AND RETROSPECTIVE ANALYSES

A wide range of sensitivity analyses were conducted during the development of the 2015 and 2019 stock assessments (Stewart and Hicks 2019a; Stewart and Martell 2016). These efforts form the primary basis for the identification of important sources of uncertainty outlined above. The most important contributors to estimates of both population trend and scale included: the sex-ratio of the commercial catch (now improved but still based on only two years of direct data, the treatment of historical selectivity in the long time-series models, and natural mortality. Several sensitivity analyses were investigated in the 2017 and 2018 stock assessment in order to update and illustrate their importance, particularly with regard to the IPHC's research program (Stewart and Hicks 2018a). Those sensitivities included trends in spawning output (due to skip spawning or changes in maturity schedules), sex ratio of the commercial landings, and the effects of unobserved mortality of spawning biomass scale and trends. The results of those analyses illustrated the importance of ongoing research into factors influencing reproductive biology and success for Pacific halibut, the genetic analysis of commercial sex-ratios at age as well as whale depredation and discard mortality rates.

For this year's stock assessment the focus of sensitivity analyses was in understanding the transition from the previous assessment (the bridging analysis presented above) as well as effects of the newly available sex-ratio data. Only one of the four assessment models included in the ensemble (the short coastwide model) currently allows for time-vary male selectivity at the oldest ages. In previous assessments the temporal change in this parameter has been weakly informed by the sexes-aggregated age data and the assumption of female natural mortality (0.15). Investigation of the estimation of these deviation parameters has indicated that they are imprecisely estimated, and several of the earliest deviations were fixed for the 2019 stock assessment as the introduction of the sex-ratio data caused them to drop to the lower biological bound. In order to determine how important the time-varying component of male selectivity is to estimated stock dynamics, a sensitivity analysis was conducted where male fishery selectivity for the oldest ages was estimated but time-invariant (as in the other three models comprising the ensemble). Results showed only small differences in spawning biomass and recruitment estimates (Figure 18). For future model stability, it may be preferable to investigate fixing historical deviations. Even though the sex-ratio data from the commercial fishery in 2017 and 2018 were very similar, it seems likely that additional temporal variability in future male selectivity may be necessary and will be evaluated as new data arise to inform the changes.

Unlike the 2018 stock assessment, the 2011 and 2012 cohorts are now informed by two years of data, including the first observations from the directed commercial fishery. The large sensitivity to the terminal years data in the 2018 assessment with regard to the strength of the 2011 and 2012 cohorts is therefore no longer a major source of uncertainty and difference among the four models comprising the ensemble. However, it is possible that the strength of these year classes may be significantly updated as additional data accrue and the precision of the estimates increases in future assessments.

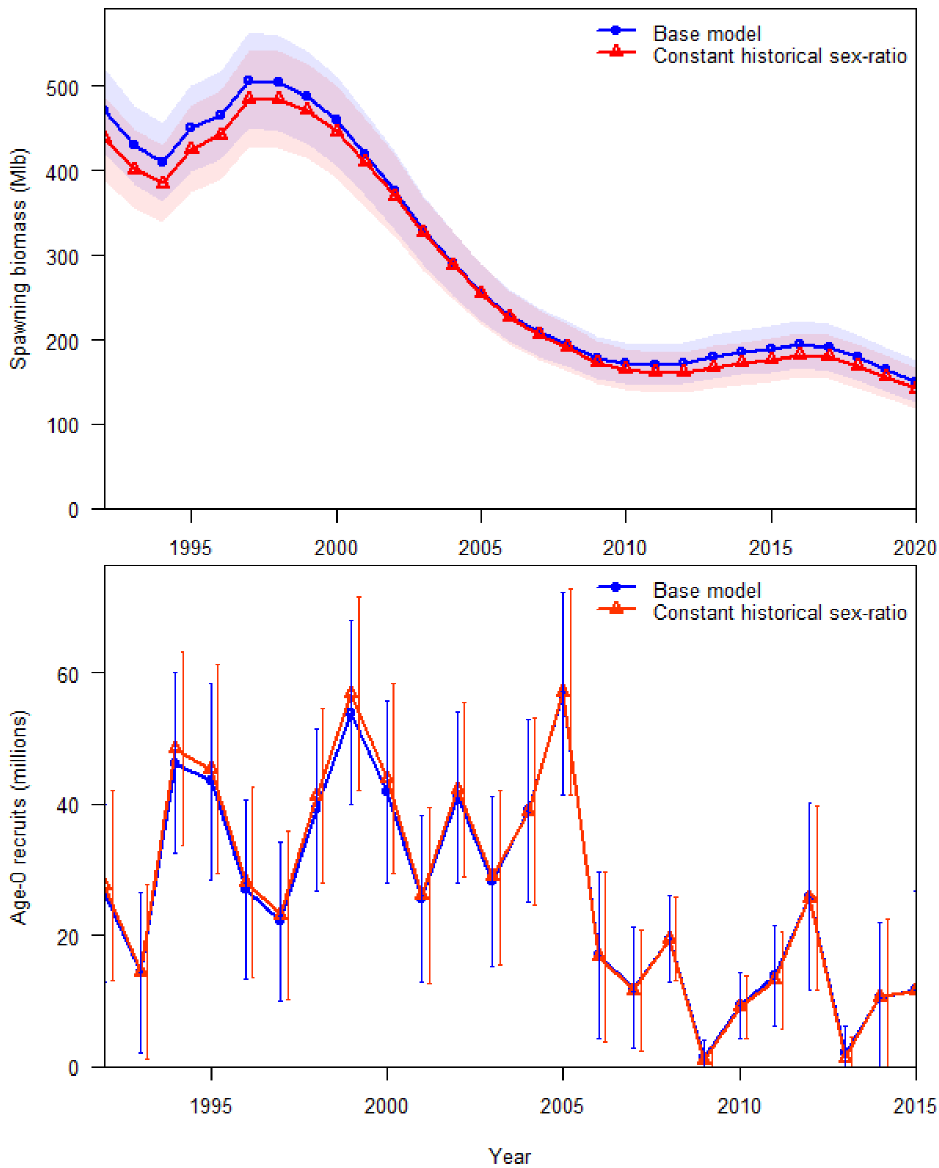

FIGURE 18. Spawning biomass (top panel) and recruitment (bottom panel) estimates from a sensitivity analysis of time-invariant male fishery selectivity at the oldest ages. Shaded regions and vertical whiskers indicate approximate 95% within-model credible intervals.

During the development of the 2019 stock assessment multiple retrospective analyses were performed for each of the individual models contributing to the assessment. This exercise consists of sequentially removing the terminal year's data and rerunning the assessment model for a total of five iterations (five years of data removed from the models) in order to investigate how that information changed the time-series estimates. The strongest effects were identified for the years 2017-2018 which contained the newly available sex-ratio data from the commercial fishery. The structural improvements to the current models are largely incompatible with estimation in the absence of these data, therefore informative retrospective analysis on the current models will be most informative after two or more years of additional sex ratio data become available. This may be included in the 2020 assessment if both 2019 and 2020 genetic assays can be completed, or the 2021 assessment.

FORECASTS AND DECISION TABLE

Stock projections were conducted using the integrated results from the stock assessment ensemble in tandem with summaries of the 2019 directed fisheries and other sources of mortality. The harvest decision table (Table 3) provides a comparison of the relative risk (in times out of 100), based on a range of stock and fishery metrics (rows), against an array of alternative harvest levels for 2020 (columns). This table differs from similarly reported metrics from the MSE in that it represents a tactical decision-making tool, reflecting the best estimates of trends and harvest levels for the next one to three years. In contrast, the risk metrics reported as part of the MSE represent strategic information about the behavior of the Pacific halibut stock over a wide range of biological and environmental conditions. Thus, the two sets of results are complementary, informing the current decision for 2020 harvest levels (assessment decision table) and informing the strategic management procedure choices most likely to optimize stock and fishery objectives (MSE metrics).

The harvest decision table rows are divided into four sections. The block of rows entitled “Stock Trend” provides for evaluation of the risks to short-term trend in spawning biomass, independent of all harvest policy calculations. The remaining rows portray risks relative to the spawning biomass reference points (“Stock Status”) and fishery performance (trend and status) relative to the approach identified in the interim management procedure. The alternatives (columns) provided include several coarsely spaced levels of mortality intended for evaluation of stock dynamics including:

- No mortality (useful to evaluate the stock trend due solely to population processes),
- A 10 million pound (~4,500 t) 2020 Total Constant Exploitation Yield (TCEY[4])
- A 50 million pound (~22,700 t) 2020 TCEY
- A 60 million pound (~27,200 t) 2020 TCEY
- The mortality at which there is a 50% chance that the spawning biomass will be smaller in three years than in 2020 (“*3-year surplus*”)
- The mortality consistent with the “Reference” SPR ($F_{46\%}$) level.
- The mortality consistent with repeating the TCEYs set for 2019 (*“status quo”*).

A grid of alternative TCEY values corresponding to SPR values from 40% to 58% is also provided. For each row of the decision table, the mortality (including all sizes and sources), the

[4] The TCEY corresponds approximately to all mortality of Pacific halibut, except non-directed discard mortality of fish less than 26 inches (66 cm) in length.

coastwide TCEY and the associated level of fishing intensity projected for 2020 (median value with the 95% credible interval below) are reported.

TABLE 3. Harvest decision table for 2020. 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.

2020 Alternative					3-Year Surplus				*Reference SPR=46%*						Status quo		
		Total mortality (M lb)	0.0	11.6	20.0	23.6	27.6	32.3	33.5	34.6	35.7	36.8	37.8	38.9	40.2	61.6	
		TCEY (M lb)	0.0	10.0	18.4	22.0	26.0	30.7	31.9	33.0	34.1	35.2	36.2	37.3	38.6	60.0	
		2020 fishing intensity	$F_{100\%}$	$F_{78\%}$	$F_{63\%}$	$F_{58\%}$	$F_{53\%}$	$F_{47\%}$	$F_{46\%}$	$F_{45\%}$	$F_{44\%}$	$F_{43\%}$	$F_{42\%}$	$F_{41\%}$	$F_{40\%}$	$F_{27\%}$	
		Fishing intensity interval	--	59-87%	44-75%	39-71%	35-67%	31-62%	30-61%	29-60%	28-59%	28-58%	27-57%	26-56%	25-56%	17-43%	
Stock Trend (spawning biomass)	In 2021	Is less than 2020	1	29	61	71	79	87	89	91	93	94	95	96	97	>99	a
		Is 5% less than 2020	<1	<1	11	23	30	42	46	50	54	58	61	64	67	98	b
	In 2022	Is less than 2020	<1	16	50	60	68	77	79	81	83	85	87	89	90	>99	c
		Is 5% less than 2020	<1	1	23	33	45	59	61	64	66	68	69	71	74	99	d
	In 2023	Is less than 2020	1	22	50	58	65	73	75	77	79	81	83	85	87	>99	e
		Is 5% less than 2020	<1	6	33	43	53	62	64	66	67	69	71	73	75	99	f
Stock Status (Spawning biomass)	In 2021	Is less than 30%	35	39	43	44	46	47	48	48	48	48	48	49	49	51	g
		Is less than 20%	<1	<1	<1	<1	<1	1	1	1	2	2	2	3	3	16	h
	In 2022	Is less than 30%	26	31	40	43	46	48	48	49	49	49	49	50	50	54	i
		Is less than 20%	<1	<1	<1	1	2	6	7	8	9	11	12	14	15	27	j
	In 2023	Is less than 30%	18	27	37	41	45	48	49	49	49	49	50	50	50	60	k
		Is less than 20%	<1	<1	<1	2	6	13	15	17	18	20	21	22	23	40	l
Fishery Trend (TCEY)	In 2021	Is less than 2020	0	<1	11	24	36	50	51	52	54	57	59	63	67	>99	m
		Is 10% less than 2020	0	<1	1	12	25	40	44	46	48	50	51	52	53	>99	n
	In 2022	Is less than 2020	0	<1	11	25	39	50	51	52	54	56	59	62	66	>99	o
		Is 10% less than 2020	0	<1	2	14	27	43	46	48	49	50	51	52	54	>99	p
	In 2023	Is less than 2020	0	<1	13	27	41	50	51	52	54	56	58	61	65	>99	q
		Is 10% less than 2020	0	<1	4	16	30	45	47	48	49	50	51	52	54	>99	r
Fishery Status (Fishing intensity)	In 2020	Is above $F_{46\%}$	0	<1	7	22	31	48	50	51	53	55	57	60	64	>99	s

The stock is projected to decrease with at least a 51% chance over the period from 2021-23 for all TCEYs greater than the “3-year surplus” of 18.4 million pounds (~8,350 t), corresponding to a projected SPR of 63% (credible interval 44-75%; Table 3, Figure 19). At the reference level (a projected SPR of 46%) the probability of spawning biomass decline to 2021 is 89%, decreasing to 75% in three years, as the 2011 and 2012 cohorts mature. At the *status quo* TCEYs (38.61 million lb, (~17,500 t), the probability of spawning biomass declines is 97 and 87% for one and three years respectively. The one-year risk of the stock dropping below $SB_{30\%}$ ranges from 43% (at the 3-year surplus level) to 49% at the *status quo* TCEYs. Over three years these probabilities range from 37% to 50% depending on the level of mortality.

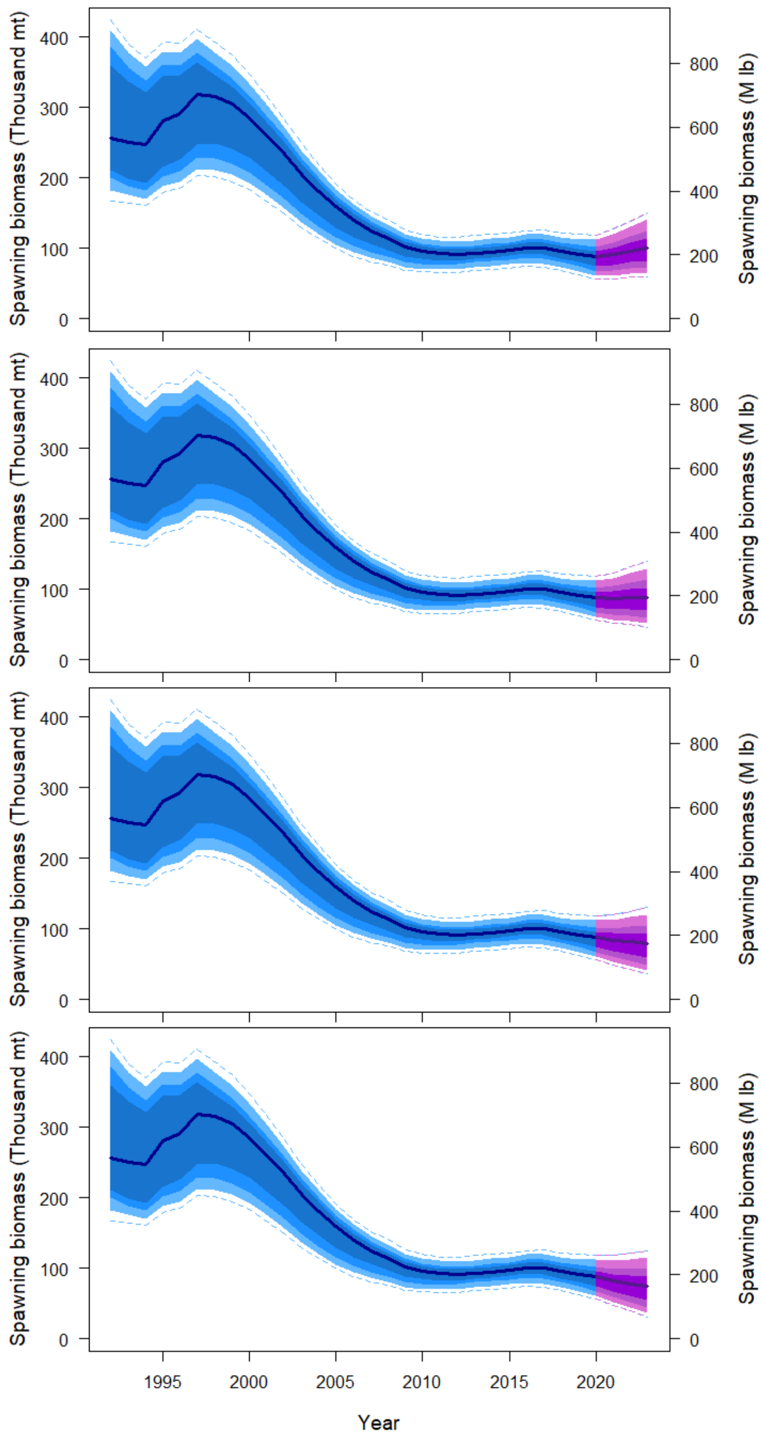

FIGURE 19. Three-year projections of stock trend under alternative levels of mortality: no fishing mortality (upper panel), the 3-year surplus (18.4 million pounds, ~8,350 t; second panel), the TCEY projected for the IPHC's interim management procedure (31.9 million pounds, 14,500 t; third panel) and a TCEY of 38.61 million pounds (~17,500 t, the *status quo* TCEYs from 2019; lower panel).

SELECT COMMENTS FROM THE EXTERNAL PEER REVIEW

Below is a list of select topics from the external independent peer review. These topics were discussed during SRB015, in some cases included in the 2019 stock assessment, and/or prioritized for future research.

Pages 7 and 22 suggest investigation of the sensitivity of the stock assessment (and MSE) to whale depredation as a potential component for the 2019 assessment (and MSE). Most recently, sensitivity to both constant and trending unobserved mortality was investigated in the 2017 and 2018 assessments, and this topic will be revisited in the future. Data on whale interactions from commercial fishery logbooks, which has been collected since 2017, will likely provide the basis for future analyses.

Pages 7 and 21 identify the high priority for the sex ratio of the 2018 commercial Pacific halibut fishery landings to supplement the 2017 used in the preliminary assessment. These data were available and included in this 2019 stock assessment. The potential for extracting sex ratio information from historical samples (prior to 2017) is also being explored.

Page 9 notes the need for documentation for the revised Fishery Independent Setline Survey (FISS) whale depredation criteria. These criteria were developed for and applied to the 2018 FISS. The IPHC's website provides tools to explore these data directly, including which historical stations were retrospectively determined to be ineffective and the specific marine mammal responsible for this determination (https://www.iphc.int/data/fiss-performance).

Page 10 discusses the relative importance of weight-at-age for 3-year projections, noting that it may be difficult to identify factors/processes leading to changes in weight-at-age to the degree that they become predictable. This and other specific research needs were included in a general discussion of research priorities for SRB015.

Pages 11 and 16 suggest reporting the tuning or data weighting applied to each data source in each model over time (Table 11 in IPHC-2019-SRB014-07). SRB015 recommended updating the data weighting for this assessment (which was completed) and providing further evaluation of alternative approaches for SRB016.

Page 12 notes that if additional models with alternative values for steepness are included in the ensemble (particularly for the coastwide long model), the weights should reflect that they are nested and not independent additions. SRB015 reviewed additional investigation of steepness and concluded that including it in the 2019 ensemble was not warranted, but requested further analyses for SRB016.

Pages 12 and 20 identify the importance of connectivity between IPHC Convention waters and those of the western Pacific (i.e. Russia). This is an important research recommendation that was inadvertently omitted from the preliminary assessment document, but is discussed in this assessment.

Page 13 suggests the potential benefits of re-developing the individual Pacific halibut models in an alternative software, perhaps coded specifically for Pacific halibut and able to utilize random effects. This suggestion raises an important consideration of the trade-offs between using a generalized stock assessment platform (in this case stock synthesis) vs. custom-developed code; both have pros and cons. IPHC Secretariat staff attended a workshop on the next generation of generalized stock assessment models (http://capamresearch.org/Next-Gen-SAM). The Secretariat will continue consideration of this topic in consultation with the SRB.

Page 14 provides support for continued development of Bayesian versions of the individual assessment models, particularly if/when the Commission transitions to a management procedure approach with a longer interval between stock assessments.

Page 15 identifies the inclusion of the deconstruction or step-by-step transition in reference point calculations from the 2018 to 2019 stock assessments as a helpful tool for understanding the changes made. This document contains an updated deconstruction, or bridging analysis.

Pages 17 to 18 suggest an informal test of the robustness of management quantities to the sequential exclusion of each individual model in the ensemble (a 'leave one out' approach) to be included in the 2019 stock assessment. This topic was discussed during SRB015 and will be revisited in the future.

Page 18 suggests that at this time it may be beneficial to minimize changes to the ensemble in order to facilitate transition to a management procedure. The Secretariat will minimize changes as possible, particularly in the updated assessments prior to the next full assessment and external review currently scheduled for 2022.

Page 19 identifies several improvements to the presentation of research recommendations, including ranking and denoting those in progress vs planned. These recommendations were discussed as part of a larger review of research priorities during SRB015. See the *Research Priorities* section below.

There are many additional comments specific to research priorities and specific data sets and analyses that will be addressed in future SRB meetings and stock assessment analyses.

RESEARCH PRIORITIES

Research priorities for the stock assessment and related analyses have been consolidated with those for the IPHC's MSE and the Biological Research program. These ranked and categorized priorities will soon be available on the IPHC's website.

ACKNOWLEDGEMENTS

We thank all of the IPHC Secretariat staff for their contributions to data collection, analysis and preparation for the stock assessment. We also thank the staff at the NMFS, DFO, ADFG, WDFW, ODFW, and CDFW for providing the annual information required for this assessment in a timely manner. The SRB, Science Advisors, and Dr. Kevin Stokes (the external peer reviewer) provided critical review and helpful guidance during the 2019 process.

REFERENCES

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., 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*). Can. J. Fish. Aquat. Sci. **56**: 242-252.

IPHC. 2019a. Report of the 14th session of the IPHC Scientific Review Board (SRB014). Seattle, Washington, U.S.A., 26-28 June 2019. IPHC-2019-SRB014-R. 16 p.

IPHC. 2019b. Report of the 15th session of the IPHC Scientific Review Board (SRB015). Seattle, Washington, U.S.A., 24-26 September 20-19. IPHC-2019-SRB015-R, 18 p.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**(0): 86-99. doi:http://dx.doi.org/10.1016/j.fishres.2012.10.012.

Stewart, I. 2017. Pacific halibut stock assessment development for 2017. IPHC-2017-SRB10-07. 14 p.

Stewart, I., and Hicks, A. 2018a. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2017. IPHC-2018-AM094-10. 25 p.

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

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

Stewart, I.J., and Martell, S.J.D. 2014. A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. Fish. Res. **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 J. Mar. Sci. **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 Hicks, A.C. 2017. 4.2 Assessment of the Pacific halibut stock at the end of 2016. IPHC Report of Assessment and Research Activities 2016: 365-394.

Stewart, I.J., and Hicks, A.C. 2018b. Interannual stability from ensemble modelling. Can. J. Fish. Aquat. Sci. **75**: 2109-2113. doi:10.1139/cjfas-2018-0238.

Stewart, I.J., Monnahan, C.C., and Martell, S. 2016. Assessment of the Pacific halibut stock at the end of 2015. IPHC Report of Assessment and Research Activities 2015. p. 188-209.

Stewart, I.J., Martell, S., Webster, R.A., Forrest, R., Ianelli, J., and Leaman, B.M. 2013. Assessment review team meeting, October 24-26, 2012. IPHC Report of Assessment and Research Activities 2012. p. 239-266.

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.

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. Fish. Res. **158**: 15-25. doi:10.1016/j.fishres.2014.01.009.