Modelling updates

PREPARED BY: IPHC SECRETARIAT (I. STEWART AND A. HICKS; 20 MAY 2018)

PURPOSE

To provide the Scientific Review Board (SRB) a summary of anticipated modelling development in support of the 2018 and 2019 stock assessment and harvest policy analyses.

INTRODUCTION

Brief history of SRB stock assessment review

The ensemble approach to modelling Pacific halibut stock dynamics (Stewart and Martell 2015) was initiated during the 2012 stock assessment in direct response to two factors: 1) a recurring retrospective bias in previous models (Stewart and Martell 2014b), and 2) an external performance review of the IPHC process recommending a more clear and transparent delineation of scientific and management aspects of the harvest policy, necessitating both a better characterization of scientific uncertainty and a decision table approach to providing annual catch 'advice.'

At that time, the IPHC did not have an established process for independent peer review, having held only one recent review specific to the stock assessment model in June 2007 (IPHC 2008). Therefore, an *ad hoc* review was conducted during October 2012 (Stewart et al. 2012). At that time, the assessment consisted of three alternative models characterized by three levels of female natural mortality (Stewart et al. 2013). The IPHC's independent Scientific Review Board (https://iphc.int/library/documents/meeting-documents/scientific-review-board-srb) was formally initiated during 2013, with a stock assessment review occurring in October 2013. The stock assessment was extended at that time to include three separate models differing in structure, data use, and programming platform (Stewart and Martell 2014a). In order to reflect the compressed timeline of data availability and the stock assessment process, a second SRB meeting (via conference call) was added to the annual process between the IPHC's Interim and Annual meetings, in December 2013. During 2014, the stock assessment ensemble underwent a full revision of data processing procedures, and another expansion to include the four models used in subsequent years. These models were evaluated by the SRB in June and October of 2014. That year marked the first cycle of what has since become the standard meeting schedule: review and research planning in June, review and final model refinements in October (shifting to late September starting in 2016), and follow-up opportunity for evaluation in December after all data sources have been finalized and initial results presented to the public.

For 2015, full technical documentation of all four models, their data components, the ensemble approach, harvest policy calculations, and projection methods was provided for SRB review during the June meeting (Stewart and Martell 2016). These four models have been subsequently updated, without major structural changes, through the SRB review processes in 2016 and 2017. All incremental changes to data sources have been reported and reviewed, with 'bridging' analyses illustrating assessment model response to each change documented each year.

The 2018 and 2019 stock assessments

This document summarizes ongoing avenues of modelling development, specific changes anticipated for inclusion into updated 2018 models and presentation material, as well as changes planned for evaluation in 2019. A full assessment evaluation, documentation and review (similar to that conducted in 2015) is planned for the 2019 SRB process, in order to capitalize on important new data anticipated to be available (described below). The 2019 process also provides the best opportunity for assessment model revision to integrate with Management Strategy Evaluation efforts shifting from consideration of coastwide to more spatially explicit management procedures, thus necessitating a change to operating models less closely linked to the tactical assessment ensemble.

ONGOING DEVELOPMENT

Model weighting methods

Equal weighting of the four models contributing to the stock assessment ensemble has been maintained since 2015. Each year, evaluation of alternative approaches has not generated any appreciable indication that different weighting is warranted based on the fit of each model to the survey index, the predictive performance of the fit to the survey, or the retrospective behavior. If or when additional models are added to the ensemble or model performance-based weighting approaches suggest differing model weights this topic may need to be revisited. A manuscript reporting alterative weighting methods remains in preparation.

Bayesian integration

Work by Cole Monnahan (PhD Thesis, University of Washington, June 2017) included development of an alternative Markov Chain Monte-Carlo (MCMC) search algorithm implemented in Automatic Differentiation Model Builder (ADMB; Monnahan et al. 2016). ADMB is the underlying code for the stock synthesis model, on which the current Pacific halibut stock assessment models are based, as well as many other stock assessment models around the world. Further work on regularization – adding informative priors and/or tactically reducing complexity to improve estimation performance – continues, using the short coastwide Pacific halibut model as one of a set of illustrative examples. An additional manuscript is in preparation (Cole Monnahan, pers. comm.). This work, in concert with some of the options available in the new stock synthesis version (see section below) may provide avenues for improved efficiency in implementing Pacific halibut models in a fully Bayesian framework.

A previously undocumented inconsistency in the ADMB software when MCMC integration is performed (but not affecting maximum likelihood estimates) was recently discovered, and pending its resolution it may not be advisable to use the posterior distributions from any model (including stock synthesis) that includes “dev_vectors” (https://github.com/admb-project/admb/issues/107). The two long-time-series Pacific halibut models use this feature. Further investigation into Bayesian integration of the Pacific halibut models remains an open avenue for development, as true probability distributions (rather than asymptotic approximations) are desirable for calculating probabilistic management results, and diagnosis of posterior convergence can be a highly informative tool for improving maximum likelihood estimation as well.

Ensemble stability

A potential and previously unexplored benefit of using a stock assessment ensemble is inter-annual stability in stock assessment results. Stability may be created via two avenues: the inclusion of new alternative models into each year's analysis while still including existing models (rather than through changes to a single base-case model), and from the buffering effect of characterizing the central tendency (or distribution) of management quantities with a set of models. This last benefit – temporal stability against the addition of new data – although logically appealing, has not been evaluated in the context of fisheries stock assessment. The IPHC Secretariat has developed a draft manuscript that evaluates the results from the International Pacific Halibut Commission's stock assessment ensemble as an example of the stability created by multiple models, and provides a simple simulation and analytical framework to explore general ensemble behavior. Counter-intuitively, we found little stability benefit in the IPHC's current ensemble due to high temporal correlations among individual models as annual data are added. However, we also found that even a small number of models with low among-model correlations could have a substantial stability benefit. We suggest that among-model temporal correlations may be a valuable ensemble diagnostic that warrants consideration by analysts developing ensembles, and also those performing sensitivity testing of single-model assessments.

DEVELOPMENT FOR 2018

Software updates

Recent Pacific halibut stock assessment models have used stock synthesis (Methot Jr and Wetzel 2013) version 3.24u. For the features that are currently included in these models there are no identified bugs that have required updating the IPHC's application to a newer version. However, a substantially revised version of Stock Synthesis (3.30.11 as of 14 May 2018) has now gone through approximately 1.5 years of code development. Several U.S. west coast and Alaska stock assessments have been conducted using the new software. To date, the IPHC has delayed full implementation of the software largely in order to avoid the efficiency costs of development and testing of new and revised features.

During 2018, the IPHC Secretariat performed extensive comparative analyses to the conversion of all major features currently implemented in the four Pacific halibut stock assessment models contributing to the ensemble. Although there are changes to the input and output structure, and many additional options for structural approaches (Methot et al. 2018), all have a backward-compatible analog to earlier versions. Sequential testing of features including the implementation of weight-at-age, the stock-recruitment equations including equilibrium offsets and environmental covariates, time-varying catchability, error distributions, and others revealed no issues. However, several key features used in the Pacific halibut models are currently not fully functional in SS3.30.11 (as of May 2018); these include the option for female selectivity offset to male selectivity (used in defining the selectivity of the commercial fishery discards), and the age-based double normal selectivity (used for all fleets in the halibut model). These issues have been reported, and are under development; it is likely that they will be fully functional in time for use during the IPHC's 2018 assessment cycle.

If implementation of all features is completed soon, the 2018 Pacific halibut stock assessment is likely to be able to update to the new software with diminutive change to the results of management quantities. This is essential to provide compatibility with MSE development during the same time period. The IPHC Secretariat will continue to work with the developers of the stock synthesis software, perhaps utilizing the IPHC's pending programming position if necessary, to ensure that a clear transition can be made. If this is not possible during 2018, it will be logical to include any remaining transition in the full analysis planned for 2019, when additional model structural evaluation and changes are already anticipated (see below).

'Replay' analyses

During recent years' annual processes a number of questions have arisen regarding the utility of creating 'replays' of the estimated historical time series of biomass under different management actions. Specifically, one recurrent interest is how the biomass trajectory would be estimated to have evolved under catch limits following exactly the IPHC's harvest strategy policy in each year. A conceptual framework for this type of analysis is represented by the following steps:

1) Begin with the maximum likelihood estimates (MLEs) from each of the models for all parameters.
2) Fix the model parameters at MLEs and substitute an alternative set of removals representing a different management decision for the first year of the 'replay.'
3) Recalculate the harvest strategy policy calculations for the second (and subsequent) years of the replay, and substitute them into the removals in a sequential fashion until the current year is reached.
4) Re-integrate the ensemble time-series under the 'replay' conditions.
5) Compare the 'replay' to the actual estimated ensemble time-series of biomass and the actual removals from each year.

In order to implement this approach, at least several implicit assumptions and caveats would be required: variance calculations would be unavailable (although they could be substituted from the actual estimated time series), spatially generated feedback mechanisms (e.g., changes in productivity due to the region in which catch was taken) are unknown, and would have to be assumed to be unimportant, and the stock-recruit relationship would be ignored.

Details yet to be worked out include the appropriate description of this analysis in terms of the trade-offs between foregone yield and stock status. Discussion of this topic during SRB12 is planned.

Phase plots and status indicators

The IPHC Secretariat introduced a number of new approaches for summarizing stock assessment results with regard to current status and recent trend during the 2017 process. These include a summary table, intended to more closely resemble those produced by other international organizations and those produced domestically in the U.S. and Canada (Table A1, Appendix A). Because the IPHC's harvest policy is currently evolving, and has never included some reference points common in other processes (e.g., an explicit overfishing limit), the choice of metrics is challenging. As the MSE process continues, it will be logical to include reference points and performance metrics developed in that context. In the interim, discussion of metrics and approaches for describing status and trend is ongoing; guidance from the SRB on this topic could be very helpful for the 2018 assessment cycle.

The SRB last discussed the use of a phase plot for Pacific halibut in October 2015. Although complex due to the quantity of information contained in current status, recent trend, and uncertainty associated with each, many processes routinely use a phase plot (sometimes recently called a ‘Kobe plot’) as part of an executive summary of stock assessment and harvest policy results.

The IPHC’s stock assessment ensemble and current harvest policy present several challenges to the generation of a ‘standard’ phase plot:

1) There is no overfishing limit, and so only one horizontal reference line.
2) The reference F_{SPR} is not formally considered a target or limit, so the implications of a level of fishing intensity that exceeds $F_{46\%}$ are unclear (and thus do not lend to unambiguous color-coding).
3) The calculation of reference points (i.e., relative fishing intensity, relative spawning biomass) is not integrated within all four of the stock assessment models. This means that although the variance of each quantity can be approximated (and therefore the variance of the ensemble value), the covariance between the two axes is unknown.
4) The level of uncertainty relative to the range of recent historical values (estimated by all four stock assessment models included in the ensemble) is very large.

An example phase plot, illustrating the adopted catch limits and estimated stock status for 2018 was produced following the methods employed by various other fisheries management bodies (Figure A2, Appendix A). In order to approximate the covariance between axes the average value from the long time series models was applied to the results from each of the four models. This leads to the potential for small differences between the marginal probabilities of exceeding a reference level (e.g., $P(SB<SB_{30\%})$), and the joint probabilities reported in the phase plot. If this approach is deemed to warrant further consideration, additional code development, particularly internal calculation of reference points, variances, and covariances could improve these approximations. Further discussion of this topic during SRB12 is planned.

Web-based projection tools

Under development for the 2018 process is an interactive tool for rapid evaluation of alternative projected catch and catch distribution. In the past, alternative projections were provided to advisory bodies and stakeholders as needed; however, this process was time-consuming and somewhat limiting in the range of options that could be provided. An interactive tool, with Catch Sharing Plans delineating all allocations among fishery components within all Regulatory Areas, a fitted (non-linear) relationship between SPR and total mortality based on a wide grid of previously produced results, as well as projection figures from those results, is anticipated to be posted to the IPHC’s website prior to the 2019 Annual Meeting (AM095).

DEVELOPMENT PROPOSED FOR 2019

Model structure

A full stock assessment analysis presented for the June SRB Meeting in 2019 will allow review of several important structural aspects of the Pacific halibut models identified in previous efforts (Stewart and Martell 2016). Of particular interest are features that will have newly available data (see IPHC-2018-SRB12-06) and/or new options for parameterization in the more recent version of stock synthesis, including:

- Data weighting, including alternative error distributions (e.g., the Dirichlet for compositional data).

- The treatment of constraints on time-varying processes such as selectivity, with the potential for explicitly estimating and including the uncertainty in the degree of temporal variability (sigmas).
- Age-based discarding and discard mortality estimation within the assessment models to better propagate uncertainty associated with these estimates.
- Incorporation of new sex-ratio information from the commercial fishery and perhaps greater estimation of related selectivity parameters (including some new parameterizations) with commensurately improved characterization of uncertainty.
- More control over the modelled timing of the catch and surveys, allowing investigation of the importance of interannual variability.

A discussion of these topics in order to expand this list for evaluation in the 2019 assessment is anticipated for SRB12.

Renewed spatial model development

The modelling with an explicitly spatial framework that was conducted by the IPHC Secretariat through 2016 will be updated and extended to provide a starting point for use as an MSE operating model. At this time, it is not anticipated that a parallel model will be developed for tactical use in the annual management process.

SUMMARY

Model development during 2018 is largely focused on refinement of existing approaches, and preparing for a full assessment and review during 2019 (Table 1). This parallels the approach taken for data sources (see IPHC-2018-SRB12-07).

As has been the practice for all recent stock assessments, any available modelling updates in preparation for the 2018-2019 annual process will be presented to the SRB at the October 2018 meeting (SRB13).

TABLE 1. Summary of model development

Improvement	Rationale	Timeline
Model weighting methods	Weighting approaches are important in determining ensemble results.	Evaluation of alternative methods will be continued; publication of these approaches is planned.
Bayesian integration	Better represents probability distributions for management use.	Ongoing.
Ensemble stability	This is a novel aspect of ensemble application.	Ongoing.
Software updates	Current stock synthesis version still being tested.	Possible inclusion in 2018, pending resolution of a small number of incomplete features.
‘Replay’ analyses	Represents a frequently asked avenue of questioning regarding the performance of recent management.	Further refinement and possible presentation in 2018.
Phase plots and status indicators	Ongoing effort to simplify and make more accessible key assessment results.	Further improvement and use in 2018.
Web-based tools	Capitalizing on the shift toward electronic support material for the IPHC process.	To be added in 2018.
Model structural investigation	New features available in the stock synthesis platform and new data anticipated for 2019 may allow improved structural assumptions.	Anticipated exploration and review in 2019.
Spatial model development	This level of model complexity will be required in order to evaluate some MSE objectives.	Continued development is planned for 2019 in support of the MSE process.

RECOMMENDATION/S

The IPHC secretariat requests that the SRB:

NOTE this document summarizing ongoing, pending and future model development efforts by the IPHC Secretariat.

NOTE any discussion occurring during SRB12, and **RECOMMEND** any conceptual or technical improvements for conducting and reporting ‘replay’ analyses.

NOTE any discussion occurring during SRB12, and **RECOMMEND** any suggestions for simple summary tools applicable to stock assessment estimates of status and trend.

RECOMMEND any specific avenues for model development in preparation for the inclusion of new data and for the full stock assessment documentation and review anticipated for 2019.

RECOMMEND any additional specific research avenues to be prioritized for inclusion in the 2019 stock assessment.

References

IPHC. 2008. International Pacific Halibut Commission stock assessment workshop. IPHC Report of Assessment and Research Activities 2007. p. 99-120.

Methot Jr, 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.

Methot, R.D., Wetzel, C., and Taylor, I. 2018. Stock synthesis user manual Version 3.30.11. NOAA Fisheries. Seattle, WA. USA. April 2, 2018. 206 p.

Monnahan, C.C., Thorson, J.T., Branch, T.A., and O'Hara, R.B. 2016. Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. Methods in Ecology and Evolution **8**(3): 339-348.

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

Stewart, I.J., and Martell, S. 2014a. Assessment of the Pacific halibut stock at the end of 2013. IPHC Report of Assessment and Research Activities 2013. p. 169-196.

Stewart, I.J., Martell, S., Webster, R.A., Forrest, R., Ianelli, J., and Leaman, B.M. 2012. Assessment review team meeting, October 24-26, 2012. 30 p.

Stewart, I.J., and Martell, S.J.D. 2014b. A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. Fish. Res. **158**: 40-49.

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.

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.

Appendices

Appendix A: Stock status and trend summary information, example phase plot of fishing intensity and spawning biomass.

APPENDIX A

Stock status and trend summary information.

TABLE A1. Status summary of Pacific halibut in the IPHC Convention Area at the end of 2017.

Indicators	Values	Trends	Status
Total mortality 2017: Retained mortality 2017: Average mortality 2013-17:	42.44 Mlbs, 19,250 t[1] 35.29 Mlbs, 11,864 t 43.34 Mlbs, 19,659 t	Mortality stable 2014-17	**2017 MORTALITY BELOW 100-YEAR AVERAGE**
SPR_{2017}: P(SPR<46%): P(SPR<limit):	40% (29-58%)[2] 75% Limit not specified	Fishing intensity increased from 2016 to 2017	**FISHING INTENSITY HIGHER THAN REFERENCE LEVEL**[3]
SB_{2018} (Mlb): SB_{2018}/SB_0: $P(SB_{2018}<SB_{30})$: $P(SB_{2018}<SB_{20})$:	202 Mlbs (148–256) 40% (26-60%) 6% <1%	SB decreased from 2017 to 2018	**NOT OVERFISHED**[4]
O32 stock distribution: All stock distribution:	See Table A2 and Figure A1.	Distribution stable 2013-17	**REGION 2 ABOVE, REGION 3 BELOW HISTORICAL VALUES**

[1] Weights in this document are reported as 'net' weights, head and guts removed; this is approximately 75% of the round (wet) weight).
[2] Ranges denote approximate 95% confidence intervals from the stock assessment ensemble.
[3] Status determined relative to the IPHC's interim reference Spawning Potential Ratio level of 46%.
[4] Status determined relative to the IPHC's interim management procedure biomass limit of $SB_{20\%}$.

TABLE A2. Recent regional stock distribution estimates based on modelling of the fishery-independent setline survey data.

	O32 stock distribution				All sizes stock distribution			
Year	Region 2 (2A, 2B, 2C)	Region 3 (3A, 3B)	Region 4 (4A, 4CDE)	Region 4B	Region 2 (2A, 2B, 2C)	Region 3 (3A, 3B)	Region 4 (4A, 4CDE)	Region 4B
2013	29.6%	45.9%	18.7%	5.8%	25.4%	50.1%	19.6%	4.9%
2014	28.8%	46.5%	19.8%	4.9%	24.2%	52.8%	19.1%	4.0%
2015	30.4%	44.2%	20.5%	4.9%	25.7%	51.4%	18.9%	4.0%
2016	30.0%	46.8%	18.6%	4.5%	25.9%	52.8%	17.4%	3.9%
2017	29.7%	45.6%	20.0%	4.8%	25.9%	50.7%	19.2%	4.2%

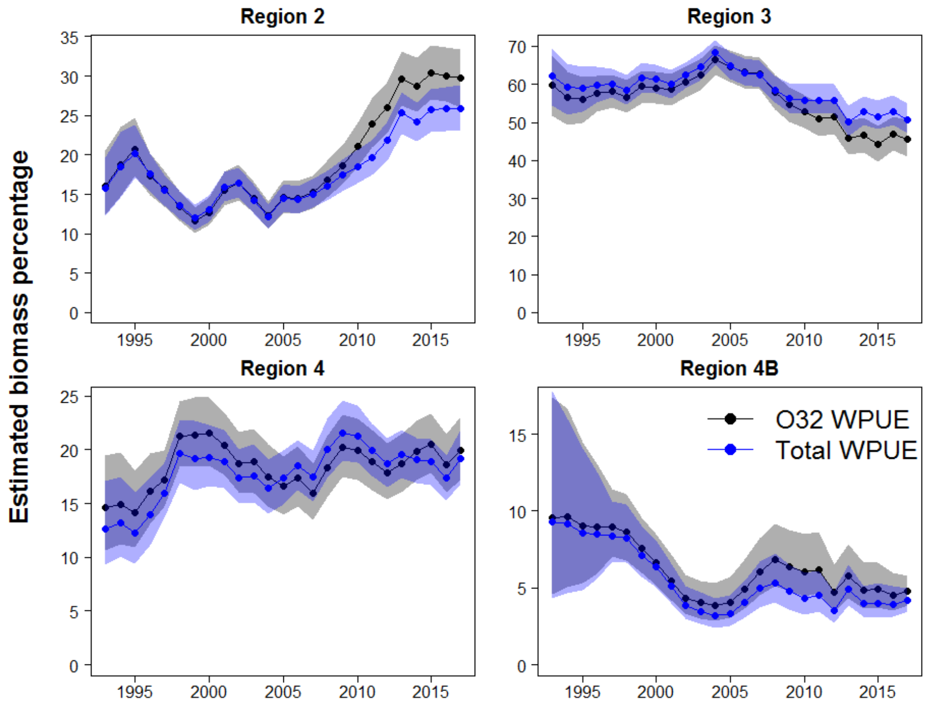

FIGURE A1. Estimated stock distribution (1993-2017) based on setline survey catch of O32 (black series) and all sizes (blue series) of Pacific halibut. Shaded zones indicate approximate 95% credibility intervals.

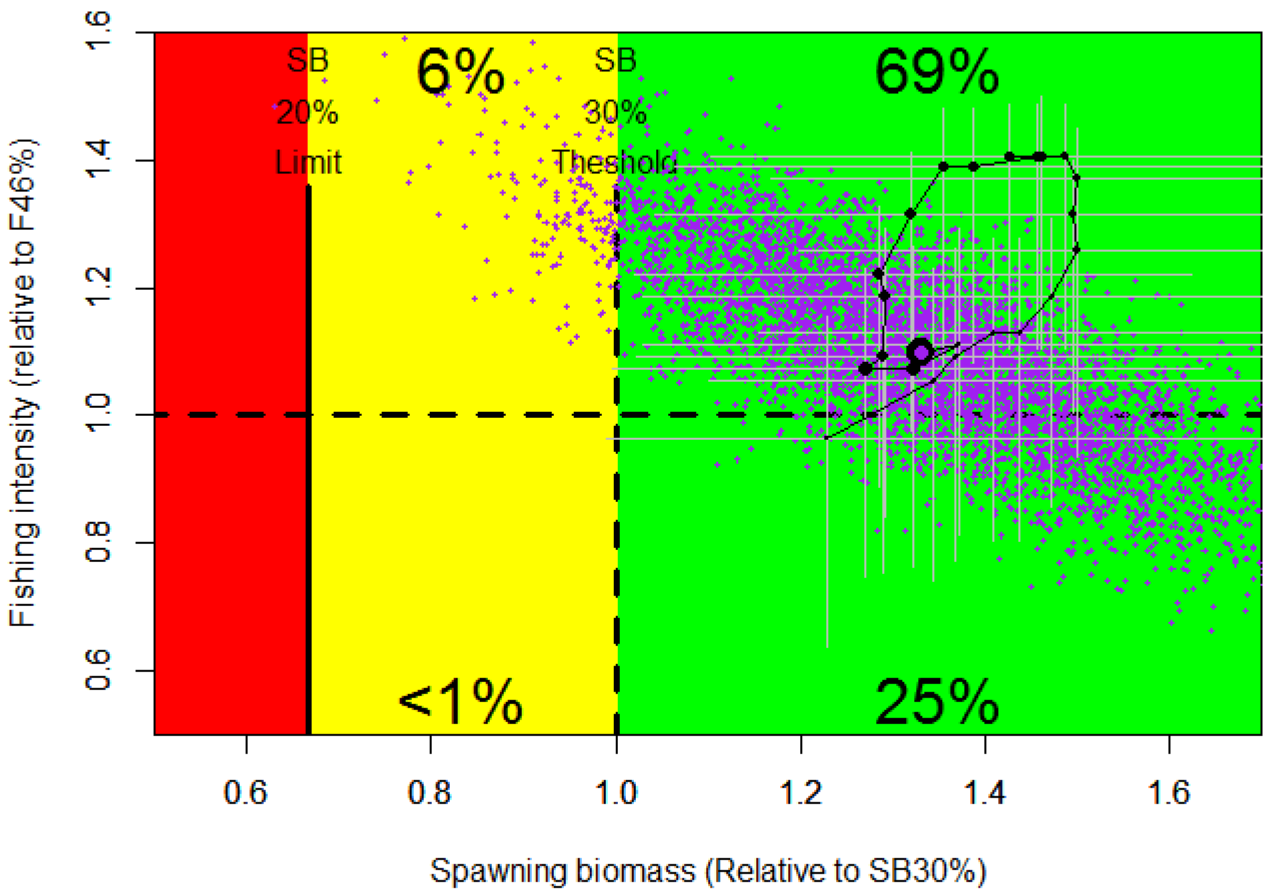

Figure A2. Example phase plot based on data from the 2017 stock assessment: Time-series of relative spawning biomass (spawning biomass divided by SB30%) and relative fishing intensity (1-SPR/1-46%). Horizontal dashed line indicates the reference SPR = 46%; vertical solid line indicates the SB20% biomass limit and vertical dashed line indicates the SB30% biomass threshold. Black points indicate the relative status in each year from 1996 through 2018 (largest point with purple center). Light lines indicate uncertainty in annual status through 2017; purple points indicate the probability distribution for the biomass and adopted catch limit in 2018. Percentages indicate the relative probability of the 2018 status falling into each quadrant.